3D-2001 Program

(and links to abstracts)

Tuesday 30 October 2001

08:30 - 10:15

Session 1: Spiral CT I

Moderated by Pierre Grangeat and Larry Zeng

1-1: Theoretically Exact FBP-Type Inversion Algorithm for Spiral CT	A Katsevich
1-2: Simple Quasi-Exact Filtered Backprojection Algorithms for Long-Object Problem in Helical Cone-Beam Tomography	H Kudo, F Noo, and M Defrise
1-3: Practical Helical Cone Beam Algorithm for the Long Object Problem	J Hu, R Johnson, and C Dawson

10:45 - 11:55

Session 2: Fast Reconstruction

Moderated by Grant T Gullberg and Jinyi Qi

2-1: Multifrequential Algorithm for Fast 3D Reconstruction	T Rodet, P Grangeat and L Desbat
2-2: Fully 3D Iterative Reconstruction of Planogram Data	PE Kinahan, D Brasse, M Defrise, R Clackdoyle, C Comtat, C Michel, and X Liu

13:30 - 15:15

Session 3: Posters I

3-1: Exact Inversion of the Exponential X-ray Transform for RSH SPECT	J-M Wagner, F Noo, and R Clackdoyle
3-2: Optimization of Penalized Block-Iterative Algorithms for Ga-67 Tumor Detection	HC Gifford, CL Byrne, MV Narayanan, and MA King
3-3: Analytical 3D Approach to Simultaneous Compensation for Photon Attenuation and Collimator Response in Quantitative Fan-Beam Collimated Brain SPECT	T Li, J You, and Z Liang
3-4: Investigation of the Dynamic SPECT (dSPECT) Method for Teboroxime Using a 4-D Kinetic Thorax Model dMCAT	A Celler, S Blinder, D Noll, T Tyler, F Duclercq, and R Harrop
3-5: Fast Computation of Statistical Uncertainty for Spatiotemporal Distributions Estimated Directly from Dynamic Cone Beam SPECT Projections	BW Reutter, GT Gullberg, and RH Huesman

3-6: A Rebinning Technique for 3D Reconstruction of Compton Camera Data	J Li, JD Valentine, JN Aarsvold, and M Khamzin
3-7: A Fully 3D Maximum Likelihood Estimator Algorithm with Attenuation and Scatter Correction Adapted to a Limited Angle Positron Camera	F Pönisch, W Enghardt and K Lauckner
3-8: Model-Based Normalization for Iterative 3D PET Image Reconstruction	B Bai, Q Li, E Asma, YC Tai, A Chatziioannou, and RM Leahy
3-9: Iterative Reconstruction Methods for High-Throughput PET Scanners	J Hamill and T Bruckbauer
3-10: A Modified Bootstrap for Importance Sampled Data	RL Harrison, J-S Kim, and TK Lewellen
3-11: A Sufficient Condition for Spiral Cone Beam Long Object Imaging via Backprojection	KC Tam
3-12: A Simple Algorithm for Increased Helical Pitch in Cone-Beam CT	MD Silver, K Taguchi, and I Hein
3-13: Practical Approach to Misalignment Correction in a Single-Circle Orbit Cone-Beam Tomography	M Karolczak, U Taubenreuther, K Engelke, A Lutz, S Schaller, and W Kalender

15:45 - 17:30

Session 4: Statistical Reconstruction

Moderated by Ronald H Huesman and Christian Michel

4-1: A Smoothing Prior with Embedded Positivity Constraint for Tomographic Reconstruction	I-T Hsiao, A Rangarajan, Y Xing, and G Gindi
4-2: 3D List-Mode Reconstruction for SPECT	L Bouwens, R Van de Walle, HC Gifford, I Lemahieu and RA Dierckx
4-3: Content-Adaptive Mesh Modeling for Tomographic Image Reconstruction	JG Brankov, Y Yang, and MN Wernick

Wednesday 31 October 2001

08:30 - 10:15

Session 5: Spiral CT II

Moderated by Jinyi Qi and David W Townsend

5-1: Segmented Multiple Plane Reconstruction - A Novel Approximate Reconstruction Scheme for Multislice Spiral CT	K Stierstorfer, T Flohr, and H Bruder
5-2: A Combination of Rebinning and Exact Reconstruction Algorithms for Helical Cone-Beam CT	M Defrise, F Noo, and H Kudo
5-3: The PI-FAST Method for Approximate Helical Cone-Beam Reconstruction	H Turbell and P-E Danielsson

10:45 - 11:55

Session 6: Keynote Address

6-1: Why Do Patients and Their Care-Givers Need 3D and 4D Imaging?	TF Budinger

13:30 - 15:15

Session 7: Posters II

7-1: Image Reconstruction Algorithm for a SPECT System with a Convergent Rotating Slat Collimator	GL Zeng and D Gagnon
7-2: Image Reconstruction in 3D Short-Scan SPECT	X Pan, C-M Kao, C Metz, and A Kiselev
7-3: 4D Processing of Gated SPECT Images Using Deformable Mesh Modeling	JG Brankov, Y Yang, and MN Wernick
7-4: A Fully 4D Expectation Maximization Algorithm Using Gaussian Diffusion Based Detector Response for Slow Camera Rotation Dynamic SPECT	TH Farncombe, MA King, AM Celler, and S Blinder
7-5: Transfer Function Restoration in 3D Electron Microscopy via Iterative Data Refinement	COS Sorzano, JJ Fernández, R Marabini, GT Herman, Y Censor and JM Carazo
7-6: Implementation of Continuous 3D Whole Body PET Scanning Using On-the-fly Fourier Rebinning	K Kitamura, K Tanaka, and T Sato
7-7: Scatter Correction for Positron Emission Mammography	J Qi and RH Huesman
7-8: Internet-Based 3D PET Image Reconstruction Using A Beowulf PC Cluster	D Shattuck, J Rapela, E Asma, A Chatziioannou, J Qi, and R Leahy
7-9: Object Dependency of Resolution and Convergence Rate in OSEM with Filtering	S Mustafovic, K Thielemans, D Hogg and P Bloomfield
7-10: Performance Evaluation of Exact Cone-Beam Algorithms for the Long-Object Problem in Spiral Computed Tomography	K Sourbelle, H Kudo, G Lauritsch, KC Tam, M Defrise and F Noo
7-11: Filtering Point Spread Function in Backprojection Cone Beam CT and its Applications in Long Object Imaging	KC Tam, G Lauritsch and K Sourbelle
7-12: Segmented Cardiac Volume Reconstruction - A Novel Reconstruction Scheme for Multislice Cardiac Spiral CT	H Bruder, K Stierstorfer, B Ohnesorge, S Schaller, and T Flohr
7-13: Lung Nodule Imaging Using Micro CT	N Niki, Y Kawata, M Fujii, E Matsui, N Moriyama, and Y Tateno

15:45 - 17:30

Session 8: 4D Reconstruction

Moderated by Jeffrey A Fessler and Gregory J Klein

8-1: Theoretical Framework for a Dynamic Cone-Beam Reconstruction Algorithm Based on a Dynamic Particle Model	P Grangeat, A Koenig, T Rodet, and S Bonnet

8-2: withdrawn	
8-3: Image to Volume Weighting Generalized ASSR for Arbitrary Pitch 3D and Phase-Correlated 4D Spiral Cone-Beam Reconstruction	M Kachelriess, T Fuchs, R Lapp, D-A Sennst, S Schaller, and W Kalender

Thursday 1 November 2001

08:30 - 10:15

Session 9: Algorithm Implementation and Evaluation

Moderated by Bryan W Reutter and Vesna Sossi

9-1: HeinzelCluster: Accelerated Reconstruction for FORE and OSEM3D	S Vollmar, C Michel, JT Treffert, D Newport, C Knöss, K Wienhard and W-D Heiss
9-2: Evaluation of an Exact Fourier Rebinning Algorithm for a Large Aperture PET Scanner	X Liu, C Michel, S Vollmar, K Wienhard, M Casey, and M Defrise
9-3: Performance of 3D RAMLA with Smooth Basis Functions on Fully 3D PET Data	S Matej, ME Daube-Witherspoon, and JS Karp

10:45 - 11:55

Session 10: Spiral CT III

Moderated by Michel Defrise and Ronald J Jaszczak

10-1: Movement Artifacts in Helical CT Cone-Beam Reconstruction	C Bontus, R Proksa, J Timmer, T Köhler, and M Grass
10-2: Longitudinal Sampling and Aliasing Properties in Multi-Slice Helical Computed Tomography	PJ La Rivière and X Pan

14:40 - 15:50

Session 11: General Reconstruction

Moderated by Margaret E Daube-Witherspoon and Gregory J Klein

11-1: Aspects of 3-D Imaging by Classical Tomography for Dual Detector PEM	MF Smith, S Majewski, AG Weisenberger, D Kieper, JD Kalen, and PP Fatouros
11-2: Diffusion Tensor MR Imaging of Principal Directions: A Tensor Tomography Approach	VY Panin, GL Zeng, M Defrise, and GT Gullberg

16:20 - 17:30

Session 12: CT Trajectories

Moderated by Paul E Kinahan and Hiroyuki Kudo

12-1: Complete Source Trajectories for C-Arm Systems and a Method for Coping with Truncated Cone-Beam Projections	H Schomberg
12-2: Computation of Unmeasured 3rd Generation VCT Views from Measured Views - Preliminary Results	SK Patch

Friday 2 November 2001

08:30 - 10:15

Session 13: SPECT

Moderated by Anna Celler and Ronald H Huesman

13-1: A Generalized Model for 3D Compton Scatter in Single Photon Emission Computed Tomography Using Slice-by-Slice Blurring and Scatter Projection Rebinning	C Bai, GL Zeng, and GT Gullberg
13-2: Computational Determination of Orlov Volumes	SD Metzler, JE Bowsher and RJ Jaszczak
13-3: Inversion of the 3-D Exponential X-ray Transform for Semi Equatorial Band	F Noo, R Clackdoyle, and J-M Wagner

10:45 - 11:55

Session 14: Iterative Techniques

Moderated by Freek Beekman and Bryan W Reutter

14-1: Accurate Iterative Reconstruction Algorithm for Sparse Objects: Application to 3-D Blood-Vessel Reconstruction from a Limited Number of Projections	M Li, H Yang, and H Kudo
14-2: ART for Helical Cone-Beam CT Reconstruction	BM Carvalho, GT Herman, and S Matej

Theoretically exact FBP-type inversion algorithm for spiral CT

Alexander Katsevich

Abstract— **Proposed is a theoretically exact formula for inversion of data obtained by a spiral CT scan with a 2-D detector array. The detector array is supposed to be of limited extent in the axial direction. The main property of the formula is that it can be implemented in a truly filtered backprojection fashion. First, one performs shift-invariant filtering of a derivative of the cone beam projections, and, second, the result is back-projected in order to form an image. Another property is that the formula solves the so-called "long object problem". Limitations of the algorithm are discussed. Results of numerical experiments are presented.**

I. INTRODUCTION

In the past decade it became clear that spiral CT can be significantly improved if one uses two-dimensional detector arrays instead of one-dimensional ones. However, accurate and efficient image reconstruction from the data provided by such scanners is very challenging because there does not exist a theoretically exact and efficient reconstruction formula. Several approaches for image reconstruction have been proposed. They can be classified into two groups: theoretically exact and approximate. See [1] for a recent review of available algorithms. Most of exact algorithms are based on computing the Radon transform for a given plane by partitioning the plane in a manner determined by the spiral path of the x-ray source [2], [3], [4], [5]. Even though exact algorithms are more accurate, they are computationally quite intensive and require keeping considerable amount of cone beam (CB) projections in memory. Approximate algorithms are much more efficient (see e.g. [6], [7], [8], [9] for several most recent techniques), but produce artifacts, which can be significant under unfavorable circumstances. Despite the progress achieved in recent years, it appears that no algorithm which would be both efficient and theoretically exact have been proposed in the literature so far.

In this paper we propose the first theoretically exact inversion formula for Spiral CT which is truly of the filtered backprojection (FBP) type. This means that the formula can be numerically implemented in two steps. First, one performs shift-invariant filtering of a derivative of the CB projections, and, second, the result is back-projected in order to form an image. The price to pay for this efficient structure is that the algorithm requires an array wider than the theoretically minimum one. Also, the algorithm is applicable if radius of support of the patient inside the gantry is not too big (not greater than ≈ 0.62× radius of gantry). Clearly, this limitation is not a big problem in many cases: for example, when one scans the head or an extremity of a patient.

II. INVERSION FORMULA

First we introduce the necessary notations. Let

$$C := \{y \in \mathbb{R}^3 : y_1 = R\cos(s), y_2 = R\sin(s), y_3 = s(h/2\pi), \\ s \in I\},\ I := [a,b], \tag{1}$$

where $h > 0, b > a$, be a spiral, and U be an open set strictly inside the spiral:

$$\overline{U} \subset \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 < r^2, a(h/2\pi) < x_3 < b(h/2\pi)\}, \tag{2}$$

$0 < r < R$, S^2 is the unit sphere in $\mathbb{R}^3$, and

$$D_f(y,\beta) := \int_0^\infty f(y+\beta t)dt,\ \beta \in S^2, \tag{3}$$

$$\beta(s,x) := \frac{x-y(s)}{|x-y(s)|}, x \in U, s \in I, \tag{4}$$

$$\Pi(x,\xi) := \{y \in \mathbb{R}^3 : (y-x)\cdot\xi = 0\}, \tag{5}$$

that is $D_f(y,\beta)$ is the CB transform of f. Given $(x,\xi) \in U \times (\mathbb{R}^3 \setminus 0)$, let $s_j = s_j(\xi, \xi\cdot x), j = 1,2,\ldots$, denote finitely many points of intersection of the plane $\Pi(x,\xi)$ with C. Also, $\dot{y}(s) := dy/ds$.

As was shown in [10], [8], any point strictly inside the spiral belongs to one and only one PI segment. Recall that a PI segment is a segment of line endpoints of which are located on the spiral and separated by less than one pitch in the axial direction. Let $s = s_b(x)$ and $s = s_t(x)$ denote values of the parameter corresponding to the endpoints of the PI segment containing x. We will call $I_{PI}(x) := [s_b(x), s_t(x)]$ the PI parametric interval. The part of the spiral corresponding to $I_{PI}(x)$ will be denoted $C_{PI}(x)$. Also, inside the PI parametric interval there exists $s_0 = s_0(x)$ such that the plane through $y(s_0)$ and parallel to $\dot{y}(s_0), \ddot{y}(s_0)$, contains x.

Fix $x \in U$. It is clear that any plane through x intersects $C_{PI}(x)$ at least at one point. Introduce the following sets:

$$\begin{aligned} Crit(x) =& \{\xi \in \mathbb{R}^3 \setminus 0 : \Pi(x,\xi) \text{ contains } y(s_b(x)), y(s_t(x)) \\ & \text{ or } \Pi(x,\xi) \text{ is tangent to } C_{PI}(x)\} \cup \{0\}, \\ \Xi_1(x) =& \{\xi \in \mathbb{R}^3 : \xi \notin Crit(x) \text{ and } \Pi(x,\xi) \cap C_{PI}(x) \\ & \text{contains one point}\}, \\ \Xi_3(x) =& \mathbb{R}^3 \setminus \{\Xi_1(x) \cup Crit(x)\}. \end{aligned} \tag{6}$$

This research was supported in part by NSF grant DMS-9704285.
The author is with the Department of Mathematics, University of Central Florida, Orlando, FL 32816-1364. E-mail address: akatsevi@pegasus.cc.ucf.edu

By construction, the sets $Crit(x), \Xi_{1,2}(x)$ are pairwise disjoint, their union is all of $\mathbb{R}^3$, $Crit(x)$ is closed and has Lebesgue measure zero, and $\Xi_{1,2}(x)$ are open.

Let $e_1(s,x)$ denote a unit vector in the plane through $y(s)$ and spanned by $\beta(s,x), \dot{y}(s)$ subject to the conditions that $e_1(s,x)$ is perpendicular to $\beta(s,x)$ and $e_1(s,x)\cdot\dot{y}(s) > 0$.

Given $y(s), s \in (s_b(x), s_t(x)) \setminus \{s_0(x)\}$, find $s_{tan} \in I_{PI}(x), s_{tan} \neq s$, such that the plane through $x, y(s)$, and $y(s_{tan})$ is tangent to $C_{PI}(x)$ at $y(s_{tan})$. For the exceptional values $s \in \{s_b(x), s_t(x), s_0(x)\}$, s_{tan} is determined by continuity. This construction defines a continuous function $s_{tan} = s_{tan}(s,x)$. One can show that $s = s_0(x)$ implies $s_{tan} = s_0(x)$.

Similarly, let $e_2(s,x)$ be a unit vector in the plane through $x, y(s)$, and tangent to $C_{PI}(x)$ at $y(s_{tan}(s,x))$. We require also that $e_2(s,x)$ is perpendicular to $\beta(s,x)$. This determines a continuous vector-valued function $e_2(s,x)$ up to a sign. The final requirement to eliminate ambiguity is $e_1(s,x) = e_2(s,x)$ when $s = s_{tan} = s_0(x)$.

For $f \in C_0^\infty(U)$ and $k = 1,2$ define

$$(\mathcal{B}_k f)(x) := -\frac{1}{2\pi^2}\int_{I_{PI}(x)} \frac{1}{|x-y(s)|} \times \int_0^{2\pi} \frac{\partial}{\partial q} D_f(y(q), \cos\gamma\beta(s,x) + \sin\gamma e_k(s,x))\bigg|_{q=s} \times \frac{d\gamma}{\sin\gamma} ds. \tag{7}$$

Our main result is the following theorem.

Theorem 1: The operators $\mathcal{B}_k, k = 1,2$, can be written in the form

$$(\mathcal{B}_k f)(x) = \frac{1}{(2\pi)^3}\int_{\mathbb{R}^3} B_k(x,\xi)\tilde{f}(\xi)e^{-i\xi\cdot x}d\xi, \tag{8}$$

where for each $x \in U$, $B_k(x,\xi) \in L^\infty(\mathbb{R}^3)$ with respect to ξ and

$$B_1(x,\xi) = \begin{cases} 1, & \xi \in \Xi_1(x), \\ 3, & \xi \in \Xi_3(x), \end{cases} \quad B_2(x,\xi) = \begin{cases} 1, & \xi \in \Xi_1(x), \\ -1, & \xi \in \Xi_3(x). \end{cases} \tag{9}$$

Since the set $Crit(x)$ has Lebesgue measure zero, (9) immediately implies the following inversion formula.

Corollary 2: Under the assumptions of Theorem 1,

$$f = \frac{1}{2}(\mathcal{B}_1 f + \mathcal{B}_2 f). \tag{10}$$

An important feature of the double integral in (7) is that for each $x \in U$ the integral with respect to s is confined only to the theoretically minimal portion of the spiral $I_{PI}(x)$. This implies that inversion formula (10) solves the so-called "long object problem" (see [8] for a definition).

III. Practical implementation and numerical experiments

In this section we discuss efficient algorithms for computing $\mathcal{B}_{1,2}f$. Fix any $y(s_0)$ on the spiral. It is assumed that the detector plane is parallel to the axis of the spiral and is tangent to the cylinder $y_1^2 + y_2^2 = R^2$ (cf. (1)) at the point opposite to the source. Thus, the distance between $y(s_0)$ and the detector plane is $2R$. Stereographic projections of the upper and lower turns of the spiral onto the detector plane are denoted by Γ_{top} and Γ_{bot}, respectively (see Figures 1 and 2). Let L_0 denote the common asymptote of Γ_{top} and Γ_{bot}. The parameter Δ shown in these figures is determined by the radius of support of the patient: $\Delta = 2\cos^{-1}(r/R)$ (cf. (2)).

Fix now any $\beta \in S^2$. By construction, all points $x \in U$ such that $\beta(s,x) = \beta$ will generate the same vectors $e_1(s,x)$ and $e_2(s,x)$. Denoting the corresponding vector-valued functions by $e_1(s,\beta)$ and $e_2(s,\beta)$, rewrite $\mathcal{B}_k f, k = 1,2$, as follows:

$$\begin{aligned}(\mathcal{B}_k f)(x) &:= -\frac{1}{2\pi^2}\int_{I_{PI}(x)} \frac{1}{|x-y(s)|}\Psi_k(s,\beta(s,x))ds, \\ \Psi_k(s,\beta) &:= \int_0^{2\pi} \frac{\partial}{\partial q} D_f(y(q), \cos\gamma\beta + \sin\gamma e_k(s,\beta))\bigg|_{q=s} \\ &\quad \times \frac{1}{\sin\gamma}d\gamma.\end{aligned} \tag{11}$$

Suppose first $k = 1$. Let $\Pi(\omega), \omega \in \mathbb{R}$, denote the family of planes containing $y(s)$ and parallel to $\dot{y}(s)$. Intersections of $\Pi(\omega)$ with the detector plane generate a family of lines $L(\omega)$ parallel to L_0 (see Figure 1). Fix any $\beta \in \Pi(\omega)$. By construction, vectors $\cos\gamma\beta + \sin\gamma e_1(s,\beta), 0 \leq \gamma < 2\pi$, belong to the same plane $\Pi(\omega)$. Here, for convenience, we think of vectors $\beta, e_1(s,\beta)$, and their linear combinations as if they are attached to $y(s)$. Let θ be a polar angle in $\Pi(\omega)$. Since $e_1(s,\beta)\cdot\beta = 0, |e_1(s,\beta)| = 1$, we can write (with abuse of notation):

$$\begin{aligned}&\beta = (\cos\theta, \sin\theta),\ e_1(s,\beta) = (-\sin\theta, \cos\theta), \\ &\beta, e_1(s,\beta) \in \Pi(\omega).\end{aligned} \tag{12}$$

Therefore,

$$\begin{aligned}\Psi_1(s,\beta) = \int_0^{2\pi} \frac{\partial}{\partial q} D_f(y(q), (\cos(\theta+\gamma), \sin(\theta+\gamma)))\bigg|_{q=s} \\ \times \frac{1}{\sin\gamma}d\gamma,\ \beta \in \Pi(\omega).\end{aligned} \tag{13}$$

Equation (13) is of convolution type. Hence, one application of Fast Fourier Transform (FFT) to the integral in (13) gives values of $\Psi_1(s,\beta)$ for all $\beta \in \Pi(\omega)$ at once.

Calculation of $\mathcal{B}_2 f$ can be arranged in a similar way. Fix $s_{tan} \in [s - 2\pi + \Delta, s + 2\pi - \Delta], s_{tan} \neq s$, and let $\Pi(s_{tan})$ denote the plane through $y(s), y(s_{tan})$, and containing $\dot{y}(s_{tan})$. If $s_{tan} = s$, $\Pi(s_{tan})$ is determined by continuity and coincides with the plane through $y(s)$ and parallel to $\dot{y}(s), \ddot{y}(s)$. The family of lines $L(s_{tan})$ obtained by intersecting $\Pi(s_{tan})$ with the detector plane is shown in Figure 2. By construction, given any $x \in U$ with $\beta(s,x) \in \Pi(s_{tan})$, s_{tan} used here is precisely the

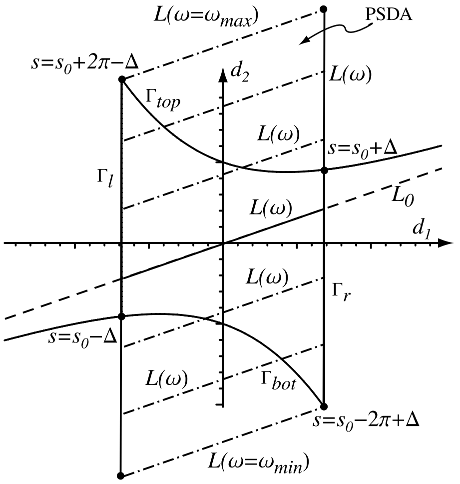


Fig. 1. Illustration of the one-parametric family of lines $L(\omega)$

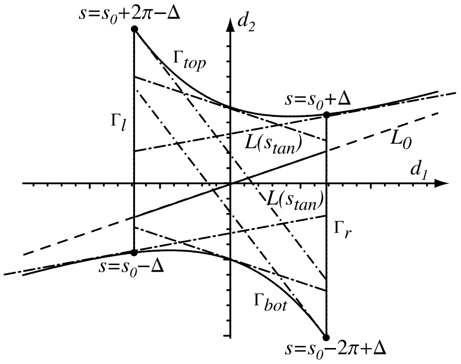


Fig. 2. Illustration of the one-parametric family of lines $L(s_{tan})$

same as s_{tan} used in the definition of $e_2(s,x)$. Since $e_2(s,\beta)\cdot\beta = 0, |e_2(s,\beta)| = 1$, we can write (with abuse of notation):

$$\begin{aligned}&\beta = (\cos\theta, \sin\theta),\ e_2(s,\beta) = (-\sin\theta, \cos\theta),\\ &\beta, e_2(s,\beta) \in \Pi(s_{tan}).\end{aligned} \tag{14}$$

Therefore,

$$\begin{aligned}\Psi_2(s,\beta) = \int_0^{2\pi} \frac{\partial}{\partial q} D_f(y(q), (\cos(\theta+\gamma), \sin(\theta+\gamma)))\Big|_{q=s}\\ \times \frac{1}{\sin\gamma} d\gamma,\ \beta \in \Pi(s_{tan}).\end{aligned} \tag{15}$$

Equation (15) is of convolution type and one application of FFT gives values of $\Psi_2(s,\beta)$ for all $\beta \in \Pi(s_{tan})$ at once.

Equations (11), (13), and (15) imply that the resulting algorithm is of the filtered-backprojection type. First, one computes shift-invariant filtering of a derivative of CB projections using (13) for all required ω: $\omega_{min} \le \omega \le \omega_{max}$ (cf. Figure 1), and using (15) - for all $s_{tan} \in [s - 2\pi + \Delta, s + 2\pi - \Delta]$ (cf. Figure 2). The second step is backprojection according to the first equation in (11). Since $\partial/\partial q$ in (13) and (15) is a local operation, each CB projection is stored in memory as soon as it has been acquired for a short period of time for computing this derivative at a few nearby points and is never used later.

This discussion shows that for the algorithm to work the following two conditions must be satisfied. First, the detector array should be large enough to contain the parallelogram formed by the lines Γ_l, Γ_r and $L(\omega_{min}), L(\omega_{max})$. We will call this parallelogram the parallelogram-shaped detector array (PSDA) and its area will be denoted A_{PSDA}. Thus, the size of the detector array required for the algorithm is greater than the theoretically minimum one, which is bounded by Γ_l, Γ_r and $\Gamma_{top}, \Gamma_{bot}$. Its area will be denoted A_{min}. The ratio of the two areas is independent of the pitch h, but grows as $r \to R$. For example, $A_{PSDA}/A_{min} = 1.53$ if $r/R = 1/3$ and $A_{PSDA}/A_{min} = 1.93$ if $r/R = 0.5$. Second, the segments of lines tangent to Γ_{top} and Γ_{bot} at $s = s + 2\pi - \Delta$ and $s - 2\pi + \Delta$, respectively, and located between Γ_l and Γ_r should be inside the detector array. This requirement leads to the restriction $r/R \le \cos(\Delta_0/2) \approx 0.62$, where Δ_0 is the unique solution to the equation $\tan(2\pi - \Delta) = 2\pi - \Delta$ on the interval $\pi/2 < \Delta < \pi$.

Consider now two numerical experiments. Parameters of the data collection protocols are given in Table I.

	Shepp phantom	disk phantom
R (radius of the spiral)	3	
h (pitch of the spiral)	0.5	
axial span of the detector array	1.02	0.96
transverse span of the detector array	4.74	4.26
number of detector rows	50	
number of detectors per row	500	
number of source positions per one turn of the spiral	1500	

TABLE I

Parameters of the data collection protocols

In Figure 3 we show the results of reconstructing the 3-D low contrast Shepp phantom (see Table 1 in [11]). In the top panel we see a vertical slice through the reconstructed image at $x_1 = -0.25$, and in the bottom panel - the graphs of exact (dashed line) and computed (solid line) values of f along a vertical line $x_1 = -0.25, x_2 = 0$. We used the grey scale window $[1.01, 1.03]$ to make low-contrast features visible.

In Figure 4 we see the results of reconstructing the disk phantom, which consists of six identical flattened ellipsoids (lengths of half-axes: 0.75, 0.75, and 0.04, distance between centers of neighboring ellipsoids: 0.16). In the top panel we see a vertical slice through the reconstructed image at $x_1 = 0$, and in the bottom panel - the graphs of exact (dashed line) and computed (solid line) values of f along a vertical line $x_1 = 0, x_2 = 0$.

As one can see, the algorithm still suffers from artifacts that are due to discretization and/or sampling errors. How-

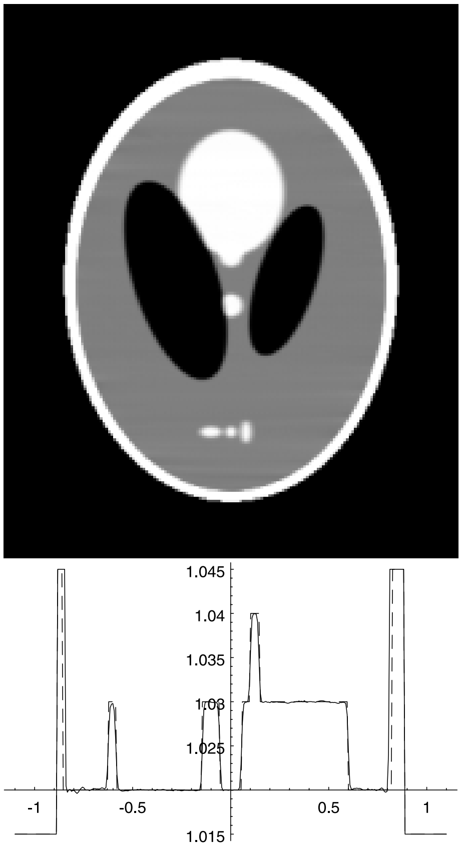


Fig. 3. Reconstruction of the 3-D Shepp phantom

ever, there are no artifacts that could be caused by non-exactness of a reconstruction scheme. Such artifacts have been theoretically studied and demonstrated numerically in [9] in the case of one approximate reconstruction algorithm. In numerical experiments presented in [9] these artifacts appear as nearly horizontal lines tangent to the ellipsoids.

References

[1] H. Turbell and P.-E. Danielsson, "Helical cone beam tomography," *Int. J. of Imaging Syst. and Technology*, vol. 11, pp. 91–100, 2000.

[2] K. C. Tam, "Method and apparatus for converting cone beam x-ray projection data to planar integral and reconstructing a three-dimensional computerized tomography (CT) image of an object," US Patent 5,257,183, October 1995.

[3] K. C. Tam, "Cone-beam imaging of a section of a long object with a short detector," in *Information processing in medical imaging*, J. S. Duncan and G. R. Gindi, Eds., vol. 1230 of *Lecture Notes in Computer Science*, pp. 525–530. Springer, Berlin, 1997.

[4] H. Kudo and T. Saito, "An extended cone-beam reconstruction using Radon transform," in *1996 IEEE Med. Imag. Conf. Record*. 1997, pp. 1693–1697, IEEE.

[5] S. Schaller, F. Noo, F. Sauer, et al., "Exact Radon rebinning algorithm for the long object problem in helical cone-beam CT," *IEEE Trans. on Medical Imaging*, vol. 19, pp. 361–375, 2000.

[6] H. Kudo, F. Noo, and M. Defrise, "Cone-beam filtered-backprojection algorithm for truncated helical data," *Phys. Med. Biol.*, vol. 43, pp. 2885–2909, 1998.

[7] F. Noo, H. Kudo, and M. Defrise, "Approximate short-scan filtered-backprojection for helical CB reconstruction," in *Conf. Rec. 1998 IEEE Nuclear Science Symposium*, Toronto, Ont., Canada, 1998, vol. 3, pp. 2073–2077, IEEE, Piscataway, NJ.

[8] M. Defrise, F. Noo, and H. Kudo, "A solution to the long-object problem in helical cone-beam tomography," *Physics in Medicine and Biology*, vol. 45, pp. 623–643, 2000.

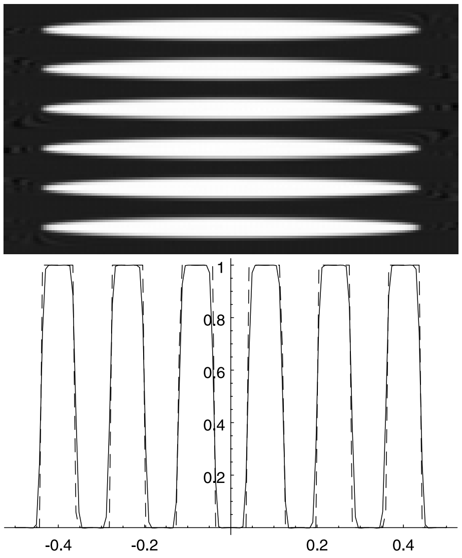


Fig. 4. Reconstruction of the disk phantom

[9] A. Katsevich, "Microlocal analysis of an FBP algorithm for truncated spiral cone beam data," 2000, (submitted).

[10] P. E. Danielsson et al., "Towards exact reconstruction for helical cone-beam scanning of long objects. A new detector arrangement and a new completeness condition," in *Proc. 1997 Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine (Pittsburgh)*, D. W. Townsend and P. E. Kinahan, Eds., 1997, pp. 141–144.

[11] H. Kudo, N. Miyagi, and T. Saito, "A new approach to exact cone-beam reconstruction without Radon transform," in *1998 IEEE Nuclear Science Symposium Conference Record*. 1998, pp. 1636–1643, IEEE.

Simple Quasi-Exact Filtered Backprojection Algorithms for Long-Object Problem in Helical Cone-Beam Tomography

Hiroyuki Kudo[1], Frédéric Noo[2], and Michel Defrise[3]
[1]Institute of Information Sciences and Electronics, University of Tsukuba, Tsukuba 305-8573, Japan
[2]Institute of Electricity Montefiore, University of Liege, Belgium
[3]Department of Nuclear Medicine, Free University of Brussels, Belgium

I. INTRODUCTION

Since the development of multi-slice helical CT, cone-beam reconstruction with a helical vertex path is receiving increasing attention. The current multi-slice CT scanners are adopting a class of approximate reconstruction algorithms which can be viewed an extension of the 2-D FBP algorithm. However, these approximate algorithms break down as the pitch of helix becomes large. In fact, major companies are beginning to develop a large area detector having more than 100 detector rows which allows a helical scan with a quite large pitch. This limitation of approximate algorithms motivated researchers to develop a class of exact reconstruction algorithms [1]-[6].

The reconstruction problems in helical cone-beam CT can be divided into the short-object (SO) problem and the long-object (LO) problem. The SO problem aims at reconstructing an object having a finite axial support when the helix is long enough to cover the object support. The LO problem aims at reconstructing a central region of interest (ROI) of a long object when the helix is long enough to cover the ROI but too short to cover the whole object. In 1998, Tam [1] and Kudo *et al.* [2] developed quasi-exact algorithms for the SO problem. In 2000, Schaller *et al.* [3], Defrise *et al.* [4], and Kudo *et al.* [5] extended the solutions to the SO problem to the LO problem. The resulting algorithms are called the Local-ROI (L-ROI) method [3],[6], the Zero-Boundary (ZB) method [4], and the Virtual-Circle (VC) method [5] in the literature. However, it is fair to say that these algorithms need rather complicated modifications to the quasi-exact algorithms for the SO problem.

The purpose of this paper is to develop new quasi-exact FBP algorithms to the LO problem. Similarly to the VC method [5], the algorithms can be viewed an extension of the quasi-exact FBP algorithm for the SO problem developed by Kudo *et al.* [2]. The main advantage of the proposed algorithms is their simplicity compared with the existing quasi-exact algorithms (L-ROI, ZB, and VC). The algorithms need only a few small changes to the quasi-exact FBP algorithm for the SO problem although their derivation involves rather different mathematical logic. We show simulation results which demonstrate that the proposed algorithms allow to reconstruct high-quality images indistinguishable from those by the VC method [5].

II. PROPOSED ALGORITHMS

A. FBP Algorithm for Short-Object Problem

We first review the quasi-exact FBP algorithm for the SO problem derived by Kudo *et al.* [2] which is the basis of proposed algorithms for the LO problem. This review also helps to convince readers what modifications are necessary to convert the algorithm for the SO problem into the algorithms for the LO problem.

We use the same notations as in our previous papers [2],[5]. Let $f(\vec{r})$ denote an object supported inside a cylinder $\Omega = \{\vec{r} \mid x^2 + y^2 \leq Q^2\}$ where $\vec{r} = (x, y, z)^T$. We assume that cone-beam projections are measured along a finite segment of helix:

$$\vec{a}(\lambda) = (R\cos\lambda, R\sin\lambda, h\lambda)^T;\ \lambda_{\min} \leq \lambda \leq \lambda_{\max} \tag{1}$$

where $2\pi h$ is the pitch and the range $[\lambda_{\min}, \lambda_{\max}]$ defines the length of helix. Let $g(u, v, \lambda)$ denote a cone-beam projection measured from the source point $\vec{a}(\lambda)$ where (u, v) denote the detector coordinates defined such that the u-axis coincides with the tangential direction of the helical path $\vec{a}'(\lambda)$. Let $D(\lambda)$ denote the detector plane corresponding to $\vec{a}(\lambda)$ where $D(\lambda)$ contains the z-axis. We assume that each cone-beam projection is measured over the finite region on the detector plane B which is bounded by the cone-beam projection of the upper turn of helix onto $D(\lambda)$ and by the cone-beam projection of the lower turn onto $D(\lambda)$. The explicit expression of B can be found in [1],[2]. This region B is known to be the minimum detector area compatible with exact reconstruction. Let $\chi_B(u, v)$ denote the indicator function of B. We also use a notation $g^{\mu}(s, t, \lambda)$ below to denote a cone-beam projection rotated by an angle μ on the detector plane where (s, t) denote the coordinates rotated by the angle μ. Kudo *et al.* [2] derived the FBP algorithm for the SO problem. This algorithm is quasi-exact when the axial support of the object is small compared with the length of helix such that $g(u, v, \lambda_{\min})$ and $g(u, v, \lambda_{\max})$ vanish over the region B. The algorithm is summarized as follows.

<Algorithm for SO Problem (Algorithm SO)>

[STEP 1] Weighting

$$g_1(u, v, \lambda) = \frac{R}{\sqrt{R^2 + u^2 + v^2}} g(u, v, \lambda) \tag{2}$$

[STEP 2] Computation of Ramp Filtering Term

$$g^{Fr}(u,v,\lambda)=\frac{1}{2\pi}\int_{-\infty}^{\infty}du'h(u-u')\chi_B(u',v)g_1(u',v,\lambda) \tag{3}$$

where $h(\cdot)$ denotes the kernel of the ramp filter.

[STEP 3] Computation of Boundary Correction Term

$$S_\Delta^{(b)}(s,\mu,\lambda)=c_U(s,\mu)g_1^\mu(s,t_U,\lambda)+c_L(s,\mu)g_1^\mu(s,t_L,\lambda) \tag{4}$$

$$g^{Fb}(u,v,\lambda)=\frac{1}{4\pi^2}\frac{\partial}{\partial u}\int_{-\pi/2}^{\pi/2}d\mu S_\Delta^{(b)}(u\cos\mu+v\sin\mu,\mu,\lambda) \tag{5}$$

where t_U (or t_L) denotes the t-coordinate of the point on which the straight line $s=u\cos\mu+v\sin\mu$ intersects the upper (or lower) boundary of B. The explicit expressions of $c_U(s,\mu)$ and $c_L(s,\mu)$ can be found in Eq. (30) of [5].

[STEP 4] Backprojection

$$f(\vec{r})=\int_{\lambda_{\min}}^{\lambda_{\max}}d\lambda\frac{R\parallel\vec{a}'(\lambda)\parallel}{[(\vec{r}-\vec{a}(\lambda))\cdot\vec{1}_w]^2}(g^{Fr}(u,v,\lambda)+g^{Fb}(u,v,\lambda)) \tag{6}$$

where $\vec{1}_w$ denotes the unit vector which is directed toward the detector center from $\vec{a}(\lambda)$.

The key feature of the above algorithm is that the filtered projection $g^{Fr}+g^{Fb}$ can be computed as a sum of the ramp filtering term applied to the truncated projection $\chi_B g_1$ and the boundary correction term which only depends on values of g_1 along the boundary of region B.

B. FBP Algorithms for Long-Object Problem

A blind use of the above FBP algorithm to the long object having a large axial support compared with the length of helix produces severe low-frequency artifacts [5]. This section describes two quasi-exact FBP algorithms which are derived from the mathematical logic outlined in Section II-C.

To describe the proposed algorithms, we need to define a region on the detector plane called the region $A(\lambda)$. Its definition is as follows. Let $P_{\min}=(u_{\min}(\lambda),v_{\min}(\lambda))$ (or $P_{\max}=(u_{\max}(\lambda),v_{\max}(\lambda))$) denote the cone-beam projection of the end point of helix $\vec{a}(\lambda_{\min})$ (or $\vec{a}(\lambda_{\max})$) onto the detector plane $D(\lambda)$. We define the region $A(\lambda)$ as a rectangular region bounded below by the straight line $v=v_{\min}(\lambda)$ and bounded above by the straight line $v=v_{\max}(\lambda)$. Let $\chi_{A(\lambda)}(u,v)$ denote the indicator function of $A(\lambda)$. In addition to the region B introduced to solve the SO problem, this region $A(\lambda)$ plays an important role in the proposed quasi-exact FBP algorithms for the LO problem. By using $A(\lambda)$, the first proposed algorithm is summarized as follows.

<Algorithm for LO Problem 1 (Algorithm LO-1)>

[STEP 1] Weighting (Eq. (2))

[STEP 2] Computation of Ramp Filtering Term (Eq. (3))

[STEP 3] Computation of Boundary Correction Term

$$g_2(u,v,\lambda)=\chi_{A(\lambda)}(u,v)g_1(u,v,\lambda) \tag{7}$$

$$S_\Delta^{(b)}(s,\mu,\lambda)=c_U(s,\mu)g_2^\mu(s,t_U,\lambda)+c_L(s,\mu)g_2^\mu(s,t_L,\lambda) \tag{8}$$

$$g^{Fb}(u,v,\lambda)=\frac{1}{4\pi^2}\frac{\partial}{\partial u}\int_{-\pi/2}^{\pi/2}d\mu S_\Delta^{(b)}(u\cos\mu+v\sin\mu,\mu,\lambda) \tag{9}$$

[STEP 4] Backprojection (Eq. (6))

The comparison between the Algorithm SO and the Algorithm LO-1 shows that we only need to multiply each cone-beam projection g_1 by the indicator function $\chi_{A(\lambda)}$ before computing the boundary correction term g^{Fb} in the LO problem. This can be done quickly in numerical implementation because $\chi_{A(\lambda)}$ does not depend on u and the explicit expression of $\chi_{A(\lambda)}$ is simple. Other parts (weighting, ramp filtering, and backprojection) do not need to be changed at all. Therefore, the Algorithm LO-1 can be easily implemented if one has a program for the Algorithm SO. Note that all the existing algorithms (L-ROI, ZB and VC) need rather complicated modifications in the structure of algorithms to deal with the long-object.

An alternative form of the Algorithm LO-1 can be obtained by performing the multiplication by $\chi_{A(\lambda)}$ before computing both the ramp filtering term and the boundary correction term. This leads to the following algorithm.

<Algorithm for LO Problem 2 (Algorithm LO-2)>

[STEP 1] Weighting

$$g_1(u,v,\lambda)=\frac{R}{\sqrt{R^2+u^2+v^2}}\chi_{A(\lambda)}(u,v)g(u,v,\lambda) \tag{10}$$

[STEP 2] Computation of Ramp Filtering Term (Eq. (3))

[STEP 3] Computation of Boundary Correction Term (Eqs. (4) and (5))

[STEP 4] Backprojection (Eq. (6))

Note that it is easy to verify that the Algorithm LO-1 and the Algorithm LO-2 produce a same reconstruction in the central ROI covered by the helix. This is thanks to the locality of ramp filtering. However, they produce different reconstructions in the region close to the end points of helix. The Algorithm LO-2 is also very simple. The comparison between the Algorithm SO and the Algorithm LO-2 shows that we only need to multiply each cone-beam projection g by $\chi_{A(\lambda)}$ before using the Algorithm SO to deal with the long-object.

C. Outline of Algorithm Derivation

Due to the lack of space, we only outline the derivation of the proposed algorithms. The algorithm derivation is along the similar line to the derivation of the VC method [5]. It proceeds according to the following two steps. The first step is to construct the Radon algorithm by using the modified Grangeat formula (the Grangeat formula [7] which allows to combine triangular patches and half-planes to compute the 3-D Radon derivative). The second step is to reduce the Radon algorithm to the FBP form to verify that unmeasurable parts of projections used in the triangulation are unnecessary to reconstruct the central ROI when the helix is long enough to cover the ROI.

Let us consider the reconstruction of single point $\vec{r}_0$ in the central ROI. Let $\Pi(\vec{\xi}, l)$ denote a plane having the unit normal $\vec{\xi}$ and the radial distance l. Let $p'(\vec{\xi}, l)$ denote the 3-D Radon derivative over $\Pi(\vec{\xi}, l)$. From the inversion formula of 3-D Radon transform, we know that the reconstruction of $f(\vec{r}_0)$ needs $p'(\vec{\xi}, l)$ over a limited set of planes:

$$P = \{\Pi(\vec{\xi}, l) \mid l - \epsilon < \vec{r}_0 \cdot \vec{\xi} < l + \epsilon,\ \vec{\xi} \in \mathrm{S}^2,\ l \in \mathrm{R}\} \quad (11)$$

where ϵ is a small positive number [7]. By using this fact, the procedure for the algorithm derivation is outlined as follows.

[STEP 1] The triangulation for each $\Pi(\vec{\xi}, l) \in P$ is performed in the following way. Assume that the plane $\Pi(\vec{\xi}, l)$ intersects the helix at N points $\vec{a}(\lambda_1), \ldots, \vec{a}(\lambda_N)$. As shown in Fig. 1(a), we consider that the first source point $\vec{a}(\lambda_1)$ covers the lower half-plane which is below the line connecting $\vec{a}(\lambda_1)$ with $\vec{a}(\lambda_2)$, the last source point $\vec{a}(\lambda_N)$ covers the upper half-plane which is above the line connecting $\vec{a}(\lambda_{N-1})$ with $\vec{a}(\lambda_N)$, and other source points cover the triangular patches as in [2],[5].

[STEP 2] The mask function $\chi_{C(\lambda)}(u, v)$ on the detector plane $D(\lambda)$ to achieve the triangulation defined in [STEP 1] is derived. The resulting mask function depends on the source point $\vec{a}(\lambda)$. The forms of the mask function $\chi_{C(\lambda)}$ are illustrated in Fig. 1(b).

[STEP 3] We use the mathematical logic similar to [5] (reduction of the Radon algorithm to the FBP form) to show that unmeasurable parts of projections (which exceed the region B for the source points close to the end points of helix) are unnecessary to reconstruct the point $\vec{r}_0$. This clarifies that the filtered projections can be computed according to the following steps. Projections measured from the central part of helix are processed in the same way as in the Algorithm SO because $C(\lambda) = B$ for these projections. However, for projections measured from the source points close to the end point of helix $\vec{a}(\lambda_{\min})$ or $\vec{a}(\lambda_{\max})$, the boundary correction term must be included only from the subset $A(\lambda) \cap \partial B$. This is because $C(\lambda)$ for these projections is different from B and the corresponding boundary ∂B stops at some point $P_{\min}$ (or $P_{\max}$) defined by the cone-beam projection of the end point of helix $\vec{a}(\lambda_{\min})$ (or $\vec{a}(\lambda_{\max})$) onto the detector plane $D(\lambda)$ (Fig. 1(b)). Therefore, the boundary term from the missing boundary part must be excluded from the computation when processing these projections. This corresponds to truncate g_1 with the indicator function $\chi_{A(\lambda)}$ in the Algorithms LO-1 and LO-2. This is the rationale behind the use of new mask function $\chi_{A(\lambda)}$ in addition to χ_B in the LO problem.

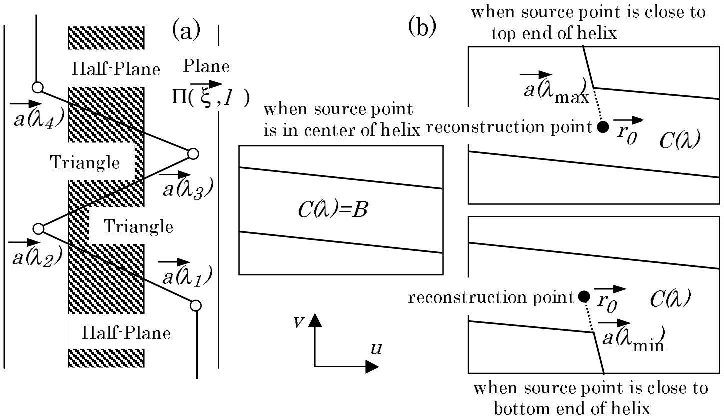


Fig 1: Illustration of triangulation for plane $\Pi(\vec{\xi}, l)$ and the corresponding mask function $\chi_{C(\lambda)}$.

III. Simulation Studies

We compared the proposed algorithms with the VC method [5] in terms of image quality, computational time, and simplicity of implementation. The same numerical phantoms as in [5] are used. The first phantom is the 3-D Shepp-Logan phantom to confirm how each algorithm reconstructs low-contrast objects. The second phantom is the disk phantom to confirm how each algorithm reconstructs high-contrast objects. These phantoms are supported inside the cylinder of radius 100 (mm). The number of helical turns is 2, the axial length of helix is 150 (mm) (pitch $2\pi h$ is 75 (mm)), and the radius of helix is 350 (mm). The number of cone-beam projections is 500 per turn and each projection consists of 256×93 detector pixels. Reconstructed images have 256×256×256 pixels. We have implemented the proposed algorithms as follows. As remarked in [2],[5], all the above algorithms are numerically unstable because they force to apply the ramp filtering across the boundary of region B. To overcome this numerical problem, we used the numerical stabilization technique based on smoothing the indicator function χ_B [2],[5]. This technique computes the filtered projection $g^{Fr} + g^{Fb}$ in the proposed algorithms as a sum of the three terms $g^F_{\mathrm{ramp}} + g^F_{\mathrm{fou}} + g^F_{\mathrm{bound}}$ which can be computed in a stable way [2],[5]. Furthermore, we implemented both the single-ROI and multi-ROI algorithms proposed in [5].

Longitudinal slices of reconstructed images with the proposed algorithms (LO-1,LO-2) and the VC method are shown in Fig. 2. For the Algorithm LO-1, we also show reconstructed images with the multi-ROI implementation. The proposed algorithms could reconstruct high-quality images which are indistinguishable from those with the VC method. The main difference among the Algorithms

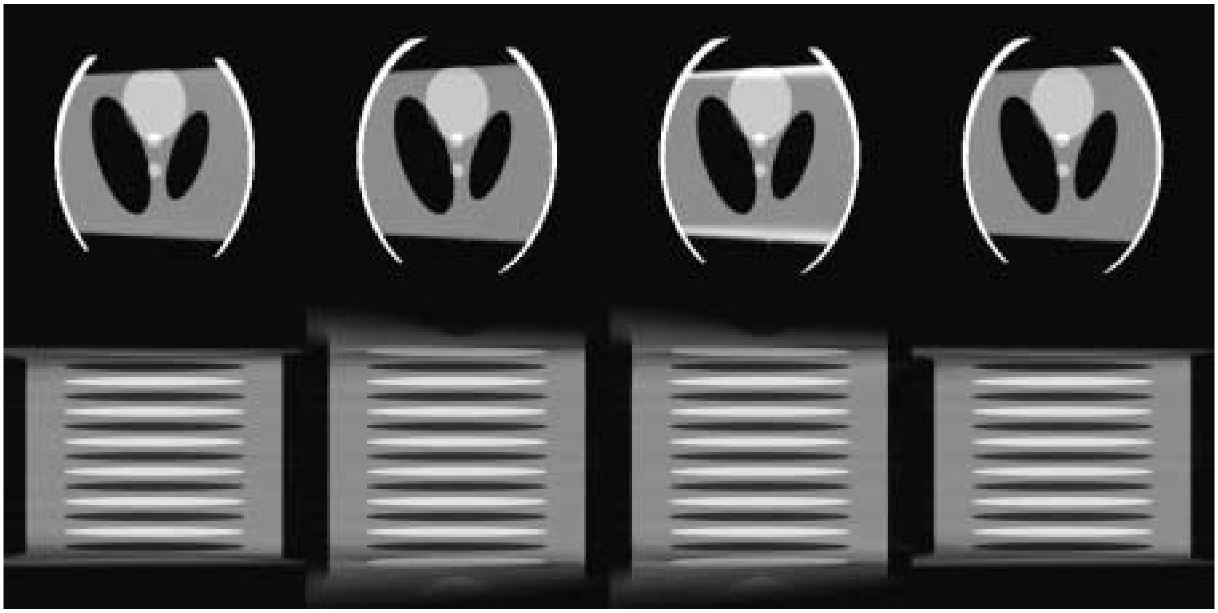


Fig 2: Reconstructed images with the VC single-ROI (first column), the LO-1 single-ROI (second column), the LO-2 single-ROI (third column), and the LO-1 multi-ROI (fourth column).

LO-1, LO-2, and VC seems to be the region on which accurate reconstruction can be achieved. Figure 2 shows that the LO-1 can obtain accurate images on a slightly larger region compared with the LO-2 and VC. This is thanks to the fact that the LO-1 discards less data compared with the LO-2 and VC because the multiplication by $\chi_{A(\lambda)}$ is applied only to the boundary correction term. Computational times are summarized in Table 1 (time for the approximate Feldkamp algorithm is also shown for comparison). The difference of reconstruction times among the LO-1, LO-2, and VC is not so much mainly because the dominant computation in all the algorithms is the 3-D backprojection which is common to all of them. In terms of simplicity, the proposed algorithms needed rather simple programming to implement compared with the VC method. In fact, a quite complicated routine to compute the boundary correction term in the VC method (Appendix C of [5]) could be completely eliminated. Thus, we believe that the proposed algorithms succeeded in dramatically reducing the complexity of implementation compared with the existing algorithms (L-ROI, ZB, and VC).

Table 1
Actual computational times measured by a SUN SPARC ULTRA-1 workstation with 256 M-byte memory.

Feldkamp (Full-Scan)	VC (Single)	LO-1 (Single)	LO-2 (Single)	LO-1 (Multi)
120(min)	412(min)	301(min)	296(min)	263(min)

IV. Conclusions

The proposed quasi-exact FBP algorithms for the LO problem are rather simple compared with the existing algorithms (L-ROI, ZB, and VC). The algorithms need only a few small changes to the quasi-exact FBP algorithm for the SO problem derived by Kudo *et al* [2]. We will present additional simulation results with more challenging Schaller's head phantom at the conference.

V. References

[1] K.C.Tam, S.Samarasekera, and F.Sauer ”Exact cone-beam CT with a spiral scan,” *Phys.Med.Biol.*, Vol.43, pp.1015-1024, 1998

[2] H.Kudo, F.Noo, and M.Defrise ”Cone-beam filtered backprojection algorithm for truncated helical data,” *Phys.Med.Biol.*, Vol.43, pp.2885-2909, 1998

[3] S.Schaller, F.Noo, F.Sauer, K.C.Tam, G.Lauritsch, and T.Flohr ”Exact radon rebinning algorithms using local region-of-interest for helical cone-beam CT,” *IEEE Trans.Med.Imaging*, Vol.19, pp.361-375, 2000

[4] M.Defrise, F.Noo, and H.Kudo ”A solution to the long-object problem in helical cone-beam tomography,” *Phys.Med.Biol.*, Vol.45, pp.623-643, 2000

[5] H.Kudo, F.Noo, and M.Defrise ”Quasi-exact filtered backprojection algorithm for long-object problem in helical cone-beam tomography,” *IEEE Trans.Med.Imaging*, Vol.19, pp.902-921, 2000

[6] K.C.Tam ”Exact local regions-of-interest reconstruction in spiral cone-beam filtered-backprojection CT: theory,” *Proc. SPIE Medical Imaging*, Vol.3979, pp.506-519, 2000

[7] P.Grangeat ”Mathematical framework of cone-beam 3D reconstruction via the first derivative of the Radon transform,” in Mathematical Methods in Tomography, G.T.Herman, A.K.Louis, and F.Natterer Eds., Springer, 1991, pp.66-97

PRACTICAL HELICAL CONE BEAM ALGORITHM FOR THE LONG OBJECT PROBLEM

JICUN HU, ROGER JOHNSON, CHRISTOPHER DAWSON

DEPARTMENT OF BIOMEDICAL ENGINEERING, MARQUETTE UNIVERSITY, USA

INTRODUCTION

Cone beam computed tomography (CT) based on non-planar orbits has been an active area of research toward the goal of producing an exact volumetric reconstruction. To date, most reconstruction algorithms for non-planar orbits have been based on the theoretical framework of Tuy[1], Smith [2] and Grangeat[3]. Of all the investigated non-planar orbits, the helical scanning geometry is most promising for clinical application since it is easy to implement and natural for volume scanning of the body.

Currently, multi-row detectors with only a few rows (e.g. four) are widely used in clinical helical CT, and the reconstruction algorithms involved differ only slightly from the traditional fan beam methods. This approximation produces minimal artifact in the reconstruction because the cone angle is very small. In order to increase the volume scanning speed and reduce motion artifacts, and to make more efficient use of the x-ray tube output, more and more detector rows will be used. In the near future a point will be reached where the z-divergence of the beam will become non-negligible. There is a need, therefore, to develop efficient reconstruction algorithms which account properly for the cone beam scanning geometry.

Some approximate algorithms for helical cone beam tomography have been developed[4,5,6]. They are efficient and provide good temporal resolution because only full scan or half scan data are used to reconstruct a slice and the ramp filtered data remain bounded in the z direction, keeping the computational intensity low. They are approximate in nature however, and will produce artifacts when the cone angle is increased.

Practical exact helical cone beam tomography algorithms are made possible by the discovery that only the truncated data within the region on the detector bounded by the projections of the adjacent upper and lower turns of the helix are required to obtain an exact reconstruction [7,8], not the full data set as utilized in the theory developed by Tuy, Smith and Grangeat. Quasi-exact algorithms outperform approximate methods in terms of reconstructed image quality, yet some researchers have argued that exact algorithms have the disadvantages of inferior temporal resolution and increased computational intensity. We show that an exact algorithm can deliver computational efficiency and temporal resolution comparable to that achieved by approximate methods. With continued research in this area, the advantages of exact over approximate algorithms will become increasingly evident.

Exact helical cone beam tomography algorithms can be categorized as addressing the short object problem or the long object problem. For the short object problem, the axial extent of the helix is sufficient to cover the entire object, providing adequate data for a comparatively simple solution. For the more complex long object problem, the helix extends only slightly beyond the ROI. It is much more difficult to solve than the short object problem due to the data contamination issue. From a clinical standpoint, a solution to the long object problem is required since only a portion of the patient should be scanned to provide accurate images of the finite volume of interest. Tam[7] first provided a solution for the long object problem, but his algorithm required two circular orbits at the end of helix which is undesirable in practice. Several exact algorithms for the long object problem have been developed which do not require the circular orbits, but reconstruct the ROI using only the helical orbit data[9,10,11].

In this paper, we present a new solution to the long object problem using helical data only. It is based on the method developed by Kudo, Noo and Defrise[8]. We invoke the concept of accessory paths with upper and lower virtual detectors having infinite axial extent. We show that our approach has the advantages of ease of implementation, good temporal resolution and computational efficiency. The algorithm possesses the filtered backprojection structure, which is very desirable for practical implementation.

ALGORITHM

We propose an algorithm to solve the long object problem for helical cone beam tomography using accessory paths with virtual detectors of infinite axial extent.

Fig. 1 illustrates the data acquisition geometry. The source path $\vec{a}(\theta)$ is a short helical segment of pitch $2\pi h$ and radius R defined by (1), containing a primary helix (the solid path), a top accessory helix (top dotted path) and a bottom accessory helix (bottom dotted path) around a long object $f(x,y,z)$ (the cylinder).

$$\vec{a}(\theta) = (R\cos\theta, R\sin\theta, h\theta)^T, \qquad (1)$$

where θ is the rotation angle of the helix.

The detector is normalized to the iso-center and its coordinate system is defined by unit vectors $\vec{\xi}_u$ and $\vec{\xi}_v$:

$$\vec{\xi}_u = (-\cos\eta\sin\theta, \cos\eta\cos\theta, \sin\eta)^T \qquad (2)$$

$$\vec{\xi}_v = (\sin\eta\sin\theta, -\sin\eta\cos\theta, \cos\eta)^T \qquad (3)$$

$$\eta = \arctan(h/R) \qquad (4)$$

$\vec{\xi}_u$ is parallel to the helix tangent.

Our objective is to reconstruct a slice (the shaded ellipse in Fig. 1) from the data collected on the source helix. From short object problem theory, we know that it is impossible to exactly reconstruct the slice without scanning the whole object due to the data contamination problem.

Our approach employs three different types of detector. For the data collected from the primary path, we have a masked detector (Fig. 2b) with the top and bottom boundaries given by[8]:

$$V_r(u_r) = h(1 + \frac{u_r^2}{R^2})[\frac{\pi}{2} - \arctan(\frac{u_r}{R})] \qquad (5)$$

$$V_r(u_r) = -h(1 + \frac{u_r^2}{R^2})[\frac{\pi}{2} + \arctan(\frac{u_r}{R})] \qquad (6)$$

where $(u_{r,} v_r)$ is the rotated coordinate system defined by:

$$u_r = u\cos\eta - v\sin\eta \qquad (7)$$

$$v_r = u\sin\eta + v\cos\eta \qquad (8)$$

so that u_r is horizontal.

For the data collected from the top accessory path, we use a detector (Fig. 2a) with no boundary on the top and with a boundary at the bottom defined by (6) in order to capture the entire upper portion of the object. For the data collected from the bottom accessory path, we use a detector (Fig. 2c) unbounded at the bottom with a boundary at the top defined by (5) in order to capture the entire lower portion of the object.

Before explaining our algorithm, we revisit an important and remarkable property of the filtered masked projection discovered by Kudo *et al.* [8]. Namely, the result of filtering the masked data can be represented as:

$$g^F(u,v,\theta) = g^F_{bounded}(u,v,\theta) + g^F_{unbounded}(u,v,\theta). \qquad (9)$$

The bounded term results from ramp filtering the data within the mask and the unbounded term results from the unbounded filtering of the boundary data. This property is the key to an understanding of our algorithm. According to the above property, we know that for the central detector (the standard PI detector, Fig. 2b), data on both boundaries will contribute to the unbounded part of the filtering result, but for the upper detector (Fig. 2a), only the data on the bottom boundary will contribute to the unbounded term and for the lower detector (Fig. 2c), only the data on the top boundary will contribute to the unbounded term.

The following conditions on the primary path and accessory paths must be met in order for our algorithm to succeed. The primary path must be long enough so that the ramp filtered part of the projections collected from the accessory paths do not contribute to the reconstructed slice. The accessory paths (the dotted paths in Fig. 1) must be of length $(\pi + 2\arcsin(\frac{r}{R}))$ at both ends of the primary path, where r is the FOV radius. When these conditions are met, we can obtain the Radon derivatives needed to reconstruct every point in the slice based on the whole object. A typical data combination is shown in Fig. 3. Notice that for points on the accessory paths, one end of the source rays' envelope opens to infinity because we employ a virtual detector of infinite extent above and below the slice. But remember, the infinite detector is virtual and the unlimited data is not available in practice. Fortunately, because of the filtering property we mentioned earlier, only the data on the bottom boundary of the upper virtual detector and the data on the top boundary of the lower virtual detector will become unbounded after filtering and contribute to the reconstruction of the slice. The ramp filtered portion of the filtered projection from the accessory paths will not contribute to the reconstructed slice. Therefore, only the data on the boundaries of the upper and lower virtual detectors are required for exact reconstruction of the slice. These data are available from the primary path itself due to the property that the integral over a PI line can be accessed from either of the two source points that define the PI line (Fig. 4). Thus, the slice of interest can be reconstructed from the primary path data alone.

Explicitly, we can reconstruct the slice using the following three steps:

- Fully filter the projections (bounded term + unbounded term) obtained from the primary path and backproject the filtered results into the slice.
- Obtain the boundary data for the accessory paths, do the filtering (unbounded term only) and backproject the result into the slice.
- Add these two contributions to get a fully reconstructed slice.

If we need to reconstruct an ROI in a long object, we can reconstruct every slice in the ROI using the above approach.
Our algorithm has the following advantages:

- It is local. Every filtered projection contributes only to a few slices and it is not necessary to extend it to cover the entire ROI. This property makes the computational efficiency comparable to the approximate methods.
- It has good temporal resolution: In order to fully reconstruct a slice with the parameters used in medical CT, we require only about one turn of helical data.
- It is relatively easy to implement, requiring only slight modifications to the short object implementation.
- Only helical data is needed.
- It possesses the filtered backprojection structure.

The algorithm can be made global if the filtered projection is extended sufficiently to contribute to the reconstruction of the whole ROI, but this increases the computational intensity and reduces the temporal resolution.

The pitch of the accessory paths can vary. Our algorithm is equivalent to Kudo's virtual circle algorithm [10] when the pitch of the accessory paths becomes 0 although our derivation differs from that of Kudo. We calculate Radon derivatives based on the whole object while Kudo calculated Radon derivatives based on the ROI defined by the two virtual circles.

In our implementation we chose the pitch of the accessory paths to be the same as that of the primary path making the mask B invariant which results in a single form of the boundary term. This makes it easier to localize the reconstruction and to implement the algorithm than is the case using Kudo's virtual circle approach [10].

SIMULATION RESULTS

We reconstructed the 3D Shepp-Logan phantom[12] (Fig. 5, display window [1.01 1.04]) and a disk phantom[8] (Fig. 6, display window [0.3 1.7]) to test the ability of the algorithm to reconstruct low contrast objects and objects with high frequency components in the z direction, respectively. We added two homogeneous cylinders at both ends of the phantoms to simulate the long objects. As demonstrated in the figures, both phantoms were reconstructed with satisfactory results.

ACKNOWLEDGEMENTS

We acknowledge support from the Whitaker Foundation, the Keck Foundation, the Falk Medical Trust, the Rose E. Bagozzi Professorship, NHLBI HL-19298, and the Veterans Administration.

REFERENCES

[1] H.K. Tuy, "An inversion formula for cone-beam reconstruction", SIAM J. Appl. Math., vol. 43, pp. 546-554,1983.

[2] B.D. Smith, "Image Reconstruction from cone-beam projections: necessary and sufficient conditions and reconstruction methods", IEEE Trans. Med. Imaging, vol. MI-4, pp.14-25,1985.

[3] P. Grangeat, "Mathematical framework of cone beam 3D reconstruction via the first derivative of the Radon transform" in Mathematical Methods in Tomography, Lecture Notes in Mathematics 1497, G.T. Herman, A.K. Louis, F. Natterer, eds. New York: Springer-Verlag, 1991, pp. 66-97.

[4] G. Wang, Y.Liu, T.H. Lin, P-C. Chen and D.M. Shinozaki, "A general cone-beam reconstruction algorithm", IEEE Trans. Med. Imaging, vol. 12, pp.486-496, 1993.

[5] H. Turbell and P-E. Danielsson, "Non-redundant data capture and highly efficient reconstruction for helical cone-beam CT", Conf. Record of 1998 IEEE Medical Imaging Conference, pp.1424-1425, 1998.

[6] F. Noo, M. Defrise and R. Clackdoyle, "Single-slice rebinning method for helical cone-beam CT", Phys. Med. Biol., vol. 44, pp. 561-570, 1999.

[7] K.C. Tam, S. Samarasekera and F. Sauer, "Exact cone-beam CT with a spiral scan", Phys. Med. Biol., vol. 43, pp. 1015-1024, 1998.

[8] H. Kudo, F. Noo and M. Defrise, "Cone-beam filtered-backprojection algorithm for truncated helical data", Phys. Med. Biol., vol. 43, pp. 2885-2909, 1998.

[9] M. Defrise, F. Noo and H. Kudo, "A solution to the long object problem in helical cone beam tomography", Phys. Med. Biol., vol. 45, pp. 1-21, 2000

[10] H.Kudo, F. Noo and M. Defrise, "Quasi-exact filtered backprojection algorithm for long object problem in helical cone beam tomography", IEEE Trans. Med. Imaging, vol.1, pp. 902-921, 2000

[11] S. Schaller, F. Noo, F.Sauer *et al.*, "Exact Radon rebinning algorithm for the long object problem in helical cone beam CT", IEEE Trans. Med. Imaging, vol. 19, pp. 361-375, 2000.

[12] C. Jacobson, "Fourier Method in 3D reconstruction from cone-beam data", Ph.D. Thesis, Linkoping University, 1996.

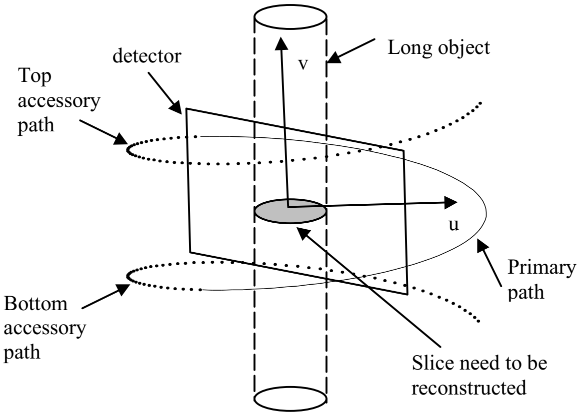


Fig. 1

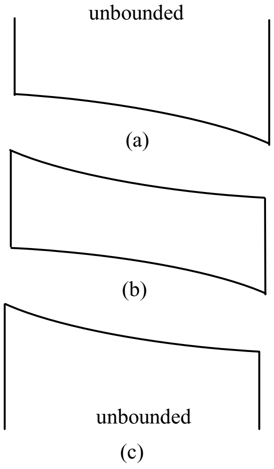


Fig. 2

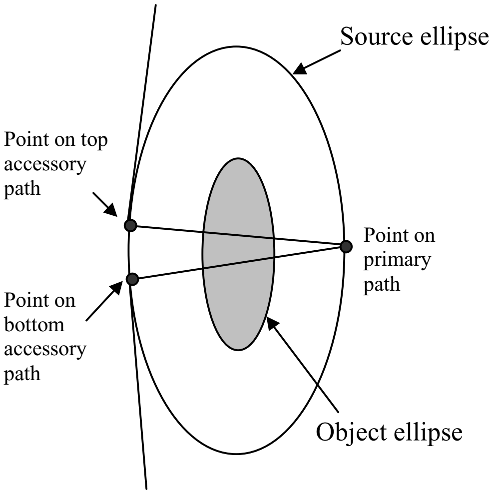


Fig. 3

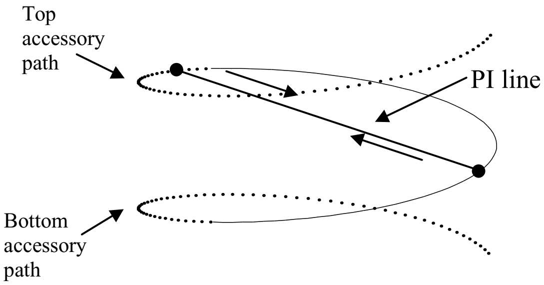


Fig. 4

Fig. 5

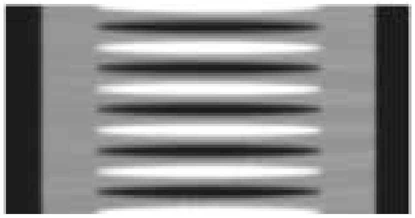


Fig. 6

Multifrequential Algorithm for Fast 3D Reconstruction

Thomas Rodet, Pierre Grangeat and Laurent Desbat

Abstract— **Some recent medical imaging applications, such as functional imaging (PET and SPECT) or interventional imaging (CT fluoroscopy) involves dynamic data. The image reconstruction time must be reduced. For that purpose, we developed a new fast algorithm for dynamic reconstruction based on frequential hierarchical reconstruction.**

Our algorithm performs an indirect subband decomposition of the image f to be reconstructed ($f = \sum f_j$) through the filtering of the projection Rf. The subband images f_j can be reconstructed on an undersampled grid without information suppression. In order to reduce the computation time, we undersample the number of projections and we choose them in accordance with the undersampled grid. But image compression can also be made directly in our algorithm by elimination of some frequential components with low information content.

Keywords— **Fast reconstruction algorithm, frequential decomposition, angular undersampling, computation compression.**

I. Introduction

CT acquisition systems involve increasing amounts of data. Moreover, for many medical applications, the computation time must be lowered. There are even some applications for which real time is required. Two main fields of medical imaging need dynamic reconstructions, namely functional imaging (PET and SPECT) and interventional imaging (CT fluoroscopy, 3D guidance) [1], [2], [3]. Fast tomographic reconstruction is currently a very active research domain [4], [5], [6].

In this paper, we present a fast algorithm for dynamic reconstruction. In our reconstruction algorithm, we implement the two following ideas to speed-up the reconstruction time: the first is to reconstruct some frequentials components with reduced number of projections, the second is to achieve computation compression. The frequential reconstruction allows to reconstruct each frequential component on an undersampled grid, and to use only a lower number of undersampled projections. Our algorithm associates this undersampling approach with computation compression (see [7]). The computation compression is based on the data compression principle. Our algorithm can generate directly compressed data.

In the first section, we introduce our notations. In the second section, we present the frequential reconstruction principle. Then, we expose the speed up factors of our algorithm: the angular undersampling, the computation compression and the block processing. The last section shows numerical applications of our algorithm on both a static 3D phantom and a dynamic 2D phantom.

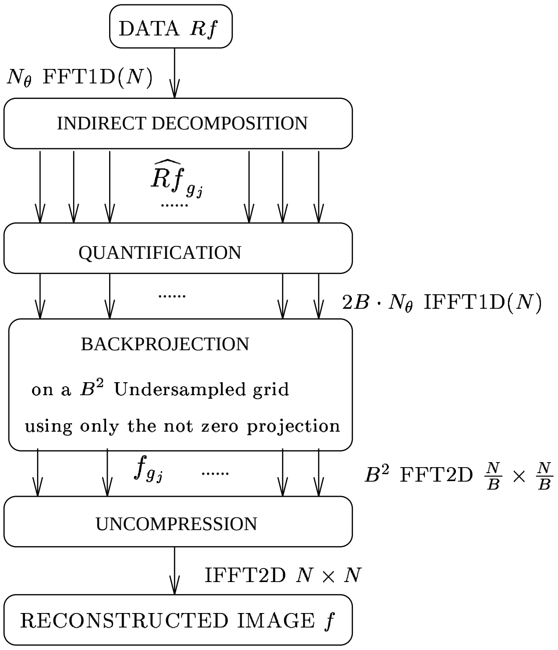


Fig. 1. *Our algorithm*

II. Notations

We define the Fourier transform $\hat{f}$ of a function $f \in L^1(\mathbb{R}^2)$ by:

$$\hat{f}(\nu) = \frac{1}{2\pi} \int_{\mathbb{R}^2} f(t) e^{-i\nu.t} dt$$

Let $s \in \mathbb{R}$, and $\theta \in S^1$, where S^1 is the unit circle, we denote the Radon transform R by:

$$Rf(\theta, s) = \int_{y \in \theta^\perp} f(s\theta + y) dy.$$

We denote by $R^\sharp$ the backprojection operator and by I the ramp filter. We define the functions $g_j \in \mathcal{S}(\mathbb{R}^2)$ as the inverse Fourier transform of adjacent band indicators such that $\sum_{j=1}^{n} \hat{g}_j$ is the indicator function of a domain covering the essential support of the function $\hat{f}$. Then, we denote each elementary frequential component of f: $f_{g_1} = f * g_1, \ldots, f_{g_n} = f * g_n$.

T.Rodet and L.Desbat are with the TIMC-IMAG, UMR CNRS 5525, IAB, Faculte de medecine, UJF 38706 La Tronche Cedex, FRANCE. E-mail:Laurent.Desbat@imag.fr

P.Grangeat are with the Laboratoire d'Electronique et de Technologie de l'Information (LETI), Departement Systemes pour l'Information et la Sante (DSIS), Direction des Recherches Technologique (DRT), Commissariat a l'Energie Atomique (CEA), 17 Rue des Martyrs, F 38054 GRENOBLE Cedex 9, FRANCE. E-mail: Thomas.Rodet@cea.fr, Pierre.Grangeat@cea.fr

III. RECONSTRUCTION ALGORITHM USING A DECOMPOSITION STEP

The main idea of our approach is to reconstruct B^2 frequential component f_{g_j} of f independently using $f_{g_j} = f * g_j = R^\sharp I(Rf * Rg_j)$ (see [7]). This allows us to undersample the number of projection and to compress the computation like in data compression.
We want to reconstruct B^2 frequential components of f without knowing the image f, but only its projections Rf. Also we must do an indirect decomposition of f through Rf. The projection slice theorem (see [8]) yields a direct relation between $\widehat{f}$ and $\widehat{Rf}$. Thus we can find a filter Rg_j to apply to Rf which leads to the decomposition of f into the sum of f_{g_j}(for a theoretical justification see [7]).

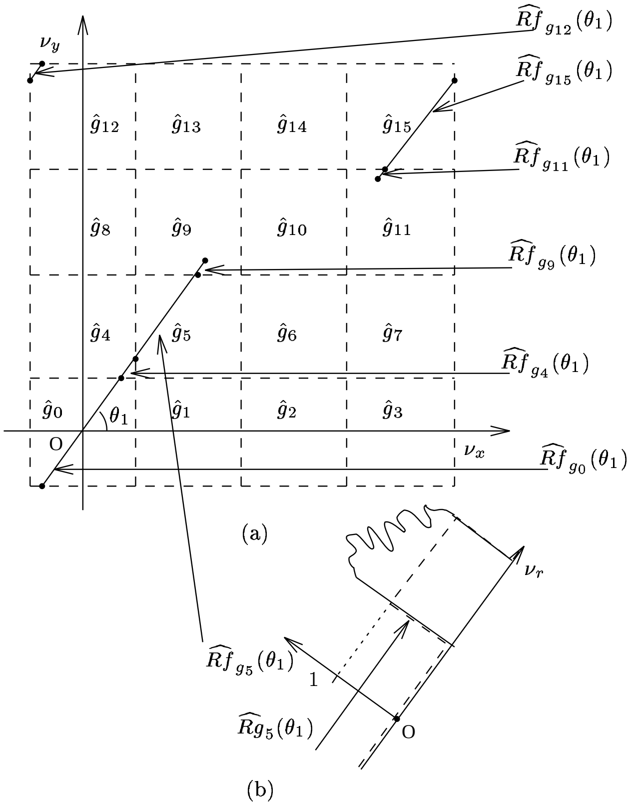


Fig. 2. *Indirect decomposition scheme*

A. Indirect decomposition step

The step of indirect decomposition (see figure 1) consists of generating the Rf_{g_j} with $j \in [0, B^2]$ by filtering Rf. We implement this filtering in the Fourier space to reduce the computation time of this step. We can see on the figure 2a the Fourier transform of one projection ($\widehat{Rf}(\theta_1)$) is localized on a line in the Fourier space. The decomposition of this projection on the indicator function $\hat{g}_j$ is the intersection between the line and the square $\hat{g}_j$ (see the figure 2a). On the figure 2b , we can see the result of the indirect decomposition $\widehat{Rf}_{g_j}$. Let us underline that $\widehat{Rf}_{g_j}(\theta)$ is equal to zero when the line $\widehat{Rf}(\theta)$ do not intersect the square $\hat{g}_j$. We will use this property in the next section to reduce the computation time.

B. Backprojection step

In this step we backproject the B^2 sets of projection Rf_{g_j} to obtain the frequential components f_{g_j}. According to the sampling theory, the function f_{g_j} can be represented by B^2 times less pixel than f. Indeed the undersampling by B^2 of f_{g_j} in the direct space leads to a periodisation of $\hat{g}_j$ in the Fourier space. The function $\hat{g}_j$ is localized on a little square. When the sampling rate is equal to B pixels along each axis, squares do not overlap. Thus the undersampling by B^2 preserves all the information contained in the function f_{g_j}. Thus we can backproject Rf_{g_j} on an undersampled grid. Finally, as we make B^2 backprojections on B^2 undersampled grids, the backprojection step computation cost is the same as for a classical algorithm.

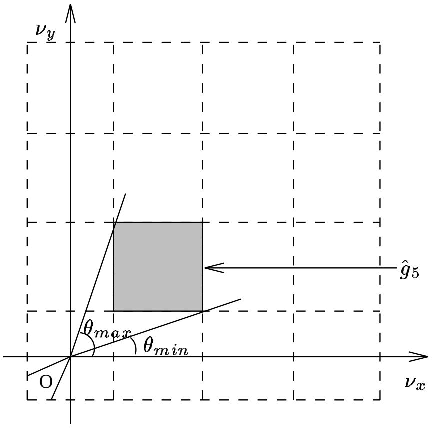


Fig. 3. *Angular interval*

IV. ANALYZE OF THE SPEED UP

A. Angular undersampling

The projections of the frequential components f_{g_j} are usually equal to zero for θ outside of an relative small interval $\theta \notin [\theta_{min}, \theta_{max}]$ (see figure 3, $\forall \theta \notin [\theta_{min}, \theta_{max}], \widehat{f_{g_5}} = 0$). We use this property to reduce the backprojection time. In our algorithm when we reconstruct the frequential components f_{g_j}, we backproject only the angle between the corresponding θ_{min} and θ_{max}. The speed up obtained depends on the frequential component. For example, we backproject all angle for f_{g_0} and only approximatively $\frac{N\theta}{B}$ angles for the frequential component $f_{g_{B-1}}$, if B is large (N_θ denotes the number of projections). To analyze the speed up factor obtained by this angular undersampling we count the contribution of the projection to the backprojection of each frequential component. Each projection $Rf(\theta)$ contributes to less than $2B$ frequential components (see figure 2). Thus the number of contributions per bloc is less than $2BN_\theta/B^2 = 2N_\theta/B$. Thus we win at least a factor $\frac{B}{2}$ compare to the backprojection seen in the previous paragraph. The elimination of the zero projection can be considered as an angular undersampling. If we respect the sampling condition to reconstruct f, and if we reconstruct the frequential component f_{g_j} on a B^2 undersampled grid, we can backproject B times less projections.

B. Computation compression

The computation compression will have the same fundamental steps as the data compression. We want also to preserve only the pertinent information and we want to code it with as few bytes as possible. To achieve this result, we decompose the image f in frequential components f_{g_j}, and we compute only the components containing pertinent informations. We denote "quantification", the elimination step of frequential components (see figure 1). This step corresponds to $||\hat{f} \cdot \hat{g}_j||_2 < \epsilon$, where the parameter $\epsilon > 0$ is given and controls the compression rate. Thus, the number of components is reduced from factor depending on the value ϵ. As for data compression, information will be lost if ϵ becomes high.

C. Block processing and parallelism

Most of uniprocessor systems have hierarchical memory. It is composed of various levels of cache memory, a main memory and disk storage. Movement of data between two levels in the hierarchy represents latency time cost. To compute efficiently, we must reduce data movement. Our approach allows to reconstruct some undersampled pictures. If an undersampled picture can be contained in cache memory, the data movement are minimized. Our approach allows to adapt the reconstruction computation to the size of the cache memory and thus to the computer. It is the well known block effect. Because our algorithm is naturally divided on B^2 elementary reconstructions, it can be easily adapted on multiprocessor systems. Owing to the block structure data movement between processor is minimized. In shared memory system, data movement between processor is made through the main memory. These access are limited band width, thus efficiency is improved by reducing memory access in this case too. Our algorithm natural block structure yields efficient computations.

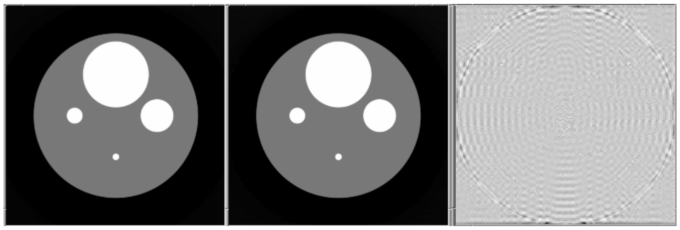


Fig. 4. *Reconstruction of a volume* $512\times512\times32$*: reference (left), our algorithm with 64 frequential components (center), difference (right).*

V. Application

In this section, we make numerical experiment on two types of phantom: a 3D static phantom to evaluate the optimal number of frequential components and a 2D dynamic phantom to obtain the maximum performance of our algorithm.

A. Static phantom

For our first test, we reconstruct a 3D static phantom f (see phantom definition Table I). The reconstructions are calculated on a $512 \times 512 \times 32$ voxel grid (with x, y and z resolution equal to 1,04 mm), using a subset of the projections collected on 512 angles, uniformly spaced over $[0, \pi[$. We compare our reconstruction with a classical filtered back projection (FBP) algorithm (fig. 4). We observe some artifact on the difference image caused by numerical approximation, but in the case of 64 frequential components, the level on these artifacts is less or equal to 1 %. To highlight the angular undersampling effect, we execute our algorithm with different numbers B^2 of frequential components without quantification. Indeed, we have seen in part IV-A that the speed up factor is proportional to B. Theoretically, the CPU time becomes shorter with big numbers of frequential components. The CPU time decreases up to 64 frequential components but increases afterwards (see figure 5). The first part of the curve is explained by angular undersampling described in part IV-A. The second part of the curve is explained by supplementary computations induced by our approach. Indeed, we can see on the figure 1, that the number of inverse Fourier transforms is equal to $N_\theta \cdot 2B$ instead of N_θ in a classical algorithm. This effect becomes dominating for $B^2 > 64$.

Sphere number	Center (mm)			Radius(mm)	Attenuation At
	x	y	z	r	
1	0.0	0.0	0.0	200	15
2	0.0	-100	0.0	10.0	35
3	-100	0.0	0.0	20	35
4	100	0.0	0.0	40	35
5	0.0	100	0.0	80.0	35

TABLE I
Definition of our 3D phantom

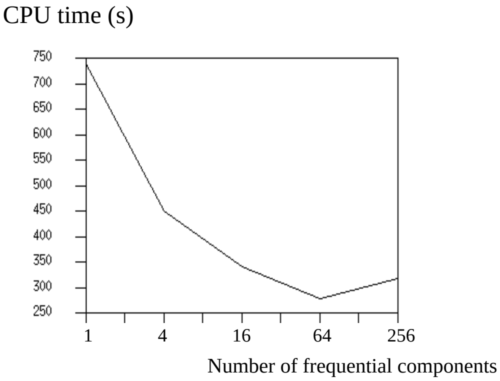


Fig. 5. *Reconstruction CPU time of a volume* $512 \times 512 \times 32$

B. 2D dynamic phantom

For our second test, we reconstruct a 2D dynamic phantom composed of 32 images f_i with $i \in [0, 31]$ (see phantom definition Table II). The phantom motion is a translation of 3 mm per frame along the y axis. In this case, the pertinent information is restricted to one oblic plane in the spatio-temporal Fourier domain [9]. Thus, this dynamic data contains a low number of pertinent information: a lot of frequential components can be eliminated. The

Circle number	Center (mm)		Radius(mm)	Attenuation
	x	y	r	At
1	0.0	-45 + 3 × i	200	15
2	0.0	-145 + 3 × i	10.0	35
3	-100	-45 + 3 × i	20	35
4	100	-45 + 3 × i	40	35
5	0.0	55 + 3 × i	80.0	35
6	0.0	-45 + 3 × i	2.0	60

TABLE II

Definition of our 2D dynamic phantom: with $i \in [0, 31]$

32 reconstructions are calculated on a 512 × 512 pixel grid (with x and y resolution equal to 1,04 mm), using the same parameter as the previous experiment. During the reconstruction we decompose each f_i in 64 frequential components. To highlight the quantification effect we reconstruct the image sequence with two configurations: the first reconstruction with 38% of components (see figure 6b and c) and the second one with 22% of components (see figure 6d and e). The results are summarized on the table III.

Types of algorithm	% of components	CPU time	speed up	relative L_1 error
FBP	100	754	1	0%
Our algorithm	100	290	2.6	1.09%
	38	200	3.8	1.17%
	22	169	4.5	1.54%

TABLE III

CPU time and quality of dynamic reconstruction

We can observe that the first configuration allows to eliminate 62% of frequential components without deterioration of the image quality. If 78% of components are eliminated, pertinent information is lost (see figure 6e): some artifacts appear in images. Even if a large number of components are eliminated, the speed up factor is not very high. The computation compression is thus only a secondary speed up factor. This can still be interesting for some real time applications, because we can adapt image quality for cpu resource.

VI. CONCLUSION

Our subband Fourier decomposition speeds up the 3D reconstruction. Moreover a compromise between speed up and image quality can be improved through the compression factor. Alternatively decompositions such as Cosinus, Wavelet or Karhunen Loève decomposition could also be used.

REFERENCES

[1] B. Daly and P. A. Templeton, "Real-time ct fluoroscopy: evolution of an interventional tool," *Radiology*, vol. 211, pp. 309–315, 1999.

[2] M. Zhang and M. Kono, "Solitary pulmonary nodules : Evaluation of blood flow paterns with dynamic ct," *Radiology*, pp. 471–478, 1997.

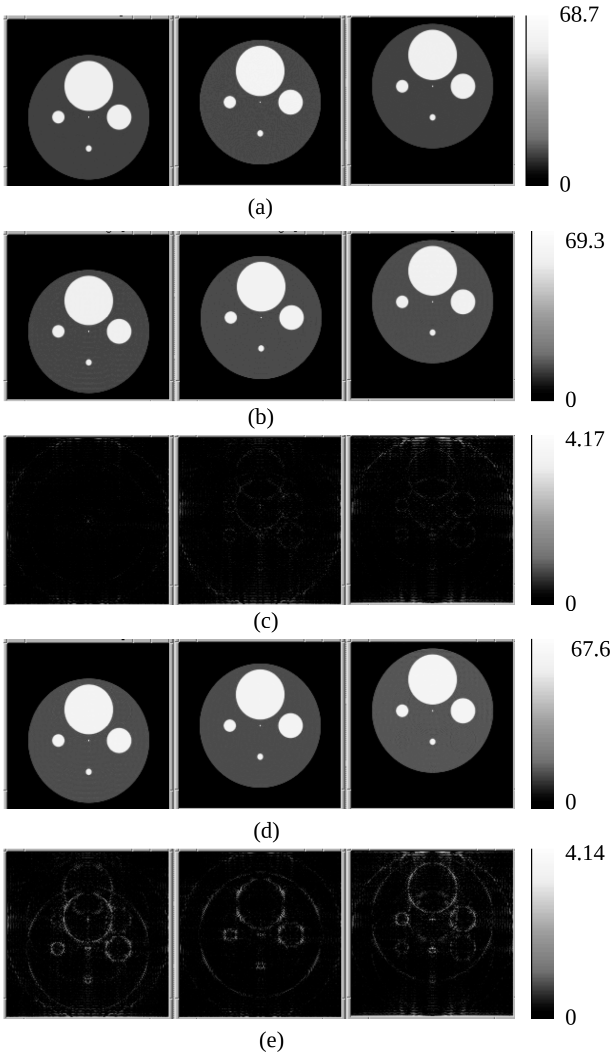


Fig. 6. *(a) Frame 0,15,31 reconstruct by classical FBP, (b) our algorithm with 38 % of frequential components, (c) absolute value of (a)-(b), (d) our algorithm with 22 % of frequential components, (e) absolute value of (c)-(d)*

[3] L. Desbat, M. Fleute, G. Champleboux, and O. Desaint, "3D reconstruction using the PIXIUM 4600 digital X-ray detector ," in *3D-1999*, 1999, pp. 149–152.

[4] M. L. Brady, "A fast discrete approximation algorithm for the radon transform," *SIAM*, vol. 27, no. 1, pp. 107–119, 1998.

[5] H. Turbell, *Three-dimensional image reconstruction in circular and helical computed tomography*, Ph.D. thesis, Linköping University, April 1999.

[6] P. E. Danielsson and M. Ingerhed, "implementation of backprojection in $o(n^2 \log n)$ time," Technical Report LiTH-ISY-R-2003, Linköping University, 1998.

[7] T. Rodet, P. Grangeat, and L. Desbat, "A new computation compression scheme based on a multifrequential approch," in *Conf. Rec. 2000 IEEE Med. Imag. Conf.*, 2000, number IEEE350.

[8] F. Natterer, *The Mathematics of Computerized Tomography*, Wiley, 1986.

[9] Bernd Jähne, *Digital Image Processing*, Springer Verlag Berlin, 1997.

Fully 3D Iterative Reconstruction of Planogram Data

P. E. Kinahan[1], D. Brasse[1], M. Defrise[2], R. Clackdoyle[3], C. Comtat[4], C. Michel[5], X Liu[2]
[1]Department of Radiology, University of Pittsburgh, Pittsburgh, PA 15213 USA
[2]Department of Medical Research, SHFJ, CEA, 91401 Orsay, France
[3]Department of Radiology, University of Utah, Salt Lake City, UT 84108 USA
[4]Division of Nuclear Medicine, University Hospital AZ-VUB, B-1090, Belgium
[5]CTI Inc., Knoxville, TN USA

Background

At the previous Fully 3D meeting we presented the concept of the 'planogram' data format for fully-3D imaging. The planogram data format is a generalization of the 2D linogram data format of Edholm et al. [1, 2]. This planogram format is based on the native acquisition geometry of planar detectors, illustrated in figure 1, where the natural rectangular detector coordinates are given by (x_{d1}, y_{d1}) and (x_{d2}, y_{d2}) for two individual detector events that determine a line of response (LOR).

We can parameterize the LOR orientation by the coordinates (u_1, u_2, v_1, v_2) where $u_1 = (x_{d1} - x_{d2})/2$, $u_2 = (y_{d1} + y_{d2})/2$, $v_1 = -(x_{d1} + x_{d2})/2$, and $v_2 = (y_{d2} - y_{d1})/2$. In this parameterization we assume the detectors are separated by unit distance so (v_1, v_2) are the tangents of the angle of the LOR projected onto the $x - y$ and $y - z$ planes, respectively relative to the y axis. We can then further parameterize the LOR w.r.t. the y-coordinate as $u_1 = x + v_1 y$ and $u_2 = z + v_2 y$.

If we regard a fixed (x, y, z), then the subset of LORs passing through that point will appear as a 2D plane in the (u_1, u_2, v_1, v_2)-space, thus the choice of the term 'planogram' for the data acquisition histogram.

In this case the measured line integral data are defined, after appropriate scaling by $((1 + v_1^2 + v_2^2)^{-1/2})$ (determined by the angle between the detector surface and the unit normal vector), as:

$$
\begin{aligned}
g_1(u_1, u_2, v_1, v_2) &= \int_{-\infty}^{\infty} f(x, y, z)\,\mathrm{d}y \\
&= \int_{-1/2}^{1/2} f(u_1 - v_1 y, y, u_2 - v_2 y)\,\mathrm{d}y
\end{aligned}
\qquad (1)
$$

where we now use $x = u_1 - v_1 y$ and $z = u_2 - v_2 y$. We can also define backprojection as:

$$
b_1(x, y, z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(x + v_1 y, z + v_2 y, v_1, v_2)\, dv_1 dv_2 \qquad (2)
$$

To adequately sample the function $f(x, y, z)$, we assume the data has been collected from a second detector position rotated about the z-axis by 90 deg, so that $u_1 = y - v_1 x$ and $u_2 = z - v_2 x$.

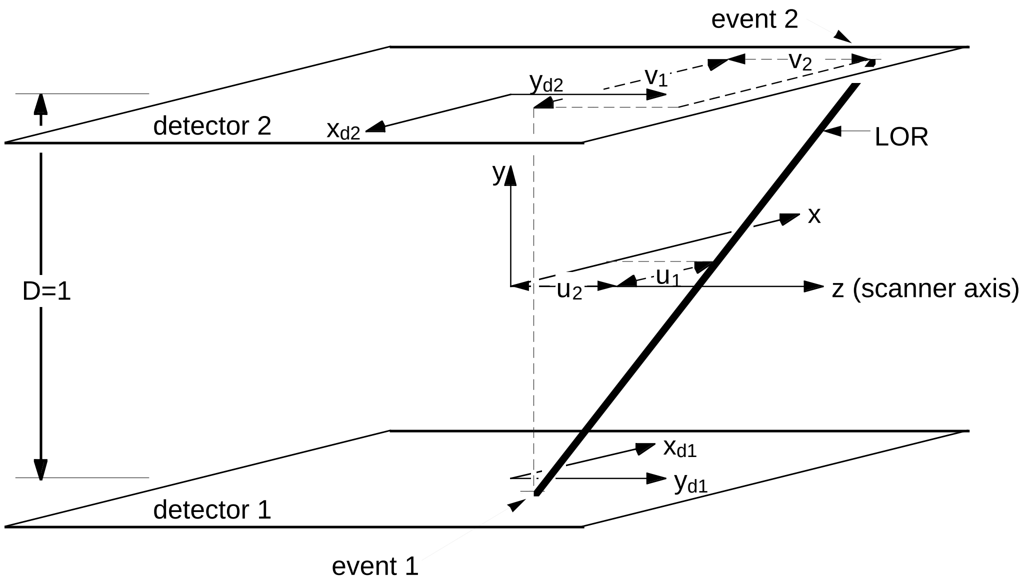


Figure 1. Coordinates used to index a LOR between two detector elements on planar detectors.

This work is supported by NIH grant CA80943.

Then by symmetry we have:

$$g_2(u_1,u_2,v_1,v_2) = \int_{-\infty}^{\infty} f(x,y,z)\,\mathrm{d}x = \int_{-1/2}^{1/2} f(x,u_1+v_1x,u_2+v_2x)\,\mathrm{d}x \tag{3}$$

where now $y = u_1 + v_1x$ and $z = u_2 + v_2x$. For this position we define backprojection as:

$$b_2(x,y,z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_2(y - v_1x, z - v_2x, v_1, v_2)\, dv_1 dv_2 \tag{4}$$

Two key properties of the planogram data formats are based on Fourier transform relations. We define:

$$G_{1,1111}(U_1,U_2,V_1,V_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(u_1,u_2,v_1,v_2) \cdot \mathrm{e}^{-2\mathrm{i}\pi(u_1U_1+u_2U_2+v_1V_1+v_2V_2)}\,\mathrm{d}u_1\mathrm{d}u_2\mathrm{d}v_1\mathrm{d}v_2 \tag{5}$$

and

$$B_{1,101}(X,y,Z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} b_1(x,y,z)\mathrm{e}^{-2\mathrm{i}\pi(x\mathrm{X}+zZ)}\,\mathrm{d}x\mathrm{dz} \tag{6}$$

where the subscript indicates w.r.t. which variable the Fourier transform has been taken.

The quadruple integral in equation (5) is over the (u_1, u_2, v_1, v_2) domain covered by the pair of rectangular detectors. The limits are implicitly extended to the whole of $\mathcal{R}^4$. Substituting equation (2) into equation (6), and by symmetry for equation (4), we have the surprising results that,

$$B_{1,101}(X,y,Z) = G_{1,1111}(X,Z,-yX,-yZ) \tag{7a}$$

and

$$B_{2,011}(x,Y,Z) = G_{2,1111}(Y,Z,xY,xZ) \tag{7b}$$

The results are a case of the 4D version of the central section theorem. Equation (7) indicates that fully 3D backprojection can be performed with only using Fourier transform operations. An algorithm for analytic backprojection is:

1. Compute the 4D Fourier transform $G_{1,1111}(U_1,U_2,V_1,V_2)$.
2. For each y_i:
   - interpolate $B_{1,101}(X,y_i,Z)$ using equation (7a).
   - compute the inverse 2D Fourier transform to obtain $b_1(x,y_i,z)$.
3. Repeat step 2 for $G_{2,1111}(U_1,U_2,V_1,V_2)$, using each x_i to interpolate $B_{2,011}(x_i,Y,Z)$ using equation (7b).

We have presented results showing that this backprojection method offers speed improvements of approximately 15 compared to standard fully 3D sinogram methods [3]. These speed improvements can be further leveraged by the use of readily available FFT processors.

Application to iterative reconstruction methods

To investigate the application of this approach to iterative methods we make two observations:

1. Equation (7) (backprojection) does not require that the data are non-truncated.
2. If we compute the 2D Fourier transform of equation (1), we derive a version of the 3D central section theorem:

$$G_{1,1100}(U_1,U_2,v_1,v_2) = F(U_1, v_1U_1 + v_2U_2, U_2) \tag{8}$$

with a similar version for $G_{2,1100}(U_1,U_2,v_1,v_2)$.

Equation (8) implies a fast method for forward-projection, so from equations (7) and (8) we thus have methods for fast forward-projection and backprojection. These are suitable for incorporation into the 3D-OSEM algorithm [4, 5], represented here in two steps:

$$g_i^{(k)} = \sum_{j'} a_{ij'} f_{j'}^{(k)}$$
$$f_j^{(k+1)} = \frac{f_j^{(k)}}{\sum_{i\in S_N} a_{ij}} \left\{ \sum_{i\in S_N} a_{ij} \frac{g_i}{g_i^{(k)}} \right\} \tag{9}$$

where $f_j^{(k)}$ is the k-th estimate of the value of image voxel j, g_i is the measured data in planogram bin i, and $g_i^{(k)}$ is the k-th estimate of the value of image voxel j. The probability of an event from image voxel j being detected in planogram bin i is given by a_{ij} using those LORs comprising the data subset S_N [6]. The first part of equation (9) can be regarded as a forward-projection, which can be computed with equation (8). Typically unmeasured (truncated) data are also estimated, which can be discarded. The portion within braces of the second part of equation (9) can be regarded as the backprojection of the ratio of the measured/estimated data. This can be computed using equation (7). We note that this approach corresponds to using a line integral model of the a_{ij}, and that using equations (7) and (8) does not strictly correspond to the same a_{ij} due to numerical discretization errors, that is we are using un-matched backprojector/forward projector pairs.

Implementation

The accuracy of using equation (8) for forward-projection is illustrated in figure 2.

The version of 3D-OSEM described by equation (9) was implemented using equations (7) and (8) for backprojection and forward-projection. The use of two detector positions leads to a simple sorting of the data into two subsets. The results of this approach are illustrated in figure 3. We note that the use of only two subsets is not an optimal implementation as figure 3 that 16 iterations are required to reach a minimum RMS error estimate.

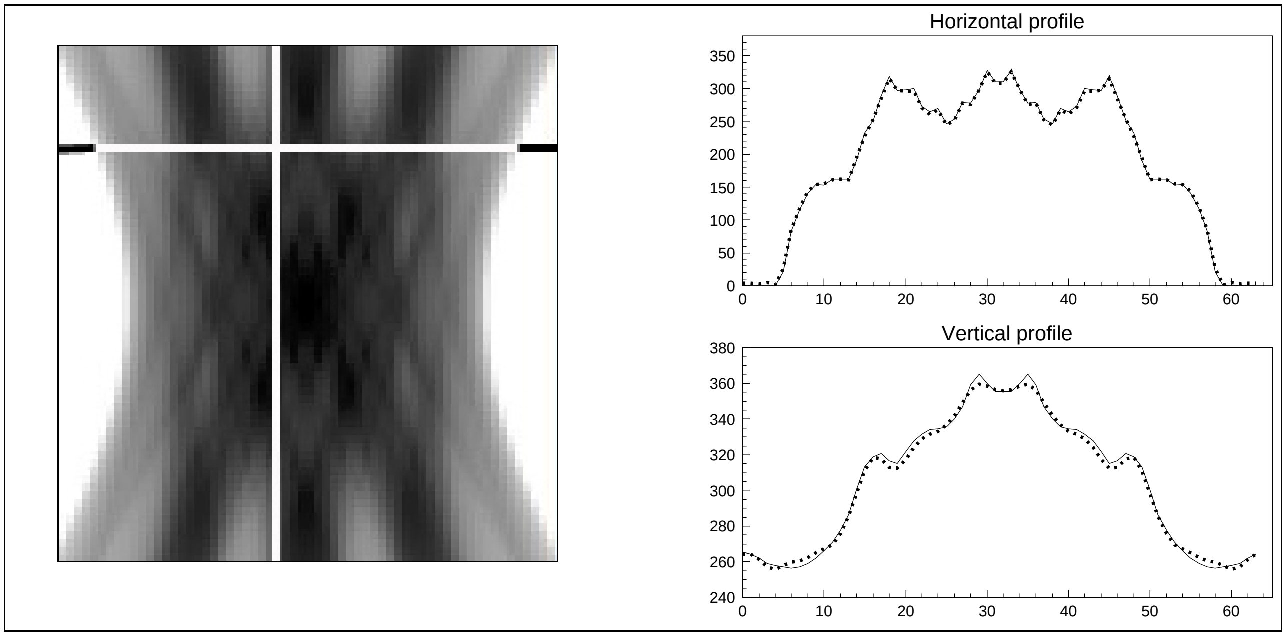
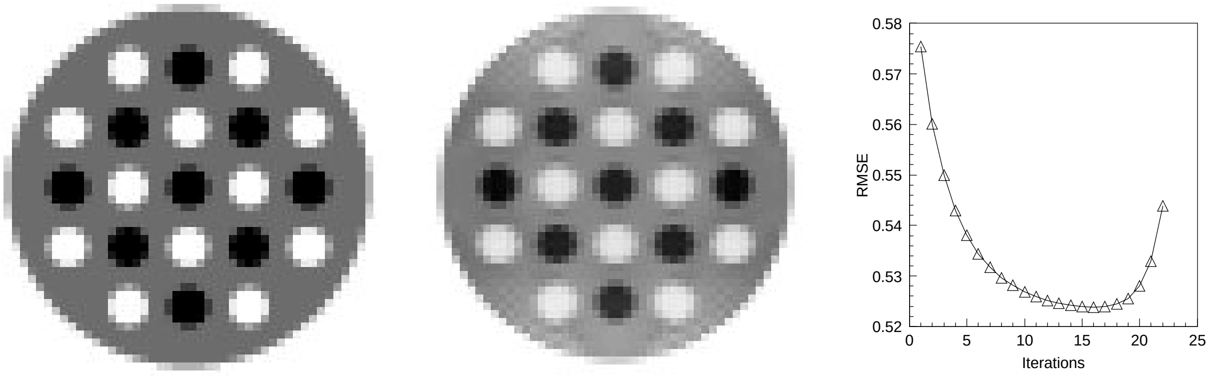


Figure 2. Forward projection planogram data using equation (9). Solid lines in the image indicate the positions of the profiles shown on the right for the Forward projected data (dotted line) compared to the original simulated planogram data (solid line).



Transverse section of original object

Estimated object after 16 iterations

RMS error between original object and estimated object at each iteration

Figure 3. 3D-OSEM method applied to a noiseless spherical object with imbedded hot and cold spheres for two subsets with up to 20 iterations.

This approach can be used for any iterative method using repeated 3D backprojection and forward-projection. We will present further results on the relative timing of this approach for iterative 3D reconstruction methods as well as the performance in the presence of statistical noise and with increasing numbers of subsets.

References

[1] Edholm PR and Herman GT, "Linograms in Image Reconstruction from Projections," *IEEE Trans Med Imag*, Vol. MI-6, pp. 301-307, 1987.

[2] Magnusson M, "Linogram and Other Direct methods for Tomographic Reconstuction," Ph.D. Thesis, Linköping University, Linköping, Sweden, 1993.

[3] Brasse D, et al., "Fast Fully 3D Image Reconstruction Using Planograms," Proceedings of: IEEE Nuclear Science Symposium and Medical Imaging Conference, Lyon, France, 2000, pp.

[4] Johnson CA, et al, "Evaluation of 3D reconstruction algorithms for a small animal PET scanner," *IEEE Trans Nuc Sci*, Vol. 44, pp. 1303-1308, 1997.

[5] Liu X, et al. "Comparison of 3D reconstruction with 3D-OSEM and with FORE+OSEM for PET," *IEEE Trans Med Imag*, Vol. , pp. (to appear), 2001.

[6] Hudson H and Larkin R, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans Med Imag*, Vol. 13, pp. 601-609, 1994.

Exact Inversion of the Exponential X-ray Transform for RSH SPECT

J.-M. Wagner,[1] F. Noo,[2] R. Clackdoyle[2] *

[1]University of Liège, Belgium [2]University of Utah, USA

Abstract

The RSH SPECT scanner provides parallel-beam attenuated projections for a fully 3D acquisition geometry. The geometry can be represented by circles on the unit sphere of projection directions, one circle for each position of the detector head. Unlike most other fully 3D geometries this one is particularly challenging because there are no 2D subsets in the data. When no attenuation is present, it is well-known that an unmeasured projection can be synthesized if it lies inside one of the measured circles. The main result of this work is that under some assumptions on the attenuation distribution, *attenuated* projections within a circle can be synthesized from available attenuated projections. One consequence is that RSH SPECT projections can be rebinned into a conventional SPECT geometry for which analytic attenuation correction techniques are available.

1 Introduction

In Single Photon Emission Computed Tomography (SPECT) imaging the objective is to visualize the concentration of a radioactive tracer within the 3D body under investigation. The limitations in SPECT are essentially due to attenuation of the photons and to the poor sensitivity of the collimator-detector system. A number of researchers have been considering alternatives to the conventional parallel-hole collimator (figure 1a).

The use of a rotating slant-hole (RSH) collimator with two (figure 1b) or four segments significantly increases the detection sensitivity by allowing a higher photon count during the same acquisition period [1]. The collimator-detector system is sucessively placed at different angular positions around the body to be studied. For each of these positions, the collimator is rotated about its center, while allowing several projections (two or four according to the collimator used) to be acquired simultaneously. This acquisition mode constitutes the RSH SPECT geometry as described in [1].

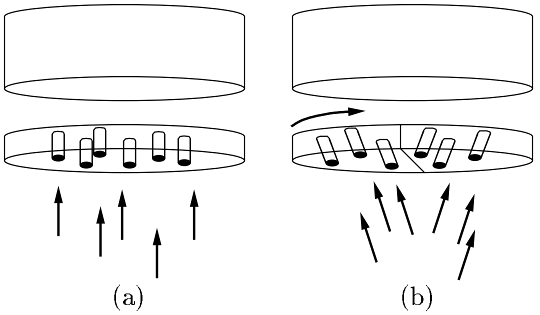


Figure 1: Different types of collimators. a) conventional parallel-hole collimator, b) RSH collimator.

The RSH SPECT scanner provides a set of attenuated projections. The exponential X-ray transform is a mathematical tool used in SPECT reconstruction for modeling and correcting for attenuation. Using the exponential X-ray transform, it is possible to reconstruct the emission map with attenuation correction, assuming the attenuation is constant in the emission region. Attenuated projections can be converted to exponential X-ray projections using a well-known point-by-point scaling [2]. Moreover, even if the attenuation map is unknown, the consistency conditions of the exponential X-ray transform provide an effective method to find the scaling coefficients [3].

In two-dimensions (2D), image reconstruction from exponential X-ray projections has been thoroughly studied over the past twenty-five years and is now well understood, especially due to the works of Tretiak & Metz [9] and Pan & Metz [10]. A very recent work [11] provides an inversion formula for the case of only 180-degrees of exponential data. The acquisition geometry for RSH SPECT is a fully 3D geometry for which an inversion formula for

---

*This work was partially supported by the National Institutes of Health, grant number R01 HL55610.

the X-ray transform has not yet been established. To our knowledge, only a few simple geometries have been treated. In [4] and [5], the exponential projections must be finely sampled on the unit sphere while the algorithm described in [6] only handles collections of projections on any subset of the unit sphere described as a union of great circles. Currently, we do not know if it is possible to obtain exact reconstructions from more general collections, such as those satisfying Orlov's condition [7] for the non-attenuated case.

We give a description of the general RSH geometry in section 2. In section 3, we generalize Orlov's result [7] and establish a rebinning technique which allows us to calculate new exponential projections (with any attenuation coefficient) from exponential projections given on a circle in the RSH geometry. Finally, in section 4, we use these new theoretical results to obtain a method of exact reconstruction from a set of complete exponential projections (complete in the sense of Orlov [7]) for the RSH SPECT geometry.

2 The RSH SPECT Geometry

Figure 2 illustrates the RSH SPECT geometry. In this figure, O is the origin of the unit sphere S^2. The orientation of the detector with respect to the origin is given by its unit normal vector $\underline{c}_i$. We assume N different positions of the detector ($i = 1, 2, \dots, N$). Also, we assume that for each position i of the detector, the collimator has a slant angle equal to α_i. The unit vector $\underline{n}$ defines the direction of photon propagation through the collimator. When it effects a rotation of 360 degrees about itself, the vector $\underline{n}$ describes, on the unit sphere, a circle C_i of angular aperture α_i and whose axis of symmetry is the vector $\underline{c}_i$. The RSH SPECT geometry is mathematically defined by the trajectory Ω on the unit sphere corresponding to the union of all the circles C_i.

After conversion from attenuated projections, the data available for image reconstruction are the exponential X-ray projections

$$p_\mu(\underline{n}, \underline{s}) = \int_{-\infty}^{+\infty} dt\, f(\underline{s} + t\underline{n})\, e^{-\mu t}, \qquad \underline{s}.\underline{n} = 0 \qquad (1)$$

for the directions $\underline{n} \in \Omega$. The 3D image to reconstruct is f while μ is the known, constant attenuation coefficient.

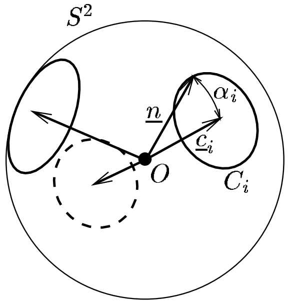


Figure 2: Illustration of the RSH SPECT geometry.

3 Theory

We show here that it is possible to calculate any parallel projection (with any finite attenuation coefficient μ_1) of whose direction $\underline{\alpha}$ is situated in the region of the unit sphere bounded by one of the circles making up Ω. In section 4 we show how the rebinning technique can be used to create, from the RSH SPECT data, a collection of exponential projections corresponding to a conventional parallel-hole SPECT geometry for which an exact inversion formula exists. An example reconstruction is given.

Let us consider one of the circles C_i making up Ω and let $\underline{\alpha}$ be a unit vector situated in the region of S^2 bounded by C_i. Figure 3 illustrates the situation. We use $\mathcal{C}(\underline{\alpha})$ to denote the great circle orthogonal to $\underline{\alpha}$ and we introduce four unit vectors $\underline{a}$, $\underline{b}$, $\underline{\theta}$ and $\underline{\theta}^\perp$ all lying on the great circle $\mathcal{C}(\underline{\alpha})$. The vectors $\underline{a}$ and $\underline{b}$ are defined mathematically by

$$\underline{a} = \frac{\underline{c}_i \times \underline{\alpha}}{||\underline{c}_i \times \underline{\alpha}||}, \quad \underline{b} = \underline{a} \times \underline{\alpha} \qquad (2)$$

while

$$\left\{ \begin{array}{rcl} \underline{\theta} & = & \cos\theta\, \underline{a} + \sin\theta\, \underline{b} \\ \underline{\theta}^\perp & = & -\sin\theta\, \underline{a} + \cos\theta\, \underline{b} \end{array} \right. \qquad (3)$$

where θ belongs to the interval $[0, 2\pi]$. For the case where the vector $\underline{\alpha}$ corresponds to $\underline{c}_i$ we choose $\underline{a}$ arbitrarily on $\mathcal{C}(\underline{\alpha})$.

The great circle orthogonal to $\underline{\theta}$, denoted $\mathcal{C}(\underline{\theta})$, cuts the circle C_i at a point $\underline{n}$ given by

$$\underline{n} = \cos\psi(\theta)\, \underline{\alpha} + \sin\psi(\theta)\, \underline{\theta}^\perp \qquad (4)$$

where $\psi(\theta) \in]0, \pi[$ for all $\theta \in [0, 2\pi[$. We show that the function $\psi(\theta)$ is given by

$$\tan\frac{\psi(\theta)}{2} = \frac{-\sin(k\alpha_i)\, \cos\theta + \sqrt{\sin^2\alpha_i - \sin^2(k\alpha_i)\, \sin^2\theta}}{\cos\alpha_i + \cos(k\alpha_i)} \qquad (5)$$

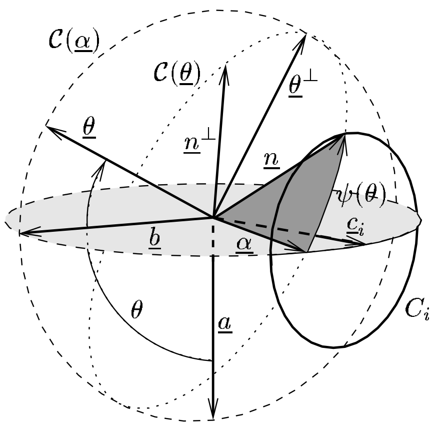


Figure 3: Rebinning from a set of exponential projections corresponding to a circle C_i.

where $(k\alpha_i)$ is the angle between $\underline{c}_i$ and $\underline{\alpha}$ ($k \in [0,1[$). Finally, we introduce the vector $\underline{n}^{\perp} = \underline{\theta} \times \underline{n} = -\sin\psi(\theta)\,\underline{\alpha} + \cos\psi(\theta)\,\underline{\theta}^{\perp}$.

From exponential projections $p_{\mu}(\underline{n}, \underline{s})$ known for $\underline{n} \in C_i$, it is possible to calculate the exponential projection $p_{\mu_1}(\underline{\alpha}, \underline{s})$ for any finite value μ_1 of the attenuation coefficient. To this end we show that

$$\int_{-\infty}^{+\infty} p_{\mu_1}(\underline{\alpha}, l\underline{\theta}+s\underline{\theta}^{\perp})e^{-\mu_2(\theta)s}ds = \int_{-\infty}^{+\infty} p_{\mu}(\underline{n}, l\underline{\theta}+t\underline{n}^{\perp})e^{-\mu_d(\theta)t}dt \tag{6}$$

where

$$\mu_2(\theta) = \frac{\mu - \mu_1\ \cos\psi(\theta)}{\sin\psi(\theta)} \quad \text{and} \quad \mu_d(\theta) = \frac{\mu\ \cos\psi(\theta) - \mu_1}{\sin\psi(\theta)} \tag{7}$$

for all $(\theta, l) \in [0, 2\pi[\, \times\,]-\infty, +\infty[$. The expression on the left of (6) is the 2D exponential Radon transform (with attenuation μ_2 depending on projection angle, and which we will denote AD-ERT for Angle Dependent Exponential Radon Transform) of the projection $p_{\mu_1}(\underline{\alpha}, \underline{s})$ while the expression on the right constitutes a sample of the 2D exponential Radon transform (with attenuation μ_d depending on angle) of the projection $p_{\mu}(\underline{n}, \underline{s})$. A sample of the AD-ERT of the projection $p_{\mu_1}(\underline{\alpha}, \underline{s})$ is therefore obtained by judiciously integrating in the plane of one of the available projections. By applying the relation (6) for all (θ, l), we obtain the AD-ERT of the projection $p_{\mu_1}(\underline{\alpha}, \underline{s})$, which we denote $g(\theta, l)$. The problem of the inversion of the AD-ERT was resolved in [8], and allows us to reconstruct $p_{\mu_1}(\underline{\alpha}, \underline{s})$ from $g(\theta, l)$:

$$p_{\mu_1}(\underline{\alpha}, \underline{s}) = \int_0^{2\pi} d\theta\, g^F(\theta, \underline{s}.\underline{\theta})\, e^{\mu_2(\theta)\underline{s}.\underline{\theta}^{\perp}} \tag{8}$$

where $g^F(\theta, l)$ is given in the Fourier domain by $G^F(\theta, \sigma) = G(\theta, \sigma)\, H(\theta, \sigma)$ and

$$H(\theta, \sigma) = \begin{cases} |\sigma| - j\,\mathrm{sgn}(\sigma)\,\mu_2'(\theta) & \text{if } |\sigma| > |\mu_2(\theta)|/(2\pi) \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

where $\mu_2'(\theta)$ is the derivative of $\mu_2(\theta)$ and $j = \sqrt{-1}$.

We note that although any value of μ_1 can be used in principle, it is safer in practice to use $\mu_1 = \mu$ to control the behavior of $\mu_2(\theta)$.

4 Simulations and Results

In this section, we study an example RSH SPECT geometry. The set Ω under consideration consists of 3 circles whose centers $\underline{c}_i$ are situated on the great circle $\mathcal{C}(\underline{e}_z)$ at regular intervals of 60 degrees and with a slant angle of 30 degrees (for $i = 1, 2, 3$). Figure 4 illustrates the situation. According to Orlov [7], this set is complete because all great circles on the unit sphere intersect Ω. An exact reconstruction is therefore possible in the non-attenuated case. We show below that exact reconstruction is also possible for the attenuated case.

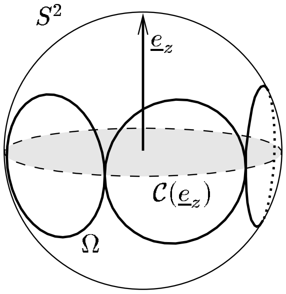


Figure 4: An example RSH SPECT geometry.

There are 3 distinct positions of the detector and we have simulated 32 angular positions of the collimator rotating about its own axis, which makes a total of $3 \times 32 = 96$ simulated attenuated projections for this RSH SPECT geometry. A phantom modeling the heart was used for the emission map. It was composed of three ellipsoids, two of which modeled the ventricles with 20% of the specific activity of the myocardium. The attenuation map was modeled with 4 ellipsoids, representing the thorax, the two lungs, and the spinal column. Figure 5 shows the emission and attenuation maps for two slices in different orientations. Each attenuated projection was sampled on a grid of 100^2 pixels of side 1.5 mm. The attenuation coefficients being constant in the emission region, we converted the attenuated projections to exponential projections with $\mu = 0.15/\text{cm}$.

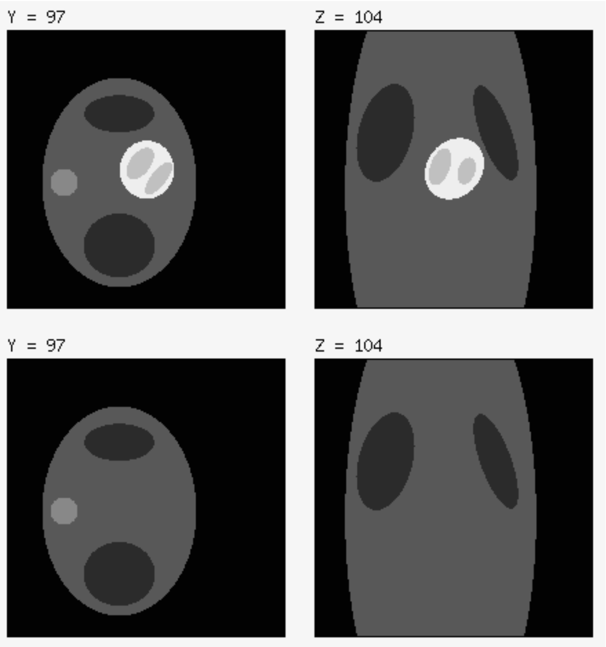


Figure 5: Top: Emission and attenuation maps superimposed. Bottom: Attenuation map only.

We applied the rebinning method (with $\mu_1 = \mu$) to calculate the exponential projections $p_\mu(\underline{\alpha}, \underline{s})$ for 181 directions $\underline{\alpha}$ uniformly sampled on the half of the great circle $\mathcal{C}(\underline{e}_z)$ with an angular step of 1 degree. The calculated projections were reconstructed on a grid of 100^2 pixels of side 1.5 mm. The details of the reconstruction algorithm will be given at the conference. Figure 6 shows one of the 181 exponential projections calculated by the rebinning method. This figure illustrates the exactness of the rebinning method.

The 181 calculated exponential projections constitute a set of projections in the conventional (180-degree) parallel-hole collimator SPECT geometry. For this configuration an exact inversion formula now exists ([11]). We have applied this algorithm to reconstruct, slice by slice perpendicular to $\underline{e}_z$, the image f on a grid of 100^3 voxels of side 1.5 mm. The quality of the reconstruction given in figure 7 illustrates the efficacy of the reconstruction method.

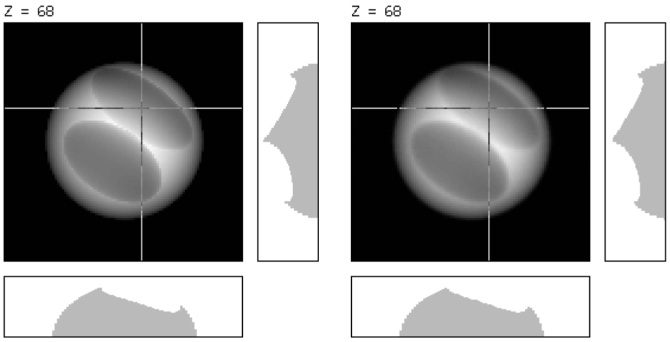


Figure 6: Illustration of the rebinning method ($k = 0.5$). Left: Ideal projection. Right: Synthesized projection.

Reconstructions with simulated noisy data have also been successful, and reconstructions with data from a prototype RSH SPECT scanner is being submitted to the 2001 IEEE MIC conference.

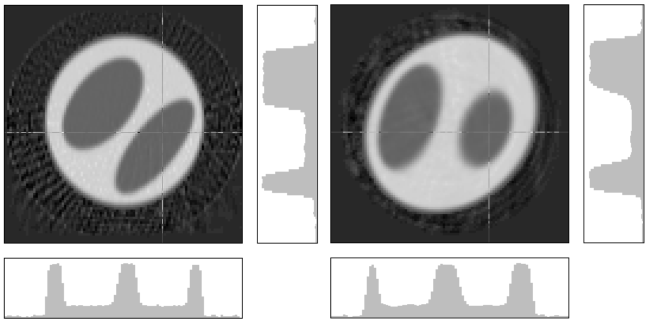


Figure 7: Heart phantom reconstruction for $\mu = 0.15/\text{cm}$.

References

[1] R. Clack, P.E. Christian, M. Defrise, A.E. Welch, "Image reconstruction for a novel SPECT system with rotating slant-hole collimators". In Conf. Rec. 1995 IEEE Med. Imag. Conf., 1948-1952, 1996.

[2] A. Markoe, "Fourier inversion of the attenuated X-ray transform", SIAM J. Math. Anal., Vol. 15(4), 718-722, 1984.

[3] C. Mennessier, F. Noo, R. Clack, G. Bal, L. Desbat, "Attenuation correction in SPECT using consistency

conditions for the exponential ray transform", Phys. Med. Biol., Vol. 44, 2483-2510, 1999.

[4] I.A. Hazou, D.C. Solmon, "Inversion of the exponential X-ray transform. I:Analysis", Math. Methods in the Applied Sciences, Vol. 10(10), 561-574 (1988).

[5] Y. Weng, G.L. Zeng, G.T. Gullberg, "Filtered backprojection algorithm for attenuated parallel and cone-beam projections sampled on a sphere", in Three-dimensional Image Reconstruction In Radiation and Nuclear Medicine, ed. P. Grangeat and J.-L. Amans (Dordrecht: Kluwer), 19-34, 1996.

[6] J.-M. Wagner, F. Noo, "TTR algorithm for the inversion of the exponential X-ray transform". In Conf. Rec. IEEE 2000 Med. Imag. Conf., to be published.

[7] S.S. Orlov, "Theory of three dimensional reconstruction. 1. Conditions of a complete set of projections.", Sov. Phys.-Crystallogr., Vol. 20, 312-314, 1975.

[8] P. Kuchment, I. Shneiberg, "Some Inversion formulas in the Single Photon Emission Computed Tomography", Appl. Anal., vol. 53, 221-231, 1994.

[9] O. Tretiak, C. Metz, "The exponential Radon transform", SIAM, J. Appl. Math., Vol. 39(2), 341-354, 1980.

[10] C.E. Metz, X. Pan, "A unified analysis of exact methods of inverting the 2D exponential Radon tranform, with implications for Noise control in SPECT", IEEE Trans. Med. Imag., vol. 14(4), 643-658, 1995.

[11] F. Noo, J.-M. Wagner, "Image reconstruction in 2D SPECT with 180-degree acquisition," submitted to Inverse Problems.

Optimization of Penalized Block-iterative Algorithms for Ga-67 Tumor Detection

H.C. Gifford[1], C.L. Byrne[2], M.V. Narayanan[1], and M.A. King[1]
[1]Department of Radiology, University of Massachusetts Medical School, Worcester, MA, 01655, USA
[2]Department of Mathematics, University of Massachusetts-Lowell, Lowell, MA, 01854, USA

Abstract— **Our purpose is to optimize penalized block-iterative algorithms for detection of Ga-67 tumors in the thorax. We illustrate some of our methods involving psychophysical studies with results from a preliminary channelized Hotelling observer (CHO) optimization for a one-step-late (OSL) version of the penalized RBI-EM algorithm that features a 3D uniform quadratic penalty function. The algorithm parameters to be optimized are iteration number and a penalty weighting parameter β. Use of the CHO is an efficient means of bounding regions of parameter space that contain the optimal parameters, although the final determination of the optimal parameters will be left to human psychophysical studies. From the CHO optimization, we found that one iteration of the penalized RBI-EM algorithm could outperform a previously optimized reconstruction strategy with RBI-EM and post-filtering that required three iterations.**

I. INTRODUCTION

By most definitions of image quality, unregularized reconstructions of SPECT data are substandard. Procedures for improving the quality of iterative reconstructions usually involve penalty functions, stopping rules, and post-reconstruction filters. Whether a particular regularization scheme actually improves an image will depend on one's image quality measure. In a series of earlier papers [1-3], we documented optimizations of the ordered-sets expectation-maximization (OSEM) [4] and rescaled block-iterative EM (RBI-EM) [5] algorithms that were aimed at determining if detection of Ga-67 tumors in the thorax could be improved by modeling the physics of the data acquisition in the inverse problem. These algorithms used a post-reconstruction Gaussian filter as a regularizer. This current abstract describes the early stages of similar optimizations of algorithms derived by incorporating 3D penalty-function regularizers into the RBI-EM algorithm and the block-iterative interior point algorithm (IPA) [6]. Such penalized-likelihood methods are often viewed in the context of maximum a posteriori (MAP) reconstruction [7].

One motivation for investigating these penalized algorithms comes from the results of our work with the OSEM and RBI-EM algorithms. Based on detection-performance metrics drawn from localization ROC (LROC) [8] studies with human observers, it was determined [2] that the OSEM algorithm with nonuniform attenuation correction (AC) and three-dimensional (3D) detector-response correction (DRC) did improve detection in comparison to the OSEM algorithm with only AC or with neither correction. Even so, it was also seen that a significant difference in performance existed between our DRC reconstructions and "ideal" reconstructions that represented an upper bound on DRC. The Gaussian post-filter may have contributed to this difference by partially negating the effect of the DRC in our reconstructions. If this were true, we might expect to find that regularization with edge-preserving penalty functions [7, 9] boosts observer performance over that obtained with the post-filter regularization.

To show that one algorithm offers improvement over another in a fair comparison first requires optimization of each algorithm with respect to its adjustable parameters. We perform these preliminary optimizations using a combination of channelized Hotelling observer (CHO) [10] ROC and human-observer LROC studies. Use of the CHO offers an efficient means of coarsely bounding regions of parameter space that contain the optimal parameters. The actual determination of those optimal parameters is then left to the LROC studies, but the workload for the human observers is reduced by having applied the CHO. In this abstract, we illustrate this approach with results from a preliminary CHO optimization for a one-step-late (OSL) version of the penalized RBI-EM algorithm that features a 3D uniform quadratic penalty function. The adjustable parameters to be optimized in this case are iteration number and a penalty weighting parameter β. This penalty function does not have edge-preserving characteristics, but allows comparison to our past studies with the Gaussian post-filter. Edge-preserving penalty functions are also to be investigated for use with the penalized RBI-EM and IPA algorithms. The inclusion of the penalized IPA algorithm in this research provides a reconstruction method that is not based on the OSL approximation.

II. METHODS

A. The Phantom

A distribution of Ga-67 citrate in the chest was simulated with one geometry of the 3D mathematical cardiac-torso (MCAT) phantom [11]. Background activity levels in the phantom were set by reference to clinical values. Tumors were modeled as 1-cm diameter spheres with a tumor-to-soft-tissue activity ratio of 20:1. Each "abnormal" case contained one tumor, randomly placed within a 3D map of likely tumor areas.

B. The Projection Data

Separate analytic projections of the phantom and tumors were created and then combined into a noise-free tumor-present projection sets. The projector modeled nonuniform attenuation in the phantom and the response for a medium-energy parallel-hole collimator. For the results presented herein, perfect scatter rejection was assumed so that comparisons with our previous optimizations of OSEM and RBI-EM could be made, but simulations that include scatter are also being done. Separate projections for 93- and 185-keV photons were obtained using energy-specific attenuation maps, and then added as a weighted sum based on the relative abundances and camera efficiencies for these energies. Projection sets consisted of 128 256 × 256 projections (pixel width of 0.1585 cm), rebinned to 128 × 128 pixels (pixel width of 0.317 cm). Poisson noise was added to form data sets of 8 million counts.

C. The Reconstruction Algorithm

For this abstract, we consider a OSL penalized version of the RBI-EM algorithm [5]. Like the OSEM algorithm, the RBI-EM algorithm is a block-iterative version of the maximum-likelihood expectation maximization (MLEM) algorithm, but unlike the OSEM algorithm, it converges in the case of consistent data for any choice of subsets. Our human-observer LROC studies have also shown that the RBI-EM algorithm is more robust than the OSEM algorithm as the number of projection data subsets is increased in order to accelerate the reconstruction [3]. To define the penalized RBI-EM algorithm, we let $\hat{\mathbf{f}}^{(j)}$ be the j^{th} iteration ($j = 0, ..., \infty$) of $\mathbf{f}$, with individual elements $\hat{f}_n^{(j)}$, and consider the detector pixels divided into R disjoint subsets $S_1, \ldots, S_R$. The convention followed here is that every cycle through the R subsets represents an iteration. An RBI-EM algorithm regularized with penalty function $U(\hat{\mathbf{f}}^{(j)})$ requires solving the equation

$$\hat{f}_n^{(j+1)} = \frac{\hat{f}_n^{(j)}}{\sum_m H_{mn} + \beta \frac{\partial U(\hat{\mathbf{f}}^{(j+1)})}{\partial \hat{f}_n^{(j+1)}}} \times \left[\sum_m H_{mn} + a_r{}^{-1}\left(\sum_{m\in S_r} H_{mn}\left(\frac{g_m}{[\mathbf{H}\hat{\mathbf{f}}^{(j)}]_m} - 1\right)\right)\right], \tag{1}$$

for $\hat{f}_n^{(j+1)}$, where $r = (j \bmod R) + 1$, β is the nonnegative penalty weight, and

$$a_r \equiv \max_n \left(\sum_{m\in S_r} H_{mn} / \sum_m H_{mn} \right). \tag{2}$$

For arbitrary U, a closed-form expression for $\hat{f}_n^{(j+1)}$ is unobtainable from Eq. (1). One way around this difficulty is to replace $U(\hat{\mathbf{f}}^{(j+1)})$ with $U(\hat{\mathbf{f}}^{(j)})$, which is referred to as the OSL approach [7].

Our uniform quadratic penalty function has the form

$$U(\hat{\mathbf{f}}^{(j)}) = \frac{1}{2} \sum_{k\in V} \sum_{i\in N_k} w_{ik} (\hat{f}_k^{(j)} - \hat{f}_i^{(j)})^2, \tag{3}$$

where V is the full set of voxels in the 3D reconstruction, N_k represents the 3 × 3 × 3 cube of voxels centered about the k^{th} voxel, and w_{ik} is the distance between the centers of the k^{th} and i^{th} voxels in N_k, normalized by the voxel width. The optimization parameters for this choice of U are the number of iterations and the magnitude of β. Each combination of iteration number and β shall be referred to as a reconstruction strategy. For a point of comparison, our past studies [3] found optimal observer detection performance with 3 iterations (32 subsets of 4 projection angles each) of the RBI-EM algorithm combined with a 3D Gaussian post-filter with a FWHM of 0.95 cm.

Reconstructions with matrix dimensions of 128 × 128 × 128 were obtained with strategies using one iteration of Eq. (1) and four projection angles per subset for values of β between 0.0 and 1.5. Both the AC and the DRC procedures that had been used in [2] were included in the iteration. The attenuation map used for the AC was the same one used in the creation of the projection sets, and in that sense was an optimal attenuation correction [12]. No post-reconstruction filter was applied. For LROC studies, 2D images through the center of the tumor are extracted from the 3D reconstructions. These are then adaptively thresholded, interpolated to 256 × 256 pixels, and then converted from floating-point to byte images for display to the human observers. For the CHO, the same 2D slices were extracted, but the thresholding, interpolation, and conversion to greyscales were not applied. Excluding these steps reduced the computing time and also allowed use of a low-noise approximation for constructing the CHO that is described in Section II-D. For OSEM-reconstructed images, we have found [13] that omitting these processes has little impact on the correlation between the CHO and average human observer.

D. Observer Studies

Our evaluation of reconstruction algorithms uses LROC methodology to measure human-observer performance of the Ga-67 tumor detection task. In a standard ROC study, an observer's response for an image is a confidence rating that a tumor is either present or absent. For an LROC study, the observer is asked to give this confidence rating and the coordinates of the maximally suspicious tumor location. By implementing this detection and localization task, LROC offers both a better approximation of clinical detection tasks and increased statistical accuracy over ROC for measuring detection performance [8]. With this improvement in accuracy comes increased statistical power for discriminating between reconstruction strategies.

The CHO is a linear discriminant function that has been shown [14-18] to correlate with humans for many "signal-known-exactly" (SKE) detection tasks, in which the observer knows the tumor location from the outset. We do not yet have a modified version of the CHO that can perform the search and detection task associated with these LROC studies and still correlate with humans, but we have found that the average CHO performance in series of SKE-ROC studies with multiple tumor location agrees with average human performance in the sense that both indicate similar significant differences between reconstruction strategies [3, 13, 19].

For a given tumor location, the response of the CHO to an image is a scalar λ that is a weighted sum of the image voxels. This sum can be viewed as the inner product between the image and a template image $\mathbf{w}_{\mathrm{cho}}$. Let $\hat{\mathbf{f}}$ be a 3D reconstruction from noisy data that, with equal likelihood, comes from an ensemble $\mathcal{H}_0$ of tumor-absent reconstructions or an ensemble $\mathcal{H}_1$ of tumor-present reconstructions. Post-reconstruction processing such as extraction of the transaxial slice of interest is performed by operator $\mathcal{S}$, so that $\mathcal{S}\hat{\mathbf{f}}$ is a 2D image intended for the observer. Then $\lambda = \mathbf{w}_{\text{cho}}^{\dagger}\mathcal{S}\hat{\mathbf{f}}$, where the † indicates a vector transpose. Comparison of λ to a threshold value determines whether the CHO decides the image contains a tumor. The CHO's overall performance can be quantified by a signal-to-noise ratio for λ [20],

$$\text{SNR}^2 = \frac{[\langle\lambda|\mathcal{H}_1\rangle - \langle\lambda|\mathcal{H}_0\rangle]^2}{\frac{1}{2}\text{var}(\lambda|\mathcal{H}_1) + \frac{1}{2}\text{var}(\lambda|\mathcal{H}_0)}, \tag{4}$$

where $\langle\lambda|\mathcal{H}_i\rangle$ and $\text{var}(\lambda|\mathcal{H}_i)$ $(i = 0, 1)$ are the mean and the variance of the distribution of λ for the i^{th} ensemble.

The CHO template is composed of a matrix $\mathbf{U}$ of impulse responses for a set of C 2D prefilters (the "channels") centered on the tumor location, a channel covariance matrix $\mathbf{K}_{\text{chan}}$, and a matched filter. For ensemble $\mathcal{H}_i$, let $\overline{\hat{\mathbf{f}}}_i$ be the mean 3D reconstruction and let $\mathbf{K}_i$ be the voxel covariance matrix. Assuming $\mathcal{S}$ applies only linear post-processing, we make the definition

$$\mathbf{K}_{\text{chan}} = \frac{1}{2}\mathbf{U}\mathcal{S}[\mathbf{K}_0 + \mathbf{K}_1]\mathcal{S}^{\dagger}\mathbf{U}^{\dagger}, \tag{5}$$

and express the CHO template as [10]

$$\mathbf{w}_{\text{cho}} = \mathbf{U}^{\dagger}\mathbf{K}_{\text{chan}}^{-1}\mathbf{U}\mathcal{S}(\overline{\hat{\mathbf{f}}}_1 - \overline{\hat{\mathbf{f}}}_0). \tag{6}$$

Substituting Eq. (6) into Eq. (4) leads to the formula

$$\text{SNR}_{\text{cho}} = \left[(\overline{\hat{\mathbf{f}}}_1 - \overline{\hat{\mathbf{f}}}_0)^{\dagger}\mathcal{S}^{\dagger}\mathbf{U}^{\dagger}\mathbf{K}_{\text{chan}}^{-1}\mathbf{U}\mathcal{S}(\overline{\hat{\mathbf{f}}}_1 - \overline{\hat{\mathbf{f}}}_0)\right]^{\frac{1}{2}} \tag{7}$$

for evaluating the CHO's performance.

The channel covariance and matched-filter portions of the CHO require knowledge of the ensemble statistics of $\mathcal{H}_0$ and $\mathcal{H}_1$, and determining these statistics accounts for the computational expense of using the CHO. Oftentimes, sets of noisy reconstructions are generated to produce sample statistics. In doing so, one makes use of the fact that the number of channels C (C=3 in our CHO model) is much smaller than the number of image voxels by computing $\mathbf{K}_{\text{chan}}$ directly from the $C \times C$ "channelized" images $\mathbf{U}\mathcal{S}\hat{\mathbf{f}}$ instead of estimating $\mathbf{K}_i$ in full. For complex iterative reconstruction algorithms, this approach is very computationally intensive [21]. An alternative is to make a low-noise assumption about the images [22-24] and thus approximate the ensemble statistics by a recursive procedure using noise-free data. Let $\hat{\mathbf{f}}^{(k)} \in \mathcal{H}_i$ be the k^{th} iteration with $\mathbf{g}$, and let $\mathbf{a}^{(k)}$ be the k^{th} iteration with the noise-free data $\overline{\mathbf{g}}$. Also, define $\mathbf{D}^{(k)}$ as the diagonal matrix with element $D_{jj}^{(k)} = a_j^{(k)}$, the j^{th} element of $\mathbf{a}^{(k)}$. When $\hat{\mathbf{f}}^{(k)} \approx \mathbf{a}^{(k)} + \boldsymbol{\epsilon}^{(k)}$ for small perturbations $\boldsymbol{\epsilon}^{(k)}$, then $\overline{\hat{\mathbf{f}}}_i \approx \mathbf{a}^{(k)}$ [22], and

$$\mathbf{K}_i \approx \mathbf{D}^{(k)}\ \mathbf{O}^{(k)}\ \mathbf{K}_g\ [\mathbf{O}^{(k)}]^{\dagger}\ \mathbf{D}^{(k)}. \tag{8}$$

The matrix $\mathbf{O}^{(k)}$ describes how noise in the data is transferred into the reconstruction through k iterations. We determined the specific form of $\mathbf{O}^{(k)}$ for the penalized RBI-EM algorithm by merging the results in [24] for MAP-type reconstuction algorithms with [22], which treats block-iterative methods such as the RBI-EM algorithm. As with the sample statistics approach described above, the channel covariance should be calculated directly by computing the product $\mathbf{A}^{\dagger}\mathbf{A}$, where

$$\mathbf{A} = \mathbf{K}_g^{\frac{1}{2}}[\mathbf{O}^{(k)}]^{\dagger}\mathbf{D}^{(k)}\mathcal{S}^{\dagger}\mathbf{U}^{\dagger}. \tag{9}$$

The SKE-ROC studies were conducted for 5 tumor locations. The overall CHO performance for a strategy was figured by converting SNR_{cho} for each location into an area under the ROC curve A_{Z} as described in [19] and then computing the average area $\langle\text{A}_{\text{Z}}\rangle$ over location. The computation time required to evaluate SNR_{cho} for a given strategy at a single location is on the order of the time required to perform $2(C + 1)$ reconstructions with that strategy.

III. RESULTS AND CONCLUSIONS

The results of the ROC studies are presented in Figure 1, in which $\langle\text{A}_{\text{Z}}\rangle$ is plotted against β for the penalized RBI-EM strategies that used a subset size of four angles per subset. The high values of $\langle\text{A}_{\text{Z}}\rangle$ seen for these ROC studies are partly due to the fact that the tumor contrast and the number of counts in the data sets were originally set for use in the LROC studies.

The solid horizontal line at $\langle\text{A}_{\text{Z}}\rangle \approx 0.98$ denotes the CHO performance using three iterations of the RBI-EM algorithm and a 3D Gaussian post-filter of FWHM = 0.95 cm. The best-performing ($\langle\text{A}_{\text{Z}}\rangle \approx 0.99$) of the strategies we tested used β=1.25, but several other strategies also outperformed the RBI-EM strategy. The effect on CHO performance of using more iterations and different numbers of subsets is currently being investigated. Also, whether the precipitous fall-off in performance above β=1.25 will find a correlation in human LROC studies remains to be seen.

IV. ACKNOWLEDGEMENTS

This work was supported in part by the National Cancer Institute (NCI) under grant number CA-42165.

References

[1] R G Wells, M A King, P Simkin, P Judy, A B Brill, H C Gifford, R Licho, P H Pretorius, P Schneider, and D W Seldin, "Comparing FBP and OSEM for small lesion detection and localization in Ga-67 SPECT," *J. Nucl. Med.*, vol. 41, pp. 1391–1399, 2000.

[2] H C Gifford, M A King, R G Wells, W G Hawkins, M V Narayanan, and P H Pretorius, "LROC analysis of detector-response compensation in SPECT imaging," *IEEE Trans. Med. Imaging*, vol. 19, pp. 463–473, 2000.

[3] H C Gifford, R G Wells, and M A King, "Effect of block-iterative acceleration on Ga-67 tumor detection in thoracic SPECT," *IEEE Trans. Nucl. Sci.*, (submitted).

[4] H M Hudson and R S Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imaging*, vol. 13, pp. 601–609, 1994.

[5] C L Byrne, "Block-iterative methods for image reconstruction from projections," *IEEE Trans. Imag. Proc.*, vol. 5, pp. 792–794, 1996.

[6] C L Byrne, "Block-iterative interior point optimization methods for image reconstruction from limited data," *Inverse Problems*, vol. 16, pp. 1405–1419, 2000.

[7] P J Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE Trans. Med. Imaging*, vol. 9, pp. 84–93, 1990.

[8] R G Swensson, "Unified measurement of observer performance in detecting and localizing target objects on images," *Med. Phys.*, vol. 23, pp. 1709–1725, 1996.

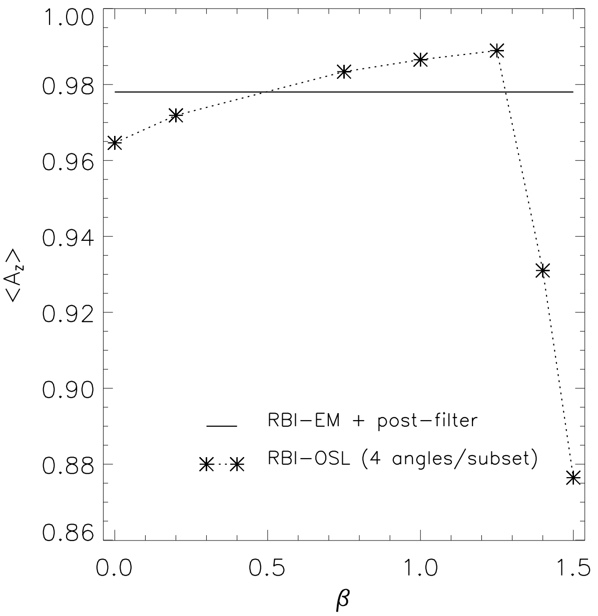


Fig. 1. A plot of $\langle \mathrm{A_Z} \rangle$ versus β for the penalized RBI-EM strategies with one iteration and four angles per subset. The horizontal line denotes CHO performance using the RBI-EM algorithm and post-filtering.

[9] D S Lalush and B M W Tsui, "Simulation evaluation of Gibbs prior distributions for use in maximum a posteriori SPECT reconstructions," *IEEE Trans. Med. Imaging*, vol. 11, pp. 267–275, 1992.

[10] K J Myers and H H Barrett, "Addition of a channel mechanism to the ideal-observer model," *J. Opt. Soc. Am. A*, vol. 4, pp. 2447–2457, 1987.

[11] B M W Tsui, X D Zhao, G K Gregoriou, J Li, D L Lalush, and R L Eisner, "Quantitative cardiac SPECT reconstruction with reduced image degradation due to patient anatomy," *IEEE Trans. Nucl. Sci.*, vol. 41, pp. 2838–2848, 1994.

[12] R G Wells, H C Gifford, P H Pretorius, T Farncombe, and M A King, "The impact of noisy attenuation maps and patient motion on human-observer performance at Ga-67 lesion detection in SPECT," in *2000 IEEE NSS-MIC Conference Record*, to appear.

[13] H C Gifford, E J Soares, R G Wells, and M A King, "Evaluating attenuation correction in Ga-67 SPECT image reconstruction through numerical observer ROC and human observer LROC," in *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2000.

[14] J P Rolland and H H Barrett, "Effect of random background inhomogeneity on observer detection performance," *J. Opt. Soc. Am.*, vol. 9, pp. 649–658, 1992.

[15] C K Abbey, H H Barrett, and D W Wilson, "Observer signal-to-noise ratio for the ML-EM algorithm," *Proc. SPIE*, vol. 2712, pp. 47–58, 1996.

[16] M P Eckstein, C K Abbey, and J S Whiting, "Human vs. model observers in anatomical backgrounds," *Proc. SPIE*, vol. 3340, pp. 16–26, 1998.

[17] S D Wollenweber, B M W Tsui, E C Frey, D S Lalush, and K J LaCroix, "Comparison of human and channelized Hotelling observers in myocardial defect detection in SPECT," *J. Nucl. Med.*, vol. 39, pp. 170P, 1998 (abstract).

[18] H C Gifford, M A King, D J de Vries, and E J Soares, "Channelized Hotelling and human observer correlation for lesion detection in hepatic SPECT imaging," *J. Nucl. Med.*, vol. 41, pp. 514–521, 2000.

[19] H C Gifford, R G Wells, and M A King, "A comparison of human observer LROC and numerical observer ROC for tumor detection in SPECT images," *IEEE Trans. Nucl. Sci.*, vol. 46, pp. 1032–1037, 1999.

[20] H H Barrett, "Objective assessment of image quality: effects of quantum noise and object variability," *J. Opt. Soc. Am. A*, vol. 7, pp. 1266–1278, 1990.

[21] E C Frey, K J LaCroix, and B M W Tsui, "Application of task-based measures of image quality to optimization and evaluation of three-dimensional reconstruction-based scatter compensation in SPECT," in *Proc. Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, Egmond aan Zee, The Netherlands, 1999, pp. 289–291.

[22] E J Soares, C L Byrne, and S J Glick, "Noise characterization of block-iterative reconstruction algorithms: I. Theory," *IEEE Trans. Med. Imaging*, pp. 261–270, 2000.

[23] H H Barrett, D W Wilson, and B M W Tsui, "Noise properties of the EM algorithm: I. Theory," *Phys. Med. Biol.*, vol. 39, pp. 833–846, 1994.

[24] W Wang and G Gindi, "Noise analysis of MAP-EM algorithms for emission tomography," *Phys. Med. Biol.*, vol. 42, pp. 2215–2232, 1997.

Analytical 3D Approach to Simultaneous Compensation for Photon Attenuation and Collimator Response in Quantitative Fan-Beam Collimated Brain SPECT

Tianfang Li, Jiansheng You, and Zhengrong Liang*

Abstract – **Fan-beam collimation provides the optimal geometry for data collection for brain SPECT imaging. For a non-parallel projection geometry, there is no symmetry on the projection rays, except the periodical property of the projection angle. This property is well defined by circular harmonic decomposition (CHD). This paper utilizes the CHD to explore the fan-beam collimation geometry, including photon attenuation and collimator response, for quantitative brain SPECT in three dimensions. An analytical solution compensating simultaneously for both of the photon attenuation and collimator responses is presented. An efficient algorithm for the solution is formulated and implemented by fast Fourier transforms. The reconstruction has been validated by experiments on the Shepp-Logan and Hoffman brain phantoms at various noise levels.**

***Index terms*—circular harmonic decomposition, point spread function, fan-beam collimation.**

I. INTRODUCTION

Single photon emission computed tomography (SPECT) can provide quantitative information of the tissue functionality in three-dimensions (3D). However, two major problems that are causing difficulties in image reconstruction for quantitative SPECT arise from the absorption of γ-rays by the body and the non-stationary point spread response of the collimator (the scatter of the γ-rays can be treated by other means and will not be discussed here). For brain SPECT, the attenuation problem can be simplified, because there are negligible radioactive nuclides distributed inside the skull and scalp and furthermore the attenuation of the skull and scalp can be equivalent to that of an enlarged brain tissue of a constant attenuation coefficient, as a compensation for uniform attenuation in a convex region[1]. Collimator blurring makes the image reconstruction more complicated, even for the uniform attenuation and parallel-hole collimators[2]. There does not exist a complete analytical algorithm by now that compensates accurately both the attenuation and the point spread response effects simultaneously, especially for the fan beam collimator geometry in 3D. L. van Elmbt and S. Walrand considered the problem for parallel geometry with approximated algorithm[3], while E. J. Soares et al attacked the problem of the same geometry for some particular resolution variation functions, such as the Cauchy model[4]. Other researchers correct either for the non-stationary resolution variation or for the constant attenuation, but not both. And most of these work are either for 2D applications or for parallel hole collimators[5-9].

Although many iterative reconstruction algorithms can do the job and are flexible to be applied for many kinds of complicated collimator geometry[10-12], the computing burden is always the drawback. If an analytical inversion formula can be derived for the solution of the projection equation, it is an interesting research topic for further investigation for practical use. The derivation itself is also an interesting research topic.

Fan-beam collimation is an optimal geometry for brain SPECT. It offers no symmetry for the projections, except the periodical property of projection angle. Circular harmonic decomposition (CHD) has been widely used to explore the property. In this paper, we present an analytical inversion solution that simultaneously compensates for both photon attenuation and collimator response of the fan-beam collimated SPECT system. This method considers the collimators blurring effect and intrinsic detector response together as a system point spread function (PSF). It doesn't limit the PSF in certain forms, but only based on the assumption that the PSF is valid to each individual collimator hole. Our method can be applied to parallel-hole, fan-beam, or varying focal-length fan-beam collimator geometry. It greatly reduced the burden of computational complexity. It was validated under a more realistic case of a Gaussian response function whose FWHM (full width at half maximum) is a function of relative distances (lateral and normal to the collimator surface) on both the Shepp-Logan ellipse phantom and Hoffman brain phantom.

II. THEORY

A typical fan-beam collimator geometry is shown by Figure 1, restricted on a single slice. The object activity distribution $f(x,y,z)$ is of our interest.

With the inclusion of uniform attenuation of the brain tissues and an 3D collimator response, the measured projection data $p_m(s,z,\theta)$ is expressed as:

$$p_m(s,\theta,z) = \iiint f(x',y',z') e^{-\mu d(s,\theta,z;x',y',z')} k(s,\theta,z;x',y',z') dx'dy'dz' \quad (1)$$

---

*T. Li is with the Department of Physics and Astronomy, State University of New York (SUNY), Stony Brook, NY 11790, USA.

J. You is with MD OnLine, Inc, Lexington, MA 02421, USA.

Z. Liang is with the Departments of Radiology and Computer Science, SUNY at Stony Brook.

where $k(.)$ is the system point spread function. Let the lateral distance to each individual collimator hole being labeled as s, see Figure 1, and a Cartesian coordinate system of l-t is chosen such that the t axis is parallel to the focusing direction of this hole, then the system response function $k(.)$ can be assumed as:

$$k = k\left(l' - l(s), z' - z, t' + R\right). \tag{2}$$

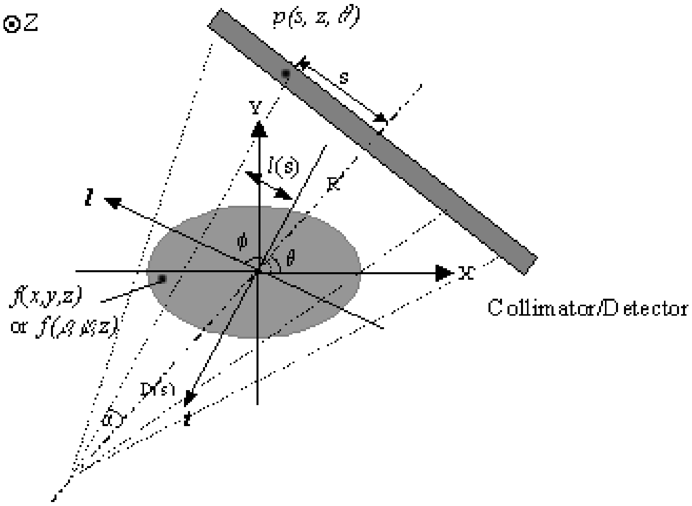


Figure 1. Fan-beam geometry and the coordinate systems**.**

This assumption simply means that the system response is a function of the distances from source activity point to the collimator hole in 3D. This PSF assumption has been investigated by many researchers and is consistent in those references [13-18].

Following the method as described by Bellini et al [19] in dealing with the constant attenuation, we define $L(s,\theta)$ as the distance from point (s,θ) on detector to the boundary of the object (s,t) with t>0, and let $p_\mu(s,z,\theta) = \exp[-\mu L(s,\theta)+\mu R]\, p(s,z,\theta)$, then (1) become:

$$\begin{aligned} &p_\mu(s,\theta,z) \\ &= \iiint f(x',y',z') e^{-\mu t'} k(l(s) - l', z - z', t' + R)\, dx' dy' dz' \end{aligned} \tag{3}$$

By polar coordinate system, (2) can be rewritten as:

$$\begin{aligned} &p_\mu(s,\theta,z) \\ &= \int_0^R \rho d\rho \int_0^{2\pi} d\varphi \int_{-\infty}^{\infty} dz' f'(\rho,\varphi,z') k\left(\rho\sin(\varphi-\theta-\alpha) - l(s), z'-z, \rho\cos(\varphi-\theta-\alpha)\right) \\ &= \int_0^R d\rho \int_0^{2\pi} d\varphi \int_{-\infty}^{\infty} dz' f'(\rho,\varphi,z') k'\left(\rho, l(s); (\varphi-\theta), z'-z\right) \end{aligned} \tag{4}$$

where the relations held as follows:

$$\phi = \theta + \alpha + \pi/2 \ , \qquad l(s) = \frac{D \cdot s}{\sqrt{(D+R)^2 + s^2}},$$

$$\alpha = \alpha(s) = \tan^{-1}\left(\frac{R+D}{s}\right) \ , \qquad \begin{pmatrix} l \\ t \end{pmatrix} = \begin{pmatrix} \rho\sin(\varphi-\theta-\alpha) \\ -\rho\cos(\varphi-\theta-\alpha) \end{pmatrix} \tag{5}$$

In general we may not have a closed inversion form for formula (4). The purpose of this paper is to derive an optimal estimation of $f(x,y,z)$ or $f(\rho,\varphi,z)$ based on the equation. In the following, we will use the CHD technique to solve this simultaneous compensation problem. The theoretical procedure is given as follows.

By Fourier transform on (4) with respect to variable z, we have:

$$\tilde{p}(s,\theta,\xi) = \int_0^R d\rho \int_0^{2\pi} d\varphi\, \tilde{f}(\rho,\varphi,\xi) \tilde{k}\left(l(s), \rho; \theta - \varphi, \xi\right) \tag{6}$$

where $\tilde{k}$ denotes the Fourier transform of k' for the variable z, and $\tilde{p}$ and $\tilde{f}$ have the similar meanings. Note that the image property $\tilde{f}(\rho,\varphi,\xi)$ can be estimated slice-by-slice for ξ, indicating that the compensation for the 3D collimator response may be uncoupled to a 2D compensation. This is because that the PSF is shift invariant along the rotation z-axis. By computing the Fourier series expansion of θ on both sides of (6) (this is usually called circular harmonic decomposition of a function in mathematics), we have the following formula:

$$\hat{p}(s,n,\xi) = \int_0^R d\rho\ \hat{f}^*(\rho,n,\xi) \hat{k}\left(l(s), \rho; n, \xi\right) \tag{7}$$

where

$$\hat{p}(s,n,\xi) = \frac{1}{2\pi} \int_0^{2\pi} d\theta\ \tilde{p}(\rho,\theta,\xi) e^{-in\theta}$$

$$\hat{f}(s,n,\xi) = \frac{1}{2\pi} \int_0^{2\pi} d\varphi\ \tilde{f}(\rho,\varphi,\xi) e^{-in\varphi}, \tag{8}$$

and notation * stands for conjugate operation. Obviously, (7) is only a CHD expression of (6). Now the reconstruction problem becomes a task of solving a linear algebraic equation for each n and ξ:

$$\begin{pmatrix} \tilde{p}(s_0) \\ \tilde{p}(s_1) \\ \tilde{p}(s_2) \\ \ldots\ldots \end{pmatrix} = \begin{pmatrix} M(s_0\rho_0) & M(s_0\rho_1) & M(s_0\rho_2) & \ldots\ldots \\ M(s_1\rho_0) & M(s_1\rho_1) & M(s_1\rho_2) & \ldots\ldots \\ M(s_2\rho_0) & M(s_2\rho_1) & M(s_2\rho_2) & \ldots\ldots \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots \end{pmatrix} \begin{pmatrix} \tilde{f}(\rho_0) \\ \tilde{f}(\rho_1) \\ \tilde{f}(\rho_2) \\ \ldots\ldots \end{pmatrix}$$

$$\boldsymbol{P=Mf} \tag{8}$$

Given p and M, and looking for f, this is a typical linear problem that could be solved by several methods. Here we are using the conjugate gradient algorithm. Since the dimension is significantly reduced from 3D to 2D, efficient calculation is expected.

III. IMPLEMENTATION

Fan-beam projections of the Shepp-Logan ellipse phantom and Hoffman brain phantom with different $^{99}\mathrm{Tc}^m$ activity concentrations were simulated on a circular orbit of 64 evenly spaced views, taking into account the photon attenuation and collimator response effects. Both the noise-free projection data and their noisy versions with Poisson noise were used to test the reconstruction algorithm.

The reconstruction procedure is as follows:

1. Perform the transforms on PSF matrix *M*. This task is a pre-calculated step given the PSF, before image reconstruction.
2. Perform Fourier transform on the projection data for variable *z*, and Fourier series expansion for variable θ. Note that since the PSF is periodic about θ, so zero padding is not necessary. But for variable *z* (the shift invariant characteristics), zero padding may be needed.
3. Utilize the conjugate gradient algorithm to solve the linear equation (8).
4. After the $\tilde{f}(\rho, n, \xi)$ is obtained, perform two times of Fourier inversion transform, relating to *z* and θ, to find the source image.

In our experiment, the response function was chosen as the most general Gaussian function. Because of the CHD method, transforms between Cartesian coordinate system and polar coordinate system is needed.

IV. RESULTS AND DISCUSSION

The simulation results of different reconstruction procedures are shown in Figure 2, which gives four arbitrary slices for illustration and comparison purposes. Similar results were obtained for the Shepp-Logan phantom. The first row contains slices from the original Hoffman brain phantom. The second row shows the reconstructed results from the blurred projection data containing the photon attenuation and collimator response effects. The third row represents the reconstructed images from projection data with Poisson noise at a noise level similar to a typical clinic study. For a 64x64x32 image, the reconstruction time was less than one minute on a PC platform with Pentium III 550 processor.

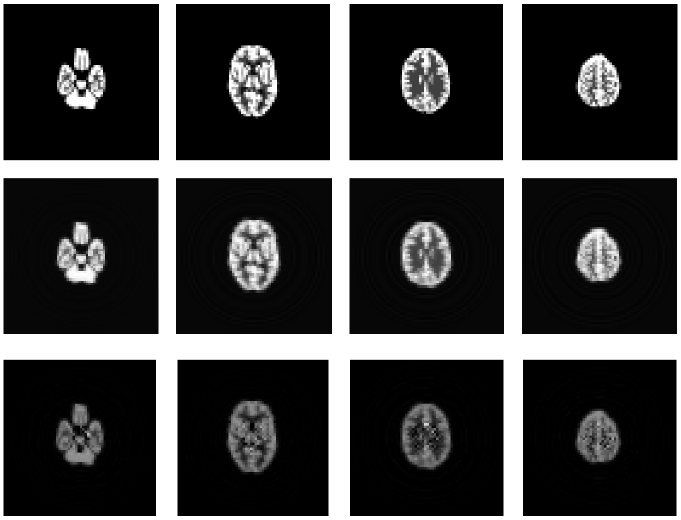


Figure 2. Hoffman brain phantom results. Top row represents the original Hoffman phantom slices. Second row is the reconstructed image from noise-free projection data with attenuation and PSF effects included. Bottom row is the reconstructed image from projection data containing Poisson noises.

V. CONCLUSION

First, our method inherits the idea about angular expression in our previous work[20,21], which is beneficial for non-parallel ray collimator geometry. It is not a back-projection method of source and projections with compensation for photon attenuation and resolution variation. It constructs a new relationship between source image and projection data using CHD technique. In this relationship, the photon attenuation and collimator response can be simultaneously considered in 3D. The compensation can be implemented slice-by-slice along the rotation direction after Fourier transform, because of the shift invariance along that direction. Second, estimation of source image is efficient and accurate via the conjugate gradient algorithm, which converges in finite iterations, as demonstrated by Figure 2. The simulation study revealed this optimization of the proposed algorithm. The order of matrix *M* in (8) is only *N*x*N* size, and the calculation for *N* slices can be performed in parallel at the same time, so the computation is very efficient.

This method considers the periodical property of projection rays of non-parallel-hole collimator geometry and the shift invariant characteristics of fan-beam configuration along the rotation direction. The periodic property is explored by the CHD technique. The shift invariant characteristics are efficiently utilized in the Fourier space. This strategy reduces the 3D PSF treatment into 2D task in the Fourier space and, therefore, improves the computing efficiency. For the reduced matrix size, the conjugate gradient method is a choice for the calculation. The computer simulation is encouraging. Further validation by physical phantom experiments is under progress.

Acknowledgment. This work was supported by NIH Grant HL-51466 and Grant NS-33853, and by Established Investigator Award of the American Heart Association.

References

1. Z. Liang, J. Ye and D. P. Harrington, “Quantitative brain SPECT in three dimensions: an analytical approach without transmission scans”, *Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine,* Kluwer Academic Publishers, pp. 117-132, 1996.
2. J. Li, Z. Liang, J. Ye, and G.. Han, "Implementation and preliminary investigation of analytical methods for correction of distance-dependent resolution variation and uniform attenuation in 3D brain SPECT", *IEEE Trans. Nucl. Sc*i. **46**: 2162-2171, 1999.
3. L van Elmbt and S. Walrand, “Simultaneous correction of attenuation and distance-dependent resolution in SPECT: an analytical approach”, *Phys. Med. Biol.* **38**: 1207-1217, 1993.
4. E. J. Soares, C. L. Byrne, S. J. Glick, C. R. Appledorn and M. A. King, “Implementation and evaluation of an analytical solution to the photon attenuation and non-stationary resolution reconstruction problem in SPECT”,

IEEE Trans. Nucl. Sci. **40**: 1231-1237, 1993.

5. G. L. Zeng and G. T. Gullberg, "Frequency domain implementation of the 3D geometric point response correction in SPECT imaging", *IEEE Trans. Nucl. Sci.* **39**: 1444-1453, 1992.
6. Y. Weng, G. L. Zeng, G. T. Gullberg, "Analytical inversion formula for uniformly attenuated fan-beam projections", *IEEE Trans. Nucl. Sci*, **44**: 243-249, 1997.
7. Z. Cao and B. M. W. Tsui, "A filtering Technique to compensate for detector response in converging-beam SPECT reconstruction", *Phys. Med. Biol.* **39**: 1281-1293, 1994.
8. X. Pan and C. E. Metz, "Analytical approaches for image reconstruction in 3D SPECT", *Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, Kluwer Academic Publishers, pp. 103-116, 1996.
9. G. T. Gullberg and T. F. Budinger, "The use of filtering methods to compensate for constant attenuation in SPECT", *IEEE Trans. Bio. Eng*. **28**: 142-157, 1981.
10. B. M. W. Tsui, H. B. Hu, D. R. Gilland, and G. T. Gullberg, "Implementation of simulation and detector response correction in SPECT", *IEEE Trans. Nucl. Sci.* **35**: 778-783, 1988.
11. D. R. Gilland, R. J. Jaszczak, J. E. Bowsher, T. G. Turkington, Z. Liang, K. L. Greer, and R. E. Coleman, "Quantitative SPECT brain imaging: effects of attenuation and detector response", *IEEE Trans. Nucl. Sci.* **40**: 1723-1727,1993.
12. G. L. Zeng, G. T. Gullberg, B. M. W. Tsui, and J. A. Terry, "3D iterative reconstruction algorithms with attenuation and geometric points response correction", *IEEE Trans. Nucl. Sci.* **38**: 693-702, 1991.
13. C. E. Metz, F. B. Atkins and R. N. Beck, "The geometric transfer function component for scintillation camera collimators with straight parallel holes", *Phys. Med. Biol.* **25**: 1059-1070, 1980.
14. E. C. Frey and B. M. W. Tusi, "Parameterization of the scatter response function in SPECT imaging using Monte Carlo simulation", *IEEE Trans. Nucl. Sci.* **37**: 1308-1315, 1990.
15. B. M. W. Tsui and G. T. Gullberg, "The geometric transfer function for cone and fan beam collimators", *Phys. Med. Biol*. **35**: 81-93, 1990.
16. K. Ogawa, S. Paek, M. Nakajima, et al, "Correction of collimator aperture using a shift-variant deconvolution filter in gamma camera emission CT", *SPIE Medical Imaging II*, **914**, pp.699, 1988.
17. M. J. Yzuel, S. Millan, and S. Miracle, "Evaluation and comparison between the collimator and the detector responses in a scintillation camera", *OPTICA ACTA*, **29**: 197-208, 1982.
18. S. Miracle, M. J. Yzuel, and S. Millan, "A study of the point spread function in scitillation camera collimators based on Fourier analysis", *Phys. Med. Biol*. **24**: 372-384, 1979.
19. S. Bellini, M. Piacentini, C. Cafforio, and F. Rocca, "Compernastion of tissue absorption in emission tomography", *IEEE Trans. ASSP*, **27**: 213-218, 1979
20. J. You, S. Bao, and Z. Liang, "Benefits of angular expression to reconstruction algorithms for collimators with spatially varying focal lengths", *IEEE Trans. Med. Imag*. **16**: 527-531, 1997.
21. J. You, Z. Liang, and G. L. Zeng, "A unified reconstruction framework for both parallel-hole and variable focal length fan-beam colliamtors by a Cormack-type inversion of exponential Radon transform", *IEEE Trans. Med. Imag*. **18**: 59-65, 1999.

Investigation of the dynamic SPECT (dSPECT) method for Teboroxime using a 4-D kinetic thorax model dMCAT

A. Celler[†], S.Blinder[†], D. Noll[‡], T.Tyler[§], F.Duclercq[‡], R. Harrop[†]
[†]Div. of Nuclear Medicine, Vancouver Hospital and Health Sciences Centre, Vancouver, B.C., Canada
[‡]Mathématiques pour l'Industrie et la Physique (MIP), Université Paul Sabatier, Toulouse, France
[§]Department of Engineering Physics, Simon Fraser University, Burnaby, B.C., Canada

Abstract

Tomographic SPECT allows us to visualize, in 3D, the distribution of radiopharmaceuticals in the body which are reconstructed from a series of two dimensional measurements performed around the investigated object. The dSPECT method extends these capabilities into the fourth dimension and creates a series of dynamic 3D images which can be obtained from a single set of dynamic data acquired using only a single rotation of the camera. In order to test the performance of the method for Teboroxime cardiac images, we have modified the 3D MCAT digital phantom to model dynamic changes of activity distributions in different organs of the thorax. Experiments using different acquisition protocols were performed and the resulting data were reconstructed using static (FBP and OSEM) and dynamic (dEM) methods. The images obtained by summing the data from dynamic reconstructions over all time-frames were as good as or better than those from static reconstructions (for the scans which begun 60 or 120 seconds post injection i.e. after the bolus of activity left the heart). Including the bolus in the reconstructions produced distorted images. An important advantage of the dSPECT method is that it reconstructs images in a form of 3D movies that may be used to extract additional diagnostic information, mainly temporal information which is not available in static images and 3D spatial information not present in planar studies.

I. INTRODUCTION

Single Photon Emission Computed Tomography (SPECT) is well recognized as a powerful diagnostic tool to investigate organ function rather than anatomy since it allows us to image, in three dimensions, the bio-distribution of radiolabelled tracers within the body. Standard clinical SPECT methods, however, image only stationary activity distributions. It is believed that nuclear medicine procedures which could trace changing distributions of radio-labelled substances would provide important information for both diagnosis and research (see for example [1, 2]).

To this end we have proposed a dynamic SPECT (dSPECT) method [3] that allows us to obtain quantitative information about kinetic processes in the body from the data acquired using a standard clinical acquisition protocol, one with a single rotation of a tomographic camera. The method can be used with all standard, currently available SPECT systems including single, double and triple head cameras. The result of the dSPECT reconstruction, which includes attenuation and resolution recovery corrections, is a 4D data set, composed of a time-series of 3-D SPECT images (3D movie). The dSPECT reconstruction is based on a mathematical optimization procedure where all the dynamic projections are being considered simultaneously, resulting in images with better signal to noise ratio than in the ”fast-rotation” method, where each data set is reconstructed separately. An important feature of the method is that each dynamic voxel (doxel) is reconstructed independently and, as a result, it is possible for an object to contain an assortment of doxels where the activity may increase, decrease, increase and decrease or remain constant over the acquisition time.

Over the last few years we have investigated the performance of the method using 2D computer simulations, phantom experiments and patient studies [3, 4, 5]. The goal of these tests has been to verify the method for a broad range of kinetic parameters, to evaluate its accuracy and to optimize acquisition protocols. At the present stage, however, we are focusing our research on a few practical applications of the method. One of such applications is investigation of cardiac studies using Teboroxime Tc-99m.

It has been shown that Teboroxime extraction reflects the true blood flow better than other myocardial perfusion agents such as Tc-99m MIBI or Tl-201 [6, 7]. Its use has been limited because it is not trapped within the cell, but rather, due to its neutral charge, is rapidly washed out. Thus, using standard static imaging methods with this radiopharmaceutical is difficult. Image artifacts can be produced and also, as will be shown later in this paper, the location, or even the presence of any defect can often not be visualized due to the delay of blood perfusion in the stenotic area.

In order to investigate these issues we have adapted the 3D mathematical cardiac torso phantom (MCAT) [8] to model temporal changes in activity distributions. Using a dynamic version of the model (dMCAT) we have simulated different clinical situations and acquisition protocols, then assessed the diagnostic content of images obtained from this data using static and dynamic reconstructions. Although in this work we present only the results of Teboroxime simulations, the dMCAT phantom is fully versatile and can be used with other radiopharmaceuticals using appropriate compartmental models.

II. METHODS

In our dynamic version of the model (dMCAT) we first simulate the time-activity concentrations of the radiotracer in different organs of the human thorax using a compartmental model approach. A set of differential equations representing the flow of activity between organs is solved with kinetic parameters being defined by the user in a MATLAB user interface. This part of the program has to be modified for each tracer so that the equations match the appropriate compartmental model. The result of this calculation is displayed on a graph as a set of time activity curves (TAC) representing concentrations in different compartments. In our simulations these activity changes were modeled assuming a bolus injection of the tracer. The activity then moves from the right ventricle to the lungs, next to the left ventricle, and to the myocardium, the muscles (representing the rest of the body) and to the liver. An additional equation is used to create a defect in the heart with dynamic parameters different from those of a healthy myocardium. Kinetic parameters used in the model were based on experimental Teboroxime studies and the shapes of the resulting simulated TAC's were adjusted to match the experimental shapes [9, 10]. Figure 1 shows an example of the curves used in our simulations. As can be seen from this figure, the injection bolus lasts for about 30 seconds in the right ventricle, then, for the next 30-60 seconds the activity in all organs rises fast and this period is followed by much slower changes, with the normal myocardium washout characteristic time in the range of 6-8minutes.

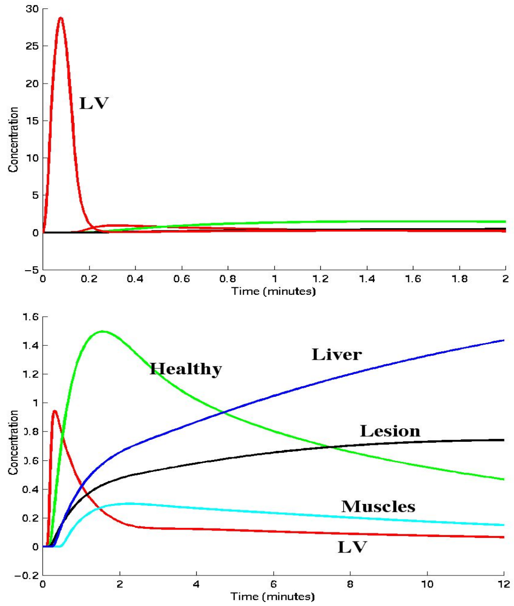


Figure 1: An example of the time activity curves (TAC) used in Teboroxime dMCAT simulations. The top part of the figure shows the injection bolus as it moves through the right ventricle and the lower part the activity concentrations in the remaining. Note the change of scale between the images.

In the next step these values of activity concentration are used to compute the 4D thorax model (the "truth") with the activity in each organ following a separate time activity change. Figure 2 shows a volume rendered image of the dMCAT with activities corresponding to 2 minutes post injection. We have simulated the healthy myocardium and the one with stenosis (heart defect) located at two different places in the heart (in the apex and close to the liver). Also, two different levels of liver uptake were modeled with the MCAT option of a liver positioned high in the thorax.

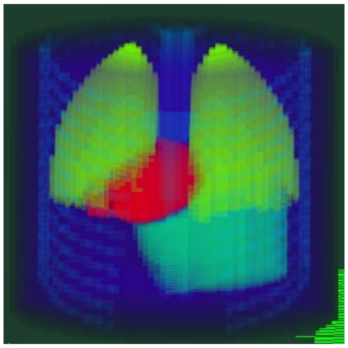


Figure 2: A volume rendered image of the dMCAT phantom with activities in the organs corresponding to 120 seconds (2 minutes) post injection.

The user then specifies the acquisition parameters using a MATLAB user interface. The number of camera heads, their starting angle, the matrix and pixel sizes, the angle of rotation, the time of each camera stop and the time between the injection and the beginning of the acquisition are specified to generate the projector without or with attenuation and without or with 2D or 3D collimator blurring. At this point the sinograms are calculated.

In our tests we have investigated all these options modeling slow acquisitions with dual and triple head systems, rotating a maximum 180° per head. In all cases the matrix size was 64 x 64 with 64 or 32 camera stops and 10 or 20 seconds per projection, respectively. Two different approaches were tested: (i) a short 3 minute scan where the activity in the object does not change too much and (ii) a longer 10 minute scan during which the activity drops by about 50%, this allows us to collect dynamic data with better statistics and with more data points. Similarly, we have tested acquisitions starting (a) at the time of injection (at 0 seconds) where fast movement of bolus through the heart is strongly influencing the projection data, (b) at 60 seconds post injection with more moderate rates of change and the with increase and decrease of activity in the organs, and (c) at 120 seconds post injection where activity changes are the slowest.

Noiseless as well as noisy projection data (according to a Poisson noise) corresponding to about 7-8 counts per second per milliliter of myocardium tissue as measured at 120 seconds post injection (based on real patient data) were created. Additionally, an experiment simulating triple-head fast rotation was performed with the camera rotating for 20 seconds over 120° per head. This study was started at 60s post injection and continued for 10 minutes (30 rotations in total).

The last part of our MATLAB interface allows us to reconstruct the data using (i) static reconstruction methods: filtered backprojection (FBP) or ordered subsets expectation maximization (OSEM) or (ii) using a dynamic approach with 2D [11] or 3D [12] dynamic expectation maximization (dEM) reconstructions. All images presented in this work were smoothed using a 3x3 gaussian filter.

III. RESULTS AND DISCUSSION

First, in order to check the results for the presence of artifacts, the static reconstructions were compared with the images obtained by summing dynamic images over all time frames. The static images obtained from the data from both acquisitions which started at the moment of injection (0-3 minutes and 0-10 minutes) were dominated by strong streak artifacts which were due to large variations in the object created by the bolus activity entering the right and later the left ventricle. As expected, the FBP images were much worse than those from the OSEM method. Dynamic reconstructions, which have the advantage of creating a whole time-series of images, displayed some artifacts in the first images corresponding to the beginning of the scan, but were surprisingly good for the remaining time frames. These results indicate that dSPECT can handle even large activity changes. Summing of the time frames in this case did not make much sense, as artifacts from these first frames dominated and spoiled the resulting images as it is illustrated on Figure 4.

Similar comparisons were performed for scans starting at 60 and 120 seconds post injection. In this case, although FBP static images were bad, the OSEM and summed over time dEM reconstructions produced very comparable images (see Figure 4). Also, there was not much difference between the acquisitions which started at 60 and 120 seconds post injection. Careful analysis of this data, however, reveals one disturbing effect. Both static and summed dynamic reconstructions tend to average the activity distributions over time. Therefore, in situations where there is a defect in the heart it is possible that both normal and stenotic myocardium will have equal levels of such an averaged activity and therefore these images will not reveal the location, or even the presence, of the defect. On the other hand, when reviewing the dynamic time series reconstructed using the dSPECT method one can clearly identify the defect and analyse its perfusion. Figure 5 presents two dynamic times-frames (80 and 260 seconds) corresponding to the same data as presented in Figure 4. Changes in relative levels of activity display clearly the location of the defect. For comparison, the top two images present the true activity distributions, in the middle the results of dSPECT reconstruction are displayed and the bottom part presents images obtained using the data acquired with the fast rotation of the triple head camera reconstructed with OSEM. Since these images correspond to only 20 seconds total acquisition time they have very poor statistics and the high noise makes their analysis practically impossible.

On the other hand, when reviewing the dSPECT reconstruction, in addition to spatial information about the activity distribution, a dynamic series of images also contains temporal information which can be used in subsequent analysis for diagnostic purposes or in order to separate different organs. Important applications of this approach would be the identification and quantitation of heart defects and the elimination of the problem of overlapping liver and heart distributions. The time activity curves for normal and diseased myocardium and for the liver are quite different (see Figure 1) and, indeed, dynamic data could be used in a such separation procedure.

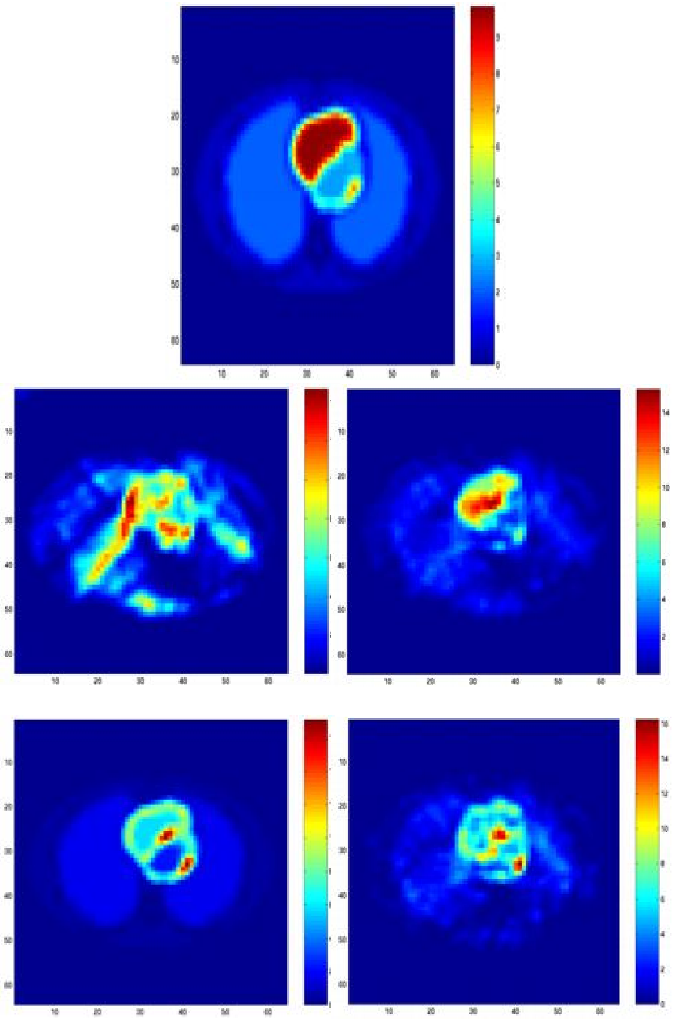


Figure 3: Images of a transaxial slice of the dMCAT phantom for 3 minutes acquisitions starting at the time of the injection. The true activity distribution (upper) and the results of static OSEM (middle left), and dynamic dEM summed over all time frames (middle right) are presented. In the bottom part of the figure the true and the dSPECT images corresponding to the 2.5 minute time frame are displayed.

IV. CONCLUSIONS

Performance of the dSPECT method has been investigated for myocardial viability studies with Teboroxime using the MCAT phantom, which was modified to model changing activity distributions in the organs. The images which were obtained by summing the images from dynamic reconstructions over all time-frames were as good and often better or even much better than those from static FBP or OSEM. An important advantage, however, of the dSPECT method is that it reconstructs images in a form of 3-D dynamic movies that may be used to extract additional diagnostic information. All these findings need to be confirmed in patient studies which are currently being performed.

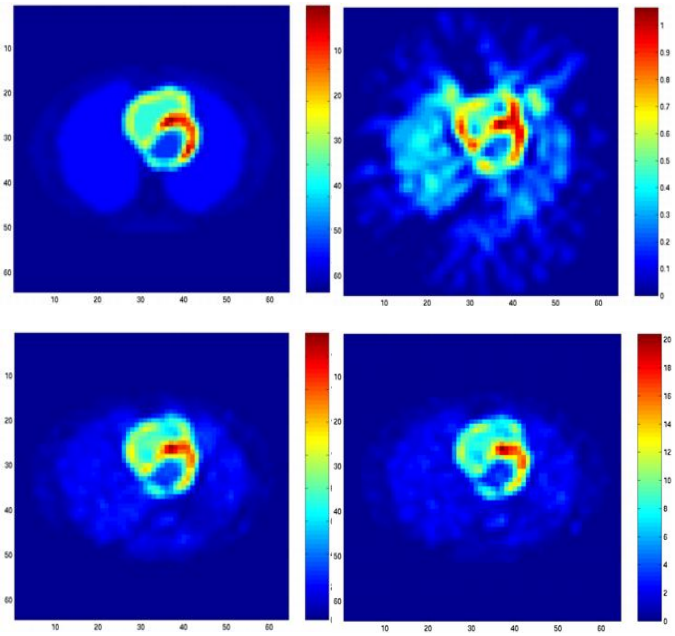


Figure 4: Images of a transaxial slice of the dMCAT phantom. The true activity distribution (upper left) and the results of static FBP (upper right), static OSEM (lower left) and dynamic dEM summed over all time frames (lower right) are presented.

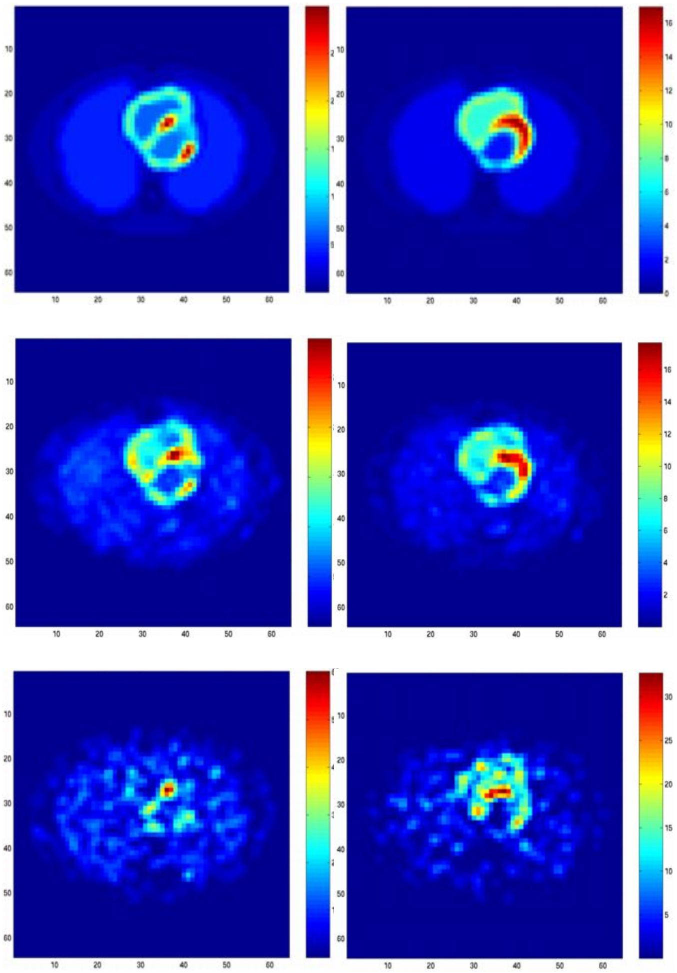


Figure 5: Images of a transaxial slice of the dMCAT phantom corresponding to the activity distribution at 80 seconds (left) and 260 seconds (right). The true activity distribution (upper row), the results of dynamic dEM (middle row), and dynamic "fast-rotation" (bottom row) are presented.

V. REFERENCES

[1] R. Stewart, B.Heyl, R.O'Rourke, R.Blumhardt, and D. Miller. Demonstration of differential post-stenotic myocardial technetium-99m-teboroxime clearance kinetics after experimental ischemia and hyperemic stress. *J. Nuc. Med.*, 32(10):2000–2008, 1991.

[2] H. Weinstein, S.T. Dahlberg, B.A. McSherry, R.C. Hendel, and J.A.Leppo. Rapid redistribution of teboroxime. *Am. J. Cardiol.*, 71:848–852, April 1999.

[3] T. Farncombe, A. Celler, D. Noll, J. Maeght, and R. Harrop. Dynamic SPECT imaging using a single camera rotation (dSPECT). *IEEE Trans. Nuc. Sci.*, 46(4):1055–1061, 1999.

[4] A.Celler, T.Farncombe, C.Bever, D.Noll, J.Maeght, R.Harrop, and D. Lyster. Performance of the dynamic single photon computed tomography (dspect) method for decreasing or increasing activity changes. *Phys. Med. Biol.*, 45:3525–3544, 2000.

[5] T. Farncombe, A. Celler, D. Noll, J. Maeght, and R. Harrop. The incorporation of organ uptake into dynamic spect (dSPECT) image reconstruction. *IEEE Trans. Nuc. Sci.*, 48(1):3–9, 2001.

[6] R. Narra, A. Nunn, B. Kuczynski, T. Feld, P. Wedeking, and W. Eckelman. A neutral technetium-99m complex for myocardial imaging. *J. Nuc. Med.*, 30:1830–1837, 1989.

[7] Q-S. Li, G. Solot, T.L.Frank, Jr H.N. Wagner, and L.C. Becker. Tomographic myocardial perfusion imaging with technetium-99m-teboroxime at rest and after dipyridamole. *J. Nuc. Med.*, 32(10):1968–1976, 1991.

[8] P.H. Pretorius, M.A. King, B.M.W.Tsui, K.J. LaCroix, and W. Xia. A mathematical model of motion of the heart for use in generating source and attenuation maps for simulating emission imaging. *Med. Phys.*, 26:2323–2332, 1999.

[9] A.M. Smith, G.T. Gullberg, P.E. Christian, and F.L. Datz. Kinetic modeling of teboroxime using dynamic SPECT imaging of a canine model. *J. Nuc. Med.*, 35:484–495, 1994.

[10] A. Sitek, E.V.T. DiBella, and G.T. Gullberg. Factor analysis with a priori knowledge - application in dynamic cardiac spect. *Phys. Med. Biol.*, 45:2619–2638, 2000.

[11] T.Farncombe, S.Blinder, A.Celler, D.Noll, J.Maeght, and R.Harrop. A dynamic expectation maximization algorithm for single camera rotation dynamic spect (dspect). *IEEE Nuclear Science Symposium Conference Record*, 2000.

[12] T.H. Farncombe, M.A. King, A. Celler, and S.Blinder. A fully 4d expectation maximization algorithm using gaussian diffusion based detector response for slow camera rotation dynamic spect. In *Proceedings of the 2001 International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, 2001. submitted.

Fast Computation of Statistical Uncertainty for Spatiotemporal Distributions Estimated Directly from Dynamic Cone Beam SPECT Projections

Bryan W. Reutter[†], Grant T. Gullberg[‡], and Ronald H. Huesman[†]

[†]Center for Functional Imaging, Lawrence Berkeley National Laboratory
University of California, Berkeley, CA 94720, USA
[‡]Medical Imaging Research Laboratory, Department of Radiology
University of Utah, Salt Lake City, UT 84108, USA

I. INTRODUCTION

THE estimation of time-activity curves and kinetic model parameters directly from projection data is potentially useful for clinical dynamic single photon emission computed tomography (SPECT) studies, particularly in those clinics that have only single-detector systems and thus are not able to perform rapid tomographic acquisitions. Because the radiopharmaceutical distribution changes while the SPECT gantry rotates, projections at different angles come from different tracer distributions. A dynamic image sequence reconstructed from the inconsistent projections acquired by a slowly rotating gantry can contain artifacts that lead to biases in kinetic parameters estimated from time-activity curves generated by overlaying regions of interest on the images. If cone beam collimators are used and the focal point of the collimators always remains in a particular transaxial plane, additional artifacts can arise in other planes reconstructed using insufficient projection samples [1]. If the projection samples truncate the patient's body, this can result in additional image artifacts. To overcome these sources of bias in conventional image based dynamic data analysis, we and others have been investigating the estimation of time-activity curves and kinetic model parameters directly from dynamic SPECT projection data by modeling the spatial and temporal distribution of the radiopharmaceutical throughout the projected field of view [2–8].

In our previous work we developed a computationally efficient method for fully four-dimensional (4-D) direct estimation of spatiotemporal distributions from dynamic SPECT projection data [5], which extended Formiconi's least squares algorithm for reconstructing temporally static distributions [9]. In addition, we studied the biases that result from modeling various orders of temporal continuity and using various time samplings [5]. In the present work, we address computational issues associated with evaluating the statistical uncertainty of spatiotemporal model parameter estimates, and use Monte Carlo simulations to validate a fast algorithm for computing the covariance matrix for the parameters. Given this covariance matrix, the covariance between the time-activity curve models for the blood input function and tissue volumes of interest can be calculated and used to estimate compartmental model kinetic parameters more precisely, using nonlinear weighted least squares [10, 11].

II. FAST COMPUTATION OF STATISTICAL UNCERTAINTY FOR SPATIOTEMPORAL DISTRIBUTIONS

Following our development in [5], time-varying activity concentrations within volumes of interest encompassing the projected SPECT field of view can be modeled by selecting a set of temporal basis functions capable of representing typical time variations and having desired smoothness properties. Similarly, the spatially nonuniform activity concentration within a particular volume of interest can be modeled by selecting an appropriate set of spatial basis functions. Given a set of temporal basis functions and sets of spatial basis functions for the volumes of interest, coefficients for the resulting spatiotemporal basis functions can be estimated directly from the SPECT projection data, along with the covariance matrix for the coefficients.

A. *Covariance Matrix for the Spatiotemporal Basis Function Coefficients*

Denoting the projection of the m^{th} spatial basis function along ray i at angle j by u_{ij}^m, and the integral of the n^{th} temporal basis function during the time interval associated with angle j of rotation k by v_{jk}^n, the projection equations can be expressed as

$$p_{ijk} = \sum_{m=1}^{M} \sum_{n=1}^{N} a_{mn} u_{ij}^m v_{jk}^n, \tag{1}$$

where the p_{ijk} are the modeled projections, the a_{mn} are the linear coefficients associated with the time integrals of the projections of the spatiotemporal basis functions, M is the number of spatial basis functions, and N is the number of temporal basis functions. The criterion which is minimized by varying the linear coefficients a_{mn} is the weighted sum of squares function

$$\chi^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{(p_{ijk}^* - p_{ijk})^2}{W_{ijk}}, \tag{2}$$

where the p_{ijk}^* are the measured projections, the W_{ijk} are weighting factors, I is the number of projection rays per angle,

---

This work was supported by the National Heart, Lung, and Blood Institute of the US Department of Health and Human Services under grants R01-HL50663 and P01-HL25840 and by the Director, Office of Science, Office of Biological and Environmental Research, Medical Sciences Division of the US Department of Energy under contract DE-AC03-76SF00098. This work was developed in part using the resources at the US Department of Energy National Energy Research Scientific Computing (NERSC) Center.

J is the number of angles per rotation, and K is the number of rotations. Typically, the weighting factors are either unity for an unweighted fit or the estimated variances of the projections for a weighted fit.

Equations (1) and (2) can be rewritten in matrix form as

$$\mathbf{p} = \mathbf{F}\mathbf{a} \tag{3}$$

and

$$\chi^2 = (\mathbf{p}^* - \mathbf{F}\mathbf{a})^{\mathrm{T}}\mathbf{W}(\mathbf{p}^* - \mathbf{F}\mathbf{a}), \tag{4}$$

respectively, where $\mathbf{p}$ is an IJK element column vector whose $[i+(j-1)I+(k-1)IJ]^{\text{th}}$ element is p_{ijk}, $\mathbf{F}$ is an $IJK \times MN$ matrix whose $\{[i+(j-1)I+(k-1)IJ],[m+(n-1)M]\}^{\text{th}}$ element is $u_{ij}^m v_{jk}^n$, $\mathbf{a}$ is an MN element column vector whose $[m+(n-1)M]^{\text{th}}$ element is a_{mn}, $\mathbf{p}^*$ is an IJK element column vector whose $[i+(j-1)I+(k-1)IJ]^{\text{th}}$ element is p_{ijk}^*, and $\mathbf{W}$ is an $IJK \times IJK$ diagonal matrix whose $[i+(j-1)I+(k-1)IJ]^{\text{th}}$ diagonal element is $1/W_{ijk}$. The criterion, χ^2, is minimized by the vector of spatiotemporal basis function coefficients

$$\hat{\mathbf{a}} = (\mathbf{F}^{\mathrm{T}}\mathbf{W}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{W}\mathbf{p}^*. \tag{5}$$

The covariance matrix for the coefficients $\hat{\mathbf{a}}$ is

$$\text{cov}(\hat{\mathbf{a}}) = (\mathbf{F}^{\mathrm{T}}\mathbf{W}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{W}\text{cov}(\mathbf{p}^*)\mathbf{W}\mathbf{F}(\mathbf{F}^{\mathrm{T}}\mathbf{W}\mathbf{F})^{-1}, \tag{6}$$

where $\text{cov}(\mathbf{p}^*)$ is the covariance matrix for the measured projections. Given an estimate of $\text{cov}(\mathbf{p}^*)$, estimates of the statistical uncertainties of the coefficients $\hat{\mathbf{a}}$ are the square roots of the diagonal elements of the covariance matrix given by equation (6) and are denoted individually by $\hat{\sigma}_{\hat{a}_{mn}}$. In general, the errors in the coefficients are correlated and the covariance matrix given by equation (6) has nonzero elements off the diagonal.

For an unweighted least squares reconstruction of the spatiotemporal basis function coefficients $\hat{\mathbf{a}}$ (i.e., for $\mathbf{W}$ an identity matrix), an estimate of the symmetric $MN \times MN$ covariance matrix $\text{cov}(\hat{\mathbf{a}})$ can be obtained quickly from equation (6) as follows. Assuming Poisson noise, the diagonal matrix having the modeled projections $\hat{\mathbf{p}} = \mathbf{F}\hat{\mathbf{a}}$ along the diagonal can be used as an estimate of the covariance matrix for the measured projections. Substituting this diagonal matrix for $\text{cov}(\mathbf{p}^*)$ and the identity matrix for $\mathbf{W}$, equation (6) can be rewritten as

$$\text{cov}(\hat{\mathbf{a}}) = (\mathbf{F}^{\mathrm{T}}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\text{diag}(\mathbf{F}\hat{\mathbf{a}})\mathbf{F}(\mathbf{F}^{\mathrm{T}}\mathbf{F})^{-1}. \tag{7}$$

We have presented a method for quickly calculating $(\mathbf{F}^{\mathrm{T}}\mathbf{F})^{-1}$ in [5]. Using a similar development, the symmetric $MN \times MN$ matrix $\mathbf{F}^{\mathrm{T}}\text{diag}(\mathbf{F}\hat{\mathbf{a}})\mathbf{F}$ can be calculated quickly as follows. Denoting the $\{[m+(n-1)M],[m'+(n'-1)M]\}^{\text{th}}$ element of $\mathbf{F}^{\mathrm{T}}\text{diag}(\mathbf{F}\hat{\mathbf{a}})\mathbf{F}$ by $\psi^{mnm'n'}$, one has

$$\begin{aligned}\psi^{mnm'n'} = {} & \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} u_{ij}^m v_{jk}^n \times \\ & \left[\sum_{m''=1}^{M}\sum_{n''=1}^{N} \hat{a}_{m''n''} u_{ij}^{m''} v_{jk}^{n''}\right] \times u_{ij}^{m'} v_{jk}^{n'},\end{aligned} \tag{8}$$

where $\hat{a}_{m''n''}$ is the $[m''+(n''-1)M]^{\text{th}}$ element of $\hat{\mathbf{a}}$, and "$\times$" denotes scalar multiplication. Rearranging the summations yields

$$\begin{aligned}\psi^{mnm'n'} &= \sum_{m''=1}^{M}\sum_{n''=1}^{N} \hat{a}_{m''n''} \times \\ &\qquad \sum_{j=1}^{J}\left[\sum_{i=1}^{I} u_{ij}^m u_{ij}^{m'} u_{ij}^{m''}\right]\left[\sum_{k=1}^{K} v_{jk}^n v_{jk}^{n'} v_{jk}^{n''}\right] \\ &= \sum_{m''=1}^{M}\sum_{n''=1}^{N} \hat{a}_{m''n''} \sum_{j=1}^{J} \alpha_j^{mm'm''} \beta_j^{nn'n''} \\ &= \sum_{m''=1}^{M}\sum_{n''=1}^{N} \hat{a}_{m''n''} \gamma^{mnm'n'm''n''},\end{aligned} \tag{9}$$

where the factors $\alpha_j^{mm'm''}$ and $\beta_j^{nn'n''}$ denote the summations $\sum_{i=1}^{I} u_{ij}^m u_{ij}^{m'} u_{ij}^{m''}$ and $\sum_{k=1}^{K} v_{jk}^n v_{jk}^{n'} v_{jk}^{n''}$, respectively, and the factor $\gamma^{mnm'n'm''n''}$ denotes the sum $\sum_{j=1}^{J} \alpha_j^{mm'm''} \beta_j^{nn'n''}$.

Using the factorization given by equation (9), it can be shown that most of the overhead associated with computing the symmetric matrix elements $\psi^{mnm'n'}$ lies in calculating the $\alpha_j^{mm'm''}$ factors and the $\gamma^{mnm'n'm''n''}$ factors. These calculations take about $[(I/N^3)+1]JQ$ multiply-and-add operations, where $Q = (MN)^2(MN+1)/2$. By comparison, relatively straightforward computation of the summations given by equation (8) takes about $IJKQ$ multiply-and-add operations. Thus, for the computer simulations described in Section III, for which I/N^3 = 1/2, the factorization given by equation (9) reduces the computation by a factor of about $(2/3)IK \approx 20{,}000$.

B. Covariance Between Integrated Time-Activity Curve Model Segments

Given an estimate of $\text{cov}(\hat{\mathbf{a}})$, the covariance matrix for the spatiotemporal basis function coefficients, estimates of the covariance between integrated segments of the time-activity curve models for the volumes of interest can be obtained as follows.

The integral of the time-activity curve model for volume of interest m, during the time interval associated with angle j of rotation k, can be expressed as $\sum_{n=1}^{N} \hat{a}_{mn} v_{jk}^n$. Thus, the covariance of this time integral with the time integral associated with volume of interest m' during angle j' of rotation k' is

$$\begin{aligned}&\text{cov}\left(\sum_{n=1}^{N} \hat{a}_{mn} v_{jk}^n, \sum_{n=1}^{N} \hat{a}_{m'n} v_{j'k'}^n\right) \\ &\qquad = \sum_{n=1}^{N}\sum_{n'=1}^{N} v_{jk}^n \,\text{cov}(\hat{a}_{mn}, \hat{a}_{m'n'})\, v_{j'k'}^{n'},\end{aligned} \tag{10}$$

and the variance of each time integral is

$$\begin{aligned}\sigma_{jkm}^2 &= \text{var}\left(\sum_{n=1}^{N} \hat{a}_{mn} v_{jk}^n\right) \\ &= \sum_{n=1}^{N}\sum_{n'=1}^{N} v_{jk}^n \,\text{cov}(\hat{a}_{mn}, \hat{a}_{mn'})\, v_{jk}^{n'}.\end{aligned} \tag{11}$$

As a figure of merit related to the global precision of the time-activity curve model for volume of interest m, the following expression yields a squared noise-to-signal ratio calculated as the mean (over all of the time segments) of the expected values of the squared errors between the integrated segments of the "true" and modeled curves, normalized by the mean square value of the integrated segments of the "true" curve:

$$\xi_m^2 = \frac{\sum_{j=1}^{J} \sum_{k=1}^{K} \sigma_{jkm}^2}{\sum_{j=1}^{J} \sum_{k=1}^{K} \left[\sum_{n=1}^{N} \hat{a}_{mn} v_{jk}^n\right]^2}. \qquad (12)$$

Substituting equation (11) into equation (12), the squared noise-to-signal ratio ξ_m^2 can be calculated quickly by rearranging the summations, precomputing the inner products of the temporal basis functions, $\nu^{nn'} = \sum_{j=1}^{J} \sum_{k=1}^{K} v_{jk}^n v_{jk}^{n'}$, and exploiting the symmetry with respect to the indices n and n':

$$\xi_m^2 = \frac{\sum_{n=1}^{N} \sum_{n'=1}^{N} \mathrm{cov}(\hat{a}_{mn}, \hat{a}_{mn'}) \nu^{nn'}}{\sum_{n=1}^{N} \sum_{n'=1}^{N} \hat{a}_{mn} \hat{a}_{mn'} \nu^{nn'}}. \qquad (13)$$

III. COMPUTER SIMULATIONS

To validate the fast algorithm presented in Section II, 1600 realizations of cone beam projection data having Poisson noise were generated using the simulation apparatus described in [5].

Simulated spatiotemporal distributions were obtained using the Mathematical Cardiac Torso (MCAT) phantom developed at the University of North Carolina [12]. The emission phantom (Fig. 1) was composed of 128 contiguous 1.75 mm-thick slices and contained the blood pool, three myocardial tissue volumes of interest (normal myocardium, septal defect, and lateral defect), liver, and background tissue. Projections were attenuated using the corresponding MCAT attenuation phantom.

The simulated time-activity curves (Fig. 2) mimicked the kinetics of teboroxime [13]. The simulated 15 min data acquisition consisted of $I = 2048$ cone beam projection rays per angle (64 transverse × 32 axial), $J = 120$ angles per revolution, and $K = 15$ revolutions, and thus yielded about 3.7 million projection samples. The projection bins were 7 mm × 7 mm at the detector, and the detector was 30 cm from the center of the field of view. The cone beam collimators had a hole diameter of 2 mm, a length of 4 cm, and were offset 1 cm from the detector. The focal length was 70 cm, which resulted in truncation of the data (Fig. 1). Attenuation and geometric point response were modeled using a ray-driven projector with line length weighting [14]. Scatter was not modeled. The amplitude of the simulated blood input function was adjusted so that about 10 million events were detected using the cone beam collimators.

The spatial basis projection factors u_{ij}^m were defined by forward projecting each of the six known emission volumes composing the MCAT phantom (Fig. 1). Each emission volume was modeled to contain spatially uniform activity, which yielded $M = 6$ sets of spatial basis projection factors.

The temporal basis integral factors v_{jk}^n were defined by integrating $N = 16$ splines spanning 15 time segments having geometrically increasing length (Fig. 3). Piecewise quadratic B-splines were used with an initial time segment length of 10 sec. The resulting curve models were continuous through their first derivative and yielded errors of less than 2% for noiseless projections, where the error was defined to be the root mean square (rms) difference between the simulated curve and the spline model, normalized by the rms value of the simulated curve [5].

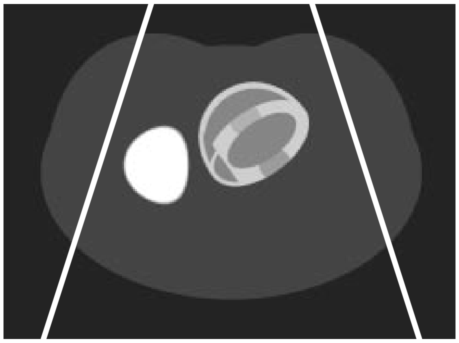


Fig. 1. Transverse cross section through the MCAT emission phantom, showing the truncation of data resulting from the use of cone beam collimators.

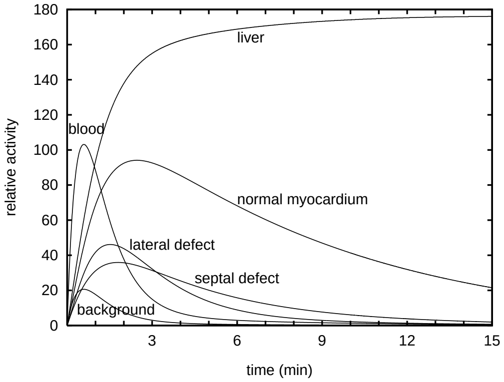


Fig. 2. Simulated time-activity curves for the volumes shown in Fig. 1.

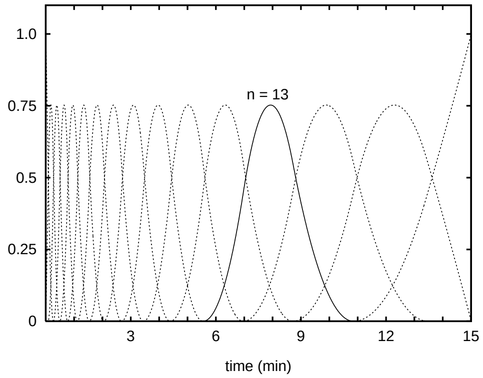


Fig. 3. Piecewise quadratic B-spline temporal basis functions. Sixteen splines are used to span 15 time segments having geometrically increasing length. The initial time segment length is 10 sec. The thirteenth spline is shown as a solid curve.

The computational benefit of using the factorization given by equation (9) to estimate the covariance matrix for the spatiotemporal basis function coefficients was evident in the simulations. The number of multiply-and-add operations used to calculate $\mathbf{F}^{\mathrm{T}}\mathrm{diag}(\mathbf{F}\hat{\mathbf{a}})\mathbf{F}$ was reduced from about 1.6 trillion to about 80 million. Using a 194-MHz R10000-based SGI workstation, it took 34 sec to estimate the 96 coefficients for the spatiotemporal basis functions, their covariance matrix, and the squared noise-to-signal ratios given by equation (13).

TABLE I

ACTUAL AND ESTIMATED STATISTICAL UNCERTAINTIES FOR SPATIOTEMPORAL BASIS FUNCTION COEFFICIENTS, FOR 1600 REALIZATIONS OF NOISY PROJECTIONS. THE SAMPLE STANDARD DEVIATIONS OF THE COEFFICIENTS (THE SECOND COLUMN IN EACH OF THE FOUR SUB-TABLES) AGREE CLOSELY WITH THE SAMPLE MEANS OF THE ESTIMATED STATISTICAL UNCERTAINTIES (THE THIRD COLUMN IN EACH OF THE FOUR SUB-TABLES).

n	blood pool			normal myocardium			septal defect			lateral defect		
	$\hat{a}_{1n}$		$\hat{\sigma}_{\hat{a}_{1n}}$	$\hat{a}_{2n}$		$\hat{\sigma}_{\hat{a}_{2n}}$	$\hat{a}_{3n}$		$\hat{\sigma}_{\hat{a}_{3n}}$	$\hat{a}_{4n}$		$\hat{\sigma}_{\hat{a}_{4n}}$
	sample mean	sample sdev	sample mean	sample mean	sample sdev	sample mean	sample mean	sample sdev	sample mean	sample mean	sample sdev	sample mean
1	0.279	0.128	0.131	0.0102	0.166	0.166	-0.0529	1.28	1.29	-0.212	0.747	0.762
2	5.20	0.138	0.140	1.05	0.171	0.174	0.559	1.33	1.35	0.979	1.47	1.46
3	7.65	0.157	0.160	2.30	0.185	0.187	1.36	1.89	1.88	1.38	3.00	2.99
4	8.20	0.172	0.175	3.75	0.331	0.333	1.89	3.04	3.05	2.42	1.95	1.95
5	7.19	0.153	0.160	5.15	0.222	0.219	2.33	1.91	1.94	3.25	1.26	1.27
6	5.40	0.105	0.107	6.30	0.148	0.145	2.70	1.20	1.24	3.58	1.75	1.73
7	3.51	0.0974	0.0953	7.07	0.141	0.140	2.87	1.50	1.54	3.59	0.876	0.864
8	2.00	0.0658	0.0668	7.36	0.108	0.109	2.67	0.886	0.904	3.11	1.08	1.09
9	1.04	0.0545	0.0543	7.21	0.0839	0.0836	2.45	0.600	0.602	2.44	0.593	0.602
10	0.535	0.0464	0.0446	6.71	0.0743	0.0735	2.02	0.560	0.562	1.69	0.418	0.411
11	0.308	0.0361	0.0362	5.98	0.0632	0.0628	1.57	0.486	0.477	1.03	0.364	0.354
12	0.208	0.0305	0.0298	5.13	0.0538	0.0530	1.13	0.397	0.399	0.592	0.296	0.300
13	0.149	0.0250	0.0245	4.20	0.0454	0.0441	0.761	0.327	0.325	0.307	0.258	0.253
14	0.103	0.0197	0.0201	3.28	0.0360	0.0365	0.478	0.268	0.270	0.179	0.206	0.207
15	0.0657	0.0167	0.0168	2.39	0.0303	0.0308	0.268	0.232	0.234	0.0853	0.165	0.170
16	0.0399	0.0169	0.0169	1.64	0.0314	0.0313	0.135	0.270	0.269	0.0651	0.161	0.164

TABLE II

ACTUAL AND ESTIMATED NOISE-TO-SIGNAL RATIOS FOR TIME-ACTIVITY CURVES, FOR 1600 REALIZATIONS OF NOISY PROJECTIONS.

	noise-to-signal ratio (%)		ξ_m (%)	
	sample mean	sample sdev	sample mean	sample sdev
blood pool	1.51	0.35	1.56	0.008
normal myocardium	1.13	0.27	1.16	0.003
septal defect	32.5	9.1	32.5	2.6
lateral defect	28.5	8.0	28.6	2.3
liver	0.167	0.031	0.170	0.0001
background	0.242	0.058	0.247	0.0002

Table I shows that the sample means of the $\hat{\sigma}_{\hat{a}_{mn}}$ (the square roots of the diagonal elements of the estimated covariance matrix) were within 5% of the sample standard deviations of the $\hat{a}_{mn}$ (the estimated spatiotemporal basis functions coefficients), for the blood pool and myocardial tissue volumes. For the liver and background tissue, the agreement was to within 4%. The coefficients of variation for the $\hat{\sigma}_{\hat{a}_{mn}}$ were less than 2%.

Table II shows that the sample means of the ξ_m (the estimated noise-to-signal ratios) were within 4% of the sample means of the rms differences between the 1600 sets of time-activity curve models and their corresponding mean curves, normalized by the rms values of the mean curves. The time-activity curves for the septal and lateral defects exhibited the largest variability, because of their small spatiotemporal support.

IV. FUTURE DIRECTIONS

The fast algorithm presented in Section II will facilitate the study of the statistical variability that results from modeling various orders of temporal continuity and using various time samplings, when estimating time-activity curves directly from dynamic cone beam SPECT projections. It may also allow compartmental model kinetic parameters to be estimated from the curves more precisely, using nonlinear weighted least squares.

REFERENCES

[1] H K Tuy, "An inversion formula for cone-beam reconstruction," *SIAM J Appl Math*, vol. 43, no. 3, pp. 546–552, 1983.

[2] R H Huesman, B W Reutter, G L Zeng, and G T Gullberg, "Kinetic parameter estimation from SPECT cone-beam projection measurements," *Phys Med Biol*, vol. 43, no. 4, pp. 973–982, 1998.

[3] G T Gullberg, R H Huesman, S G Ross, E V R Di Bella, G L Zeng, B W Reutter, P E Christian, and S A Foresti, "Dynamic cardiac single-photon emission computed tomography," in *Nuclear Cardiology: State of the Art and Future Directions*, B L Zaret and G A Beller, Eds., chapter 11, pp. 137–187. Mosby Inc, St Louis, 1999.

[4] A Sitek, E V R Di Bella, and G T Gullberg, "Reconstruction from slow rotation dynamic SPECT using a factor model," in *Information Processing in Medical Imaging: Proceedings of the Sixteenth International Conference*, A Kuba, M Šámal, and A Todd-Pokropek, Eds., 1999, pp. 436–441.

[5] B W Reutter, G T Gullberg, and R H Huesman, "Direct least-squares estimation of spatiotemporal distributions from dynamic SPECT projections using a spatial segmentation and temporal B-splines," *IEEE Trans Med Imag*, vol. 19, no. 5, pp. 434–450, 2000.

[6] J S Maltz, "Direct recovery of regional tracer kinetics from temporally inconsistent dynamic ECT projections using dimension-reduced time-activity basis," *Phys Med Biol*, vol. 45, no. 11, pp. 3413–3429, 2000.

[7] A Celler, T Farncombe, C Bever, D Noll, J Maeght, R Harrop, and D Lyster, "Performance of the dynamic single photon emission computed tomography (dSPECT) method for decreasing or increasing activity changes," *Phys Med Biol*, vol. 45, no. 12, pp. 3525–3543, 2000.

[8] T Farncombe, A Celler, C Bever, D Noll, J Maeght, and R Harrop, "The incorporation of organ uptake into dynamic spect (dSPECT) image reconstruction," *IEEE Trans Nucl Sci*, vol. 48, no. 1, pp. 3–9, 2001.

[9] A R Formiconi, "Least squares algorithm for region-of-interest evaluation in emission tomography," *IEEE Trans Med Imag*, vol. 12, no. 1, pp. 90–100, 1993.

[10] R H Huesman and B M Mazoyer, "Kinetic data analysis with a noisy input function," *Phys Med Biol*, vol. 32, no. 12, pp. 1569–1579, 1987.

[11] D J Kadrmas, E V R Di Bella, R H Huesman, and G T Gullberg, "Analytical propagation of errors in dynamic SPECT: Estimators, degrading factors, bias and noise," *Phys Med Biol*, vol. 44, no. 8, pp. 1997–2014, 1999.

[12] B M W Tsui, J A Terry, and G T Gullberg, "Evaluation of cardiac cone-beam single photon emission computed tomography using observer performance experiments and receiver operating characteristic analysis," *Invest Radiol*, vol. 28, no. 12, pp. 1101–1112, 1993.

[13] R K Narra, T Feld, and A D Nunn, "Absorbed radiation dose to humans from technetium-99m-teboroxime," *J Nucl Med*, vol. 33, no. 1, pp. 88–93, 1992.

[14] G L Zeng, G T Gullberg, B M W Tsui, and J A Terry, "Three-dimensional iterative reconstruction algorithms with attenuation and geometric point response correction," *IEEE Trans Nucl Sci*, vol. 38, no. 2, pp. 693–702, 1991.

A rebinning technique for 3D reconstruction of Compton camera data

Junqiang Li[1], John D. Valentine[1], John N. Aarsvold[2], Murat Khamzin[1],
[1]Nuclear and Radiological Engineering Program, Georgia Institute of Technology
[2]Department of Radiology, Emory University

ABSTRACT: A newly developed 3D image reconstruction technique for Compton camera data is described in this paper. For Compton cameras, the energies and positions of gamma-ray interactions in at least two detectors from a single incident photon are recorded using coincidence techniques. Based on this information, the Compton scattering formula establishes a cone surface from which the incident photon must have originated. Through an extension of the previously developed rebinning technique, instead of tracing the entire cone surface into the image space, a number of lines on the cone surface are sampled. All the lines start from the apex of the cone and are evenly distributed over the cone surface. The number of lines on each cone is determined by the desired spatial resolution. Each line is then projected to a perpendicular imaginary detector plane. The 2D Fourier transform of the line-projection data on this plane is shown to be one rotated plane of the 3D Fourier transform of the source distribution in the frequency domain. By projecting all of the sampled lines, performing Fourier transforms on all of the projected data, and summing up all the transformed data in the frequency domain, the 3D Fourier transform of the source distribution can be obtained. Interpolation and geometry normalization of data points in the 3D frequency domain will subsequently be applied. An image can then be reconstructed by a 3D inverse Fourier transform. The development of this technique will be discussed in detail.

I. INTRODUCTION

For a Compton scatter camera as illustrated in Figure 1 in which it is assumed that one emitted gamma ray Compton scatters in the primary detector then undergoes a photoelectric absorption in a secondary detector (this interaction sequence is defined as a preferred event), a cone surface can be established to determine the possible locations from where the incident photon was emitted [1-3]. The location and orientation of the cone surfaces are random in space and depend on where the two interactions occur in these two detectors, as well as the energies deposited. Thus image reconstruction is a difficult task. One potential 3D image reconstruction technique is the conventional Fourier deconvolution, which is based on projecting cone surfaces into the image space to establish a data matrix and then deconvolution of the projected data matrix with the point spread function that was previously determined for the image system [4]. In this study, a rebinning technique is being developed for the 3D image reconstruction of Compton camera data.

II. METHODOLOGY

Prior to describing the rebinning method for image reconstruction for Compton cameras, some aspects of rebinning need to be discussed. As illustrated in Figure 2, f(x,y,z) is a function in 3D space and P is an arbitrary plane that crosses the origin O. A rotated coordinate system (O,X′,Y′,Z′) is established such that the X′-Y′ plane is plane P. The relationship between the rotated coordinates and the global coordinates is simply a rotation operator R_P. Function f(x,y,z) in the rotated coordinates will be:

$$f(x', y', z') = R_P\left[f(x, y, z)\right] \tag{1}$$

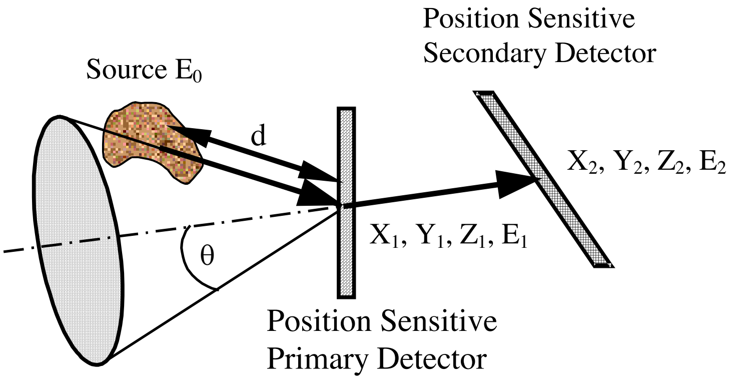


Figure 1. Principle of Compton scatter camera. E_1 and E_2 correspond to the energy deposited in the primary and secondary detectors, respectively.

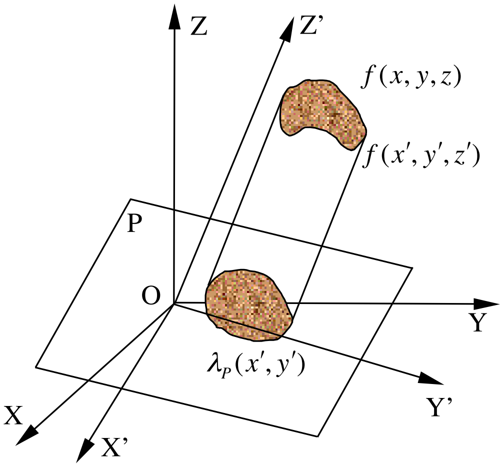


Figure 2. Projection of a 3D function f(x′,y′,z′) (in rotated coordinates) onto a plane P.

When f(x,y,z) is perpendicularly projected onto P, it is projected along the Z′ axis in the rotated coordinates. We denote this projection as $\lambda_P(x',y')$ in the rotated coordinate system:

$$\lambda_P(x',y') = \int_\infty f(x',y',z')dz'. \tag{2}$$

By taking the Fourier transform of $\lambda_P(x',y')$ on x′ and y′, we get

$$\begin{aligned}\Lambda_P(u',v') &= F_2\left[\lambda_A(x',y')\right] = \int_\infty dx' \int_\infty \lambda_P(x',y') e^{\left[-2j\pi(u'x'+v'y')\right]} dy' \\ &= \int_\infty dx' \int_\infty dy' \int_\infty f(x',y',z') e^{\left[-2j\pi(u'x'+v'y')\right]} dz',\end{aligned} \tag{3}$$

where the F_n operator is the Fourier transform operator of dimension n and *u*′ and *v*′ are the frequency components corresponding to x′ and y′, respectively.

The 3D Fourier transform of f(x′, y′, z′) in the rotated coordinate system gives

$$F'(u',v',w') = F_3\left[f(x',y',z')\right] = \int_\infty dx' \int_\infty dy' \int_\infty f(x',y',z') e^{\left[-2j\pi(u'x'+v'y'+w'z')\right]} dz'. \tag{4}$$

Comparison of equations (3) and (4) yields

$$\Lambda_P(u',v') = F'(u',v',w')\big|_{w'=0}\ . \tag{5}$$

The physical meaning for equation (5) can be stated as: the 2D Fourier transform of line-projection of a 3D function onto a plane is a directly rotated plane (u′,v′,0) through the 3D transform of the original function in the frequency domain. The rotated plane in frequency domain has the same orientation as the plane to which the function is projected. This can be seen as an extension of the central slice theorem of the Radon transform from 2D to 3D, although it is different from the 3D Radon transform [5-7].

The quickest route to calculate f(x,y,z) with the knowledge of $\lambda_P(x',y')$ is through equation (5). If a complete set of projections are available, i.e. $\lambda_P(x',y')$ are known in every plane that crosses the origin, the Fourier transform of all projections will give the complete Fourier transform of the original 3D function. So f(x,y,z) can be obtained as

$$f(x,y,z) = F_3^{-1}\left[F(u,v,w)\right] = \int_\infty du \int_\infty dv \int_\infty \sum R_P^{-1}\left[\Lambda_P(u',v')\right] \cdot e^{\left[2j\pi(ux+vy+wz)\right]} dw, \tag{6}$$

where R^{-1}_P is the rotation operator transforming Λ_P from the rotated coordinates into the original coordinates in the frequency domain. Σ sums up all of the projections to establish the complete 3D Fourier transform of the original function.

III. REBINNING COMPTON CAMERA DATA

For the Compton camera, instead of tracing the entire cone surface into the source space to reconstruct an image, we propose to sample lines on the cone surface. All lines originate from the vertex of the cone surface and are evenly distributed over the cone surface. Consequently, we have two options for sampling $\lambda_A(x',y')$: 1) predefined planes and 2) arbitrary planes.

a) Predefined Planes

A series of planes each of which covers an equivalent solid angle and covers half of the sphere (i.e., a total solid angle of 2π is thus subtended) are predefined. Each line is projected onto one and only one plane to which this line is closest to being perpendicular. All of the planes that have projections from this sampled cone will mark the projected spot with a weighting factor such that when summing up all the points projected by the lines from one cone, the total weight is 1. In this manner, each cone represents only one valid event and shall contribute equally to the image.

After all of the cones have been rebinned, the 2D Fourier transform is performed on each of the predefined planes to get its frequency components. Each Fourier transform undergoes a rotation operation, according to the orientation of the plane, into the same universal coordinates. Summation of all of the rotated data, interpolation of data points where data are missing, as well as geometry normalization operations will then be employed. By performing 3D inverse Fourier transform on the summation of all the data, the original 3D function can be constructed.

b) Arbitrary Planes

An alternative means of calculating $\lambda_A(x',y')$ can be approached as follows. Each line on the cone surface can be treated as the norm to a plane, and therefore can define a plane that passes through the origin of coordinates. Since this line serves as the norm to the projection plane, the projection will be exact, no approximation is introduced as it was in a). Theoretically there will be the same total number of planes as the total number of lines, and each plane will only have one projection point. Since the Fourier transform and rotation are both linear operations, one can perform either the summation or Fourier transform in equation (4).

The second approach is still under development. The computational load seems to be very large for this method, since there will be one 2D Fourier transform for each sampled line. However, while all the plane sizes are the same, even though there might be only one point on this plane, there are only limited combinations of how this point is distributed over this plane. Therefore a table can be constructed *a priori* for the Fourier transform of every possible projection. By looking up the corresponding Fourier transform in the table, this time-consuming step may be executed relatively quickly.

III. PRELIMINARY RESULTS

For the rebinning strategy a) in the previous section, if the planes used here are a series of planes rotating along one single axis, and each plane only covers a small portion of solid angle, the image reconstruction is similar to standard SPECT. In this case, the filtered backprojection algorithm can be directly applied to the projection data to reconstruct the image. Figure 3 is an illustration of this case. 5 point sources are simulated using MCNP with 1.6 million preferred events.

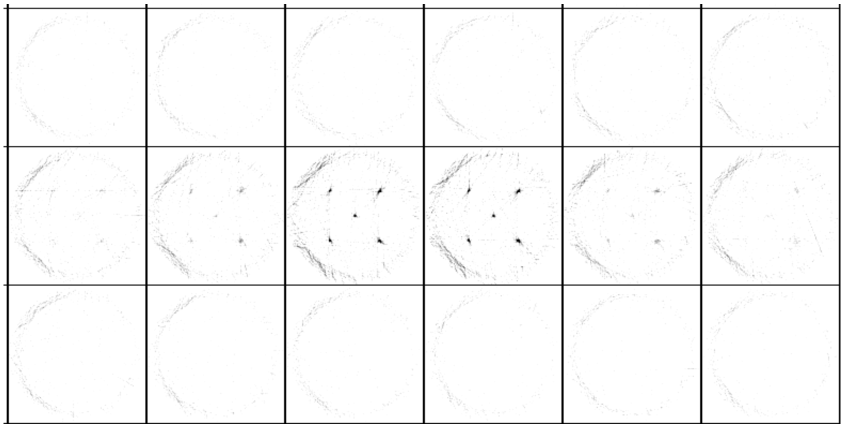


Figure 3. Reconstructed image with 5 point sources with equal activity. Only 18 out of 128 slices are shown here. 360 planes are used for projection, and they all rotate along the same axis with each one covering 0.5-degree angle in each direction.

With 360 planes rotating along the Z-axis and each covering an equal solid angle, 0.5 degrees azimuthally and 0.5 degrees in z direction, the filtered backprojection algorithm is adopted. No noise has been added to the simulation. Each slice covers a 22 cm × 22 cm area and is divided into 128 × 128 pixels. Distance between slices is the same as the space between pixels within one slice. Though only a very small portion of solid has been taken into account (~0.4%) for 1.6 million cones (i.e., only a small amount of data is used), the constructed image shows clearly the 5 point sources. Notice there are some artifacts around the edge of the image space. This kind of artifact is probably due to the fact that the lines are much denser at the vertex area, and it is possible to filter out this artifact by introducing certain types of filters during the reconstruction.

IV. SUMMARY

A rebinning technique for Compton camera image reconstruction is proposed and under development. Preliminary results show that this is a promising technique for fully 3D image reconstruction of Compton camera data. The full manuscript will provide a detailed discussion of this technique. The discussion will cover the derivation of the transform, analysis of the inverse transform, properties of this kind of transform in space domain as well as in frequency domain, and discussion on reconstructed images.

V. REFERENCES:

[1] M. Singh, "An electronically collimated gamma camera for single-photon emission computed tomography. Part I: Theoretical considerations and design criteria," *Med. Phys.*, vol. 10, No. 4, pp. 421-427, 1983.

[2] M. Singh and D. Doria, "An electronically collimated gamma camera for single photon emission computed tomography. Part II: Image reconstruction and preliminary experimental measurements," *Med. Phys.*, vol. 10, No. 4, pp. 428-435, 1983.

[3] R. C. Rohe and J. D. Valentine, "An energy subtraction Compton scatter camera design for *in vivo* medical imaging of radio-pharmaceuticals," *IEEE Trans. Nucl. Sci.* vol. 43, No. 6, 1996.

[4] P.E. Kinahan, J. G. Rogers, "Analytic 3D image reconstruction using all detected events", *IEEE Trans. Nucl. Sci.* vol. 36, No. 1, pp. 964- 968, 1989. J. G. Golsher, Fully three-dimensional positron emission tomography , *Phys. Med. Biol.,* vol. 25, No. 1, pp. 103-115, 1980.

[5] S. R. Deans, *The Radon transform and some of its applications*, John Wiley & Sons, Inc. 1983.

[6] H. H. Barrett, The Radon transform and its applications , *E. Wolf, Progress in Optics XXI*, Elsevier Science Publishers, pp. 219-286, 1984.

[7] M. Defrise, A factorization method for the 3D x-ray transform , *Inverse Problems*, vol. 11, pp. 983-994, 1995.

A Fully 3D Maximum Likelihood Estimator Algorithm with Attenuation and Scatter Correction Adapted to a Limited Angle Positron Camera

F. Pönisch[1], W. Enghardt[1], K. Lauckner[2]
[1]Forschungszentrum Rossendorf, Postfach 510119, 01314 Dresden, Germany
[2]Scientific Consulting Group GmbH, Bismarckallee 9, 79098 Freiburg, Germany

Abstract—**The in beam dual head positron camera BASTEI (Beta$^+$ Activity meaSurements at the Therapy with Energetic Ions) is used to monitor and control the applied dose distributions simultaneously to tumour irradiations with carbon ion beams at the experimental heavy ion therapy facility at GSI Darmstadt. Therefore, the PET system has been mounted directly at the treatment site. A fully 3D reconstruction algorithm based on the Maximum Likelihood Estimator algorithm has been developed and adapted to a strongly spatial varying imaging situation. The scatter and attenuation correction is applied to the measured list mode data before each iterative step. This requires an attenuation map containing the information on the tissue composition and densities. This information is derived from the X-ray computed tomograms (CT) of the patient and the patient fixation system including the head rest. The scatter correction method uses the subtraction of calculated scattered from measured data. The normalization of scattered events relative to the unscattered events is done by a global scatter fraction factor. The results are presented.**

1. INTRODUCTION

Since 1997 85 patients mainly suffering from head and neck tumours have been treated at the experimental heavy ion tumour therapy facility at GSI Darmstadt. The carbon ion therapy is favourable for treating compact, radioresistant tumours in close vicinity of organs at risk. Therefore, a dedicated in-beam PET system has been built. During the irradiation β^+-radioactive nuclei are produced by nuclear fragmentation reactions along the path of the beam. This activity can be related to the original dose distribution and imaged in-situ by means of positron emission tomography. The dedicated positron camera BASTEI has been integrated into the treatment facility, it has a dual-head geometry in order to avoid interference with the horizontally fixed beam and restrictions to the patient positioning [1]. The two large-area detector heads, which have been assembled from components of the ECAT EXACT® PET-scanner, are mounted above and below the patient couch. The measured list mode data are reconstructed by a fully 3D noise suppressing Maximum Likelihood algorithm.

Furthermore, the expected counting statistics is limited by the doses applied per therapy fraction and is usually two to three orders of magnitude lower compared with those of typical PET-scans in nuclear medicine. Nevertheless, the rather simple shape of the activity distributions alleviates the reconstruction problem.

The individual detector response of the 4 Mio. lines of response (LOR) are corrected by means of a normalisation scan. It is performed with a line source moving through the field of view (FOV) of the camera.

The process of photon scattering considerably influences the image quality in PET. For head and neck tumours about 20 % of the registered true coincidences are influenced by Compton or Rayleigh scattering, which may destroy the correspondence between the source and the reconstructed radioactivity distribution especially in highly inhomogeneous regions of the human body.

2. IMPLEMENTATION

Since the β^+-activity distribution is well localized to the irradiated volume, the scatter correction algorithm requires the information on the tissue composition and density only within and nearby the camera FOV. This is derived from the X-ray CT of the patient (created for treatment planning purpose) which is automatically combined with a CT of the head rest (Fig. 1) using characteristic landmarks. The created data set is the basis for the calculation of the attenuation correction factors. The correct position of these CTs with respect to the positron camera is derived from stereotactic coordinates used during the diagnostic CT scan and for patient positioning before the irradiation. The Fig. 1 shows the significant influence of the absorption by the head rest on the detector response leading, if not corrected, to image artefacts in the reconstruction. The dashed line denotes the acceptance cone of the tomograph

.

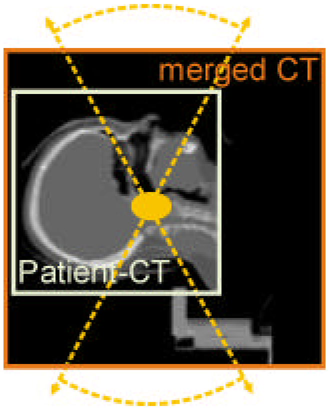


Figure 1: The density distribution used for attenuation and scatter correction obtained from the patient CT and from the CT of the patient support. A typical tumour location and the acceptance cone of the positron camera are displayed.

The reconstruction is based on the Maximum Likelihood Estimation algorithm [2] and has been extended to a three dimensional image space [3]:

$$x_j^{(n+1)} = x_j^{(n)} \cdot \sum_i \frac{a_{ij} y_i^{*(n)}}{\sum_{j'} a_{ij'} x_{j'}^{(n)}} \qquad (1)$$

$x_j^{(n)}$ denotes the activity distribution after the *n*-th iteration step, $y_i^{*(n)}$ is the modified measured projection for the *i*-th LOR and a_{ij} is the element of the transition matrix.

The aim was to model the above described non-standard imaging situation as accurate as possible within an acceptable computation time. The reconstruction is performed by using an image space with a size adapted to the irradiated volume (typical image size of 0.5 million voxels with a size of 1.6875 mm). During the treatment with a typical physical dose of 0.8 Gy per fraction around 100,000 coincidences are acquired.

The non-zero elements of the transition matrix required for a particular reconstruction are calculated at run time. In order to consider the individual crystal response, the volume of each LOR is sampled by 10 lines whose endpoints are randomly distributed over the crystal volume taking into account the interaction probability depending exponentially on the depth of interaction and the spatial orientation of the crystals.

The following pre-iterative correction for each coincidence channel *i* are applied:

$$y_i^{meas} = (y_i^p - y_i^r) \cdot \frac{w_i^p}{w_i^e} \qquad (2)$$

where y_i^p and y_i^r denotes the number of the measured prompt and random coincidences, respectively. The correction factor w_i^p depends on the length of the emission line within the image space divided by the total length of the longitudinal image space axis whereas w_i^e is the detection efficiency.

The scatter correction is applied to the measured projections before each iteration step. Thus, the events registered for LOR *i* (y_i^{meas}) can be considered as the sum of the unscattered events from this LOR ($y_i^{unscatt}$) and of events scattered into this LOR from outside (y_i^{scatt}):

$$y_i^{meas} = y_i^{unscatt} + y_i^{scatt}. \qquad (3)$$

The correction step of equation (1) has to be performed for the unscattered and attenuation corrected events, i.e.:

$$y_i^{*(n)} = \frac{y_i^{unscatt}}{a_i} = \frac{y_i^{meas} - y_i^{scatt}}{a_i}, \qquad (4)$$

where a_i is the total attenuation along LOR *i*.

To calculate the quantity y_i^{scatt} we follow the approximation of Ollinger [4] and Watson [5] denoted as the Single Scatter Simulation (SSS). This algorithm approximates the scatter coincidences in a given LOR by single Compton scattered events only. At first, a sample of around 1000 scatter points S is randomly distributed throughout the scatter volume where the density exceeds 0.1 g/cm^3 (Hounsfield > -900).

Figure 2 shows a LOR labelled *AB*. For each scatter point S there will be two distinct contributions to the single scattered coincidences, depending on which side of the scatter point the annihilation takes place (Q_1, Q_2).

The first addend in equation (6) supposes that the annihilation takes place at a point Q_1 between S and A. The photon which reaches the crystal A is unscattered. The scatter contribution for *AB* is calculated according to the following formula in a discrete notation:

$$y_i' = \sum_S \sum_{j \in AS} x_j^{(n)} \cdot e^{-\int_A^S \mu(E,s)ds} \cdot \frac{dP_{sc}(\theta(S))}{d\Omega} \cdot \Omega_{detB} \cdot e^{-\int_S^B \mu(E',s)ds} + A \leftrightarrow B \qquad (5)$$

In this expression μ is the linear attenuation coefficient (derived from the merged CT), *E* and *E'* denote the photon energy before and after the Compton scatter, respectively, depending on the scatter angle θ, $\frac{dP_{sc}(\theta(S))}{d\Omega}$ is the probability of Compton scatter by the angle θ from S into B and Ω_{detB} is the solid angle of the detector relative to S.

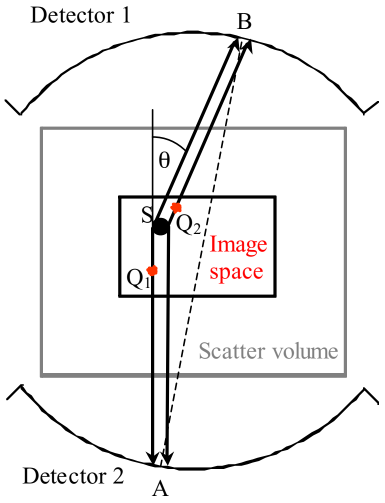


Figure 2: Trajectories of scattered photons

The total number of scattered events calculated

$$N_u^{scatt} = \sum_i y_i' \qquad (6)$$

has to be normalized to the number of acquired coincidences. Because of the limited angle geometry and the low counting statistics it is impossible to apply the normalization procedure according to Watson [5] which based on a scatter rate estimation from measured data. As an alternative we calculate by means of an additional Monte Carlo simulation the global scatter fraction SF from the *n*-th solution:

$$SF = \frac{N_{scatt}}{N_{true} + N_{scatt}} = \frac{N_{scatt}}{N_{meas}} \qquad (7)$$

where

$$N_{scatt} = \sum_i y_i^{scatt}, \quad N_{meas} = \sum_i y_i^{meas}.$$

From this follows:

$$y_i^{scatt} = y_i' \cdot \frac{SF \cdot N_{meas}}{N_u^{scatt}}. \qquad (8)$$

Introducing eq. (8) into eq. (4) yields the projection value $y_i^{*(n)}$ for each iteration step.

3. RESULTS

The algorithm has been validated using point-like ^{22}Na in a water filled cylindrical phantom (diameter 15 cm, height 15 cm).

Several geometrical configurations were simulated by means of Monte Carlo and reconstructed with the new scatter correction algorithm.

The reconstruction algorithm has been applied successfully to patient data. In dependence on the number of measured coincidences that contribute to the image space, the reconstruction time varies between 120 and 240 minutes on a R8000 processor (75 MHz) of an SGI Power Challenge. The faster converging algorithm OSEM is taken into account, but not implemented up to now.

4. REFERENCES

[1] J. Pawelke et al., "The investigation of different cameras for in-beam PET imaging", *Phys. Med. Biol.* vol. 41, pp. 279-296, 1996

[2] L.A. Shepp, Y. Vardi, "Maximum Likelihood reconstruction for emission tomography", *IEEE Trans. Med. Imag.*, MI-1 , pp.113-119, 1982.

[3] K. Lauckner, "Entwicklung eines iterativen 3D Rekonstruktionsverfahrens für die Kontrolle der Tumorbehandlung mit Schwerionen mittels der Positronen-Emissions-Tomographie", Thesis, Dresden University of Technology, 1999

[4] J.M. Ollinger, „Model based scatter correction for fully 3D PET", *Phys. Med. Biol.*, vol. 41; pp. 153-176, 1996

[5] C.C. Watson et al., „Evaluation of Simulation-Based Scatter Correction for 3-D PET Cardic Imaging", *IEEE Trans. Nuc. Sci.*, vol. 44, no. 1, pp. 90-97, 1997

Model-Based Normalization for Iterative 3D PET Image Reconstruction

[‡]B Bai, [‡]Q Li, [‡]E Asma, [†]Y C Tai, [†]A Chatziioannou, [‡]R M Leahy
[‡]Signal and Image Processing Institute, Univ. of Southern California, Los Angeles, CA 90089
[†]Crump Institute for Molecular Imaging, University of California, Los Angeles CA 90095

Abstract— **We describe a method for normalization in 3D PET for use with model-based image reconstruction methods. This approach is an extension of previous factored normalization methods in which we include separate factors for detector sensitivity, geometric response, block effects and deadtime. Since our MAP reconstruction approach already models some of the geometric factors in the forward projection, the normalization factors must be modified to account only for effects not already included in the model. We describe a maximum likelihood based approach to joint estimation of the normalization factors which we apply to data from a planar source. We then compute block-wise and block-profile deadtime correction factors using singles and coincidence data, respectively, from a multiframe cylindrical source. We have applied this method for reconstruction of data from the LSO Concorde P4 microPET scanner. Preliminary results compare favorably with those obtained using normalization based directly on cylindrical phantom measurements.**

Keywords— **3D PET, Normalization, Image Reconstruction**

I. INTRODUCTION

Accurate normalization is essential for accurate quantitative 3D PET. Inaccuracies in normalization factors can result in artifacts, poor uniformity, and increased noise in the reconstructed images. Traditional solutions to the normalization problem include direct and component-based methods. In direct methods, a known source of activity is scanned, then the normalization factors are estimated as the ratio between the ideal number of coincidences and those actually measured [1]. The main problem with this method is that it requires that a very large number of counts be detected to achieve acceptable statistical accuracy for each line or response (LOR). To maximize the number of counts over all LORs, direct approaches typically use a uniform cylindrical source. Unfortunately, this introduces its own problems since the observation model is complicated by a substantial scatter fraction.

Hoffman [2] proposed a component-based method which divides the normalization factors into detector efficiency and spatial distortion correction, which accounts for the radial mispositioning due to the geometry of the scanner. This model reduces the number of counts required by reducing the degrees of freedom in the normalization model so that the normalization factors are computed by averaging over multiple LORs. Casey [3] and Badawi [4] extended this concept to develop sophisticated models accounting for a wide variety of factors affecting detection efficiency. Casey's normalization model includes intrinsic detector efficiency, geometric factors, crystal interference and deadtime factors, Badawi added time-alignment factors and a count-dependent block-profile to this model.

These models are complex and involve the sequential estimation of multiple types of normalization factors, often from different data sets. This can lead to inconsistent estimates since the normalization models are multiplicative. While optimal estimation of individual components, e.g. the detector efficiencies [1], [5], have previously been investigated, joint estimation of all factors in the component-based models has not, to the best of our knowledge, previously been described. Here we present a unified model in which all factors are estimated simultaneously within a maximum likelihood framework. This model is specifically matched to the model-based Maximum A-Posteriori (MAP) reconstruction methods [6]. The combination of our previously described system model with matched normalization allows us to explicitly account for the imperfections in the line-integral model using an accurate physical and statistical model for coincidence detection. In this way we build on our previous model, which included effects of detector solid angle, photon pair non-colinearity and intercrystal scatter and penetration, to also include effects arising from the block design, individual detector efficiencies, geometric effects, and deadtime.

II. METHODS

A. *Normalization within a statistical image reconstruction framework*

We have developed a MAP estimation algorithm to reconstruct 3D PET images [6]. In this approach, the data are modeled as:

$$\overline{\mathbf{y}} = \mathbf{P}\mathbf{x} + \overline{\mathbf{r}} + \overline{\mathbf{s}} \tag{1}$$

where $\overline{\mathbf{y}}$ is the mean of the data, $\mathbf{x}$ is the source distribution, $\overline{\mathbf{r}}$ is the mean of the randoms, and $\overline{\mathbf{s}}$ is the mean of the scattered events. $\mathbf{P}$ is the system matrix describing the probability that an event is detected, which we factor as:

$$\mathbf{P} = \mathbf{P_{norm}}\mathbf{P_{blur}}\mathbf{P_{attn}}\mathbf{P_{geom}} \tag{2}$$

where $\mathbf{P_{geom}}$ is the geometric projection matrix describing the probability that a photon pair reaches the front faces of a detector pair in the absence of attenuation and assuming perfect photon pair colinearity. $\mathbf{P_{blur}}$ models photon

---

This work was supported by the National Cancer Institute under Grant No. R01 CA59794 and the National Foundation for Functional Brain Imaging.

pair non-colinearity, intercrystal scatter and crystal penetration, $\mathbf{P_{attn}}$ contains attenuation correction factors for each detector pair, and $\mathbf{P_{norm}}$ is a diagonal matrix containing the normalization factors.

The effects of solid angle variation at the detectors relative to the position of each voxel along a line of response and the angle between the detector surface and the LOR are accounted for in $\mathbf{P_{geom}}$. Similarly, the effects of crystal penetration that result in mispositioning of events towards the edge of the field of view is included in $\mathbf{P_{blur}}$. Consequently, these need not be included in the normalization factors as they are in previous factored methods [3]. Nevertheless, there are geometric factors that are not accounted for in (2) that we include in our normalization model as described below.

The normalization model we use here is that the diagonal matrix $\mathbf{P_{norm}}$ has elements

$$\mathbf{P_{norm}}(i,i) = \varepsilon_{d1^i}\varepsilon_{d2^i}g(l,j,k)^i\tau_{b1^i,b2^i}d_{d1^i}d_{d2^i} \quad (3)$$

where we have used i to index the LORs. The specific detectors forming this LOR are denoted $d1^i$ and $d2^i$, and $b1^i$ and $b2^i$ are the blocks containing these detectors. The components of this model are as follows:

Geometric factors $g(l,j,k)^i$: $(l,j,k)^i$ represent, respectively the radial position l, the view angle j, and the sinogram index k, associated with LOR i. The sensitivity of each LOR is a function of the position of the two detectors in the block and the distance of the LOR from the center of the field of view of the scanner. Our geometric factor essentially combines the geometric and block-interference patterns of [3] into a single factor. Since the scanner is highly symmetric both axially and transaxially, many LORs are equivalent in these respects. The average number of LORs sharing the same block and radial positions is approximately 2.3 × 168 (the number of blocks in the Concorde scanner - 336 detectors for each of 32 rings in blocks of 8 by 8). Fig. 1 illustrates the symmetries of the LORs with respect to the blocks that we use in computing the normalization: each LOR can have up to an 8-fold symmetry with respect to a single block. This pattern then repeats every 8 detectors for an 8 by 8 block design.

Detector efficiency $\varepsilon_{d1^i}, \varepsilon_{d2^i}$: these quantities describe the intrinsic efficiencies of the two detectors forming the LOR. The total number of these factors is equal to the number of detectors.

Time-alignment factor $\tau_{b1^i,b2^i}$: The time alignment factor is based on the model proposed by Badawi [4] to account for differences in timing synchronization between different blocks. As the timing windows become misaligned between any pair of blocks, so the detection efficiency will drop. We can characterize the timing properties of each block by a single delay factor. The time-alignment factor for each LOR is then a function of the difference between the delay factors for the two blocks for that LOR. The form of this function can be determined experimentally by varying the timing between a pair of blocks. We have not yet performed this experiment so that in the results presented below we instead estimate directly a separate time-alignment factor for each pair of blocks.

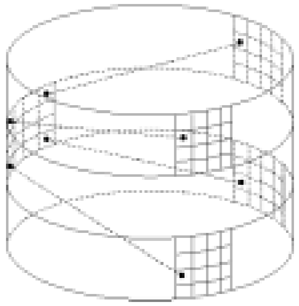


Fig. 1. The geometric factors are functions of the radial position of the LOR and the positions of the two detectors within their respective blocks. Illustrated here for a 4 by 4 block detector system is a four-fold symmetry in these factors. An additional two fold symmetry results from rotating and axially translating the LORs so that each of the four blocks on the right of the figure moves to the location of the left most block.

Deadtime factor $d_{d1^i}d_{d2^i}$: the deadtime factors are estimated separately as described in Section II-C.

B. Normalization factor estimation

We compute the normalization factors from the previous section using a joint optimization procedure. This differs from the common practice of using rotating rod sources to compute geometric factors and cylinder data to compute geometric efficiencies [3], [4]. This provides self consistent estimates of the unknown parameters. Moreover, by basing the estimation on the model (1), the normalization is matched to the specific forward projection model that we subsequently apply during reconstruction. For the plane source used in our studies, scatter is minimal and we currently ignore scatter contributions. For the Concorde microPET scanner, data is initially collected in listmode format so that we can re-sort into simultaneous prompt and delayed event sinograms. Prior to computing the normalization factors, we use a Bayesian technique to estimate the mean of randoms from the separate randoms sinogram [7].

We model the measurements as Poisson using the model (1) to give the log likelihood:

$$\begin{aligned} L(\mathbf{P_{norm}}) = & \\ & \sum_{i=1}^{N} y_i \log\{\varepsilon_{d1^i}\varepsilon_{d2^i}g(l,j,k)^i\tau_{b1^i,b2^i}[\mathbf{Px}]_\mathbf{i} + \bar{\mathbf{r}}_\mathbf{i}\} \\ & -\{(\varepsilon_{d1^i}\varepsilon_{d2^i}g(l,j,k)^i\tau_{b1^i,b2^i})[\mathbf{Px}]_\mathbf{i} + \bar{\mathbf{r}}_\mathbf{i}\} \end{aligned} \quad (4)$$

The source distribution $\mathbf{x}$ is the known plane source. We estimate the parameters by maximizing $L(\mathbf{P_{norm}})$ using a grouped coordinate ascent method, updating each of the different factors in turn using steepest ascent with a Newton-Raphson line search. We find in practice that effective convergence is reached in 5 iterations with 3 sub-iterations of line search at each main iteration.

C. Deadtime estimation

We assume for the purposes of computing the count-independent normalization factors, that the plane source is of sufficiently low activity that deadtime effects are minimal. This assumption is reasonable for the LSO detectors in the Concorde scanner for which deadtime factors are considerably lower than they would be for a BGO system.

Deadtime is affected by the properties of the PMT and detection electronics [4]. Rather than adopt the exponential model that was developed by Casey [3], we instead use an empirical quadratic correction method [4] which relates observed and true singles rates at each block by

$$\lambda_t = \frac{\lambda_a}{1 + \alpha\lambda_a + \beta\lambda_a^2} \qquad (5)$$

where λ_t is the true singles rate, λ_a is the detected count rate, and α and β are experimentally determined parameters. We allow a separate deadtime calibration of this type for each detector block in the system, based on the measured singles rate for that detector block.

Block detectors also exhibit an additional deadtime effect, characterized by gradual mispositioning of events towards the middle of the block as the count rate increases [8], [4]. This mispositioning contributes to a count dependent variation of sensitivity across the detector blocks, which we report on below. Our results indicate that these variations are significant so that we include these factors in our deadtime correction.

Thus our overall deadtime model for each detector d_i, similar to that in [4], is the product of the mispositioning deadtime $d_{mp}(d_i)$ and PMT and electronics deadtime $d_{pp}(b_i)$:

$$d_{d_i} = d_{mp}(b_i) \times d_{pp}(d_i) \qquad (6)$$

where b_i is the block containing detector d_i. The deadtime correction factor for each LOR is then the product of the factors for the two detectors forming the LOR.

We estimate the factors $d_{mp}(b_i)$ by observing the singles rate at each detector block over a series of I frames, taken as an F-18 source decays over the expected range of activities for the scanner. Since the LSO detectors have a natural background activity, we model the true activity at the block as

$$\lambda(t) = Ae^{-\phi t} + C \qquad (7)$$

where A is the initial singles rate, ϕ is the decay constant, and C is the background activity. Integrating this activity over the duration of each frame from time t_i to t_{i+T} and applying the deadtime model (5) we obtain the series of equations:

$$\frac{A}{-\phi T}\left[(e^{-\phi(t_i+T)} - e^{-\phi t_i})\right] + C = \frac{\lambda_{a,i}}{1 + \alpha\lambda_{a,i} + \beta\lambda_{a,i}^2} \qquad (8)$$

for $i = 1, \ldots I$ where $\lambda_{a,i}$ is the observed total singles rate at the ith block and α and β are the constants to be estimated. This set of equations are solved using nonlinear least squares to obtain a separate pair of parameters for each detector block. The mispositioning deadtime parameters were computed as described in [5].

III. EXPERIMENTAL STUDIES

A. Plane source experiments

Using the method described above, we estimated normalization factors for the Concorde MicroSystems P4 microPET scanner. We acquired data for a 90x200x2mm plane source of volume 38cc filled with 700 μCi of FDG. The first frame was collected for 20 minutes, the source was then rotated by 30 degree increments, with frame durations adjusted to achieve approximately equal counts in each frame. We windowed each sinogram to take only lines of response within $\pm 15^o$ of the normal to the plane. The set of six windowed sinograms were then used to estimate the normalization factors.

B. Deadtime experiments

A 221.6cc cylinder, diameter 5cm and length 15.5 cm, containing 4.80 mCi F-18 solution was scanned for 19 frames. The duration of each frame was 600s with each new frame starting 0.5 half-lives after the previous one. The average singles rate for the first frame for each block was 1.43×10^5, the average singles rate for the last frame was 827 (which was largely due to background radiation from LSO). The listmode format from the scanner allows us to acquire separate prompt and delayed sinograms and the singles rates for each individual detector block. These data were used to compute the deadtime factors.

C. Cylinder uniformity

Using the new normalization factors, we reconstructed a 5 cm diameter uniform cylinder. For comparison, we used a normalization file generated directly from a 2 hour duration frame from a uniform cylinder.

IV. RESULTS

Shown in Fig 2 are the block-wise and block-profile components of the deadtime correction factors. Fig 2(a) shows the excellent count-rate linearity for the LSO block detectors in the scanner over a wide activity range. The block-profile results indicate that these factors introduce significant count-dependent variations in sensitivity as a function of position in the block which should be included as part of the normalization process.

In Figs. 3 and 4 we show the effects of applying the model-based normalization procedure described here to reconstruction of a uniform cylinder compared to direct normalization based on a uniform cylinder. In both cases, the images were reconstructed using 30 MAP iterations with all algorithm parameters otherwise equal. Fig. 3 shows a single transaxial plane, and a profile through this plane, from the reconstructed cylinder. These results indicate some improvement in transaxial uniformity and a small reduction in noise. The latter observation was verified by computing region of interest variances which found a reduction in the percent standard deviation to mean ratio from 20% to 18%. The rise in activity towards the edge of the field of view that can be seen in both profiles is caused by the presence of scatter in the sinograms which was not corrected for in

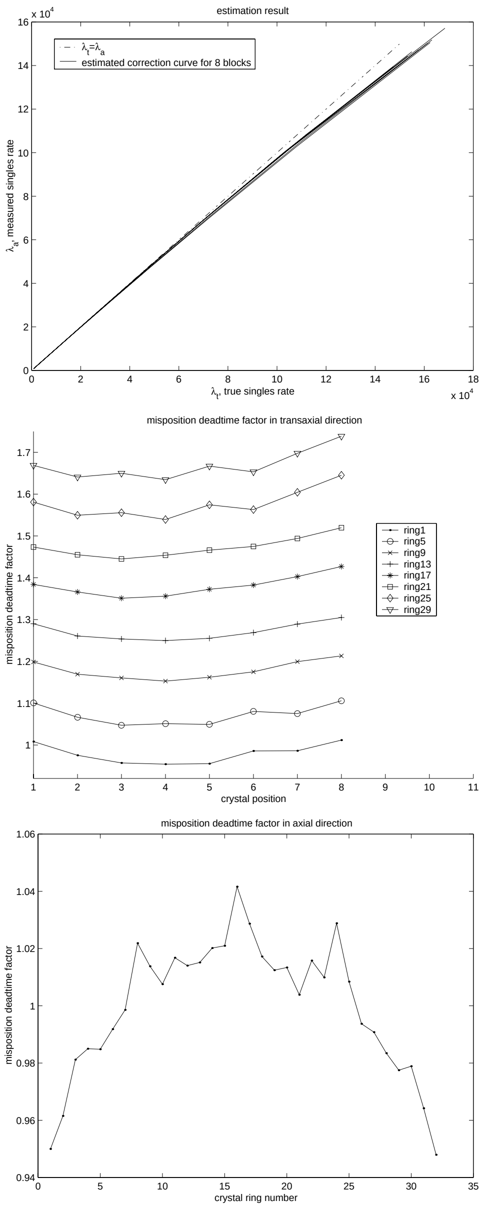


Fig. 2. Deadtime correction factors: (a) results from fitting the quadratic deadtime model to singles data from a 4.8mCi cylinder; shown are the measured vs. true singles rates for 8 different block detectors; (b) transaxial block profile factors computed from coincidence data for a measured singles rate of 20K for 8 different detector rings; (c) axial block profile from same data as (b).

these studies. Fig. 4 shows the axial uniformity as the total activity in each of the 63 reconstructed slices. Apart from over-correction in the first few planes, the model-based normalization produces improved axial uniformity compared to the direct normalization method.

The preliminary results presented show encouraging, if small, improvements in image uniformity compared to a direct normalization procedure. Further improvements should be realized as we refine our model. In order to compute the normalization factors in our approach, we must know the position of the plane source relative to the sinogram space. Currently we do this by comparing the measured plane source sinograms with the forward projection

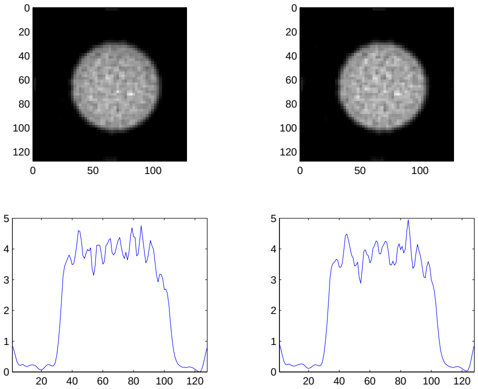


Fig. 3. Central plane of reconstructed cylinder. Left: model-based normalization; right: direct normalization.

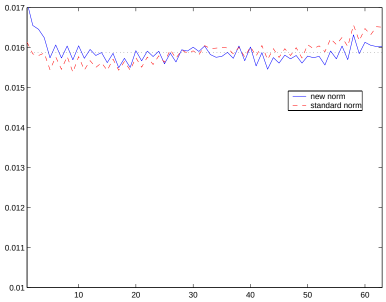


Fig. 4. Axial profiles of reconstructed cylinder using model-based and direct normalization.

of a simulated plane source, whose angle is adjusted so that the two sinograms match. Our procedure for performing this matching needs further refinement. We will also modify the time-alignment factor parameterization as we describe in Section II-A. Finally, we will include compensation for scatter to perform a full quantitative evaluation.

REFERENCES

[1] M. Defrise, D. W. Townsend, D. Bailey, A. Geissbuhler, C. Michel and T. Jones: A normalization technique for 3D PET data. Phys. Med. Biol. **36**, 1991, 939–952

[2] E. J. Hoffman, T. M. Guerrero, G. Germano, W. M. Digby, and M. Dahlbom: PET system calibrations and corrections for quantitative and spatially accurate images. IEEE Trans. Nucl. Sci. **33**, 1989, 1108–1112

[3] Casey M. E., Gadagkar H. and Newport D.: A component based method for normalization in volume PET. Proc. 3rd Int. Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, 1995 , 67–71

[4] R.D.Badawi, P. K. Marsden: Developments in component-based normalization for 3D PET. Phys. Med. Biol. **44**, 1999, 571–594

[5] R.D.Badawi: Aspects of optimisation and quantification in three-dimensional Positron Emission Tomography. Ph.D thesis, University of London

[6] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou and T. H. Farquhar: High resolution 3D Bayesian image reconstruction using the microPET small-animal scanner. Phys. in Med. Biol. **43**, 1998, 1001-1013

[7] E. U. Mumcuoglu, R. M. Leahy, S. R. Cherry, and Z. Zhou: Fast Gradient-Based Methods for Bayesian Reconstruction of Transmission and Emission PET Images. IEEE Trans. Med. Imag. **13**, 1994, 687–701

[8] G. Germano, E. J. Hoffman: A study of data loss and mispositioning due to pileup in 2-D detectors in PET. IEEE Trans. Nucl. Sci. **37** 1990, 671–675.

Iterative Reconstruction Methods for High-Throughput PET Scanners

James Hamill and Thomas Bruckbauer
CTI PET Systems, Inc., 810 Innovation Drive, Knoxville, TN 37932 USA
e mail: hamill@cti-pet.com

Abstract. **Iterative reconstructions for clinical PET must run fast. We describe a clinical processing method based on sinogram rebinning, Fourier rebinning for the 3D to 2D data reduction, and iterative reconstruction using the attenuation-weighted OSEM method with a projector based on a gaussian pixel model. When this approach is used, multi-bed clinical oncology scans can be ready for diagnosis within minutes.**

Introduction

As PET becomes used increasingly in the clinic, the need for high throughput in whole-body oncology is being met in several ways. First, PET scanners with septa to block radiation on oblique lines of response (two-dimensional, or 2D, scanners) are being replaced by scanners with removable septa (3D capable scanners), and more recently by scanners with no septa (fully 3D scanners) to make use of the increased sensitivity in 3D mode. Second, detectors with a higher intrinsic count rate capability are being developed to allow more patients to be scanned in one day, for example, scanners with LSO scintillators instead of BGO. Third, algorithms are being developed to improve image quality. The most promising of these algorithms are iterative reconstruction methods.

The iterative approach to image reconstruction is computationally intensive. One could hope that the arrival of faster computers every few months will allow the use of increasingly more sophisticated algorithms, yet it has been observed [1] that, over the last two decades, the number of lines of response in PET scanners has grown at a rate that outpaces Moore's Law which describes the density with which foundries are able to pack transistors onto a microchip. Processing algorithms must also be coded to run very fast.

In this paper, we describe an approach to generating clinical images with the new generation of 3D PET scanners. The approach is the same one described in [2] but applied to 3D. Sinogram rebinning is an important feature of our approach, in which we reduce the size of the input data sets early in the reconstruction procedure.

Methods

Images of one position of the patient bed are derived from the 3D emission sinogram, a set of normalization coefficients, and attenuation correction factors (ACF). The sinograms represent an estimate of true coincidences, since random events are subtracted from the sinograms during acquisition. The ACF's may be taken directly from a transmission measurement, or they may result from a segmentation procedure.

To prepare the sinograms for reconstruction, we first multiply them by normalization coefficients. [3] Next, each plane in each oblique segment of the high-resolution sinogram is rebinned radially and angularly. Clinical users often use 128 sinogram bins and 128 angular ones. After rebinning, the sinogram contains all quanta in the normalized sinogram, but the arc distortion has been removed and the sinogram bins represent equally spaced, parallel lines of response. The rebinning algorithm is indicated in Figure 1, where: "source" coordinates are those of the normalized sinogram with arc distortion; "destination" coordinates are those of the rebinned sinogram; and left and right edges are determined by the coordinate transformation which changes the sinogram bin size and describes the arc distortion. The algorithm integrates the interpolation function (which is a step function in the case shown) over the needed range of source coordinate positions, thereby preserving the counts.

The transformation from 3D to 2D sinogram is performed with the Fourier rebinning method. [4] Before Fourier rebinning is performed, we make the sinograms more consistent by applying the segment-0 attenuation correction to all sinogram segments. After Fourier rebinning, we form an estimate of a 2D sinogram of scattered radiation, using Watson's method. [5] To create the inputs required by this method, the 2D emission sinogram is reconstructed and smoothed to create a low-resolution image which is known to be corrupted by scatter; and the logarithm of the ACF is reconstructed to obtain an image of attenuation coefficients. Both of these reconstruction steps use a filtered backprojection technique which is implemented in frequency space. Next, the attenua-

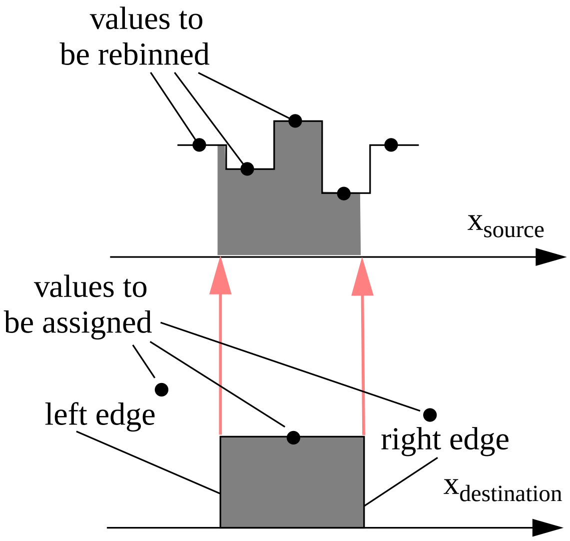


FIGURE 1. Rebinning method

tion coefficients are removed from the emission sinogram, and the scatter estimate is subtracted from it.

The reconstruction itself, that is, the inversion of the Radon transform, is performed with the attenuation-weighted OSEM method (AW-OSEM). [6] For the forward projector and the matched backprojector, we use a gaussian voxel model. [7] Negative sinogram values are set to zero before the iterative reconstruction begins. Our initial image estimate has the value 1.0 in all voxels. Clinical users often select the option to use eight subsets and to stop after two iterations. After reconstruction, the images are filtered transaxially and axially.

The convergence properties of this algorithm were presented in [2]. In this paper we present new clinical results from a fully 3D LSO scanner. Although it is not the thrust of this work, we compare AW-OSEM images to those from another, older OSEM reconstruction method, [8] in which attenuation factors are applied to sinograms before reconstruction. We call the old method UW- OSEM, where UW denotes unweighted.

Clinical Performance and Patient Images

This approach is designed to run fast. In a typical clinical acquistion, images are available about 3 min after all patient data is acquired in fully interleaved mode when the reconstruction is paralleled to the acquisition. In the case discussed in this article, the images were available when the patient left the scanning room. An analysis of the time involved indicates that about 45 seconds are used to read data over the network from the acquisition computer; 50 seconds to set up the 3D normalization matrix; about 20 seconds for the scatter correction; and approximately 20 seconds for every iteration on a typical computer configuration (Sun Ultra-60 running at 450 MHz).

The comparison of the UW-OSEM and AW-OSEM algorithms is shown in Figure 2. Clinicians normally prefer the results from the new, faster, attenuation weighted algorithm.

Image quality is demonstrated in Figure 3. The sinograms for this scan were acquired in 3D on the ECAT Accel, a new generation of LSO based PET scanners, in 45 min (9 bed positions, 3 min in emission scan time and 2 min transmission scan time each).

Conclusions and Future Work

We have realized an acceptable and, we think, high level of image quality with iterative reconstructions that run in clinically realistic times. The keys to doing the calculation fast enough are: sinogram rebinning and Fourier rebinning as techniques for reducing the number of sinogram bins; and use of the attenuation-weighted OSEM algorithm. The result is a procedure that allows routine 3D reconstruction of clinical images of as many patient scans as can be acquired during the day.

This model does not treat the background of random events accurately. Among the techniques to be explored in the future is NEC scaling. [9]

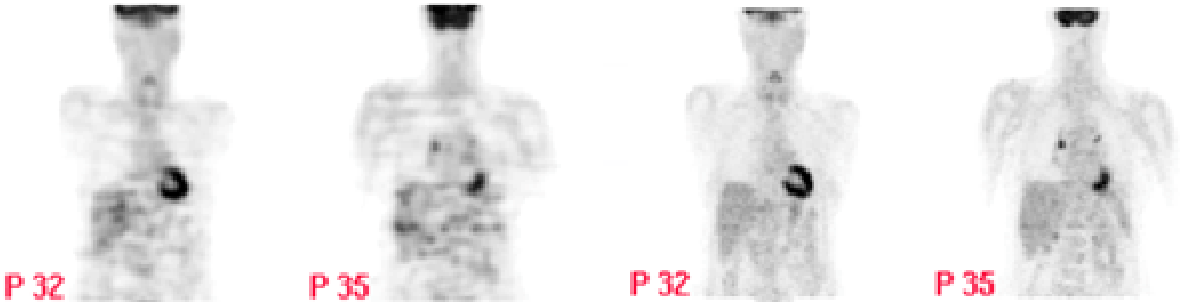


FIGURE 2. The left side shows two coronal slices through a typical data set reconstructed with the previous version of the iterative reconstruction using pre-corrected sinograms and 1 iteration, 30 subsets, and a 10 mm gaussian filter. On the right side, the same data set is reconstructed with the new algorithm and 2 iterations and 8 subsets. The parameters were chosen to generate images preferred by the clinicians for each algorithm

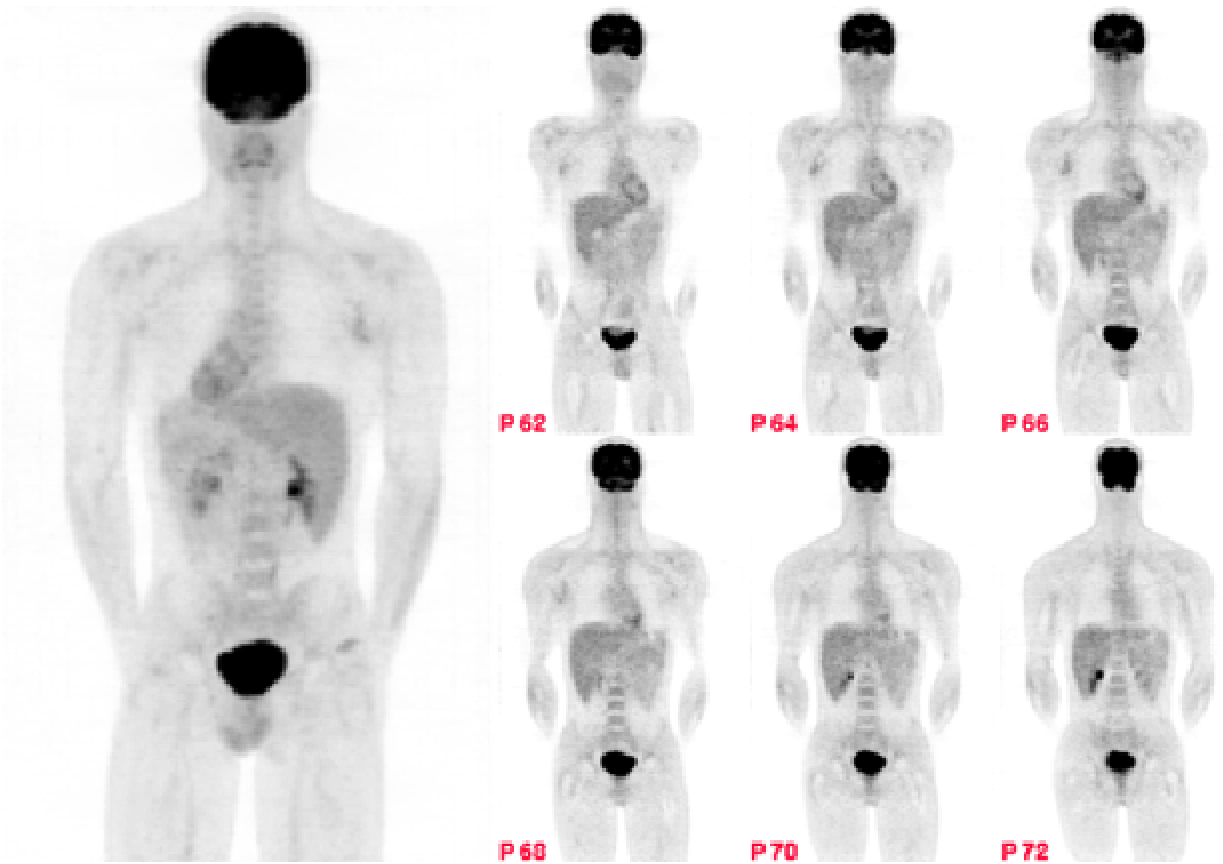


FIGURE 3. Image of a 25 year old, 71 kg, male melanoma patient 45 min post injection of 16 mCi FDG, 9 bed positions acquired for 3 min emission and 2 min transmission each.

The gaussian voxel model should be well suited to fully iterative 3D reconstruction. Schmand showed that the line spread function of high-resolution LSO and GSO brain scanners is well modeled as a gaussian. [10] See Figure 4. We are looking for ways to extend our algorithm to that reconstruction problem.

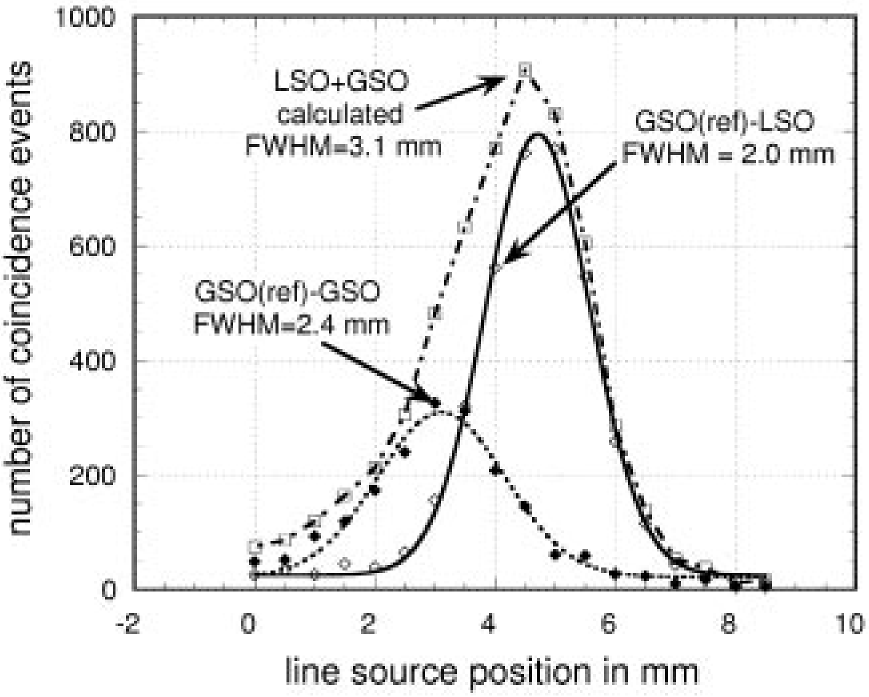
Acknowledgment

Patient images are presented by arrangement with Dr. Peter Valk, Northern California PET Imaging Center, whose assistance is gratefully acknowledged.

References

1. Ronald Nutt, invited talk, IEEE Medical Imaging Conference, 1993
2. T.Bruckbauer et al., “Implementation of weighted OSEM for Whole-Body PET,” *J Nucl. Med.* vol 40 no 5S, May, 1999 (abstract)
3. M. E. Casey, H. Gadagkar, D. Newport, “A Component Based Method for Normalization in Volume PET,” Conf. Record of the 1995 Conference on Fully Three Dimensional Reconstruction.
4. M. Defrise, P.E. Kinahan, and D.W. Townsend, “A New Rebinning Algorithm for 3D PET: Principle, Implementation, and Performance,” 1995 International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine.
5. C. C. Watson, “New, faster image-based scatter correction for 3D PET,” *IEEE Trans. Nucl. Sci.*, vol 47 no 4, Aug 2000
6. C. Michel *et al.*, “Preserving Poisson characteristics of PET data with OSEM reconstruction,” Conf. Record of the 1998 Medical Imaging Conference.
7. J. W. Wallis, T. R. Miller, “An optimal rotator for iterative reconstruction,” *IEEE Trans. Med. Im.*, vol 16 no 1, Feb. 1997, p. 118-123.
8. S. R. Meikle *et al.* “Accelerated EM reconstruction for total body PET: Potential for improving tumor detectability,” *J Phys. Med. Biol.*, vol 39, p. 1689-1704

FIGURE 4. Line spread function for 45-degree incidence in the 2-mm crystals of a high resolution brain tomograph with 2-mm LSO and GSO crystals. From [10]



9. J. Nuyts, C. Michel, P. Dupont, "Maximum-likelihood expectation-maximization reconstruction of sinograms with arbitrary noise distribution using NEC-transformations," *IEEE Trans Med. Im* (in press)

10. M. Schmand, Ph.D. thesis, "Higher resolution PET by means of a new scintillator LSO," University of Aachen, Germany (1999)

A modified bootstrap for importance sampled data[1]

R.L. Harrison, J.-S. Kim, T.K. Lewellen
Dept. of Radiology, University of Washington, Seattle, WA 98195-6004

Abstract

We are testing a modified bootstrap technique for importance sampled data from SimSET (a simulation system for emission tomography). The bootstrap allows us to simultaneously produce multiple raw data sets from a single simulation while at the same time reducing the weight variation caused by importance sampling. This combination may greatly reduce the CPU time required to produce the multiple image realizations needed for ROC studies. Initial testing indicates that the mean and variance of medium and high count bins in the raw data and high-activity regions of interest (ROIs) are reproduced relatively accurately.

I. INTRODUCTION

We are developing a modified bootstrap technique to apply to emission tomography simulations using SimSET [1].

Despite the use of importance sampling (IS) [2] and the ever-increasing speed of computers, many problems remain unwieldy or intractable to study using simulation. One such problem is the generation of multiple realizations of the same image for ROC studies. These studies often require hundreds of image realizations from each of several different activity distributions. Such data sets can take months or even years of CPU time to generate.

The problem of generating multiple samples from a probability distribution when only one data set is available, or when multiple data sets are difficult to produce, is often attacked using the bootstrap technique [3, 4]. In this technique, the available data set is used as an estimate of the underlying distribution. New data sets are generated by sampling data points from the original data with replacement.

We have attempted to modify the bootstrap to ameliorate a problem inherent to the use of importance sampled data: because the events have different weights, the variance has different properties than that of the equivalent analog (non-importance sampled) distribution. The idea underlying our modified bootstrap is that by sampling events with frequency proportional to their weight, we can reduce or eliminate the variation in weight. One can imagine several ways of realizing this goal: the one we chose is described below in Section II.d.

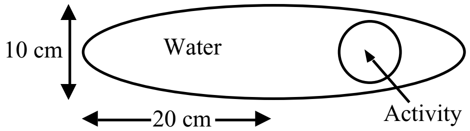


Figure 1: The simulated phantom.

We have tested this technique by generating 99 positron emission tomography (PET) data sets and comparing them to 99 data sets from an analog simulation, both in data space and image space. The results raise as many questions as they answer, but give substantial cause for optimism.

II. METHODS

A. Simulation Setup

We performed 99 analog simulations of 4 million decays each. We simulated an elliptical cylinder of water, 40 cm major axis, 10 cm minor axis, and 3 cm axially. A 6 cm diameter circular cylinder of activity was centered on the major axis, 10 cm from the center of the phantom (Figure 1). No collimators or detectors were simulated; instead, all photons reaching a target cylinder with energy greater than 400 keV were considered detected. The target cylinder was centered at the same point as the elliptical cylinder, with 90 cm diameter and 3 cm axial extent.

Coincident events were binned into a 64 by 64 distance-angle array, with the distance bins spanning -20 cm to 20 cm.

B. Bootstrap Implementation/Importance Sampled Simulation

We created 99 data sets using a modified bootstrap technique and the data from one simulation. We used the same simulation setup as described above, except with SimSET's stratification, forced detection, and forced non-absorption features turned on.

Often when bootstrap sampling is applied, it is used to create multiple realizations with N detected event from a single simulation or scan with N detected events [5]. We thought that the resulting overlap in the detected events from data set to data set might cause significant correlation between artifacts in the images reconstructed from the resulting data sets. For this reason, we chose to simulate 664,230 decays—using a short training run, we determined this would produce an equivalent number of detected events to ten 4-million decay analog runs.

Our bootstrap technique sampled from this simulation on-the-fly. Each time a detected event was produced, we sampled a random number, R, for each of the 99 data sets from the Poisson distribution with parameter

$$p = \min\left(0.1\frac{w_{event}}{w_{mean}}, 0.5\right) \qquad (1)$$

where w_{event} is the weight of the current event, w_{mean} is the average event weight, and the factor 0.1 is to compensate for the fact that we were simulating ten times as many decays as needed for one data set. The maximum value of 0.5 is set to keep the Poisson parameter small, so that a event does not appear in too many of the data sets. When the sampled R is 0,

---

[1] Supported by PHS grant CA42593

the event is not binned in that data set. Otherwise the event is binned with its weight adjusted to

$$w_{output} = \frac{R * w_{event}}{p} \tag{2}$$

In a more typical bootstrap, w_{output} would not include the factor R; instead, R events would be generated. We chose to create a single event to simplify the modifications needed to the SimSET software. In the long run, however, we believe it may be better to generate R separate events.

C. Image Reconstruction

We reconstructed images from the analog and bootstrap simulation data sets using filtered backprojection with a Hamming window, cutoff frequency 0.6.

D. Statistical Analysis in Data Space

The 99 data sets from the analog simulations were compared to the 99 data sets from the bootstrap simulation using the Student t-test and the f-test [6]. We applied these tests bin-by-bin, and grouped the results according to the mean number of counts in the analog simulation data sets.

E. Region-of-Interest Analysis in Image Space

Two square ROIs, 5 by 5 pixels (3.1 cm by 3.1 cm) were placed on each image. The first ROI was centered in the cylinder of activity, the other in a zero-activity area on the other side of the phantom. We computed the sample mean and variance for the analog images and the bootstrapped images. We then computed Student t-test and f-test statistics.

III. RESULTS

A. Simulation Run Statistics

Some simulation efficiency statistics are given in Table 1. The bootstrap/IS method significantly reduced the cost per realization—though we must note that the situation we simulated is one where the IS can help a great deal, something that is not always true.

Table 1
Simulation Efficiency Statistics

	average analog simulation	average bootstrap realization
CPU time (seconds)	1320	7.8
Total detected counts	19,905	22,385
Total detected weight	938,647	916,355
Total detected weight-squared	44.26 million	39.62 million
Quality factor (QF)	1.0	0.95
$\frac{counts * QF}{CPU\_time}$	15.1	2752

B. Data Space Analysis

The results of the Student t-test and the f-test applied bin-by-bin are shown in Tables 2 and 3. The last four columns show the percentage of analog-bootstrap bin pairs with $p > 0.5$, 0.05, 0.01 and ≤ 0.01—the last three being of interest as common levels for the rejection of the null hypothesis. No results are shown for 1223 bins for which one data sets (analog and/or bootstrap) had no counts in any realization. (The analog data sets had 659 such bins, bootstrap 1191.) The rest of the results are sorted by the mean number of counts in the analog data bin. The line for average counts between 0 and 1 should, perhaps, have been omitted, as many of these bins had a total, over all the realization, of 10 or less events. We usually wouldn't apply these tests to Poisson data with less than 20-30 counts.

However, even in the high count bins, we are seeing significantly different means and variances for a high number of bins. The distributions are close, as shown in the fact that there are also many p-values greater than 0.5 and 0.05. However, overall the distribution of data is clearly different for the analog and bootstrap realizations.

Table 2
Bin-by-Bin T-test Comparison

mean counts	# of bins	% p > 0.5	% p > 0.05	% p > 0.01	% p ≤ 0.01
0	1223	-	-	-	-
0 to 1	2209	16.3	48.2	60.8	39.3
1 to 5	97	35.1	69.1	81.4	18.6
5 to 10	67	37.3	71.6	80.6	19.4
10 to 30	185	22.2	58.9	77.3	22.7
30 to 50	183	20.8	54.6	70.0	30.1
50 to 70	132	15.2	49.2	59.9	40.2

Table 3
Bin-by-Bin F-test Comparison

mean counts	# of bins	% p > 0.5	% p > 0.05	% p > 0.01	% p ≤ 0.01
0	1223	-	-	-	-
0 to 1	2209	8.1	25.3	34.2	65.8
1 to 5	97	23.7	70.1	82.5	17.5
5 to 10	67	26.9	85.1	95.5	4.5
10 to 30	185	37.8	79.5	90.3	9.7
30 to 50	183	26.2	78.7	90.2	9.9
50 to 70	132	30.3	81.1	95.5	4.6

C. Image Space Analysis

Table 4 shows the sample mean and variance for the two image ROIs. For both ROIs, the means are very close. The sample variance make clear, however, that this was a matter of luck. The variances for the ROI with activity have an f-test p-value of 0.29, so they are also reasonably close. However, the background variances are significantly different (p = 0.0004).

Table 4
ROI Mean and Variance

	Analog mean	Analog variance	Bootstrap mean	Bootstrap variance
Active ROI	707.2	150.2	707.8	121.3
Background ROI	2.18	73.39	0.19	35.6

IV. DISCUSSION

This is preliminary work, and as such raises far more questions than it answers. The modified bootstrap technique helps to speed up the generation of multiple realizations tremendously, but we have not shown that these realizations are close enough to truly independent realizations to allow the technique to be used.

However, there are a number of ways we can address the problems. A better way to assign the number of decays for the bootstrap simulation would help—our use of a short training run for that purpose will have propagated the noise from the short run through our entire data set. There are other ways to implement the bootstrap algorithm that might help. And, given the tremendous speed-up we achieved, we could experiment with generating more events.

We do not really expect the bootstrap to perfectly reproduce the bin-by-bin statistics shown above. As seen above the performance on image statistics can be very good even in situations where the underlying data sets do not perfectly represent the true distributions. We plan more tests for this technique in image space. The ultimate test will be the performance of the bootstrap on ROC data sets.

Finally, we note that SimSET is available free of charge for non-commercial use. However, the bootstrap algorithm is not in the distributed package, and will not be until we have carried this investigation considerably further. Contact simset@u.washington.edu for more information.

VI. REFERENCES

[1] T. K. Lewellen, R. L. Harrison, and S. Vannoy, "The SimSET program," in *Monte Carlo Calculations in Nuclear Medicine*, *Medical Science Series*, M. Llungberg, S.-E. Strand, and M. A. King, Eds. Bristol: Institute of Physics Publishing, 1998, pp. 77-92.

[2] D. R. Haynor, R. L. Harrison, and T. K. Lewellen, "The use of importance sampling techniques to improve the efficiency of photon tracking in emission tomography simulations," *Med Phys*, vol. 18, pp. 990-1001, 1991

[3] Efron, B., *The Jackknife, the Bootstrap, and other Resampling Plans*. 1982, Philadelphia: Society of Industrial and Applied Mathematics.

[4] Politis, D., Romano, J., and Wolf, M., *Subsampling.* 1999, New York: Springer.

[5] Haynor, D. and Woods, S., *Resampling estimates of precision in emission tomography*. IEEE Trans Med Imag, 1989. **8**(4): p. 337-343.

[6] Press, *Numerical Recipes in C, 2nd ed*. Numerical Recipes in C, 2nd ed. 1992, Cambridge: Cambridge University Press.

.

A sufficient condition for spiral cone beam long object imaging via backprojection

K. C. Tam

Siemens Corporate Research, Inc., Princeton, NJ, USA

Abstract

The response of a point object in cone beam spiral scan is analysed. Based on the result a sufficient condition for the spiral scan long object problem employing backprojection is formulated. By making use of the sufficient condition a general class of exact, backprojection based reconstruction algorithms for spiral scan cone beam CT is developed which are capable of reconstructing a sectional ROI of the long object without contamination from overlaying materials using spiral scan cone beam data irradiating the particular ROI and its immediate vicinity only. Also, at each source position the minimum size of the region on the detector plane required for 3D backprojection is reduced, which in term brings about reduction in the amount of 3D backprojection computation.

I. 2D filtering and masking

Spiral scan computed tomography with large area detectors is of increasing interest for rapidly scanning spacious volumes. As the cone angle increases the artifacts generated in the reconstructed images by the approximate reconstruction algorithms will become more and more serious, and exact reconstruction algorithms are required. It is known that if the spiral path is long enough so that every plane intersecting the object also intersects the spiral path, the object can be reconstructed. For long objects, however, it is highly desirable to scan only the portion of the object that is of interest, for the sake of reduction in scan time as well as radiation protection of the patient in medical imaging. However, as a consequence of the divergent nature of the X-ray cone-beams different regions of the object are correlated. To reconstruct only a region-of-interest (ROI) from a spiral scan which covers the particular ROI and its immediate vicinity only poses a challenge for the imaging community. This is referred to as the long object problem in the literature.

The first solution to the long object problem in spiral cone beam CT is the Radon space driven (spiral + 2 circles) algorithm reported in [1,2]. A key part of the reconstruction algorithm is the data-combination technique in which the radial Radon derivative for each plane intersecting the ROI is obtained by combining the partial results computed from the cone beam data at the various source positions that the plane intersects. The method is illustrated in Figure 1 which represents a plane Q intersecting the ROI and the scan path. Since the partial planes do not overlap and together they completely cover the portion of plane Q that lies within the ROI, the Radon derivative for plane Q can be obtained exactly by summing the Radon derivatives for the partial planes. From Figure 1 it is evident that the portions of the object outside the ROI do not need to be irradiated. Therefore during scanning collimators can be used to block off radiation from reaching those portions.

Restricting the cone beam projection data to the appropriate angular range for data combination can be accomplished by a masking process. The mask consists of a top curve and a bottom curve formed by projecting on the detector the spiral turn above and the turn below from the current source position. It can be easily seen that such masking procedure corresponds exactly to the angular range bound by the prior and the subsequent source positions as indicated in Figure 1. We shall refer to this mask as the data-combination mask. For a flat detector located at the rotation axis such that the line connecting the source to the detector origin is normal to the

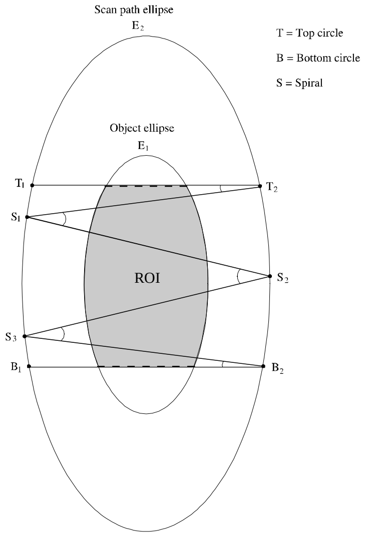


Figure 1. A typical integration plane covering the ROI defined by the source positions. Other integration planes may have more or less spiral scan path intersections, and may not intersect either the top or the bottom circle scan paths.

detector plane, the equation for the top curve for the spiral scan is given by:

$$
\begin{aligned}
v &= \frac{h}{2\pi}\tan^{-1}\left(\frac{R}{u}\right)\left(1+\frac{u^2}{R^2}\right) && u \geq 0 \\
v &= \frac{h}{2\pi}\left[\pi+\tan^{-1}\left(\frac{R}{u}\right)\right]\left(1+\frac{u^2}{R^2}\right) && u < 0
\end{aligned}
\tag{1}
$$

where u and v are the Cartesian coordinate axes of the detector with the v axis coinciding with the rotation axis, R is the radius of the spiral, and h is the distance between adjacent spiral turns (the pitch). The bottom curve is the reflection of the top curve about the origin, i.e. (u,v) -> $(-u,-v)$. The shape of the spiral mask is shown in Figure 2. The figure assumes right-handed spiral rotation.

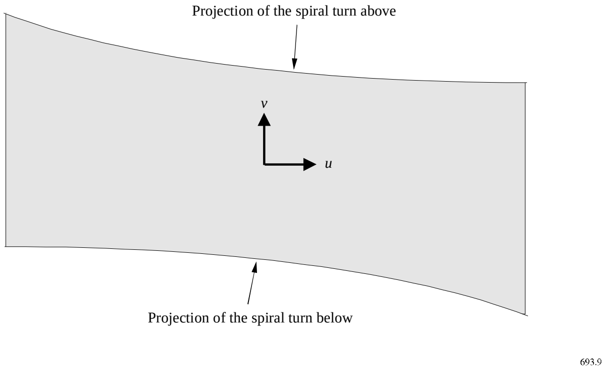


Figure 2. The mask on the flat detector which defines the desired partial plane for Radon derivative computation. For any plane of integration, the portion of its intersection line with the detector within the mask is the desired partial plane.

In the backprojection version of the (spiral + 2 circles) algorithm [3], the masked cone beam data are 2D filtered and then 3D backprojected. The 2D filtering is carried out in 2 different manners for different parts of the cone beam data: the data inside the mask are line-by-line ramp filtered in the direction of the projected scan path direction, whereas those on the mask boundary are processed with a 2D filter which includes 2D backprojection at all angles on the detector plane. By virtue of the Radon inversion formula the 2D backprojection operation should be extended to the entire detector plane extended to infinity; in practice it is extended to the extent sufficient to cover the entire ROI. Through the line-by-line ramp filtering in the direction of the projected scan path direction the data inside the mask boundary only affect a localized portion of the reconstruction volume. On the other hand the data on the mask boundary affect the entire ROI because of the long range of the 2D backprojection. This long range correlation caused by the mask boundary data is the crux of the long object problem employing backprojection driven algorithms.

II. Spiral scan long object problem

Recently a number of approaches solving the long object problem with only the spiral scan appeared in the literature. In the virtual circle (VC) method reported by Kudo et al [4] it is found that by utilizing the unique property of the PI lines [5], removing the circles in the (spiral + 2 circles) algorithm contaminates only a localized portion at each end of the ROI, and thus the remaining portion of the ROI can still be reconstructed without contamination from overlaying materials. In the zero boundary (ZB) method reported by Defrise et al [6], the unique property of the PI lines is also utilized to remove the long range correlation between different regions of the object caused by the data on the mask boundary. In the local ROI (LR) method developed by Sauer et al [7] and later implemented by Schaller et al [8], de-correlation between different regions of the object is achieved on the ϕ-planes, which are the planes which contain the *z* axis in the Radon space. Subsequently the backprojection version of the local ROI method was developed by Tam [9] and implemented by Lauritsch et al [10]. Unlike the (spiral + 2 circles) algorithm, with only spiral scan it is necessary to scan some portions of the object adjacent to the ROI in order to reconstruct the ROI without contamination; the spiral path required beyond the ROI is referred to as overscan in the literature. For comparison the overscan for the (spiral + 2 circles) algorithm is zero.

All approaches are theoretically exact solutions to the long object problem. However, even among the backprojection driven algorithms very different methodologies are employed in reconstructing the ROI without contamination, and the overscan required by each algorithm is substantially different from the others. It is not apparent that the three methods have any features in common.

III. A sufficient condition

In this paper a sufficient condition for backprojection driven image reconstruction algorithms for the long object problem with spiral scan is derived. The analysis is based on the analysis of the response of a point object enclosed inside the spiral path. It is found that the support of the contribution to the reconstruction volume from the cone beam data on the mask boundary becomes localized when certain condition is satisfied. Each mask boundary data point corresponds to a PI line, as illustrated in Figure 3, which intersects two source positions. At each of the 2 source positions the mask boundary data point, after some processing, is 2D backprojected along each line intersecting the data point, and then 3D backprojected onto the 3D backprojection planes defined by the source position and each 2D backprojection line through the data point. Thus each 3D backprojection plane intersects the line connecting the source position and the data point, which is the PI line corresponding to the data point. Since the 2 source positions that acquire the mask boundary data point have the same PI line, it follows that the 2 source positions have the same set of 3D backprojection planes when processing the data point.

Consider a fixed mask boundary data point. If for each 3D backprojection plane the data point is processed, which includes filtering and 3D backprojection, to the same extent at the 2 source positions that acquire the data point, then the support of the contribution to the reconstruction volume from the data point can be shown to be localized. The minimum size of the region on the detector plane required for 2D backprojection and the subsequent 3D backprojection for these mask boundary data can be prescribed using projective geometry, and is found to be smaller than the minimum size according to current understanding, viz. the size required to cover the entire ROI. The extent to which the detector size is reduced depends on the spiral pitch, and is substantial for small pitch. The reduction in the detector size is important for the reduction in the amount of 3D backprojection computation.

Among the three above-mentioned long object backprojection driven reconstruction algorithms, the VC method and the backprojection LR method are found to satisfy the sufficient condition, but not the ZB method. Based on the sufficient condition a general class of exact, backprojection driven reconstruction algorithms for long object imaging in spiral scan cone beam CT is developed. It is found that the VC method is a special case of this class of algorithms.

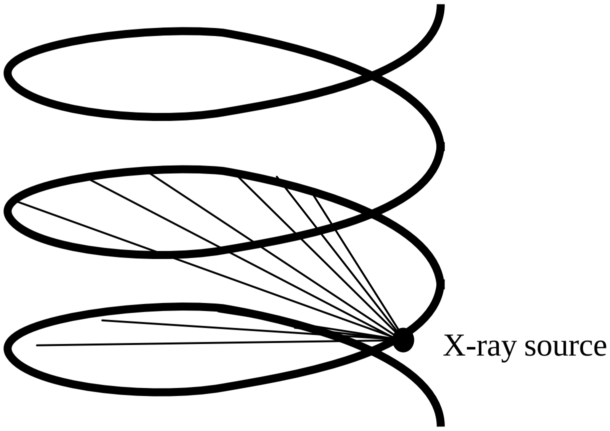


Figure 3 Mask boundary data and the corresponding 2 source positions on the PI lines.

IV . References

1. K.C. Tam, “Helical and circle scan region of interest computerized tomography”, *US. Patent* 5,463,666, Oct 31, 1995.
2. K.C. Tam, S. Samarasekera, and F. Sauer, “Exact Cone Beam CT with A Spiral Scan”, *Phys. Med. Biol.*, 43, pp. 1015-1024, 1998.
3. K.C. Tam, B. Ladendorf, F. Sauer, G. Lauritsch, and A. Steinmetz, “Backprojection spiral scan region-of-interest cone beam CT”, *Proc. SPIE Medical Imaging 1999: Physics of Medical Imaging*, pp. 433-441, 1999.
4. H. Kudo, F. Noo, and M. Defrise, “Quasi-exact reconstruction for long-object problem in helical cone-beam tomography,” *Proceedings of the 1999 International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, pp. 127-130, 1999.
5. P. E. Danielsson, P. Edholm, J. Eriksson, M. Seger, “Towards exact 3D-reconstruction for helical cone-beam scanning of long objects. A new detector arrangement and a new completeness condition.” *Proceedings of the 1997 International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, pp. 141-144, 1997.
6. M. Defrise, F. Noo, and H. Kudo, “A solution to the long-object problem in helical cone-beam tomography,” *Phys. Med. Biol.*, 45, pp. 623-643, 2000.
7. F. Sauer, S. Samarasekera, and K.C. Tam, “Practical cone-beam image reconstruction using local regions-of-interest”, U.S. patent 6,009,142, December 28, 1999
8. S. Schaller, F. Noo, F. Sauer. K. C. Tam, G. Lauritsch, and T. Flohr. "Exact Radon rebinning algorithms using local regions-of-interest for helical cone-beam CT," *Proceedings of the 1999 International Meeting on Fully Three-dimensional Image Reconstruction in Radiology and Nuclear Medicine,* pp.11-14, 1999
9. K.C. Tam, “Exact local regions-of-interest reconstruction in spiral cone-beam filtered-backprojection CT: theory”, *Proc. of SPIE Medical Imaging Conf.,* vol. 3979, pp. 506-519, 2000
10. G. Lauritsch, K. C. Tam, K. Sourbelle, and S. Schaller, “Exact Local Regions-of-Interest in Spiral Cone-Beam Filtered-Backprojection CT: numerical implementation and first image results,” *Proc. of SPIE Medical Imaging Conf.,* 3979, pp. 520-532, 2000.

A Simple Algorithm for Increased Helical Pitch in Cone-Beam CT

Michael D. Silver, Katsuyuki Taguchi, Ilmar Hein*

I. INTRODUCTION

A goal of cone-beam CT is to decrease the scan time needed to image a volume of the patient. Four things are required to meet this goal: (1) a detector with multiple rows of sensors, (2) higher bandwidth data acquisition system and image processor to handle the increased data rates and need for fast reconstruction, (3) high helical pitch, (4) a practical reconstruction algorithm that lends itself to fast reconstruction times. Within the last several years, CT-systems with first two, then four row detector arrays have been introduced to the marketplace [1]-[5]. The four row arrays are actually combinations of finer pitched detectors. Table I gives a comparison of the detector designs for the Toshiba Aquilion Multi, the LightSpeed QX/i from General Electric Medical Systems, the Siemens Volume Zoom and Marconi Mx8000. All the major manufacturers have announced plans for an output of 8 and/or 16 slices, which requires a new or modified detector design from Siemens and Marconi. Toshiba has presented images taken on a 256-row prototype detector array [6], [7]. Thus, detector development for wider cone-angles is underway and will lead to faster scan times. The manufacturers are working on the second requirement as computer electronics improve. This presentation looks at the last two requirements. We present a scheme for high helical pitch while maintaining a practical although approximate reconstruction algorithm.

So far, the four row CT-systems have used two-dimensional reconstruction algorithms [8]-[10]. As more rows are added, we have found that three-dimensional backprojection leads to images with better quality than the two-dimensional approximations [11], [12]. Therefore, we believe that as more rows are added to the detector, the reconstruction problem approaches the fully three-dimensional problem. However, because the cone angle is not too large, we propose an approximate algorithm that is a modification of helical Feldkamp [13], [14], where the maximum helical pitch limit is determined by two-dimensional arguments as described in the next section. Moreover, the Feldkamp approach is based on a heuristic use of one-dimensional convolution (rigorous for two-dimensional reconstruction) combined with true three-dimensional backprojection.

We first reported on these ideas at a previous Fully 3D Meeting [15], [16] and presented some results from image evaluation at the last RSNA meeting [17]. Recently [18], we presented how we relate helical pitch, number of views to reconstruct, and field-of-view based on a weighting scheme published in *Medical Physics* [19]. This will be briefly summarized here. In this study, we show how we can extend the maximum helical pitch to higher values.

II. REVIEW: HELICAL, CONE-BEAM SCANNING

A. Validity requirements

Backprojection follows the straight-line ray-sum from the focal spot of the x-ray source through a pixel of interest in the image volume and onto the two-dimensional detector array. Typically, the processed signal at this location in the detector array is weighted and added to the contents of the voxel. This is repeated for all voxels and for a range of x-ray source angles. For helical, cone-beam scanning, the relation between a voxel and a location in the detector array for a given source angle is given by

TABLE I
Comparison of Four-row CT-Scanners

	Toshiba Aquilion Multi	GE LightSpeed	Siemens Volume Zoom Marconi Mx8000
Detector Rows	34 (4 x 0.5, 30 x 1)	16 (16 x 1.25)	8 (2 x 1, 2 x 1.5, 2 x 2.5, 2 x 5)
DAS Output and nominal slice thicknesses	4 x 0.5, 4 x (1 to 8)	4 x (1.25,2.5,3.75,5)	2 x 0.5, 4 x (1,2.5,5)

Widths of detector rows are given in mm as projected at isocenter. Nominal slice thicknesses are in mm.

$$\gamma(\beta,x,y) = \sin^{-1}\frac{x\cos\beta + y\sin\beta}{L(\beta,x,y)} \tag{1}$$

$$\alpha(\beta,x,y,z) = \tan^{-1}\frac{[\beta-\beta_0(z)]H}{2\pi L(\beta,x,y)} \tag{2}$$

where γ is the fan angle (in the *x-y* plane) of the ray-sum,

α is the cone angle of the ray-sum,

β is the x-ray source angle,

$\beta_0(z)$ is the source angle when the focal spot is in the image slice at *z*,

x,y,z is the coordinates of an image pixel,

H is the helical pitch: table travel per rotation of the source,

and $L(\beta,x,y) = \sqrt{(R\sin\beta + x)^2 + (R\cos\beta - y)^2}$. (3)

L is the distance from the focal spot to the pixel *x,y,z* times the cosine of the cone angle, *R* is the radial distance of the focal spot to isocenter. The coordinate system moves with the patient/table so that each image slice is at a fixed *z*.

Multi-row CT-scanners have detector arrays that are sections of a cylinder, focused on the source; thus equal angular increments $\Delta\gamma$ and equal axial linear increments separate the individual sensor elements. Therefore, γ is a natural coordinate for the ray-sum but (2) is changed in favor of detector rows (also known as slices or segments),

$$n(\beta,x,y,z) = \frac{[\beta-\beta_0(z)]R}{2\pi L(\beta,x,y)} r_H, \tag{4}$$

where n is the relative detector row,

$$-\tfrac{1}{2} \le n \le \tfrac{1}{2}, \tag{5}$$

and $$r_H = \frac{H}{W} \tag{6}$$

with *W* as the full axial height of the detector array as projected at isocenter. Thus, r_H is the normalized helical pitch ratio.

A ray-sum is valid—that is, measured—if γ and n are locations within the physical detector array. For helical, cone-beam CT, the key validity equation is (5) combined with (4). We can solve (4) with $n = \pm\tfrac{1}{2}$ for the surfaces of $\beta_{1,2}$ as a function of voxel position *x,y,z*; $\beta_{1,2}$ represents the source angular position when the voxel *x,y,z* enters and then leaves the cone-beam. These surfaces are warped, depending on helical pitch. The difference of the two surfaces shows that up to a normalized helical pitch ratio of 2, all voxels have at least 180° of coverage. Unfortunately, the voxels don't have the same 180° range of coverage. This implies that a reconstruction algorithm could exist that uses only valid ray-sums up to a normalized helical pitch ratio of 2. However, such as algorithm could have additional computational complexities.

B. Weighting Scheme

The helical pitch determines how long a given voxel is irradiated. Obviously, the higher the pitch, the less the voxel is within the rotating cone-beam and vice versa. We adapt the weighting scheme of [19] to helical cone-beam scanning. Consider the sinogram from a single detector row. Suppose all the voxels of a given slice are in the cone-beam for at least $\pi+\Delta\beta$ as shown in Fig. 1a. The weighting scheme, based on Parker's half-scan method [20], introduces a virtual fan-angle, 2Γ, such that $2\Gamma = \Delta\beta$, as shown in Fig. 1b. Use the same weights as in half-scan [20], [21] for the redundant (in two-dimensions) triangles but with 2Γ as the fan-angle.

C. Helical Pitch Limits

Under the condition that we use valid rays-sums, as described in Section II.a, and that all voxels in a slice are reconstructed over the same angular range of the source, then the relation between the angular range of the source used in the reconstruction, the helical pitch, and the field-of-view *FOV* (*R* is the source radius) can be shown to be [17]:

$$\beta_2 - \beta_1 = \pi + 2\Gamma = \frac{2\pi}{r_H}\times\left(1 - \frac{FOV}{2R}\right) \tag{7}$$

The minimum and maximum values for the helical pitch correspond to when the virtual fan-angle approaches π and $2\gamma_m$, the true fan-angle, respectively:

$$r_H(\min, FOV) = 1 - \frac{FOV}{2R} \tag{8}$$

$$r_H(\max, FOV) = \frac{2\pi}{\pi + 2\gamma_m} r_H(\min, FOV). \tag{9}$$

III. INCREASE THE HELICAL PITCH

Keeping the helical pitch below the limit given in (9) assures that all ray-sums that go into making the image are valid ray-sums in a three-dimensional sense, although we use weighting derived from two-dimensional arguments. Consider again Fig. 1b. For the smaller FOV's, the ray-sum values in the region

$$\sin^{-1}\left(\frac{FOV}{2R}\right) < |\gamma| < \gamma_m \tag{10}$$

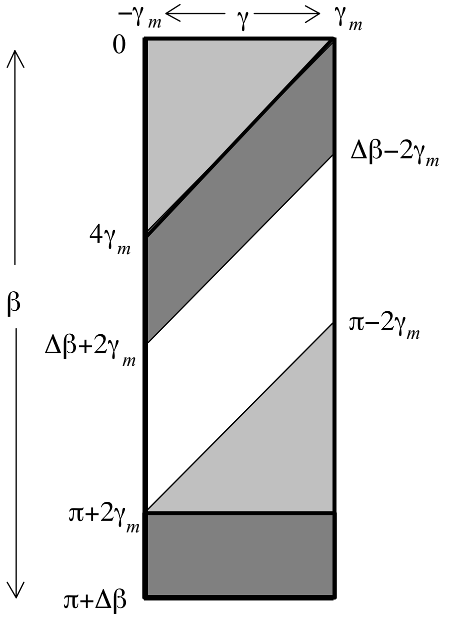


Fig. 1a. Sinogram made from one row of the detector. Similarly shaded regions contain redundant information (ignoring the cone-angle).

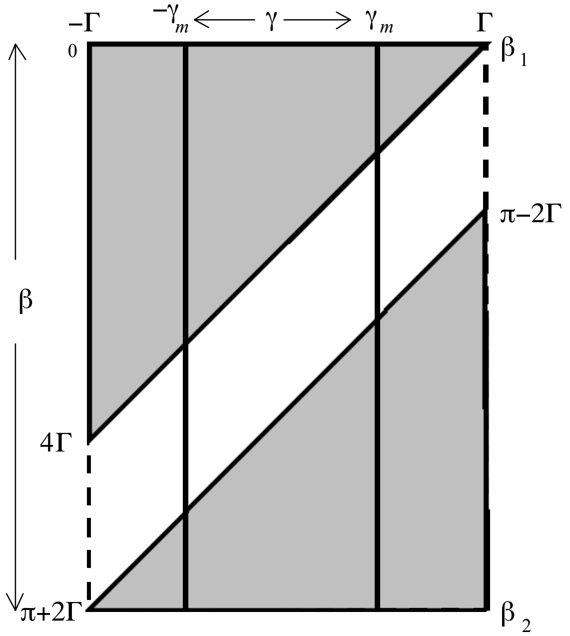


Fig. 1b. Sinogram including two strips of virtual (zeros) data. Fan angle is increased from $2\gamma_m$ to 2Γ. The shaded regions contain redundant information (ignoring the cone angle).

are zero. Note that $\gamma_m = \sin^{-1}\left(\frac{FOV_{\max}}{2R}\right)$, where $FOV_{\max}$ is the maximum field-of-view for the scanner. This implies that the range of validity for (7) is increased. Instead of $r_H\left(\max, FOV\right)$ being given by when the virtual fan-angle reaches $2\gamma_m$, it is given by when the virtual fan-angle reaches $2\sin^{-1}\left(\frac{FOV}{2R}\right)$. Therefore, (9) becomes

$$r'_H(\max, FOV) = \frac{2\pi}{\pi + 2\sin^{-1}\left(\frac{FOV}{2R}\right)} r_H\left(\min, FOV\right). \quad (11)$$

The ratio of the increase is

$$\frac{r'_H\left(\max, FOV\right)}{r_H(\max, FOV)} = \frac{\pi + 2\sin^{-1}\left(\frac{FOV_{\max}}{2R}\right)}{\pi + 2\sin^{-1}\left(\frac{FOV}{2R}\right)}. \quad (12)$$

Fig. 2 compares the new maximum for the helical pitch with the previous calculation as a function of *FOV*. We demonstrate the efficacy of the higher helical pitch with a computer simulation of a 16-row cone-beam CT-scanner using clinical images.

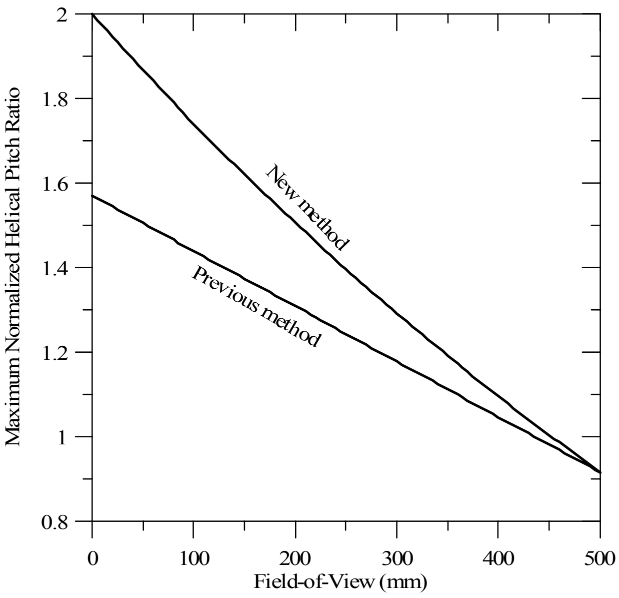


Fig. 2. Maximum helical pitch as a function of *FOV*. Both curves require valid ray-sums and a simple reconstruction algorithm. The previous method uses (9) while the new method uses (11). The curves assume a source radius of 600 mm, a maximum *FOV* of 500 mm.

REFERENCES

[1] Y. Liang and R. A. Kruger, “Dual-slice spiral versus single-slice spiral scanning: comparison of the physical performance of two computed tomography scanners,” *Med. Phys.* **23**, 205 – 220 (1996).

[2] C. N. McCollough and F. E. Zink, “Performance evaluation of a multi-slice CT system,” *Med. Phys.* **26**, 2223 – 2230 (1999).

[3] H. Hu, H. D. He, W. D. Foley, S. Fox, “Four multidetector-row helical CT: image quality and volume coverage speed,” *Radiology*, **215**, 55 – 62 (2000).

[4] W. A. Kalender and T. O. J. Fuchs, “Principles and performance of single-and multi-slice spiral CT,” in L. W. Goldman, J. B. Fowlkes, eds., *Categorical Course in Diagnostic Radiology Physics: CT and US Cross-sectional Imaging,* 127 – 142, RSNA, Oakbrook, IL, 2000.

[5] W. A. Kalender, *Computed Tomography*, Publicis MCD Verlag, Munich, 2000.

[6] Y. Saito, H. Aradate, H. Miyazaki, K. Igarashi, H. Ide, “Development of a large area two-dimensional detector for real-time 3-dimensional CT (4D CT),” *Radiology,* **217(p)**, 405 (2000).

[7] Y. Saito, H. Aradate, H. Miyazaki, K. Igarashi, H. Ide, “Large area two-dimensional detector system for real-time three-dimensional CT (4D CT),” *Proc. SPIE*, **4320** (2001).

[8] K. Taguchi and H. Aradate, “Algorithm for image reconstruction in multi-slice helical CT,” *Med. Phys.* **25**, 550 – 561 (1998).

[9] H. Hu, “Multi-slice helical CT: Scan and reconstruction,” *Med. Phys.* **26**, 5-8 (1999).

[10] S. Schaller, T. Flohr, K. Klingenbeck, J. Krause, T. Fuchs, and W. A. Kalender, “Spiral interpolation algorithm for multi-slice spiral CT – Part I: Theory,” *IEEE Trans. Med. Img.* **19**, 822 – 930 (2000).

[11] K. Taguchi, S. Saito, M. Silver, I. Hein, I. Mori, “Evaluation of image reconstruction methods for 8- and 16-slice helical,” submitted to *RSNA* 2001.

[12] I. Mori, S. Saito, M. Silver, I. Hein, K. Taguchi, “Eight-slice and 16-slice helical CT: Comparison of image reconstruction methods,” submitted to *RSNA* 2001.

[13] L. A. Feldkamp, L. C. Davis, and J. W. Kress, “Practical cone-beam algorithm,” *J. Opt. Soc. Am. A,* **1**, 612 – 619 (1984).

[14] H. Kudo and T. Saito, “Three-dimensional helical-scan computed tomography using cone-beam projections,” *J. Electron. Information Commun. Soc. Japan,* **J74-D-II**, 1108 – 1114 (1991).

[15] M. D. Silver, “Practical limits to high helical pitch, cone-beam computed tomography,” *Proc. 1997 Int. Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine,* 44-47 (1997).

[16] M. D. Silver, “High-helical-pitch, cone-beam computed tomography,” *Phys. Med. Biol.,* **43**, 847 – 855 (1998).

[17] M. D. Silver and K. S. Han, “Application of a new half-scan reconstruction method to helical multi-slice CT (MSCT),” *Radiology,* **217(p)**, 565 (2000).

[18] M. D. Silver, K. Taguchi, and K. Han, “Field-of-view dependent helical pitch in multi-slice CT,” *Proc. SPIE*, **4320** (2001).

[19] M. D. Silver, “A method for including redundant data in computed tomography,” *Med. Phys.,* **27**, 773 – 774 (2000).

[20] D. L. Parker, “Optimal short scan convolution reconstruction for fan beam CT,” *Med. Phys.,* **9**, 254 – 257 (1982).

[21] C. R. Crawford and K. F. King, “Computed tomography scanning with simultaneous patient translation,” *Med. Phys.,* **17**, 967 – 982 (1990).

[*] Michael D. Silver and Ilmar Hein are with Bio-Imaging Research, 425 Barclay Blvd., Lincolnshire, IL 60062, USA (telephone: 847-634-6425, e-mail: msilver@birinc.com and ihein@birinc.com).

Katsuyuki Taguchi is with Medical Systems Company, Toshiba Corporation, Otawara, Tochigi 324-8550, Japan (telephone +81-287-26-6471, e-mail: ktaguchi@mel.nasu.toshiba.co.jp).

Practical approach to misalignment correction in a single-circle orbit cone-beam tomography

Marek Karolczak[1], Ulrike Taubenreuther[1],
Klaus Engelke[1], Andreas Lutz[1],
Stefan Schaller[2], Willi Kalender[1]

[1]Institute of Medical Physics, University of Erlangen-Nürnberg, Germany
[2]Siemens AG, Medical Engineering Group, Forchheim, Germany

Abstract

This contribution addresses the reduction of artifacts in cone-beam tomography (CT and SPECT/PET) that are caused by imperfect scanner mechanical alignment (i.e. misalignment). Such artifacts may show up as double contours, split edges, smearing, loss of resolution, change in magnification etc. Misalignment is a problem commonly recognized in SPECT/PET (see for example [12]) whereas in CT its importance has only recently been emphasized with the advent of cone beam geometry [5][9][11], although residual misalignment can be responsible for stair-step artifacts even in single-row helical scanners [7].
In order to improve the quality of reconstructed images it is therefore necessary to take care of misalignment errors, either by proper scanner alignment or by *measuring* the errors in advance and forwarding them to the *reconstruction* algorithm. As high precision mechanical components are very expensive it may be more promising to take the latter approach but the accurate measurement of misalignment parameters still remains a great problem, in particular in the case of high resolution microCT (µCT).
Here we investigated a practical approach: Is it possible to achieve satisfactory artifact reduction by correcting only a few vital misalignment parameters that can be measured easily. We used an modified Feldkamp-based circular scan reconstruction algorithm that we have developed for our µCT cone-beam scanner [1]. This algorithm incorporates corrections for all possible misalignment errors. Misalignment compensation is embedded in the backprojection.

Cone-beam geometry and scanner misalignment

The key components of a cone beam CT scanner are the X-ray source, a two-dimensional X-ray detector and a sample positioner located between them. In µCT typically the tube and detector remain fixed and the object is rotated while in medical CT the object is at rest. However cone beam CT with rotating source and detector has not been realized yet. In SPECT the detector (gamma camera) is equipped with a converging collimator, which focuses at some point behind the object. The detector rotates about the patient along a circular path. For convenience in this contribution we will refer to a CT scanner, but our analysis remains valid for SPECT as well.

In an ideal case the scanner is perfectly aligned, i.e.:

- the straight line between the X-ray focal spot and the center of the detector is normal to the detector surface; this is called the *central ray* that together with central row of the detector defines the *midplane*
- the axis-of-rotation (AOR) is parallel to the detector columns and is projected onto the central column.

There are several reasons why in practice residual misalignments are unavoidable:

- fine adjustment of the scanner requires high precision positioning mechanics, which might be often too expensive to be worth building into the scanner.
- misalignment may result from an unstable X-ray focal spot position, which is usually the case in X-ray tubes with a very small focus size.

There are several degrees of freedom for deviation from the ideal geometry. If we arbitrarily take the *central ray* and the *midplane* as a reference then the misalignment errors can be defined as follows (Fig. 1):

- deviation of the AOR from ideal orientation and position
    - ⇨ *tilt* (inclination) towards the X-ray tube
    - ⇨ *skew* (rotation) around the central ray
    - ⇨ *horizontal transversal* off-center *shift*, i.e. along detector rows
    - ⇨ *horizontal longitudinal shift*, i.e. deviation from the ideal position between the X-ray tube and the detector
- deviation of the X-ray source location from the ideal position
    - ⇨ *vertical shift* from the midplane
    - ⇨ *horizontal transversal shift* from the central ray
    - ⇨ *horizontal longitudinal shift*, i.e. deviation of the X-ray tube ↔ detector distance from the assumed value.

Other possible errors (AOR wobble etc.) are assumed to be negligible or not present at all. For spiral/helical and other more complex acquisition paths more errors have to be appended to the list. For example the direction of object translation (*table feed* in medical CT scanners) may not be parallel to the AOR. The scanned object is then incrementally shifted off-center while being advanced to the next projection position on the spiral path.

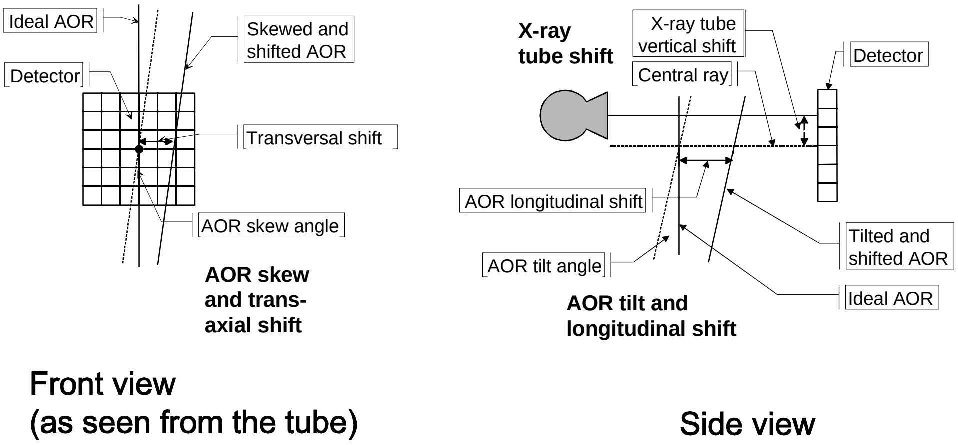


Fig. 1. Definitions of misalignment errors

Reconstruction algorithm with misalignment correction

The original Feldkamp reconstruction algorithm consists of the three main steps:

1) for every detector row and for every projection the projection pixels are independently *weighted* and then *filtered*, with the same one-dimensional filter being applied to each projection row. *Weighting* is done under the assumption that the midplane is perpendicular to the AOR and *filtration* is done along the projection lines *perpendicular* to the AOR. In the case of AOR *skew* weighting and filtration direction do not coincide with detector rows.
2) cone-beam backprojection.

An ideal algorithm incorporating misalignment correction should consist of the following steps:

1) weighting of projections with coefficients corrected for all errors
2) filtration of weighted projections, rotated to correct for the AOR *skew*
3) cone-beam backprojection corrected for tube position and tube ↔ object ↔ detector distances.

In practice such an algorithm has disadvantages. *Before* the entire volume is generated usually single preview slices are reconstructed, in order to select a region of interest for the final volume reconstruction and to possibly interactively tune misalignment correction. In case of the ideal algorithm a complete weighting and filtration of all projections is required to reconstruct even a single preview slice. Thus the reconstruction time would be heavily dominated by the filtration.

Therefore we developed an algorithm with built in misalignment correction where weighting and filtration is done on uncorrected projections. Afterwards the projections can repeatedly be used for backprojections without repeated weighting and filtration. The reconstruction algorithm, which has been described earlier [3][4], uses homogeneous coordinates for the system description and employs an incremental, pixel-driven backprojection.

In the case of a perfectly aligned scanner our algorithm produces results identical to the Feldkamp algorithm otherwise additional artifacts are introduced with our method. However, if AOR *tilt, skew* and *shift* errors are small we expect only a minor additional quality degradation of the final images. In practice they are significantly smaller than those inherent to the *approximate* Feldkamp algorithm.

Influence of misalignment on image quality

Vertical X-ray tube *shift* can be interpreted as if the tube were located on a new, vertically shifted *"midplane"*. As a consequence the sharp slice of the reconstructed volume moves up/down. The quality of reconstructed images is not be affected otherwise.

Horizontal X-ray tube *transversal shift* can be interpreted as if the object were shifted across the field of view and the AOR were *transversally shifted* and the detector slightly rotated. But with proper AOR *shift*-only correction the artifacts would be almost eliminated.

Horizontal X-ray tube *longitudinal shift* changes the object magnification. In the consequence images will be reconstructed "not to scale" and therefore cannot be used for quantitative evaluation of tomograms, but will be otherwise artifact-free.

AOR *tilt* moves the object central plane, i.e. the plane being perpendicular to the AOR and containing the tube focus. In case of tilt this plane intersects the detector above/below its central row. As a consequence weighting coefficients for projection pixels are modified. The visible effect of the *tilt* is that slice smearing towards top/bottom of the object will become unsymmetrical.

AOR *skew* has the most severe impact on the image quality in our algorithm, as we filter not only incorrectly weighted projections (weighting coefficients change their values due to the AOR *skew*), but also take slanted rows, instead of rows perpendicular to the AOR. This leads to crosstalk between slices. However, if the *skew* angle is small and if the object structure changes slowly along the AOR then the actually filtered and the required rows do not differ very much.

AOR horizontal *transversal shift* (off center shift) is responsible for double contours in reconstructed images. It can be efficiently eliminated in the backprojection.

AOR *longitudinal shift* – see X-ray tube.

Misalignment measurement

The critical problem for an effective implementation of our algorithm is proper measurement of scanner misalignment. Various procedures have been propose for this purpose [6][8][9]. Some of them take into account all misalignment errors, some are restricted to only a few, assuming others to be negligible. A good overview has been given in [7] along with their own method, in which several equiangular circular projections of a metal sphere phantom are acquired. Two spheres – one below and one above the *midplane* – are mounted away from the AOR. In a superimposed image their shadows form two ellipses – one in the upper and one in the lower half of the image. Scanner misalignment parameters are calculated from the ellipses' positions, orientations, and sizes.

We have evaluated the 'two sphere method' with respect to cone beam µCT using simulations and experiments. We simulated the CT acquisition with known scanner misalignments for cone angles ranging from 5° to 40°. The AOR *skew* error could always be determined with at least 1% accuracy. This was, however, not the case for the AOR *shift.* Accuracy errors in this parameter are linearly dependent on the error in the determination of the center of the spheres. A 1% error in the localization of the sphere center resulted in an AOR *shift* error of 0.4 pixel for a 40° cone angle and of 12 pixel for an 5° cone angle. The other misalignment errors could also determined with sufficient accuracy so that artifacts in the reconstructed images were negligible, apart from the AOR *tilt* error, which cannot be determined with this method at all. In practice even a 1% accuracy error in the detection of the sphere centers will be extremely difficult to achieve due to noise and potential distortions in the detector geometry.

Reconstruction algorithm verification

To evaluate our reconstruction algorithm and its potential to correct for misalignment errors we performed simulations of the numerical *Defrise phantom* and we scanned a physical phantom using typical acquisition parameters (512^2 pixel detector matrix, 2 x 5.7° cone angle, 720 equidistant projections acquired over 360°). The following projection sets were simulated: misalignment-free as a reference, AOR *tilt* of 0.2°, AOR *skew* of 0.75°, AOR transversal *shift* of 20 pixels and a combination of all three errors, i.e. errors equal to the actually measured for the physical Defrise phantom.

The Defrise phantom has been well established as the test object for demonstrating artifacts incurred by approximate reconstruction algorithms. However, it may not optimal for investigating the effect of scanner misalignment. Also it may overemphasize artifacts compared to samples typically investigated with µCT. Other phantoms, more suitable to this task need to be developed

Reconstruction artifacts in simulated Defrise phantom

The slice reconstructed with our algorithm in the case of no misalignment (Fig. 2a) contains characteristic Feldkamp artifacts. In case of a combined AOR misalignment the uncorrected image (Fig. 2.b) is heavily distorted. The image corrected for all three errors (Fig. 2.c) contains only residual artifacts, resulting from (consciously) imperfect handling of weighting/filtration in our algorithm. The sharp part of the image moved down slightly, which results from improper AOR tilt corrections. Slight horizontal streaks originating at the ends of the disks result from the AOR skew error. The AOR shift has been eliminated and leaves no artifacts. We then restricted our misalignment corrections to two errors: AOR skew and transversal shift. The fourth image (Fig. 2.d) shows an image reconstructed without tilt correction. There are no significant differences between the c. and d. images, e.g. traces of the tilt error.

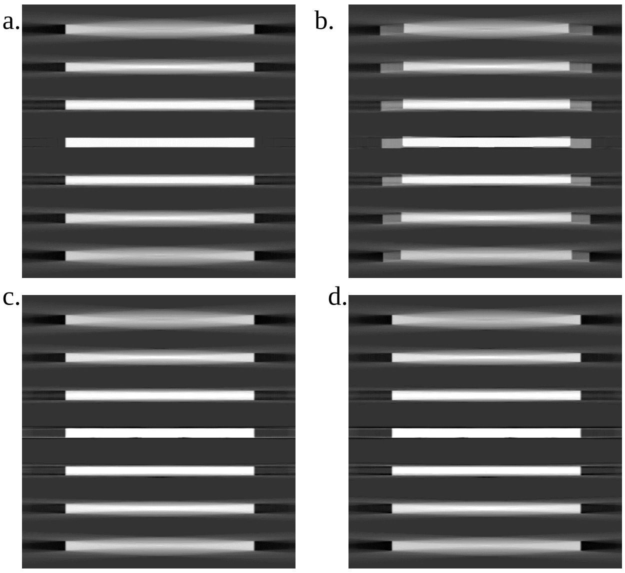


Fig. 2. Sagittal slices reconstructed from simulated projections: a – ideal case (no misalignment), b – AOR *tilt*, *skew* and *shift* uncorrected, c – all three corrected, d – AOR *skew* and *shift* corrected only (e.g. no *tilt* correction)

Images of a physical (scanned) Defrise phantom

Reconstructed images of the physical phantom demonstrate the influence of misalignment errors on the image quality in a measured phantom. It is clear that in the real case all seven misalignment errors are present. The uncorrected image is heavily distorted. However, the image corrected for AOR *shift*, *skew* and *tilt*, e.g. assuming other four errors to be negligible (as actually determined by the dual-ellipse method), contains only residual artifacts. Furthermore, the image with AOR *shift* and *skew* correction only (e.g. not *tilt* correction) does not look qualitatively worse. Hence the influence of the *tilt* error is negligible.

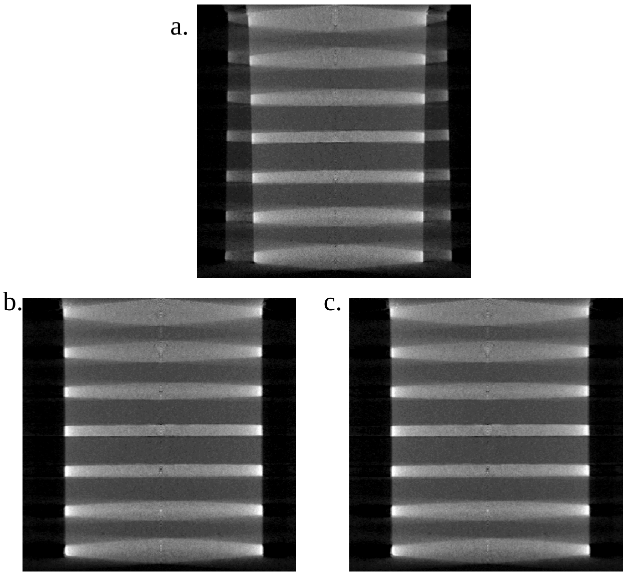


Fig. 3. Sagittal slices reconstructed from measured projections: a – uncorrected misalignment, b – AOR *skew*, *shift* and *tilt* corrected, c – AOR *skew* and *shift* corrected (no *tilt* correction)

Conclusions

This paper concerns the problem of scanner mechanical misalignment in cone-beam tomography. We have defined seven misalignment errors which may be present in a real-world scanner and analyzed their influence on the reconstructed image quality. We then proposed an effective approximate implementation of a Feldkamp-based reconstruction algorithm capable of correcting for all misalignment errors. In the algorithm a practical approach is used, with weighting and filtration performed on uncorrected projections and misalignment correction built into backprojection. The concept of storing weighted and filtered projections to be later used in backprojection proved to be advantageous in terms of saving processing time.

The problem o reliable measurement of all misalignment errors in a given scanner remains unsolved, especially for high resolution µCT scanners with small cone angles. Existing procedures did not prove to be robust enough to provide satisfactory results. Therefore we restricted misalignment correction to two main errors (AOR *shift* and *skew*) to obtain good quality tomograms. We showed that in a practical case artifacts caused by misalignment are therewith significantly reduced. However, the algorithm still preserves its potential to correct for all errors, should they be known for a given scanner.

The algorithm has been validated on simulated and physical Defrise phantom. Real data were acquired on a µCT scanner, whose misalignment had been measured. Images reconstructed with the proposed algorithm do not suffer from significant additional (apart from Feldkamp-specific) artifacts.

References

1. **K. Engelke, M. Karolczak, A. Lutz, U. Seibert (Taubenreuther), S. Schaller, W. Kalender,** *Mikro-CT. Technologie und Applikationen zur Erfassung von Knochenarchitektur,* Der Radiologe, No. 39, pp. 203-212, 1999
2. L. A. Feldkamp, . C. Davis, J. W. Kress, *Practical cone-beam algorithm,* J. Opt. Soc. Am., A/Vol. 1, No. 6, pp. 612-619, June 1984
3. **S. Schaller, M. Karolczak, K. Engelke**, K. Wiesent, *Implementation of a fast cone-beam backprojection algorithm for µCT using homogeneous coordinates,* Radiology, November 1998, (RSNA Congress, 1998)
4. **M. Karolczak, S. Schaller, K. Engelke, A. Lutz, U. Seibert (Taubenreuther),** K. Wiesent, **W. Kalender**, *Implementation of a cone-beam reconstruction algorithm for the single-circle source orbit with embedded misalignment correction using homogeneous coordinates*, submitted to Medical Physics
5. F. Noo, S. Schaller, R. Clackdoyle, T. A. White, T. J. Roney, *Image Reconstruction from Misaligned Truncated Helical Cone-beam Data,* IEEE Conference Record of the 1999 Nuclear Science Symposium and Medical Imaging Conference, Seattle, USA, 1999
6. A. V. Bronnikov, *Virtual alignment of x-ray cone-beam tomography system using two calibration aperture measurements,* Optical Engineering, Vol. 38(2), pp. 381-386, 1999
7. J. Hsieh, *Three-Dimensional Artifact Induced by Projection Weighting and Misalignment,* IEEE Transactions on Medical Imaging, Vol. 18, No. 4, pp. 364-368, April 1999
8. F. Noo, R. Clackdoyle, C. Mennessier, T. A. White, T. J. Roney, *An analytic method based on identification of ellipse parameters for scanner calibration in cone-beam tomography,* Physics in Medicine and Biology, 2000
9. **U. Taubenreuther**, *Correction and Calibration Methods for Cone-Beam MicroCT,* PhD Thesis, University of Erlangen, 2001
10. A. Rougée, C. Picard, C. Ponchut, Y. Trousset, *Geometrical calibration of x-ray imaging chains for three-dimensional reconstruction,* Computerized Medical Imaging and Graphics, Vol. 17(4/5), pp. 295-300, 1993
11. T. A. White, T. J. Roney, R. J. Pink, F. Noo, R. Clackdoyle, M. Smith, W. F. Jones, *Comparison of fan- and cone-beam imaging capabilities on a portable x-ray imaging system. In: Developments in X-ray tomography II, Ulrich Bonse,* Proceedings of SPIE, Vol. 3772, pp. 138-146, 1999
12. G. T. Gullberg, B. M. W. Tsui, C. R. Crawford, J. G. Ballard, J. T. Hagius, *Estimation of geometrical parameters and collimator evaluation for cone-beam tomography,* Medical Physics, Vol. 17(2), pp. 264-272, 1990

A Smoothing Prior with Embedded Positivity Constraint for Tomographic Reconstruction

Ing-Tsung Hsiao[1], Anand Rangarajan[2], Yuxiang Xing[1], Gene Gindi[1]
[1]Departments of Electrical & Computer Engineering and Radiology,
SUNY Stony Brook, Stony Brook, NY 11784, USA
[2]Department of Computer and Information Science and Engineering,
University of Florida, Gainesville, FL 32611, USA

I. Introduction

Positivity in regularized emission computed tomography (ECT) reconstruction is important for quantitative accuracy, especially for low-count data. However, it is often difficult to impose positivity on the reconstruction without suffering some other drawback, such as speed or lack of analyzability of the algorithm. A general framework for positivity-constrained ECT reconstruction has been the formulation and possibly constrained minimization of an objective function comprising a data penalty (usually log likelihood) and penalty term (a.k.a. "prior" in Bayesian terms). Here we propose to embed a positivity constraint via a novel prior for 3D ECT reconstruction.

The prior generalizes a notion of *I*-divergence proposed by Csiszár [1], and also bears a superficial similarity to formulations previously used in ECT [2], [3], [4]. However, unlike previous formulations, our new formulation also includes a notion of object *smoothness* in addition to positivity. In addition, the priors are convex and that makes the reconstruction independent of initial conditions.

In Sec.II, we describe the new priors mathematically and prove convexity. In Sec.III, we present initial results showing 2D and 3D reconstructions using our new priors, and compare these anecdotally to reconstructions with a conventional smoothing prior. In Sec.IV, we compare our work with other relevant work.

II. Theory

A. Regularized Likelihood Reconstruction

Let $\mathbf{g} = \{g_i; i = 1, .., M\}$ and $\mathbf{f} = \{f_n; n = 1, .., N\}$ be the emission data and object, respectively. Here, $\mathbf{f}$ and $\mathbf{g}$ are 1D vectors obtained by lexicographic ordering 3D entities into 1D vector. The projection data $\mathbf{g}$ has an independent Poisson distribution with the (negative) log-likelihood function

$$\Phi_L(\mathbf{g}; \mathbf{f}) = -\sum_i \{g_i \log \bar{g}_i - \bar{g}_i\} \tag{1}$$

where $\bar{\mathbf{g}} = \mathcal{H}\mathbf{f} + \bar{\mathbf{r}}$ is the Poisson mean of $\mathbf{g}$ with mean background events $\bar{\mathbf{r}}$. Note that $\mathcal{H}$ is the $M \times N$ system matrix with element $\mathcal{H}_{in}$ indicating the probability of a photon from pixel n detected in sinogram bin i.

From the Bayes theorem, the maximum *a posteriori* (MAP) reconstruction estimate $\hat{\mathbf{f}}$ of the object $\mathbf{f}$ is obtained by optimizing a regularized likelihood objective $\Phi(\mathbf{f}; \mathbf{g})$ with prior (or penalty) objective $\Phi_P(\mathbf{f})$. Therefore, the MAP reconstruction $\hat{\mathbf{f}}$ is computed by

$$\hat{\mathbf{f}} = \arg\min_{\mathbf{f} \geq 0} \{\Phi_L(\mathbf{g}; \mathbf{f}) + \lambda \Phi_P(\mathbf{f})\}. \tag{2}$$

where $\lambda > 0$ is the global weight controlling the influence of the prior.

B. Conventional Smoothing Prior

The objective function for a conventional smoothing prior usually takes the following form

$$\Phi_P^s(\mathbf{f}) = \sum_n \sum_{n' \in \mathcal{N}(n)} w_{nn'} \phi(f_n - f_{n'}) \tag{3}$$

where $\mathcal{N}(n)$ indicates the neighborhood system, and $w_{nn'} \geq 0$ the associated weight. The potential function $\phi(f_n - f_{n'})$ penalizes the difference between neighborhood pixels. For example, for a quadratic membrane prior [5], $\phi(f_n - f_{n'}) = 1/2(f_n - f_{n'})^2$. Note that neither Φ_L nor Φ_P^s provides a natural positivity enforcement in Eq.(2). Thus, Eq.(2) is a possibly difficult constrained optimization problem since one needs to enforce the non-negativity constraint in the reconstruction.

C. Priors Based on Cross-Entropy

As a stepping stone towards our new priors, we first consider another, older form based on *I*-divergence. The *I*-divergence between two positive vectors $\mathbf{a}$ and $\mathbf{b}$ is $S(\mathbf{a}, \mathbf{b}) = \sum_n (a_n \log \frac{a_n}{b_n} - a_n + b_n)$. $S(\mathbf{a}, \mathbf{b})$ is also termed "cross-entropy" in [3], [4] though definitions of cross-entropy vary. In [3], and [4], priors of the form $\Phi_P^{ce}(\mathbf{f}; \mathbf{m}) \equiv S(\mathbf{f}, \mathbf{m})$ or $S(\mathbf{m}, \mathbf{f})$ have been proposed. Here, $\mathbf{m}$ is an external "reference" vector chosen empirically. Each f_n is attracted towards its corresponding m_n, and positivity of $\mathbf{f}$ is maintained by the form of S. The "reference vector" $\mathbf{m}$ must be determined empirically! (For all m_n's equal, the cross-entropy prior becomes a max entropy prior.) In addition to the problem of determining $\mathbf{m}$, there is no explicit smoothing of $\mathbf{f}$ implicit in $\Phi_P^{ce}(\mathbf{f};\mathbf{m})$. Thus $\Phi_P^{ce}(\mathbf{f};\mathbf{m})$ suffers from two problems.

D. A New Prior: Smoothed I-Divergence

To solve these problems while retaining the positivity, convexity and differentiability desiderata of $S(\mathbf{f},\mathbf{m})$ priors, we define a new prior, termed a smoothed *I*-divergence prior. In this definition, $\mathbf{m}$ is a variable to be estimated, and the m_n's are defined on a lattice coincident with the f_n. The definition of our first form of the prior is:

$$\begin{aligned}\Phi_P^{fm}(\mathbf{f},\mathbf{m}) &= \sum_n \sum_{n'\in\mathcal{N}(n)} w_{nn'}\phi(f_n, m_{n'}) \\ &= \sum_n \sum_{n'\in\mathcal{N}(n)} w_{nn'}\{f_n \log\frac{f_n}{m_{n'}} - f_n + m_{n'}\} \qquad (4)\end{aligned}$$

where $\mathcal{N}(n)$ again defines a neighborhood of n, and $w_{nn'} \geq 0$ are positive weights. The form of Eq.(4) embeds positivity in $\mathbf{f}$, and the prior is also differentiable.

To explore Φ_P^{fm} further, consider its role in MAP reconstruction. Since $\mathbf{m}$ is now a variable, the optimization in Eq.(2) (now *un*constrained) becomes a joint estimation:

$$\hat{\mathbf{f}}, \hat{\mathbf{m}} = \arg\min_{\mathbf{f},\mathbf{m}}\{\Phi_L(\mathbf{g};\mathbf{f}) + \lambda\Phi_P^{fm}(\mathbf{f},\mathbf{m})\}. \qquad (5)$$

We propose to implement this joint estimation by a form of alternating (on $\mathbf{f}$ and $\mathbf{m}$) descent, which for iteration k becomes

$$\begin{aligned}\hat{\mathbf{f}}^{k+1} &= \arg\min_{\mathbf{f}}\{\Phi_L(\mathbf{g};\mathbf{f}) + \lambda\Phi_P^{fm}(\mathbf{f},\hat{\mathbf{m}}^k)\} \qquad (6) \\ \hat{\mathbf{m}}^{k+1} &= \arg\min_{\mathbf{m}}\{\Phi_P^{fm}(\hat{\mathbf{f}}^{k+1},\mathbf{m})\} \qquad (7)\end{aligned}$$

Equation (6), with $\hat{\mathbf{m}}^k$ fixed can be carried out by a suitable unconstrained gradient method. Equation (7) can be solved in closed form and has the interesting solution

$$\hat{m}_{n'}^k = \frac{\sum_{n\in\mathcal{N}(n')} w_{nn'}\hat{f}_n^k}{\sum_{n\in\mathcal{N}(n')} w_{nn'}}, \qquad (8)$$

that is, a weighted arithmetic mean of the $\hat{f}_n^k$'s in a neighborhood surrounding n'. Plugging Eq.(8) into Eq.(6), one thus sees the smoothing nature of the new prior, thus solving the first problem. The second problem has also been solved: vector $\mathbf{m}$ need not be determined empirically, and its formulation yields an appealing interpretation.

A second form of the new prior may be obtained by swapping $\mathbf{f}$, $\mathbf{m}$ in Eq.(4) to obtain

$$\begin{aligned}\Phi_P^{mf}(\mathbf{f},\mathbf{m}) &= \sum_n \sum_{n'\in\mathcal{N}(n)} w_{nn'}\phi(m_{n'}, f_n) \\ &= \sum_n \sum_{n'\in\mathcal{N}(n)} w_{nn'}\{m_{n'} \log\frac{m_{n'}}{f_n} - m_{n'} + f_n\} \qquad (9)\end{aligned}$$

Again, Φ_P^{mf} maintains positivity and is differentiable. When plugged into the alternation Eqs.(6)(7), however, the update Eq.(8) becomes something different:

$$\hat{m}_{n'}^k = \exp\left[\frac{\sum_{n\in\mathcal{N}(n')} w_{nn'}\log\hat{f}_n^k}{\sum_{n\in\mathcal{N}(n')} w_{nn'}}\right], \qquad (10)$$

which is a weighted *geometric* mean of the neighborhood pixels $\mathcal{N}(n')$. Again, we get a prescription for $\mathbf{m}$, and a novel form of smoothing for $\mathbf{f}$. For convenience, we refer to Eq.(4) as the FM prior, and Eq.(9) as the MF prior, with the quadratic version of Eq.(3) the MM prior.

E. Convexity

We would like to show the global convexity of the regularized likelihood objective with the proposed smoothing priors. Since the Poisson likelihood is convex, it is sufficient to show that the prior is convex (w.r.t. both $\mathbf{f}$ and $\mathbf{m}$).

The second derivatives of the MF prior objective are

$$\begin{aligned}\frac{\partial^2\Phi_P^{mf}}{\partial f_n^2} &= \sum_{n'} w_{nn'}\frac{m_{n'}}{f_n^2} \qquad (11) \\ \frac{\partial^2\Phi_P^{mf}}{\partial m_{n'}^2} &= \sum_n w_{nn'}\frac{1}{m_{n'}} \qquad (12) \\ \frac{\partial^2\Phi_P^{mf}}{\partial m_{n'}\partial f_n} &= -w_{nn'}\frac{1}{f_n} \qquad (13)\end{aligned}$$

It follows, for all y_n, $z_{n'}$ that

$$\begin{aligned}&\sum_n y_n^2\frac{\partial^2\Phi_P^{mf}}{\partial f_n^2} + 2\sum_{n,n'} y_n z_{n'}\frac{\partial^2\Phi_P^{mf}}{\partial m_{n'}\partial f_n} + \sum_n z_{n'}^2\frac{\partial^2\Phi_P^{mf}}{\partial m_{n'}^2} \\ &= \sum_{n,n'}\frac{w_{nn'}}{f_n^2 m_{n'}}(y_n m_{n'} - z_{n'}f_n)^2 \geq 0. \qquad (14)\end{aligned}$$

The Hessian matrix is positive semi-definite (w.r.t. $\mathbf{f}$,$\mathbf{m}$), and thus the MF prior is convex. For the convexity of the FM prior, the proof follows by symmetry. Convexity of the prior or likelihood alone does not guarantee a unique solution, but when the likelihood and prior are combined as in Eq.(2), the solution is unique.

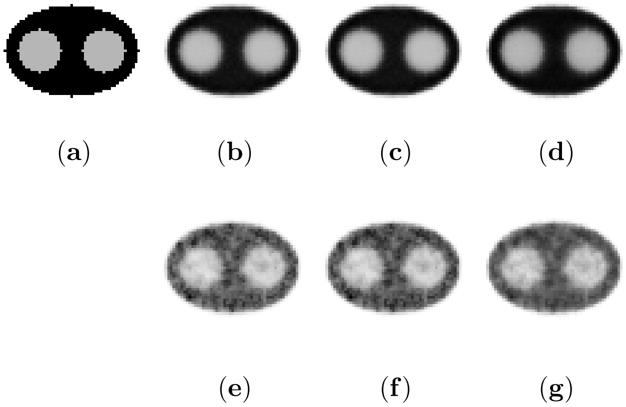
III. RESULTS

In this section, we anecdotally explore 2D and 3D reconstructions using Φ_P^{mf} and Φ_P^{fm} and show that the reconstructions are qualitatively similar to those obtained with a (positivity constrained) conventional quadratic smoothing prior.

A. Optimization Details

Since the new priors embed positivity constraints, any unconstrained method can be used to optimize Eq.(6). Here, we use the Polak-Ribiere form of the preconditioned conjugate-gradient (PCG) method [6] with a simple diagonal preconditioner. For the line search, we use the method of cubic interpolation [6]. We use, for 2D, a 4 nearest-neighborhood (4NN) system with weights=1.0 for each neighborhood pixel, and 4.0 for the center pixel. For 3D, we use a 6NN system with weights=1.0 for each neighborhood voxel, and 6.0 for the center voxel.

For comparison, we also performed a series of 2D reconstructions using the conventional quadratic smoothing Φ_P^s. We again use a 4NN neighborhood with weights as above. To maintain positivity, we used the slow EM-MAP-ICM algorithm presented in [5].


(a) **(b)** **(c)** **(d)**

(e) **(f)** **(g)**

Fig. 1. (a) The 64×64 phantom used in the 2-D simulations. The noiseless anecdotal reconstructions with priors of (b) MF, (c) FM, and (d) MM. The anecdotal reconstructions of 100K counts data using regularized likelihood reconstruction with priors of (e) MF, (f) FM, and (g) MM.

B. *Anecdotal Reconstructions*

To test the proposed new priors, we first generated noiseless and noisy sinograms using the 2D 64×64 phantom shown in Fig.1(a). The 2D phantom has an intensity ratio of (3.3:1:0) for (ellipse:circle:background). The projection data had dimensions of 65 angles by 96 detector bins. Only Poisson noise is simulated and no other physical or geometrical blurring effects or background events are modeled here. The noisy sinogram is simulated with 100K counts.

The sinograms are then reconstructed using MF, FM and MM priors as described above, with the same smoothing parameter value $\lambda = 1.2$ for the noiseless case, and $\lambda = 2.0$ for the noisy case. The noiseless reconstructions are displayed in Fig.1 for (b) MF, (c) FM, and (d) MM, respectively, while Fig.1(e)(f)(g) are for noisy reconstructions of MF, FM, and MM. A profile plot along the center row of the noisy reconstructions in Fig.1(e)(f)(g) is illustrated in Fig.2 along with the phantom. The profile result demonstrates the similar behavior of the new prior to that of a positivity-constrained quadratic smoothing prior.

A few performance measures are shown in Fig.3. Plots of log posterior vs. iteration, and *rmse* vs. iteration appear in Fig.3(a) and (b), respectively, for the FM results of Fig.1(f), but we draw no conclusions from these since the simulations are not yet sufficiently realistic. We also display the bias and STD (standard deviation) images for the FM prior using 200 noisy trials of 100K counts from the same 2D phantom. The bias and STD images are computed by $bias = \frac{1}{200}\sum_l(\hat{\mathbf{f}}^l - \mathbf{f})$ and $STD = \sqrt{\frac{1}{199}\sum_l(\hat{\mathbf{f}}^l - \bar{\hat{\mathbf{f}}})^2}$ where $\bar{\hat{\mathbf{f}}}$ is the mean reconstruction over 200 trials, $\mathbf{f}$ the phantom, and $\hat{\mathbf{f}}^l$ is the lth reconstruction. The results for MF and MM priors are similar. As λ increases, the variance as seen in Fig.3(d) tends to become more uniform.

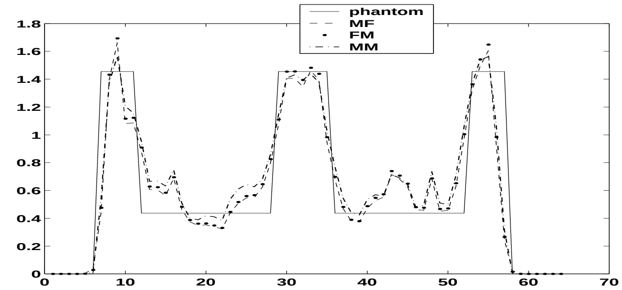
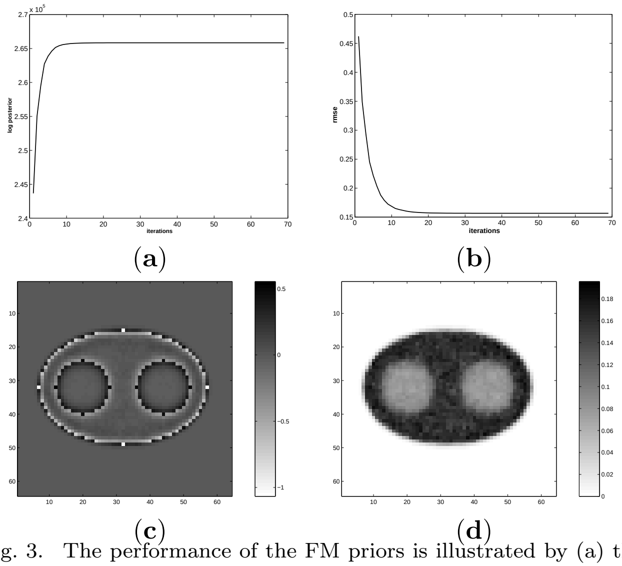


Fig. 2. Profile plots along the center row of phantom, and regularized likelihood reconstructions in Figs.1 (e) MF, (f) FM, and (g) MM priors for the noisy data of 100K counts.



(a) **(b)**

(c) **(d)**

Fig. 3. The performance of the FM priors is illustrated by (a) the log posterior and (b) the *rmse* vs. iterations. Also, the bias and standard deviation images over 200 noise trials for the FM prior regularized reconstructions are shown in (c) and (d), respectively.

We also generated a 3D phantom with 10 slices of 64×64 each. Each slice has two small hot and cold circles with an ellipse background. The intensity ratio is (hot:ellipse:cold)=(1.5:1:0.6). The hot and cold circles in the top slice have diameters of 10 and 12 voxels, respectively, while each slice below contains hot and cold circles of reducing diameters at 1 voxel per slice. A noisy sinogram is generated with 1000K total counts and reconstructed with a 3D FM prior of 6NN. Here, we show only 4 slices of the 3D phantom in the top row of Fig.4, and the FM reconstructions with smoothing parameter $\lambda = 5.0$ in the bottom row of Fig.4. The 3D results for the MF and MM priors are qualitatively similar.

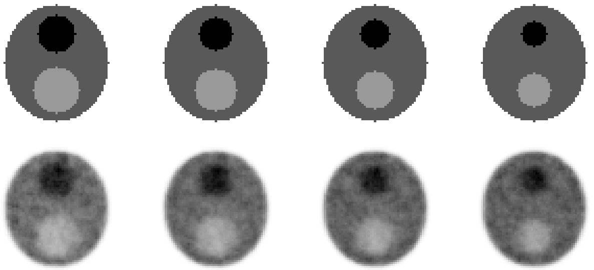


Fig. 4. This figure shows the anecdotal regularized likelihood reconstruction of a 3D phantom (10 slices and each with the size of 64×64) using the FM smoothing prior. The top row shows 4 consecutive slices of the phantom, and the bottom row shows their corresponding FM reconstructions. The projection data is simulated with a noisy level of 1 million total counts. The smoothing parameter for this 3D case is λ=5.

IV. Discussion

To locate our new work within the galaxy of ECT reconstruction methods, we take a distant (and *very* incomplete!) overview of objective-function based methods in the context of the following desiderata: (a) speed, (b) need to specify additional parameters beyond a global weight λ, (c) whether a notion of object smoothness is contained in the objective. We consider 3 categories of methods, (where "method" = objective + algorithm.) We do not consider any cases for which positivity is not imposed or the objective is non-convex.

The original ECT EM-ML algorithm and its variants led to a natural positivity imposition, but, in terms of (b) required a stopping criterion and initial condition specification. True EM-based methods are also notoriously slow, thus failing on point (a). Regularization helps, and EM-MAP approaches [7], [5] incorporating smoothing, satisfied (c), helped (b), but still suffered in terms of speed (a). Our own method, if using gradient based methods as in Sec. III, is likely to be faster than an EM version.

A second (non-EM) category of methods includes constrained gradient and coordinate-descent methods for optimization. Imposing positivity for a gradient method is often complex and difficult, and requires judicious specification of algorithm-specific parameters [8], [9] to attain a good result, thus failing on point (b), though excelling on point (a) and satisfying (c). To a great extent, the failing on (b) can be overcome [10] with improved methods. Sequential Gauss-Seidel coordinate descent methods [11], [12] incorporate positivity easily, and can be fast.

A third category of approaches embeds positivity directly into the prior. Our approach, and ones based on cross-entropy [2], [4], [3] and max-entropy [13] apply here. A very similar prior is the independent gamma prior proposed in [14]. The cross-entropy and gamma formulations fail badly on point (b), requiring empirical specification of a pointwise "reference image" as outlined in Sec. II-C, and no explicit notion of object smoothness is incorporated in these approaches. (In [2], [4], smoothing is indeed introduced in an empirical fashion.) Our own priors thus surmount the difficulties listed above.

A final, nice feature of our formulation is that the performance of the method (i.e. mean and covariance of $\hat{\mathbf{f}}$) may be analyzable using the theoretical methods in [15]. To qualify for [15], the objective should be (i) smooth, (ii) nearly quadratic, and (iii) be used in an unconstrained optimization. We are currently investigating this issue.

V. Acknowledgments

This work was supported by NIH NS32879.

References

[1] I. Csiszár, "MAXENT, mathematics, and information theory," in *Proc. 15th Inter. Work. Maximum Entropy and Bayesian Methods*, K. Hanson and R. Silver, Eds., 1995, pp. 35–50.

[2] J. Liang, R. Jaszczak, and K. Greer, "On Bayesian image reconstruction from projections: Uniform and nonuniform *a priori* source information," *IEEE Trans. Med. Imag.*, vol. 8, no. 3, pp. 227–235, 1989.

[3] C. L. Byrne, "Iterative image reconstruction algorithms based on cross-entropy minimization," *IEEE Tran. Imag. Proc.*, vol. 2, no. 1, pp. 96–103, Jan. 1993.

[4] B. A. Ardekani, M. Braun, B. F. Hutton, I. Kanno, and H. Iida, "Minimum cross-entropy reconstruction of PET images using prior anatomical information," *Phys. Med. Biol.*, vol. 41, no. 11, pp. 2497–2517, Nov. 1996.

[5] S. J. Lee, A. Rangarajan, and G. R. Gindi, "Bayesian image reconstruction in SPECT using higher order mechanical models as priors," *IEEE Trans. Med. Imag.*, vol. vol. 14, pp. 669–680, 1995.

[6] D. G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley Publishing Company, second edition, 1984.

[7] T. Hebert and R. Leahy, "A generalized EM algorithm for 3-D Bayesian reconstruction for Poisson data using Gibbs priors," *IEEE Trans. on Medical Imaging*, vol. 8, no. 2, pp. 194–202, June 1989.

[8] E. U. Mumcuoglu, R. Leahy, S. R. Cherry, and Z. Zhou, "Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 687–701, Dec. 1994.

[9] C. Johnson, J. Seidel, and A. Sofer, "Interior-point methodology for 3-D PET reconstruction," *IEEE Trans. Med. Imag.*, vol. 19, no. 4, pp. 271–285, Apr. 2000.

[10] E. U. Mumcuoglu, R. M Leahy, and S. R. Cherry, "Bayesian reconstruction of PET images: Methodology and performance analysis," *Phys. Med. Bio.*, vol. 41, pp. 1777–1807, 1996.

[11] J. A. Fessler, "Penalized weighted least-squares image reconstruction for positron emission tomography," *IEEE Trans. Med. Imag.*, vol. 13, no. 2, pp. 290–300, June 1994.

[12] C. A. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Trans. Imag. Proc.*, vol. 5, no. 3, pp. 480–492, Mar. 1996.

[13] J. Nunez and J. Llacer, "A fast Bayesian reconstruction algorithm for emission tomography with entropy prior converging to feasible images," *IEEE Trans. Med. Imag.*, vol. 9, no. 2, pp. 159–171, 1990.

[14] K. Lange, M. Bahn, and R. Little, "A theoretical study of some maximum likelihood algorithms for emission and transmission tomography," *IEEE Trans. Med. Imag.*, vol. 6, no. 2, pp. 106–114, June 1987.

[15] J. Fessler, "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography," *IEEE Trans. Imag. Proc.*, vol. 5, no. 3, pp. 493–506, Mar. 1996.

3D list-mode reconstruction for SPECT

L Bouwens §, R Van de Walle§, H Gifford‡, I Lemahieu§ and R A Dierckx†

§ ELIS Department, MEDISIP, IBITECH, Ghent University
Sint-Pietersnieuwstraat 41 B-9000 Ghent, Belgium
† Nuclear Medicine Dept., Ghent University Hospital
De Pintelaan 185 B-9000 Ghent, Belgium
‡ University of Massachusetts Medical School, Division of Nuclear Medicine
55 Lake Ave. North, Worcester, MA, USA

Abstract— **In previous work we developed a list-mode iterative reconstruction algorithm (LMIRA) for SPECT applied to two dimensional image reconstruction. In the 2D case we calculated the emission radiance distribution on the circular enclosure of the object. This part is the forward projection. In the backprojection we sampled the radiance on the intermediate layer using the geometric properties of the collimator. In 3D reconstruction we have to calculate the emission radiance distribution on the cylindrical enclosure. To characterize the radiance on the cylinder we need additional parameters to localize the position on the layer and the direction of the escaping photons, this results in a 4 dimensional function. The probabilities used in the LMIRA update equation are calculated from sampling the radiance with the cylindrical hole structure of the collimator.**

Keywords— **List-mode, reconstruction, iterative.**

I. Introduction

In binned mode PET the acquired data is often rebinned to 2D information using single slice rebinning or for example Fourier rebinning. In Single Photon Emission Computed Tomography (SPECT) reconstruction it is also possible not to use the 3D information and perform a 2D slice by slice reconstruction using the 2D reconstruction methods. 3D list-mode reconstruction has previously been the subject of research in positron emission tomography (PET) [1] and Compton scatter imaging [2]. Fully 3D reconstruction is useful to incorporate the complete blurring model to be able to correct for the distance dependent resolution [3], which is a three dimensional effect. Recently we developed a 2D list-mode reconstruction method for SPECT applying a new projector backprojector pair. The list-mode reconstruction approach differs in several ways from the bin-mode methods. Acquiring the data in list-mode format, one can store the interaction location to a high degree of accuracy (2k by 2k for example) with greater efficiency than achievable with frame mode acquisition. The gantry angles do not have to be binned into predefined frames, but one can record the actual angle thereby removing the impact of angular blurring with continuous acquisition. The actual energy of the interaction can be recorded instead of attributing the event's energy to one of a limited number of pre-defined windows. It is obvious that when increasing the dimensionality in this way it is no longer possible to bin the data into a matrix. Finally, list mode can also store gating signals without the need for temporal framing of the data before this information is completely available. The result is a significant increase in the fidelity of recording the projection data with list-mode acquisition, without a tremendous increase in storage space. In bin-mode acquisition of SPECT data the number of detector locations is predefined and rarely exceeds $256x256$ detector bins acquired over 120 angular steps. This means data will be grouped together thereby loosing detailed information. For each bin i it is possible to obtain the probability of the i-th outcome, given an emission at the j-th voxel within the object. When the number of possible detector bins is limited one can calculate and store the entire transition matrix, giving the relationship between the object space and the detector space. With this information the probability of detecting an emission from the j-th voxel can be calculated. To calculate the maximum likelihood solution to this probabilistic problem, iterative methods such as expectation maximization maximum likelihood (MLEM) [4][5] have been presented. In list-mode the set of possible outcomes is so large that few of the outcomes occur maximum more than once and most of them never occur at all. In binned format this would mean that most of the elements of the, very large, sinogram are zero. Therefore it is more efficient to acquire the detected events in a list, together with their detection parameters such as detector location, gantry angle, energy, time stamp, etc. Since it is no longer possible to store the transition matrix, it will be necessary to calculate the probabilities, used in the reconstruction algorithm, on the fly. In our previous work this was done using a new approach: the intermediate layer. In previous work we developed a list-mode iterative reconstruction algorithm (LMIRA) [6] for SPECT applied to two dimensional image reconstruction. This reconstruction method used the geometric collimator model to apply the resolution recovery in a single slice [7]. In this work we deploy a modification of the projector and backprojector to fit the three dimensional reconstruction problem. First we will explain the transition of the calculation of the emission radiance distribution from 2D to 3D. Secondly we detail the sampling of the obtained radiance distribution with the three dimensional collimator structure will be described. Finally we will present the 3D list-mode reconstruction algorithm.

E-mail: Luc.Bouwens@rug.ac.be

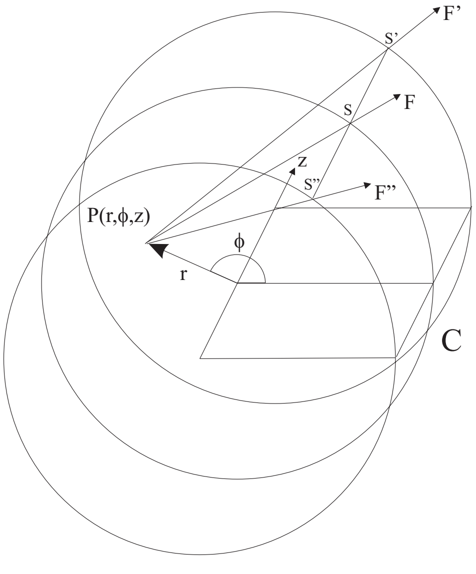


Fig. 1. Forward projection of the source distribution onto the intermediate layer.

II. Method

A. 3D emission radiance distribution

In 2D list-mode reconstruction a circular enclosure was used surrounding the object, called the intermediate layer. For the 3D problem, using for example a parallel hole or a fan beam collimator, a cylindrical enclosure will be best suited. For the imaging of small object with a pinhole or a cone beam collimator it may be that a sphere is more appropriate for this specific problem. In this work we discuss the case of a parallel hole collimator and thus use a cylindrical intermediate layer.

To calculate the radiance distribution on the intermediate layer an isotropic emission is generated at each point in the source distribution. For an isotropic emission generated at point $P(r,\phi,z)$ the contribution to the radiance in each point on the cylinder has to be calculated. For the location $S(\psi,z)$ on the cylinder, the emission from the source element P will be given by the vector $F_P(\psi,z,\theta,\kappa)$ (image 1). To obtain the radiance distribution from P at S one can use the known expression:

$$F_P(\psi,z,\theta,\kappa) = E_P.\exp\left(-\int_P^S \mu(r,\Psi,z).ds\right), \tag{1}$$

With E_P the value of the source distribution at P, and $\mu(r,\Psi,z)$ the continuous attenuation coefficient distribution. The total emission radiance distribution $F(\psi,z,\theta,\kappa)$ in point S can be calculated by integration of the function F over the line through the object defined by S, θ and κ.

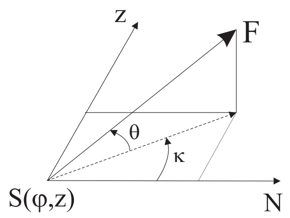


Fig. 2. Parameterizing the direction of the escape vector F.

In the expression the parameters ψ and z indicate the location on the cylinder face. The parameters θ is the angle of the projection of F on the trans-axial slice through the source point P relative to horizontal x-axis and κ is the angle, obtianed when projection F in the plane given by the axial line (z-axis) through the location point S and the normal vector N in S, relative to N. These parameters give the direction of the escape vector F. To incorporate the detector response in the z direction we will need the additional parameter κ (image 2).

When projecting the source distribution to the intermediate layer it is possible to build in the following restriction. In the axial direction the projection can be limited by the known geometric parameters of the collimator. Since photons leaving the intermediate layer, with an angle κ greater than the acceptance angle of the collimator, will never be acceptable by the collimator holes, there is no need to calculate the projection beyond this point. In this way the axial projection can be restricted. The angle θ can not be restricted since the detector is rotating in this direction. Image 3 shows the emission radiance distribution for a point source. The right image is the surface plot of the left image. The top image is the complete distribution and the bottom image is the restricted radiance for a High Resolution parallel Hole (HR) Marconi collimator.

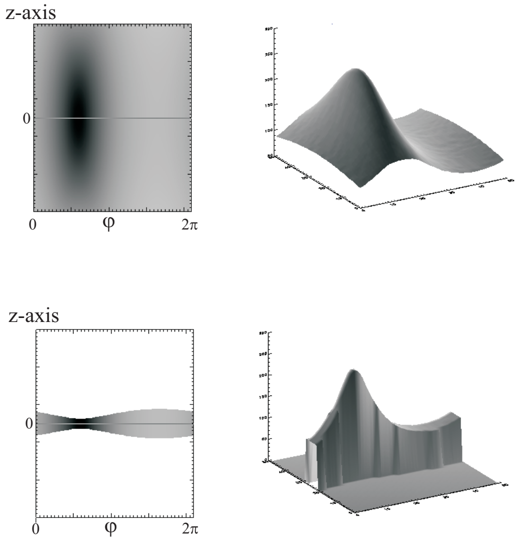


Fig. 3. Emission radiance distribution for a point source out of center. Top: complete radiance distribution. Bottom: radiance distribution restricted by the collimator axial acceptance angle.

B. Sampling the radiance distribution with a collimated detector

Consider data acquired from an object into list-mode information with the following parameters: axial location of incidence a, trans-axial location t and gantry angle δ. The collimator discussed primarily in this paper is a parallel hole collimator. For each single list-mode event n in the reconstruction we place a collimator hole over the center of the detector location (a, t). From the bottom area of the collimator hole towards the intermediate layer, (i.e. the cylinder shown in figure 1), we will see a fraction of the cylinder. Lines drawn from the outermost parts of the collimator hole, over the edges of the collimator septa, to the cylinder restrict the solid angle seen (figure 3).

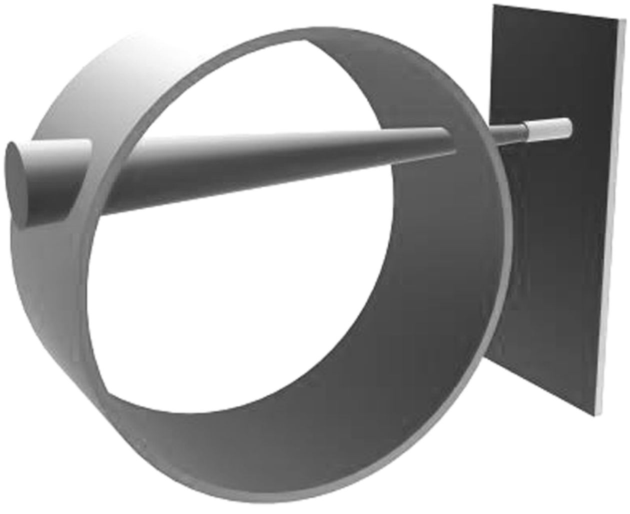


Fig. 4. Backprojection of the cone from a collimator hole through the cylinder.

When we consider the intersection of the cone coming from the collimator hole and the cylinder face, all elements on the cylinder with possible contribution of the radiance to this specific list-mode event are selected, these elements are called scanning points. In every scanning point the emission radiance distribution is parameterized by the angles θ and κ. The face of the cone will put restrictions on these angles since the vector of the escaping radiance has to lay within the cone (image 5). In a first step one has to obtain the area on the cylinder face which is the intersection of the cone and the cylinder. Secondly for each scanning point in this area the restriction on θ and κ have to be applied. This will give the four dimensional region $A_n(\psi, z, \theta, \kappa)$

When we integrate the emission radiance distribution $F(\phi, z, \theta, \kappa)$ over the region $A_n(\phi, z, \theta, \kappa)$, we get the contribution of the emitted photons from the total source distribution into the specific detector location during projection. In the backprojection step, A_n defines the region of the object into which the counts in the detector bin will be backprojected and therefor will depict how the detected photons, for a specific detector location, will contribute to the backprojected image. It is also possible to use an expression for the geometric point response in terms of the autocorrelation of the collimator aperture function for one collimator hole [8]. This operator can be used to calculate the individual voxel contributions. When we do not want to model the effects of attenuation and scatter in the backprojection, but only the distance dependent resolution, we can use the basic collimator aperture function. The effect of the mismatched projector/backprojector pair has to be investigated [9].

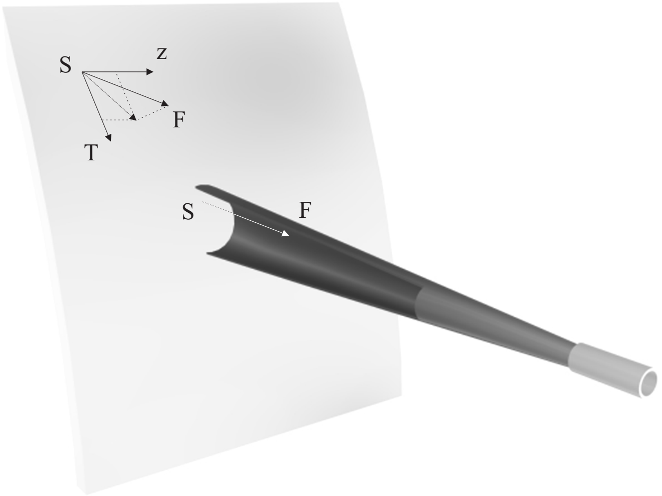


Fig. 5. Selection of the vectors from the emission radiance distribution in the acceptance cone.

C. List-mode reconstruction

Consider a data set of N list-mode events with attributes being the coordinates, gantry angle, energy level, etc. If we want to use the ML algorithm we need to calculate $P_j(l_n)$, for $n = 1, 2, ..., N$ and $j = 1, 2, ..., J$, the probability density that an item l_n would occur in the list, given there was an emission of a photon in the j-th voxel.

A list-mode maximum-likelihood reconstruction algorithm for PET was previously developed by Parra et. al. [10], [11] :

$$x_j^{(k+1)} = \frac{1}{N} \cdot \sum_{n=1}^{N} \frac{P(l_n|j).x_j^{(k)}}{\sum_{j=1}^{J} P(l_n|j).s_j.x_j^{(k)}} \tag{2}$$

In this equation $x_j^{(k)}$ is the expected number of photons emitted from source bin j per unit of time for the k^{th} iteration, and s_j is the sensitivity for that source bin. N is the number of list-mode events. J is the number of source bins.

The nominator in the summation over the list-mode events N is probability density given above. This can be seen as the probability that an emission in voxel j will lead to the detection of the list-mode event l_n. The denominator is the total over all voxel of the probability density functions given the detection of event l_n. When we look at the radiance distribution this will give the probability density functions of photons coming from the source distribution, leaving the intermediate layer in a specific direction. In the previous section we discussed how to restrict the area on the intermediate layer for a given list-mode event. Doing this we can obtain the contribution of the radiance distribution to a specific event. This is equal to the denominator in equation. The nominator can be calculated from the radiance for a singe voxel or by using the collimator aperture function. Therefore, the equation can be rewritten as follows:

$$Y_{jn}^{(k+1)} = \frac{\oint_{A_n} F_j^{(k)}(\phi, z, \theta, \kappa)}{\oint_{A_n} F^{(k)}(\phi, z, \theta, \kappa)} \tag{3}$$

where A_n is the list-mode specific acceptance area. $F_j^{(k)}$ is the radiance distribution for voxel j for the k^{th} estimate, and $F^{(k)}$ is the total radiance distribution for the k^{th} source estimate. Since one list-mode event is the detection of a single photon, the backprojected value is normalized to 1.

$$y_{jn}^{(k+1)} = \frac{Y_{jn}^{(k+1)}}{\sum_{j=1}^{J} Y_{jn}^{(k+1)}} \tag{4}$$

$$x_j^{(k+1)} = \sum_{n=1}^{N} y_{jn}^{(k+1)} \tag{5}$$

The algorithm updates the estimate of the source distribution for each list-mode event and sums the individual results to the final new estimate for the source distribution. In this way the integrals can be calculated on the fly and do not have to be stored. The reconstruction start from a uniform source distribution to calculate the primary radiance. After a run through the list-mode data set, the emission radiance distribution is recalculated and used in the next iteration.

III. Results

Preliminary test of the reconstruction algorithm were performed on a simple three dimensional Gaussian distributed sphere with FWHM of a central cross section equal to 12.5 mm, placed in the center of a 32 cube matrix with voxel size 3,125 mm. List-mode information for this object was derived from Monte-Carlo simulation, where the data was acquired in projections of 512x512 detector bins over 360 angles. A high-resolution parallel-hole collimator was specified, with hole size 1.4 mm, length 27 mm and radius of rotation equal to 112 mm. This results in the maximum diameter of the backprojected cone being 14.4 mm or less than 5 voxels. The initial emission radiance distribution was derived from a uniform cylinder with radius 10 cm within the 32x32x32 matrix. As with the 2D LMIRA, the three dimensional reconstruction is resolving the object over the different iterations. Further study is needed to optimize the reconstruction procedure.

IV. Discussion and conclusion

In this study we propose an extention of the 2D list-mode reconstruction towards fully three dimensional image reconstruction. Calculating the emission radiance distribution it is possible to incorporate the effects of attenuation and scatter. This is can be done by using ray-tracing as implemented in the 2D LMIRA. A different approach could be using a model based deformation of the radiance based on the attenuation and electron density maps. The calculation of the 4D radiance function is a computational intensive task. In our approach this calculation has to be done once every iteration. The calculation of the intersection of the cone coming from the collimator hole and the cylinder could be used more efficiently to define the voxel within the object which can possibly contribute to the list-mode event. In that way the number of voxel which have to be taken into account in the reconstruction of a single event can be reduced significantly.

Acknowledgments

This work was supported by "Institute for the Promotion of Innovation by Science and Technology in Flanders" (IWT, Belgium). Rik Van de Walle is a post-doctoral fellow of the Fund for Scientific Research-Flanders (FWO, Belgium). We thank Marconi Medical Systems for their support on the list-acquisition with the Irix gamma camera.

References

[1] AJ Reader, K Erlandsson, and RJ Ott, "Attenuation and scatter correction of list-mode data driven iterative and analytic image reconstruction algorithms for rotating 3D PET systems," *IEEE Transactions on Nuclear Science*, vol. 46, pp. 2218–2226, December 1999.

[2] SJ Wilderman, WL Rogers, GFKnoll, and EJCngdahl, "Fast algorithm for list mode back-projection of compton scatter camera data," *IEEE Transactions on Nuclear Science*, vol. 45(3), pp. 957–962, 1998.

[3] L Bouwens, R Van de Walle, J Nuyts, M Koole, Y D'Asseler, S Vandenberghe, RA Dierckx, and I Lemahieu, "Image-correction techniques in spect," *Computerized Medical Imaging and Graphics*, vol. 25, pp. 117–126, March-April 2001.

[4] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Transactions on Medical Imaging*, vol. 1, pp. 113–122, Oct. 1982.

[5] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," *J. Comput. Assist. Tomogr.*, vol. 8, pp. 306–316, April 1984.

[6] L Bouwens, R Van de Walle, H Gifford, MA King, I Lemahieu, and RA Dierckx, "LMIRA: List mode iterative reconstruction algorithm for SPECT," *submitted for publication*, 2001.

[7] L Bouwens, H Gifford, R Van de Walle, MA King, I Lemahieu, and RA Dierckx, "Resolution recovery for list mode reconstruction in SPECT," *submitted for publication*, 2001.

[8] GT Gullberg, BM Tsui, CR Crawford, JG Ballard, and JT Hagius, "Estimation of geometrical parameters and collimator evaluation for cone beam tomography," *Medical Physics*, vol. 17(2), pp. 264–272, 1990.

[9] SJ Glick and EJ Soares, "Noise characteristics of SPECT iterative reconstruction with a mismatched projector-backprojector pair," *IEEE Transactions on Nuclear Science*, vol. 45, pp. 2183–2188, August 1998.

[10] HH Barrett, T White, and LC Parra, "List-mode likelihood," *Journal of the Optical Society of America A-Optics Image Science and Vision*, vol. 14, pp. 2914–2923, November 1997.

[11] L. Parra and H.H. Barrett, "List-mode likelihood : EM algorithm and image quality estimation demonstrated on 2-D PET," *IEEE Transactions on Medical Imaging*, vol. 17, pp. 228–235, April 1998.

Content-Adaptive Mesh Modeling for Tomographic Image Reconstruction[1]

Jovan G. Brankov, Yongyi Yang, and Miles N. Wernick

Illinois Institute of Technology
Department of Electrical and Computer Engineering
3301 S. Dearborn St., Chicago, IL 60616, USA

ABSTRACT

In this paper we propose the use of a content-adaptive mesh model (CAMM) for tomographic image reconstruction. In the proposed framework, the image to be reconstructed is first modeled by an efficient mesh representation. The image is then obtained through estimation of the nodal values from the measured data. The use of a CAMM can greatly alleviate the ill-posed nature of the reconstruction problem, thereby leading to improved quality in the reconstructed images. In addition, it reduces the data storage requirement, particularly for fully three-dimensional (3D) image reconstruction, resulting in efficient algorithms. The proposed methods are tested using gated cardiac-perfusion images. Initial results demonstrate that the proposed approach achieves the best performance when compared to several commonly used methods for image reconstruction, and produces results very rapidly.

1. INTRODUCTION

A great many methods have been developed for reconstructing tomographic images. Most of these methods are based on pixel (or voxel) image representations. Bayesian priors (e.g., [1]) or regularization terms (e.g., [2]) are typically used to combat the effect of noise.

Alternative model-based reconstruction approaches have also been proposed. For example, cylindrical models were proposed in [3] and surface models were used in [4,5].

In this work we propose a new content-adaptive mesh modeling approach for image reconstruction. In this approach, a customized basis representation is computed for the image, then the parameters of this representation are estimated from the data.

In a mesh model, the image domain is subdivided into a collection of mesh elements, the vertices of which are called *nodes*. The image function is then obtained over each element by interpolation from the values of these nodes [6]. In a content-adaptive mesh model (CAMM), the mesh elements are placed in a fashion that is adapted to the local content of the image. A CAMM provides an efficient representation of the image in that the number of parameters (i.e., mesh nodes) is typically much less than the number of required pixels/voxels. In addition, a mesh model can also be used for motion tracking in an image sequence, by allowing the mesh to deform over time [7].

The potential benefits of using a CAMM for image reconstruction are: 1) a CAMM greatly reduces the number of unknowns, thus alleviating both the underdetermined nature of the reconstruction problem and the data storage requirement, particularly for the case of 3D reconstruction; 2) this reduction in the number of unknowns can lead to a fast computation; 3) a CAMM provides a natural spatially-adaptive smoothness mechanism, eliminating the need for regularization terms in the cost function; and 4) the CAMM provides a natural framework for reconstruction of moving image sequences.

2. METHODS

2.1 Mesh Tomography Model

Let $f(\mathbf{x})$ denote the image function defined over a domain D, which can be either 2D or 3D in this study. In a mesh model, D is partitioned into M non-overlapping mesh elements, denoted by $D_m,\ m=1,2,\cdots,M$. The image function is represented as

$$f(\mathbf{x})=\sum_{n=1}^{N}\varphi_n(x)f(\mathbf{x}_n), \qquad (1)$$

where $\mathbf{x}_n$ is the nth mesh node, $\varphi_n(\mathbf{x})$ is the interpolation basis function associated with $\mathbf{x}_n$, and N is the total number of mesh nodes used. Note that the support of each basis function $\varphi_n(\mathbf{x})$ is limited to those elements D_m attached to the node n. In practice, mesh elements with simple geometry such as 2D triangular or 3D tetrahedral elements are often used.

Now let $\mathbf{n}$ denote a vector formed by the nodal values of the mesh model, i.e.,

$$\mathbf{n}\equiv\left[f(\mathbf{x}_1), f(\mathbf{x}_2),\cdots f(\mathbf{x}_n)\right]^{\mathbf{T}}. \qquad (2)$$

If $\mathbf{f}$ denotes the voxel representation of the image function $f(\mathbf{x})$ over D, then from (1) and (2) one can obtain

---

[1] This research was supported by the Whitaker Foundation and by the National Institutes of Health under grant HL65425.

$$\mathbf{f} = \Phi\mathbf{n}, \tag{3}$$

where Φ is a matrix, composed from the interpolation functions $\varphi_n(\mathbf{x})$ in (1), that forms the interpolation operator from a mesh representation to the pixel representation.

For tomographic image reconstruction, the imaging equation is typically written in terms of the voxel representation $\mathbf{f}$ as

$$E[\mathbf{g}] = \mathbf{Hf}, \tag{4}$$

where $\mathbf{g}$ contains the measured data, $E[\cdot]$ is the expectation operator, and $\mathbf{H}$ is a matrix describing the imaging system.

Substituting (3) into (4), we obtain the mesh-domain imaging equation:

$$E[\mathbf{g}] = \mathbf{H}\Phi\mathbf{n} \equiv \mathbf{An}, \tag{5}$$

where $\mathbf{A} = \mathbf{H}\Phi$.

The reconstruction problem becomes that of estimating $\mathbf{n}$ from the given data $\mathbf{g}$. The image $\mathbf{f}$ can then be obtained from (3).

2.2 Reconstruction Algorithms

In this paper we investigate maximum-likelihood and least-squares estimates of the nodal values in $\mathbf{n}$.

A. Maximum-Likelihood Estimate

The maximum-likelihood (ML) estimate is obtained as

$$\hat{\mathbf{n}}_{ML} = \arg\max_{\mathbf{n}} \left\{ \log\left[p\left(\mathbf{g};\mathbf{n}\right) \right] \right\}, \tag{6}$$

where $p(\mathbf{g};\mathbf{n})$ is the likelihood function of $\mathbf{g}$ parameterized by $\mathbf{n}$. In this paper, we assume a Poisson likelihood, which characterizes emission tomography

The ML estimate can be computed by using the following expectation-maximization (EM) algorithm [8]:

$$\mathbf{n}_s^{(j+1)} = \frac{\mathbf{n}_s^{(j)}}{\sum_t \mathbf{A}_{ts}} \sum_t \mathbf{A}_{ts} \left(\frac{\mathbf{g}_t}{\sum_k \mathbf{A}_{tk}\mathbf{n}_k^{(j)}} \right), \tag{7}$$

where $\mathbf{n}_s^{(k)}$ is the value of node *s* in iteration *j*, $\mathbf{g}_t$ is the recorded count for observation t, and $\mathbf{A}_{ts}$ is the *ts* entry of matrix $\mathbf{A}$. We refer to this algorithm throughout as MESH-EM.

B. Least-Squares Estimate

The least-squares estimate is obtained as the solution of the following optimization problem:

$$\hat{\mathbf{n}}_{LS} = \arg\min_{\mathbf{n}} \left\| \mathbf{g} - \mathbf{An} \right\|^2, \tag{8}$$

where $\|\cdot\|$ is the Euclidean norm. This quadratic objective function has a unique solution, provided that $\mathbf{A}$ is of full rank. In this study, we used the conjugate gradient algorithm [9] to perform the optimization. We refer to this reconstruction algorithm as MESH-LS.

3. RESULTS

A. Evaluation Data

The proposed CAMM-based reconstruction algorithms were tested using the 4D gated mathematical cardiac-torso gMCAT D1.01 phantom [10], which is a time sequence of 16 3-D images. The field of view was 36 cm; the pixel size was 5.625mm. Poisson noise, at a level of 4 million total counts per 3D time-frame image, was introduced into the projections to simulate a clinical Tc^{99m} study. To validate the concept of our proposed new approach, in our initial study a single slice (No.70) was chosen, which has approximately 55,000 counts per frame and a total of 16 frames. No attenuation correction was used. Each image frame was reconstructed separately, and a single mesh structure was used for all frames.

B. Reconstruction Methods Considered

In addition to the two proposed reconstruction algorithms, we also considered three well-known reconstruction procedures for comparison purposes: (1) filtered back projection (FBP); (2) pixel-based ML-EM reconstruction [8] with spatial post-filtering; and (3) a pixel-based MAP method with a spatial Gibbs prior [1,11].

The coefficients used for the spatial Gibbs prior are $\alpha = 1$, $\beta = 1.2$, $\delta = 3$, $\gamma = 0.35$ [11]. For the spatial post-filtering a 2D Butterworth spatial filter with a cutoff frequency of 0.2 cycles/pixel was used. For consistency in the comparison, the same post-filtering was also applied to MESH-EM and MESH-LS methods in the final results. Each of the iterative reconstruction algorithms was run for 30 iterations.

C. Mesh generation

The mesh structure was estimated from the projection data using the following procedure. First, the projection data were summed over the 16 frames. From these summed projections an image was reconstructed using FBP. The resulting image, denoted by $\bar{f}(\mathbf{x})$, provides a rough estimate of the heart summed over all 16 frames.

Based on $\bar{f}(\mathbf{x})$, we generated a mesh structure using a procedure similar to the one we proposed in [12]. In that paper we proposed a very fast and effective method for mesh generation, in which error-diffusion halftoning of a gradient-magnitude image is used to generate mesh nodes whose spatial density is proportional to the local rate of intensity change in the image.

The method reported in [12] was presented as an *ad hoc* approach, but we have since derived a theoretical basis for this concept, which shows that the correct image to use in place of the gradient magnitude is the following:

$$\vartheta(\mathbf{x}) = \max\left(\left|\nabla^2_{xx}\bar{f}(\mathbf{x})\right|, \left|\nabla^2_{xy}\bar{f}(\mathbf{x})\right|, \left|\nabla^2_{yy}\bar{f}(\mathbf{x})\right|\right). \quad (9)$$

From this image, we compute a feature map as follows:

$$\sigma(\mathbf{x}) = \begin{cases} \vartheta(\mathbf{x})^{0.475} & \mathbf{x} \in \textit{Heart region} \\ \vartheta(\mathbf{x})^{0.95} & \mathbf{x} \in \textit{Background} \end{cases}. \quad (10)$$

In our preliminary experiments, the "*Heart region*" and "*Background*" were estimated using a simple intensity-based segmentation procedure.

The mesh node locations are obtained from this feature map [12] by error-diffusion halftoning, from which the mesh structure is obtained by Delaunay triangulation (see Figure 1). A total of 609 mesh nodes are used in the mesh shown in Figure 1, only about one-seventh the number of pixels. Note that the algorithm places mesh nodes densely in the important heart regions, and sparsely in the background. This mesh was used as a basis on which to reconstruct each of the image frames in the sequence. In another paper submitted to this conference, we describe how the mesh can be used to track motion in a 4D smoothing algorithm for gated studies.

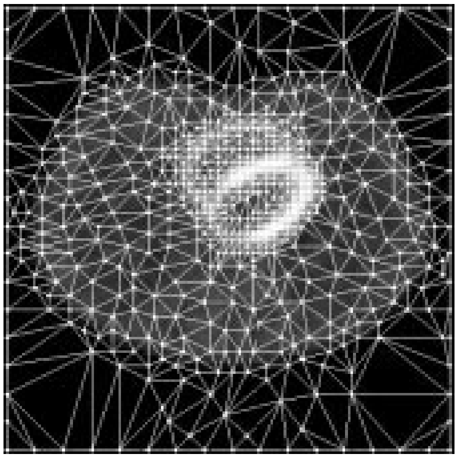


Figure 1. Content-adaptive mesh model of the torso, including the heart, using 609 mesh nodes.

D. Results

For visual comparison, images of frame 14, obtained by different reconstruction methods, are presented in Figure 2. The MESH-EM algorithm appears to produce the best images, accurately capturing the heart wall and applying appropriate smoothing in the background. The MESH-LS algorithm does not perform as well, possibly because it is based on a suboptimal statistical representation of the noise.

In Figure 3 we show the peak-signal-to-noise ratio (PSNR) versus the frame number. In Table 1 we summarize the execution time, memory requirement and PSNR averaged over all frames for various algorithms. According to all of these criteria the MESH-EM algorithm exhibits the best performance.

We also tested the proposed methods using a much coarser mesh structure (only 353 nodes). In this case, the speed of MESH-EM is further improved (reduced from 4.5 seconds to 3.9 seconds in runtime), but the image quality is almost preserved (average PSNR reduced from 27.4 dB to 26.9 dB, which is still better than that of the other methods in Table 1).

A final note is that the MESH-EM algorithm can be further accelerated using an ordered-subsets framework [13,14]. Preliminary results indicated that our mesh based ordered-subsets EM algorithm is about 4 times faster when compared to a voxel-based ordered-subsets EM algorithm [14]. We plan to furnish detailed, complete results of these comparisons by the time of the conference.

5. REFERENCES

[1] S. Geman and D. Geman, "Stohastic relaxation, Gibbs distributions, and Bayesian restoration of images," *IEEE Trans. Patt. Anal. Mach. Intell*., vol. 6, pp. 721-741, 1984.

[2] J. Fessler, "Penalized Weighted Least-Squares Reconstruction for Positron Emission tomography," *IEEE Trans. Med. Imaging*, vol. 13, pp. 290-300, 1994.

[3] Y. Bresler, J.A. Fessler, and A. Macovski, "A Bayesian approach to reconstruction form incomplete projections of a multiple object 3D domains," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 11, pp. 840-858, 1989.

[4] G. S. Cunningham, K. M. Hanson, and X. L. Battle, "Three dimensional reconstruction from low-count SPECT data using deformable models," *IEEE. Med. Imaging Conf.*, 1997.

[5] G. R. Jennings and D. R. Wolf, "Tomographic reconstruction based on flexible geometric models," *IEEE Int. Conf. on Image Proc*., 1994.

[6] Y. Wang and L. O., "Use of two-dimensional deformable mesh structures for video coding .I. The synthesis problem: mesh-based function approximation and mapping," *IEEE Trans. Circuits Syst. for Video Tech.*, vol. 6, pp. 636 -646, 1996.

[7] Y. Altunbasak and A. M. Tekalp, "Occlusion-adaptive, content-based mesh design and forward tracking," *IEEE Trans. Image Proc.*, vol. 6, pp. 1270-1280, 1997.

[8] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Sect*., vol. 39, pp. 1-38, 1977.

[9] E. K. P. Chong and S. H. Zak, *An Introduction to Optimization*. New York: John Wiley & Sons, Inc., 1996.

[10] P. H. Pretorius, M. A. K. W. Xia, B. M. W. Tsui, T. S. Pan, and B. J. Villegas, "Evaluation of right and left ventricular volume and ejection fraction using a mathematical cardiac torso phantom for gated blood pool SPECT," *J. of Nucl. Med*, vol. 38, pp. 1528-1534, 1997.

[11] D. S. Lalush and B. M. W. Tsui, "Space-Time Gibbs Priors applied to Gated SPECT Myocardial Perfusion Studies," presented at *3D Image Rec. in Radiology*, Dordrecht, 1996.

[12] J. Lee, Y. Yang, and M. N. Wernick, "A New Approach for Image-Content Adaptive Mesh Generation," *ICIP-2000,* 2000.

[13] H. Hudson and R. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans Med Imag*, Vol.13, pp.601-609, 1994.

[14] H. Erdogan and J. A. Fessler, "Ordered subsets algorithms for transmission tomography," *Phys. Med. Biol*, Vol.44, 1999.

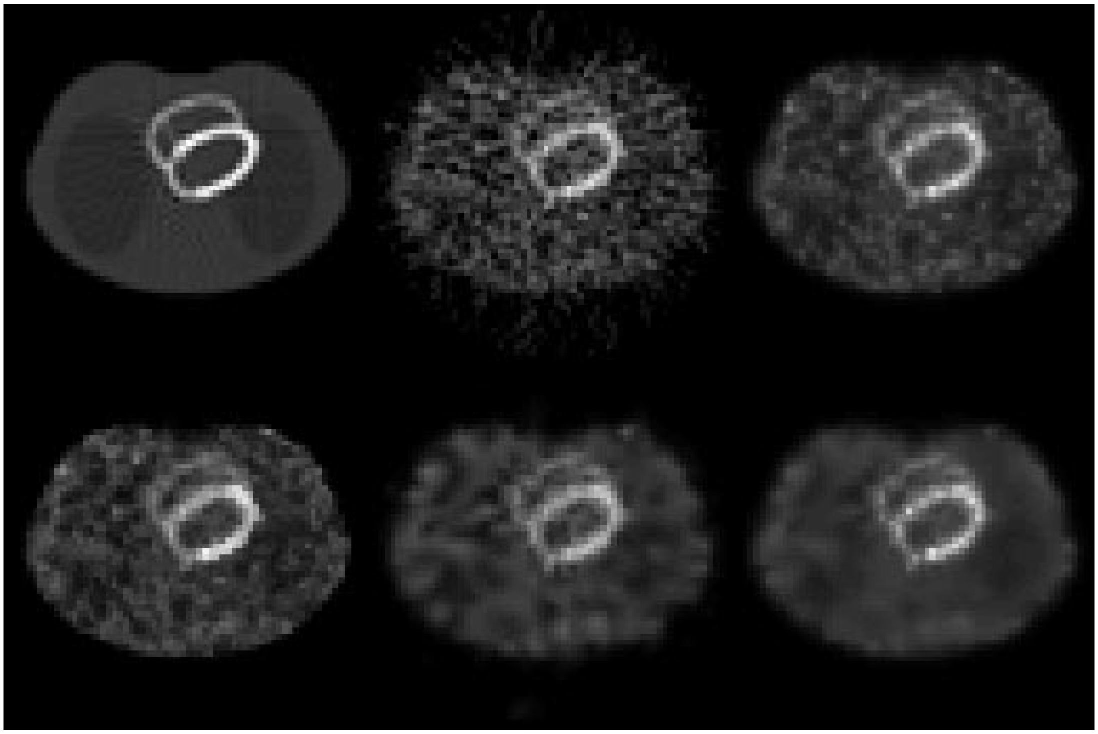


Figure 2. From left to right in top row: Original phantom, Filtered-backprojection reconstruction, ML-EM reconstruction. Bottom row: MAP reconstruction, MESH-LS reconstruction, and MESH-EM reconstruction.

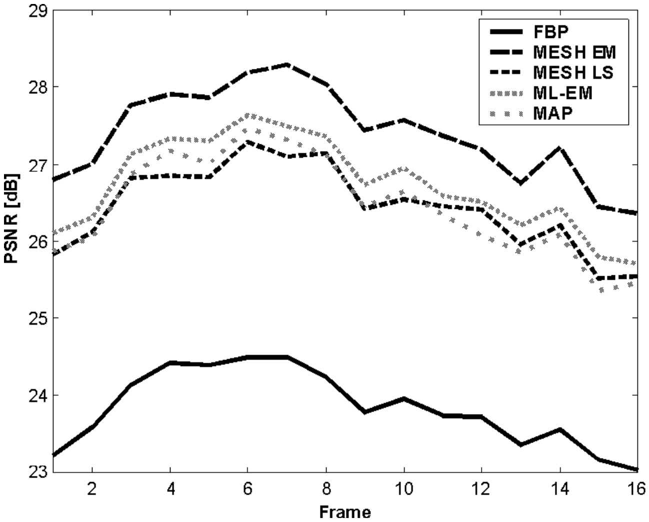


Figure 3. PSNR vs. frame number for various reconstruction methods.

TABLE 1. COMPARISON OF RECONSTRUCTION METHODS

	Execution Time [sec]	Memory requirement [Mb]	Average PSNR [dB]
FBP	0.12	0	23.8
ML-EM*	5.7	5.1 (4096x4096)	26.7
MAP*	9.3	5.1 (4096x4096)	26.5
MESH-LS*	8.3	4.2 (4096x609)	26.4
MESH-EM*	4.5	4.2 (4096x609)	27.4

* Obtained by prestoring the system matrix as a sparse matrix

Segmented Multiple Plane Reconstruction - A Novel Approximate Reconstruction Scheme for Multislice Spiral CT

Karl Stierstorfer[a], Thomas Flohr, Herbert Bruder

***Abstract*—A new reconstruction scheme for multirow spiral CT is described and results are presented. The spiral path is decomposed into small, overlapping segments which are used for a separate convolution and backprojection yielding a stack of segment images which contain only data of a partial scan. These segment image stacks are, in a second step, reformatted to the requested planes. In a third step, the reformatted segment images are added to obtain full images.**

The main benefit of the proposed algorithm is the superior image quality. The limit of the algorithm has not been probed; a 64-row dataset with pitch 80 has been reconstructed with excellent image quality.

***Index Terms*—Approximate reconstruction, cone artifact, spiral CT, multislice CT.**

I. INTRODUCTION

A. Basic idea

Lately, a new class of reconstruction algorithms has emerged which further extend the domain of approximative reconstruction for multislice spiral CT.

The underlying idea is the following simple observation: if it was possible to find an image plane so that the focus does not leave this plane during a half turn of the spiral, we would be able to choose (or interpolate), for each projection angle and each fan parameter, rays that are fully contained in the image plane. A simple 2D reconstruction of these rays would yield an *exact* reconstruction of this image plane. Of course, we know that this is impossible for a spiral scan path, but it turned out fruitful to take this idea as a starting point for approximations. All of the algorithms described shortly in the following are based on this idea.

B. Advanced Single Slice Rebinning (ASSR)

The Advanced Single Slice Rebinning (ASSR) reconstruction method, ([1]-[3]) tries to match image planes directly to a π segment of a spiral path. Kachelrieß et al. reconstruct these images from overlapping π intervals and reformat the tilted images to axial planes in a second step. The images are reconstructed from overlapping scan intervals of π. The main drawback of the ASSR algorithm lies in the fact that it is useful only for the maximum pitch (approximately 1.4 times the number of rows for typical CT scanners) or, alternatively, requires severe detector masking by software for lower pitch values resulting in poor dose usage. Even at the optimum pitch, the dose usage is only 70 %. Another drawback of ASSR is that the average distance of the focal spot from the tilted image plane - which is a measure of the quality of the approximation to the basic idea described in section A - increases with pitch, hence degrading image quality.

C. Adaptive Multiple Plane Reconstruction (AMPR)

The Adaptive Multiple Plane Reconstruction (AMPR) scheme [4] solves the pitch restriction problem of the ASSR by introducing a second tilt angle, with a tangent to the spiral as the hinge line. By reconstructing, for each of the overlapping reconstruction intervals, several images, rotated by different angles around this hinge, a much larger fraction of the dose can be used.

D. Segmented Multiple Plane Reconstruction (SMPR)

The Segmented Multiple Plane Reconstruction (SMPR) algorithm presented here goes one step further: it uses only a small *segment* (typically less than one eighth of a full turn) of the spiral to reconstruct a *booklet* (stack) of *segment image pages*. Since the segments are small, the pages of a booklet can be matched almost perfectly to the spiral path. Then, for all segments of a π interval, all booklets belonging to a segment and its complementary segment *in all rotations* are reformatted to the desired (e.g. axial) planes. In a last step, images from segments of a π interval are combined to a full image. Essentially, the steps of reconstruction (convolution and backprojection) and z-interpolation are interchanged with respect to conventional multislice spiral algorithms.

In principle, the match of the pages to the spiral path will be the better the shorter the segment is. For a 16-row scanner, we found that eight segments are sufficient for good image quality. The 64-row reconstructions were done with 32 segments.

II. ALGORITHM

A. Segment images

The first step is the reconstruction of segment images. For each of the segments, a fraction of $2\pi/N_{seg}$ (N_{seg} is the number of segments per rotation) of one rotation (plus some

---

[a] All authors are with Siemens Medical Solutions Group, D-91301 Forchheim, Germany. E-Mail: karl.stierstorfer@med.siemens.de.

overlapping to provide smooth transitions from one segment to the next) is used to reconstruct a *booklet* of at least N_{rows} (the number of detector rows) pages. For each page, the rays closest to the image plane are selected for convolution and 2D-backprojection. Hence, we obtain per rotation a total of $N_{seg} \cdot N_{rows}$ segment images. The segment images are reconstructed with the finally desired field of view. It can be shown that, by doing this, the information available in the data is used almost perfectly if N_{seg} is sufficiently large. Figure 1 gives an impression of a typical segment image.

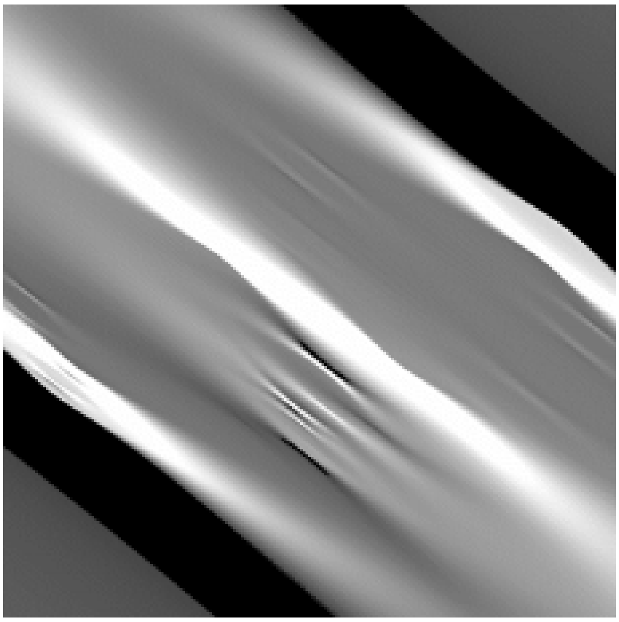


Fig. 1. A typical segment image (N_{seg}=32).

B. Stack reformation

The next step is to reformat the pages of a segment and its complementary segment from all rotations to the desired (e.g. axial) image planes. This can be done in a pixelwise fashion, in z direction only. In this step, the final image slice thickness can be adjusted by changing the width of the weight function used for reformation.

The location of the segment pages is shown in Figure 2 for the case of N_{seg}=N_{rows}=6. The pages contributing to one of these reformatted segment images are shown in the same shade.

C. Segment adding

The final step is a simple adding of the reformatted segment images of at least one half turn to a complete image.

III. TRANSITION TO FELDKAMP ALGORITHM

A theoretically interesting aspect of the SMPR approach is that, by using smaller and smaller segments, we finally end up with filtering and backprojecting only one (parallel) projection at a time onto planes which are spanned by the projection vectors of one detector row. Taking into account the second step of reformatting to axial planes and the final step of combining segments it becomes clear that this is equivalent to a spiral version of the Feldkamp algorithm which involves a filtering parallel to the spiral path followed by a 3D backprojection. Hence, the SMPR algorithm provides a smooth transition from an algorithm which utilizes a 2D-backprojection to a 3D-backprojection algorithm. A byproduct of this consideration is that a canonical filter direction for the filter step of the spiral Feldkamp reconstruction is obtained.

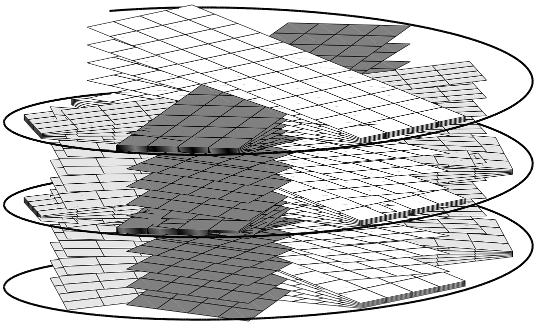


Fig. 2. Demonstration of the spiral path together with some sample booklets. Pages which are reformatted together are shown in the same shade.

IV. RESULTS

To test the algorithm we used the simulation program DRASIM (Siemens Medical, Forchheim) to produce a test data set of a thorax phantom (geometry definition by Katja Sourbelle, FORBILD project) for a fan-beam scanner with 64x1mm rows and a pitch of 80. The data were reconstructed with the SMPR using N_{seg}=32 (Figure 3) and, for comparison, also with the AMPR algorithm (Figure 4).

The SMPR image is obviously almost free of artifacts whereas the AMPR image exposes severe artifacts, particularly near the strongly tilted ribs.

V. CONCLUSION

A substantially improved approximative reconstruction algorithm for multislice spiral CT has been presented. While the limits of the algorithm have not yet been probed, results are excellent. A drawback of the algorithm is the large amount of intermediate segment images which has to be handled during the reformatting step.

From the theoretical point of view, an interesting aspect of the algorithm is that it provides a smooth transition from a 2D reconstruction approach to a 3D Feldkamp-type spiral algorithm.

Fig. 3. Typical axial image of a thorax phantom at pitch 80, reconstructed with SMPR using 32 segments. The field of view is 400 mm; the display window is 200 HU. Almost no artifacts are visible.

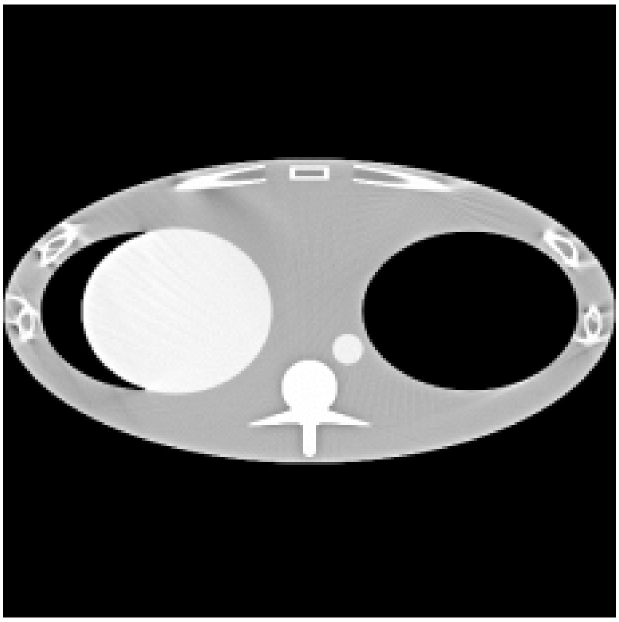


Fig. 4. Axial image of a thorax phantom at pitch 80, reconstructed with AMPR. The field of view is 400 mm; the display window is 200 HU; same slice location as figure 3. Severe artifacts are visible, particularly near the ribs.

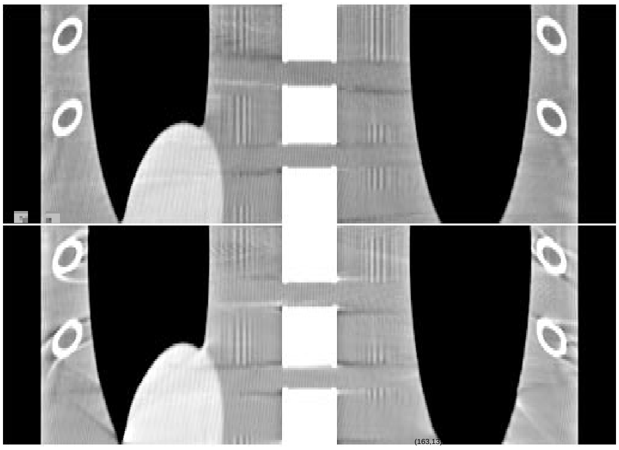


Fig. 5. MPRs of stacks of axial images, thorax phantom at pitch 80. Top: reconstructed with SMPR using 32 segments, bottom: reconstructed with AMPR. Display window 200 HU. Again, the SMPR images are almost free of cone artifacts while severe artifacts are visible in the AMPR image. The fine vertical streaks are spiral artifacts which are independent of the type of reconstruction.

REFERENCES

[1] G. L. Larson, C. C. Ruth, C. R. Crawford, US-Patent US5802134, "Nutating slice CT image reconstruction apparatus", 1998.

[2] M. Kachelrieß, S. Schaller, W. A. Kalender, "Advanced Single-Slice Rebinning in Cone-Beam Spiral CT", Med. Phys. **27**, pp. 754-772, 2000.

[3] H. Bruder, M. Kachelrieß, S. Schaller, K. Stierstorfer, T. Flohr, "Single-Slice Rebinning Reconstruction in Spiral Cone-Beam Computed Tomography", IEEE Trans. Med. Imag. **19**, pp. 873-887, 2000.

[4] S. Schaller, K. Stierstorfer, H. Bruder, M. Kachelrieß, and T. Flohr, "Novel approximate approach for high-quality image reconstruction in helical cone beam CT at arbitrary pitch", to be published in Proc. SPIE Medical Imaging Conf. 4322, 2001.

A combination of rebinning and exact reconstruction algorithms for helical cone-beam CT

Michel Defrise, Frédéric Noo, Hiroyuki Kudo

INTRODUCTION

Cone-beam X-ray transmission computerized tomography (CB-CT) has been applied to the non-destructive evaluation of industrial samples, and a small number of prototype clinical scanners have also been used for specific applications (cardiac and vascular imaging), where fast imaging is an essential requirement and only high contrast structures must be visualized. The recent introduction of multi-row scanners now allows a wide-spread utilization of CB-CT in radiology, and the continuing development of large area x-ray detectors will reinforce this trend.

The acquisition of cone-beam projections allows faster imaging than standard or spiral single-row CT, but presents challenges for 3D image reconstruction. Despite significant advances, the definition of a clinically acceptable, fast and accurate, cone-beam reconstruction algorithm is still an open problem, especially for the helical geometry where the x-ray source moves along a segment of helix relative to the patient. Accurate algorithms for helical CB-CT have been derived by discretizing exact or quasi-exact analytical inversion formulae for the 3D divergent-beam x-ray transform [4,7,8,12]. These algorithms yield accurate reconstructions even when the ratio between the pitch of the helix and the axial slice width exceeds two orders of magnitude. Unfortunately, the numerical complexity of these algorithms significantly increases the reconstruction time and the discretization errors and affects the resolution. These limitations of the exact algorithms motivate an active research for approximate algorithms for helical CB-CT.

In this paper we introduce a new quasi-exact algorithm for the long-object problem in helical CB-CT, which combines a rebinning method with the quasi-exact ZB method [4].

Rebinning algorithms are based on the factorization of the 3D reconstruction into a set of independent 2D reconstructions. This factorization can be viewed as a four-step procedure:

a) *Selection* of a number of 2D surfaces, called *rebinning surfaces*, which partition the 3D field-of-view.

b) *Rebinning*: estimation of a complete 2D tomographic data set (sinogram) for each selected surface using the measured CB projections.

c) *2D reconstruction* of each selected surface from its estimated sinogram.

d) *Axial Interpolation*: 1D interpolation of the reconstructed surfaces to obtain the transaxial slices to be visualized.

Several rebinning algorithms have been proposed, which differ mainly by the type of surface for which rebinned 2D data are estimated. These surfaces can be transaxial slices orthogonal to the axis of the helix [SSRB, 10], tilted planar slices "tailored" to the slope of the helix [9,14,1,3], or even non-planar surfaces as in Heuscher [2] to further improve the possibility to select rays lying as close as possible to the rebinning surface. We present in the following section a unified derivation of the various rebinning methods.

Rebinning algorithms only involve a 1D ramp filtering of the data, and avoid the cone-beam backprojection. This numerical simplicity results in less discretization errors than the exact algorithms. The price to pay for this improvement is that rebinning is approximate and introduces cone-beam artefacts when the pitch of the helix becomes too large.

Therefore, the aim of this paper is to combine the best of the two classes of algorithms: we propose to use a quasi-exact algorithm, the ZB method, to correct a first image reconstructed using a rebinning algorithm. The additive correction is obtained by applying the ZB method to the *residual error*, i.e. to the difference between the measured projections and the cone-beam projections of the first image estimate. This approach is motivated by the fact that the discretization artefacts induced by the exact algorithm only affect the correction image, and hence are expected to be smaller than when the exact algorithm is applied directly to the measured data.

The performance of the combined algorithm is evaluated using synthetic data for mathematical phantoms and for data obtained by reprojecting high resolution CT scans.

M.D. is with the Department of Nuclear Medicine, AZ-VUB, Free University Brussels, Belgium (mdefrise@minf.vub.ac.be).
F.N. is with the Department of Radiology, University of Utah, Salt Lake City (noo@montefiore.ulg.ac.be)
H.K. is with the Department of Information Science and Electronics, University of Tsukuba, Japan (kudo@is.tsukuba.ac.jp)

PLANAR AND NON-PLANAR REBINNING ALGORITHMS

The first contribution of this paper is to give a unified and formal derivation of the various rebinning algorithms which have been proposed in the litterature [1,2,3,9,10,14,15]. We determine the mathematical properties, in particular the symmetries, of the surface allowing the best rebinning, given some reasonable quality figure. This surface is shown to be the solution to an integral equation which can be solved by means of an iterative algorithm. A proof of the convergence of this algorithm and of the unicity of the optimal rebinning surface is given. We also investigate how much can be gained by rebinning onto non-planar surfaces as opposed to planes, and conclude that rebinning onto planar surfaces is sufficiently accurate as long as the radius of the field-of-view does not exceed about half the radius of the helix (figure 1).

For the combined algorithm described in the next section, we rebin on tilted planes, the orientation of which optimize the mean square axial distance Q between the rebinning plane and the measured rays used to build the corresponding 2D data set [3]. We generate a short-scan fan-beam data set rather than a parallel-beam data set, and the mean square axial distance Q is calculated taking into account both the fan-beam redundancy weight (Parker's weight) and the distance dependent weight in the 2D fan-beam backprojection.

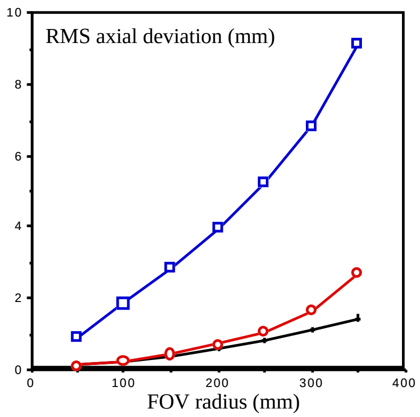


Figure 1: Root mean square axial distance (mm) between the rebinning surface and the measured rays used to build the 2D data set, versus the radius of the FOV (mm). Helix radius 400 mm. Pitch 100 mm. The 3 curves correspond from top to bottom to: SSRB (rebinning on transaxial planes), optimal tilted plane, optimal non-planar surface.

THE COMBINED ALGORITHM

We denote the unknown 3D image by f and the 3D X-ray transform (the CB projector) by X. The data are denoted as $g=Xf$.

Recall first the steps of the ZB algorithm [see 4 for details]:

a) *Weighted backprojection* of the subset of cone-beam data g located, in the detector, along the boundary of Tam's window B[1]. This backprojection operator is denoted X_B^* and is defined so as to yield an image $f_1 = X_B^* g$ the cone-beam projections of which, $g_1 = X f_1$, are equal to the data along the boundary of Tam' window: $g_1 = g$ on ∂B.

b) *Reprojection*: Calculation of the cone-beam projections of the first image estimate f_1 and subtraction from the data to get modified data $g_2 = g - g_1$. Note that $g_2 = 0$ on ∂B.

c) *Reconstruction of the modified data* by

i) setting $g_2 = 0$ outside B,

ii) applying a 1D ramp filter along the direction tangent to the helix,

iii) backprojecting as in Feldkamp's algorithm.

We denote this procedure as $F_{B\text{-}FDK}$ [6] and the result as $f_2 = F_{B\text{-}FDK}\, g_2$.

d) *Addition* of the two images after applying a smoothing filter S designed to match spatial resolutions: $f = S\ f_1 + f_2$.

The key to the quasi-exactness of the ZB method is the property that $X f_1 = g$ on ∂B. This property is satisfied by the image $f_1 = X_B^* g$ constructed in step a above, but this image has no resemblance to the true image f, and therefore the applied correction f_2 (and the associated discretization errors) may be large.

In the combined algorithm proposed in this paper, we build an alternative image f_1 which also satisfies $X f_1 = g$ on ∂B but is in addition a good approximation to f. This is done using a rebinning algorithm (described in the previous section, and denoted by F_{reb}):

a) *Initial reconstruction* by rebinning: $f_{reb} = F_{reb}\, g$

b) *Calculation of a first image estimate* as

$$f_1 = X_B^*(g - X f_{reb}) + f_{reb}$$

where X_B* is the weighted backprojection of the data along ∂B, defined above.

The last three steps are as in the ZB method:

c) *Reprojection*: calculation of the cone-beam projections $g_1 = X f_1$ of the first image estimate, and subtraction from the data: $g_2 = g - g_1$

d) *Feldkamp's reconstruction*: $f_2 = F_{B\text{-}FDK}\, g_2$

e) $f = S\ f_1 + f_2$

One easily checks that the property $X f_1 = g$ on ∂B is satisfied exactly (within discretization errors) by the image $f_1 = X_B^*(g - X f_{reb}) + f_{reb}$ used in this combined algorithm. This guarantees the quasi-exactness of the algorithm. The advantage over ZB is that this f_1 image is a good approximation of the original object, and therefore the correction term f_2, and the associated discretization errors, are

---

[1] Tam's window B is the region bounded, in the detector, by the cone-beam projection of the upper and lower turns of the helix [11,13].

smaller. The most time consuming steps in the combined algorithm are the calculation of g_1 (step c) and the Feldkamp's step (d).

For the results presented in the following section, we have implemented a simplified version of the algorithm in which we skip step b) and simply use $f_1 = f_{reb}$ instead of $f_1 = X_B^*(g - X f_{reb}) + f_{reb}$. This approximation is motivated by the fact that the image f_{reb} obtained by rebinning may be already sufficiently accurate to guarantee that $X f_{reb} \approx g$ along ∂B. A comparison between the approximate and exact versions of the combined method will be presented at the conference.

RESULTS

We have evaluated the performance of the combined algorithm using simulated data for a head phantom similar to that used by S. Schaller [5]. The phantom is contained in a sphere of radius 100 mm. Data were simulated for a helical path with 1.5 turns and a pitch P=108 mm. The radius of the helix was R=400 mm and there were 1200 vertices per helix turn. The first and last vertices were at locations z=±81 mm. As defined, the helix did not extend over the whole axial extent of the phantom, and we are therefore dealing with a *long object problem*.

Data were simulated on a virtual rectangular detector located at a distance D=400 mm from the cone vertex, i.e. at the isocenter. The detector pixel size was 0.5 x 0.5 mm. There were 400 channels and 200 detector rows.

The angle between the optimal rebinning tilted plane and the transaxial plane was η=3.4 degrees, and the planar rebinning algorithm used the central 134 rows of the detector (maximum cone angle 4.8 degrees). The maximum distance between a ray used for rebinning and the corresponding tilted plane was 2.5 mm, to be compared with 7.8 mm with the SSRB method (η=0). The difference between these two figures illustrates the benefit expected from the planar rebinning algorithms. For all reconstructions, the ramp filter was apodized with a Hamming window cut-off at the Nyquist frequency.

Figures 2, 3 and 4 show reconstructions on a grid of 400x400x200 cubic voxels of 0.5 mm, displayed with a gray scale in the range [1.0,1.1]. The results illustrate the important improvement in image quality obtained by rebinning on tilted planes instead of transaxial planes in SSRB [9,14,1,3]. A reconstruction using optimal non-planar rebinning surfaces (not shown) was practically identical to the reconstruction using tilted planes, as could be expected from figure 1.

The results in figures 2,3,4 also demonstrate that the artefacts observed with the rebinning algorithm are largely suppressed by the combined rebinning-ZB algorithm introduced in this paper. These data also confirm that the short-scan helical Feldkamp (FDK) algorithm (bottom right images) is not superior to planar rebinning, even though FDK uses a "true" cone-beam backprojection.

Figure 5 shows a comparison of the new combined algorithm with the ZB method. Some improvement is observed, especially in the longitudinal section where the streak artefacts caused by the 4 disks are suppressed by the combined method. On the other hand a new artefact is observed around the dark ellipse in the transaxial section. Note that all these results have been obtained with the simplified version of the combined method, as described in the previous section.

A more detailed study using data obtained by reprojecting a high resolution CT scan will be presented at the conference.

REFERENCES

[1] Bruder H, Kachelriess M, Schaller S, Stierstorfer K and Flohr T 2000 Single-slice rebinning reconstruction in spiral cone-beam computetomography, IEEE Trans. Med. Imag.**19** 873-887

[2] Heuscher D 1999 Helical cone beam scans using oblique 2D surface reconstruction. In International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, June 23-26 1999, Egmond aan Zee, The Netherlands, 204-207.

[3] Kachelriess M, Schaller S, and Kalender W 2000 Advanced Single-Slice Rebinning in Cone-Beam Spiral CT. Medical Physics 27 754-772

[4] Defrise M, Noo F and Kudo H 2000 A solution to the long object problem in helical cone-beam tomography Phys Med Biol **45** 623-643

[5] Schaller S 1998 Practical image reconstruction for cone-beam tomography PhD thesis, University of Erlangen

[6] Kudo H, Noo F, Defrise M 1998 Cone-beam filtered-backprojection algorithm for truncated helical data Phys Med Biol **43** 2885-2909

[7] Schaller S, Noo F, Sauer F, Tam K C, Lauritsch G, Flohr T 2000 Exact Radon rebinning algorithm for the long object problem in helical cone-beam CT, IEEE Trans. Med. Imag. **19** 361-375

[8] Kudo H, Noo F and Defrise M 2000 Quasi-exact filtered-backprojection algorithm for long-object problem in helical cone-beam tomography IEEE Trans. Med. Imag. **19** 902-921

[9] Larson G L, Ruth C C, and Crawford C R 1998 Nutating slice CT image reconstruction. Patent Application WO 98/44847.

[10] Noo F, Defrise M and Clackdoyle R 1999 Single-slice rebinning method for helical cone-beam CT Phys. Med. Biol. **44** 561-570

[11] Tam K C, Samarasekera S and Sauer F 1998 Exact cone-beam CT with a spiral scan. Phys Med Biol **43** 1015-1024

[12] Tam K C 2000: Exact local regions-of-interest reconstruction in spiral cone-beam filtered-backprojection CT: theory Proc SPIE Med Imag **3979** 506-519

[13] Turbell, H and Danielsson P-E 1998 The PI Method – Non-Redundant Data Capture and Efficient Reconstruction for Helical Cone-Beam CT. Conf. Rec. 1998 IEEE Med. Imag. Conf. (Toronto, Canada)

[14] Turbell, H and Danielsson P-E 1999 An improved PI-method for reconstruction from helical cone-beam projections. Conf. Rec. 1999 IEEE Med. Imag. Conf. (Seattle, WA)

[15] Turbell, H 2001 Cone-beam reconstruction using filtered-backprojection PhD thesis, University of Linkoping

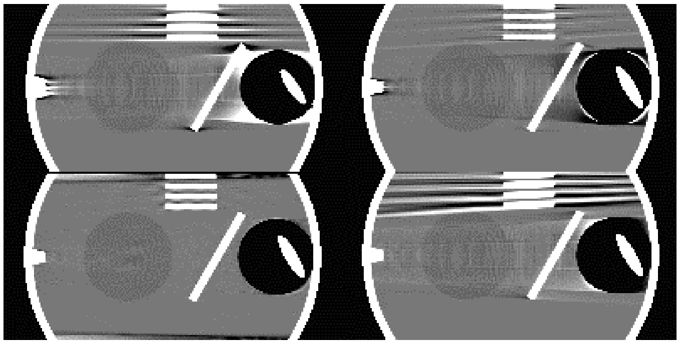


Figure 2. Central longitudinal section x=0. Gray scale [1.0,1.1]. Top left: SSRB, top right: planar rebinning, bottom left: combined rebinning+ZB, bottom right: short-scan FDK.

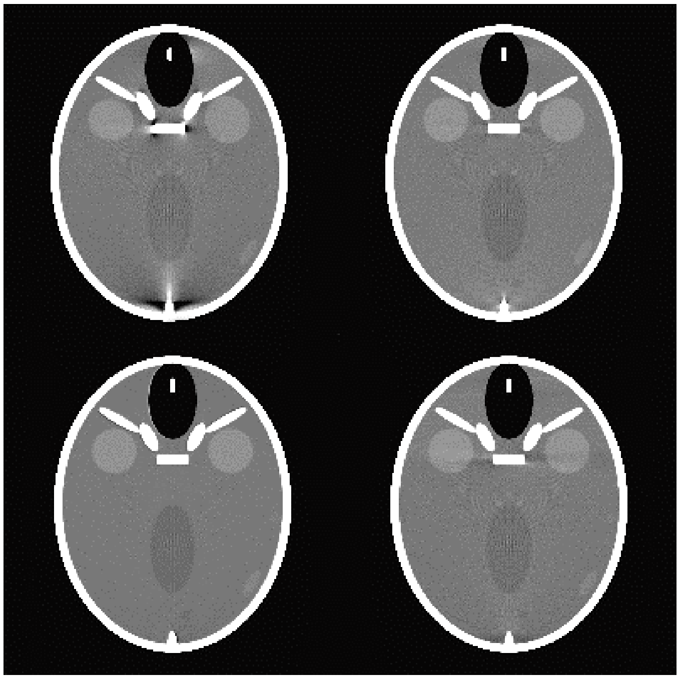


Figure 3. Central transaxial section z=0. Gray scale [1.0,1.1]. Top left: SSRB, top right: planar rebinning, bottom left: combined rebinning+ZB, bottom right: short-scan FDK.

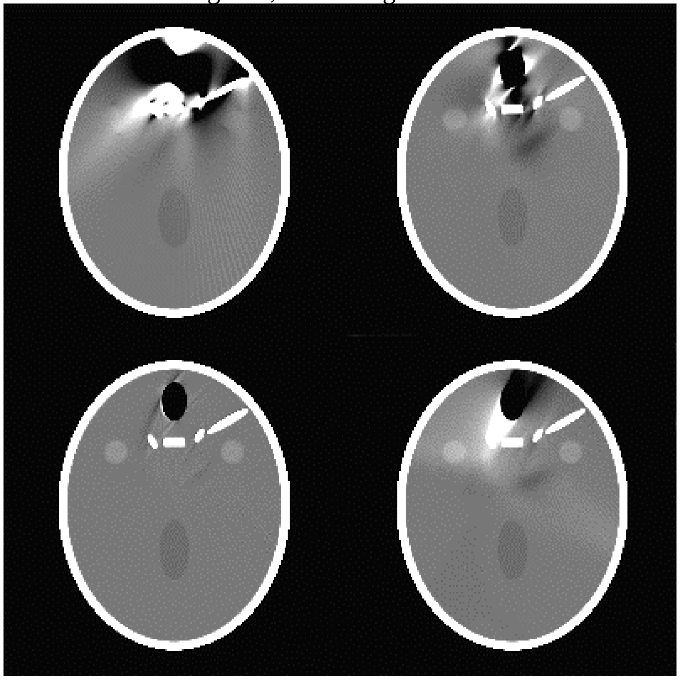


Figure 4. Transaxial section z=19 mm. Gray scale [1.0,1.1]. Top left: SSRB, top right: planar rebinning, bottom left: combined rebinning+ZB, bottom right: short-scan FDK.

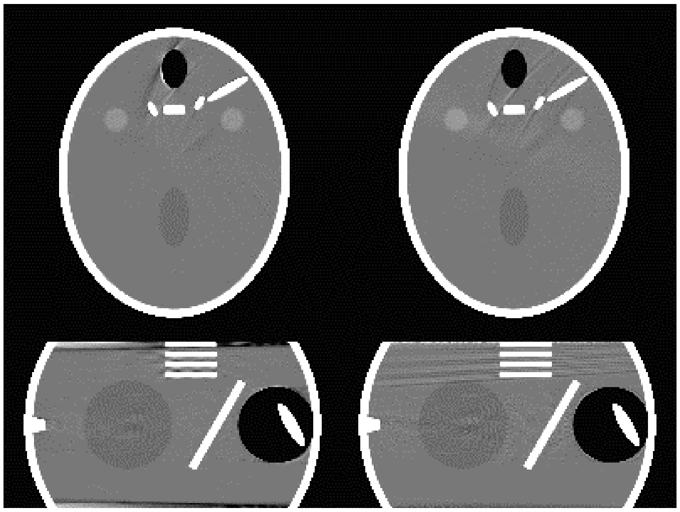


Figure 5. Comparison between the combined method (left column) and the ZB method (right column) for a transaxial section z=19 mm and for a longitudinal section (x=0). Gray scale [1.0,1.1].

The PI-FAST Method for Approximate Helical Cone-Beam Reconstruction

Henrik Turbell and Per-Erik Danielsson
Image Processing Lab, Department of Electrical Engineering
Linköping University, SE-581 83, Sweden
`turbell@isy.liu.se` `ped@isy.liu.se`

Abstract— **The original PI-method is a non-exact method for 3D-reconstruction using cone-beam projections acquired from a helical source trajectory. In the new PI-FAST method, our aim is to reduce the artifacts while keeping the algorithmic simplicity. The detector is still bounded by the Tam-window, which makes the data capture complete and non-redundant. Also, the filtering step still consists of 1D ramp-filtering but comprises the following novelty. Using the principles of fast backprojection we are able to do the backprojection in two steps interleaved with by rampfiltering. This unusual order is shown to be advantageous by exploiting frequency-distance coherence in projection data. The first backprojection step lumps together projection data over shorter angular intervals along directions which correspond to various velocities in projection space. By rampfiltering these data instead of the original projection data we obtain less unwanted interaction in the z-direction of the volume and a substantial improvement in image quality.**

I. TWO-STEP FILTERED BACKPROJECTION USING LINKS

Filtered backprojection of 2D parallel data, $p(\theta, t)$, can be expressed as filtering with the ramp-filter $g(t)$

$$\tilde{p}(\theta, t) = p(\theta, t) * g(t) \tag{1}$$

followed by backprojection

$$f(x, y) = \sum_{i=0}^{N_\theta - 1} \tilde{p}(\theta_i, y\cos\theta_i - x\sin\theta_i) \tag{2}$$

over a projection angle interval of length $\theta_{N_\theta} - \theta_0 = \pi$.

It has been shown to be efficient to perform the backprojection summation (2) recursively [1], [2] in $\log_2 N_\theta$ steps. In [3] the process is descibed as iterative summing of filtered projection values along a sinusoid in the sinogram. We will utilize a simplified version of this fast backprojection technique that performs the calculations in two steps. First, small intermediate summations along *linear* segments approximating the sinusoid are calculated and stored. We call the linear segments *links*. In the second step the values of the links are combined in building over π a sinusoid to yield the final result. The reason for the computational gain is that the value of a link will be used for several pixel values in the second step. It should be noted, however, that this potential speedup is not our main reason for using the two-step backprojection. More interestingly we will show that we may postpone the ramp filtering till the first step of the backprojection has been performed.

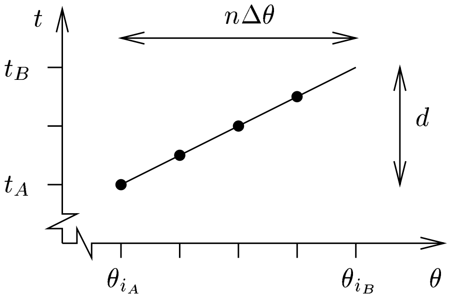


Fig. 1. The link value is a summation of projection values along the link.

Figure 1 shows an arbitrary link from (θ_{i_A}, t_A) to (θ_{i_B}, t_B) where $i_B = i_A + n$ and $t_B = t_A + d$. The link value $\tilde{I}(\theta_{i_A}, t_A, d)$ is calculated as

$$\tilde{I}(\theta_{i_A}, t_A, d) = \sum_{i=i_A}^{i_A+n-1} \tilde{p}(\theta_i, t_A + d\frac{i - i_A}{n}) \tag{3}$$

The tilde on $\tilde{I}$ indicates that the link value is a summation of filtered projection data. Although the end points (θ_{i_A}, t_A) and (θ_{i_B}, t_B) preferably are choosen as sample points in the sinogram, the intermediate values along the link require an implicit 1D-interpolation each in (3). The link length n is constant in the algorithm. It has to be short enough for the line elements not to deviate significantly from the true curved sinusoid. A reasonable choice [3] is $n \approx \sqrt{N_\theta}$.

The endpoints of a link corresponds to a ray each in the image domain. The intersection (x, y) of the two rays (θ_{i_A}, t_A) and (θ_{i_B}, t_B) is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -\sin\theta_{i_A} & \cos\theta_{i_A} \\ -\sin\theta_{i_B} & \cos\theta_{i_B} \end{pmatrix}^{-1} \begin{pmatrix} t_A \\ t_B \end{pmatrix} \tag{4}$$

A pixel value is computed by summing the values of the links that build up the sinusoid of the pixel. See Figure 2. For each step l along the θ-axis, we use values of the four links surrounding the sinusoid segment to interpolate the contribution to the pixel value. For notational simplicity we denote the four link-values $\tilde{J}_l$, $\tilde{K}_l$, $\tilde{L}_l$, and $\tilde{M}_l$. The pixel value is obtained as

$$f(x, y) = \sum_{l=1}^{N_\theta/n} \Big(w_l \left(w_{l+1}\tilde{J}_l + (1 - w_{l+1})\tilde{K}_l\right) + (1 - w_l)\left(w_{l+1}\tilde{L}_l + (1 - w_{l+1})\tilde{M}_l\right)\Big) \tag{5}$$

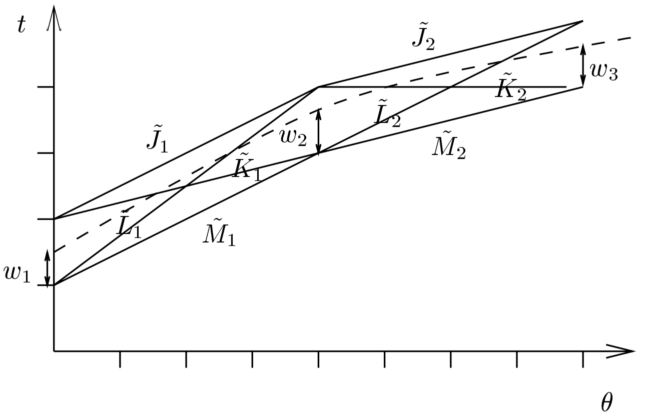


Fig. 2. A pixel value is the sum of interpolations between groups of four links.

where the interpolation weights w_l depend on the t-distance between the sinusoid and the link starting point.

By combining equations (1) and (3), the link value can be written as

$$\tilde{I}(\theta_{i_A}, t_A, d) = \sum_{i=i_A}^{i_A+n-1} \left(p(\theta_i, t_A + d\frac{i - i_A}{n}) * g(t) \right) =$$
$$\left(\sum_{i=i_A}^{i_A+n-1} p(\theta_i, t_A + d\frac{i - i_A}{n}) \right) * g(t) = I(\theta_{i_A}, t_A, d) * g(t) \tag{6}$$

where $I(\theta_{i_A}, t_A, d)$ naturally denotes the link value calculated from unfiltered projection data. The change of order between summation and convolution is possible since the links are linear and equidistantly spaced along the t-axis. Equation (6) tells us that we may indeed switch the order between rampfiltering and the first backprojection step, the link computation. The right hand side convolution should be seen as a one-dimensional convolution along the t-axis where all links of constant θ_{i_A} and d participate in one filtering event.

II. THE ORIGINAL PI-METHOD

The original PI-method [4] is an approximate reconstruction method for the helical cone-beam geometry in Figure 3. The source moves along a helical trajectory of radius R and pitch P around the z-axis. We define the pitch as the distance, measured in an arbitrary length unit, between two consecutive turns of the helix. The effective area of the 2D detector is limited in height to a beam window between two consecutive turns of the source helix. See Figure 3(b). The physical construction and geometry of the detector may vary as long as all measurements are confined to this window, the PI-window. The cone-beam projections $p^C(\beta, \gamma, s)$ and the corresponding rays inside the PI-window are parameterized so that β is the projection angle, γ is the fan-angle, and $2s$ is the detector height coordinate on a detector wrapped onto the source trajectory cylinder. The rows of this detector are aligned with the curved helix. The top boundary of the PI-window is therefore at height $s = P/4$ and the bottom boundary at $s = -P/4$.

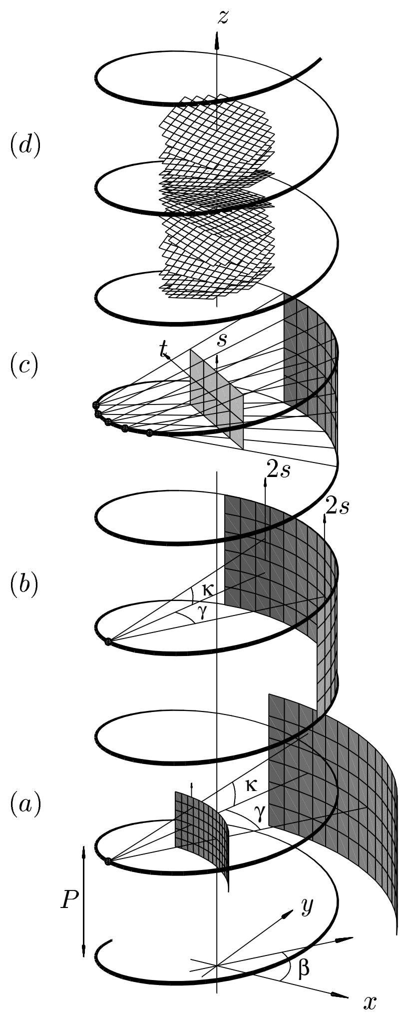


Fig. 3. Geometry for helical cone-beam acquisition. (a) Example of focus-centered detector. (b) A detector on the PI-window. (c) A pseudo-parallel beam together with the planar virtual detector. (d) Nutating PI-surfaces.

The first step of the PI-method is a re-sampling step from cone-beams to pseudo-parallel beams according to

$$p(\theta, t, s) = p^C(\beta, \gamma, s) = p^C(\theta - \arcsin\frac{t}{R}, \arcsin\frac{t}{R}, s) \tag{7}$$

Note that the detector height coordinate s is left unchanged. The re-sampling can thus be performed row-by-row and is then identical to the well-known 2D procedure known as parallel rebinning. We observe that the resulting beam shown in Figure 3(c) is divergent when seen from the side, but parallel when seen from along the rotation axis. Also shown is a virtual planar detector orthogonal to the projection direction placed on the rotation axis. The pseudo parallel beam intersects this planar detector on a perfectly Cartesian grid contained in a rectangle. The horizontal rows of this detector have constant value of s.

The PI-method then proceeds with a pre-weighting with the cosine of the cone-angle for each ray. This is followed by 1D ramp-filtering of data along the horisontal lines of the virtual planar detector. The voxel values are finally obtained by 3D backprojection of the filtered data.

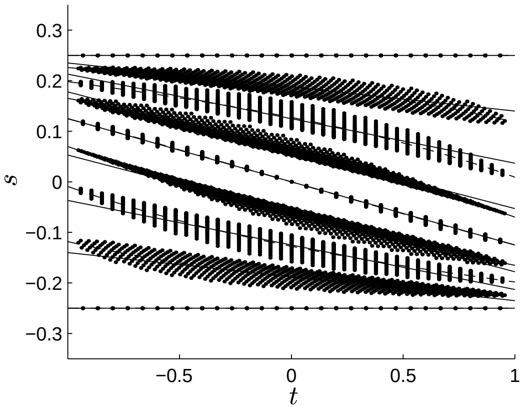


Fig. 4. Point on a PI-surface of a helix of pitch $P = 1$ as projected onto the planar virtual detector at different projection angles.

III. APPROXIMATIONS IN SOME PREVIOUS HELICAL CONE-BEAM ALGORITHMS

In order to track the inexactness of the PI-method, we will now study the set of object points entering the rectangular window of the planar detector at the same time. These points lie on a surface, which we will call a PI-surface. It can be shown that each object point belongs to one and only one PI-surface [5]. The complete set of PI-surfaces has a nutation around the rotation axis and fills up the complete volume to be reconstructed. See Figure 3(d).

We assume that the projection system is rotating upwards in which case the points of a PI-surface enter the rectangular detector window on a perfect line on the upper boundary simultaneously. See Figure 4. As the rotation of the projection system continues, the projection of the PI-surface moves downwards on the detector. Unfortunately, it starts immediately to deviate from the line shape and occupies instead an elongated area with a non-horizontal mid-line. However, after a rotation of exactly 180°, all points on the PI-surface are again lined up horizontally, now along the lower boundary of the window, and exit the rectangular window simultaneously.

The algorithms found in [6], [7], [4] utilize the observation that the object points on nutating surfaces are concentrated along slanted lines in-between entrance and exit. The 1D ramp-filtering in these algorithms is performed along these slanted lines or curves. The filtered data is then backprojected in two dimensions onto the nutating surface or in three dimensions into the volume. The fact that some of the object points on the nutating surface are projected above such a slanted line and some below is not handled in the filtering step of these algorithms, but will be addressed by the new PI-FAST method.

Consider a neighbourhood of points on a PI-surface. They are projected onto a neighbourhood on the detector. Furthermore, they have similar velocities in the t-direction on the detector. Unfortunately, the projection values of this detector neighbourhood are contaminated by object points belonging to other PI-surfaces. However, on the detector, the projection of these contaminating points have different velocities in the t-direction. Only links of the same slope $\frac{d}{n}$ have the same projected t-velocity [8]. By not filtering links of different slopes together, PI-FAST aims to decrease unwanted interaction between neighbouring PI-surfaces during the filtering event.

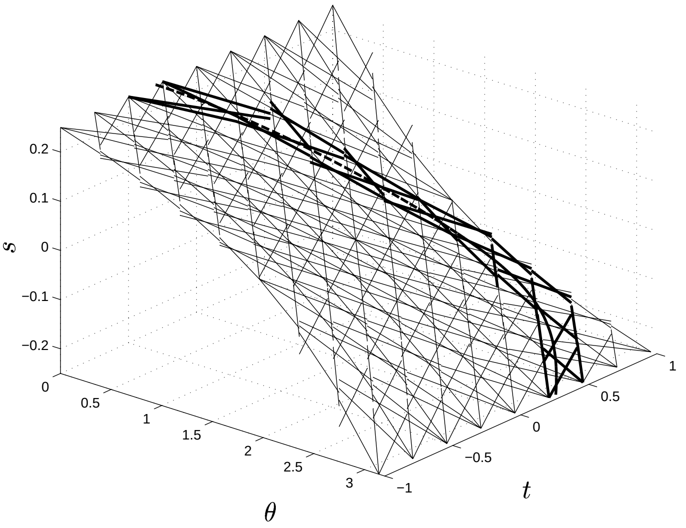


Fig. 5. The links needed for a PI-surface of a helix of pitch $P = 1$. Note that the number of links in both θ- and t-direction is small compared to a typical case to simplify the illustration. This makes the links sprawl out in the s-direction more than in a typical case.

IV. THE PI-FAST RECONSTRUCTION METHOD

For each PI-surface let us construct a complete set of link values $I(\theta, t, d)$. The links are positioned in the 3D (θ, t, s) projection space. The (θ, t)-coordinates of the links are identical to the 2D case since we may regard the reconstruction as 2D when seen from along the rotation axis. The s-coordinates require some further analysis.

Given a PI-surface and the values of θ_{i_A}, t_A, θ_{i_B}, and t_B we use (4) to compute the (x, y)-coordinate of the point on the PI-surface corresponding to the link. Knowing (x, y), the z-coordinate of the point is uniquely given by the equation of the PI-surface. The link endpoint coordinates s_A and s_B can then be calculated by projecting the point (x, y, z) onto the virtual planar detector from projection angles θ_{i_A} and θ_{i_B} respectively. Analytical expressions of these geometry calculations are found in [9].

The complete set of links for a PI-surface will approximately follow the projection tracks of the points in the PI-surface from their movements in the projection space. See Figure 5. At the entrance projection angle $\theta_{i_{\text{in}}}$ of the PI-surface, the links will all start at $s = P/4$ and at the final projection angle $\theta_{i_{\text{in}}} + \pi$ the links will all end at $s = -P/4$. In-between these two angles, the links will spread out in the s-direction.

When the end-points of the links have been calculated it is possible to calculate each link value along the line segment between the link end point as

$$I(\theta_{i_A}, t_A, d) = \sum_{i=i_A}^{i_A+n-1} p(\theta_i, t_A + d\frac{i-i_A}{n}, s_A + (s_B - s_A)\frac{i-i_A}{n}) \tag{8}$$

The implicit 1D interpolation in the t-direction for the 2D case in (3) has here become an implicit 2D interpolation on each (t, s)-plane. If we insert the re-sampling step in (7) into (8), we may construct the link values directly from the cone-beam data $p^C(\beta, \gamma, s)$ without the intermediate re-sampling step in (7). There is furthermore no need to initially re-sample the original cone-beam data to rows of constant s since the mapping between the actual physical detector sampling points and the (β, γ, s)-system also can be inserted into (8). The sampling points of the terms of the summation in (8) will then not coincide with the projection data sampling points even in the projection angle direction. An implicit 3D-interpolation is thus necessary for each term in (8). To avoid aliasing, it may be necessary to have a smaller step size in the summation along the θ-direction than the sampling distance step used in (8).

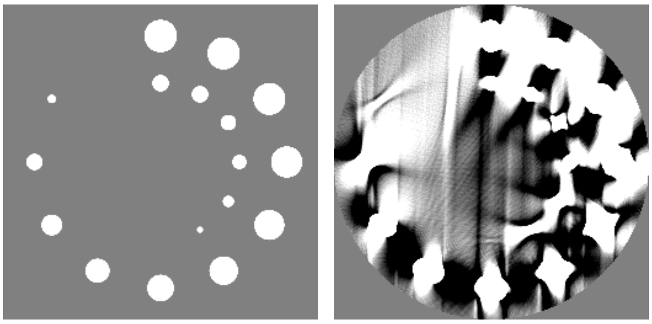
Once the link values of a PI-surface have been computed they may be ramp-filtered as described in the right hand side of (6). The filtered links are then combined into pixel values according to (5). This backprojection step is identical to the 2D case, but will nevertheless perform a 3D backprojection. It does not have to consider the link positions in the s-direction, since this information is already taken care of in the first backprojection step in (8). The pixels on the PI-surfaces are finally interpolated onto a suitable final 3D sampling grid, such as a cubic grid, for presentation and analysis. Note that this interpolation only is performed along the z-direction, since the pixels on the PI-surfaces already are positioned on a Cartesian (x, y)-grid.

We summarize the computation steps in the PI-FAST algorithm as

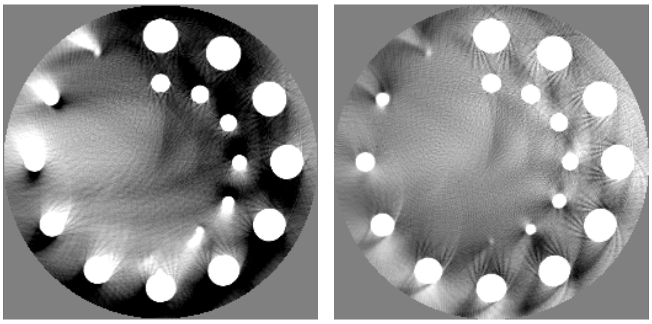
```
Pre-weight the projection data with the cosine of the cone-angle of each ray.
for all PI-surfaces do
    Compute the link values according to (8) using the appropriate mappings to the acquisition geometry of the original data.
    Ramp-filter the link values as in (6).
    Compute the pixel values of the PI-surface using (5).
end for
Resample the pixels of the PI-surfaces in the $z$-direction onto a Cartesian grid.
```

V. EXPERIMENTAL RESULTS

Figure 6 shows the reconstruction results of a phantom consisting of homogeneous spheres. For comparison the result of a so called multi-slice method with 2D backprojection [10] found in present 4-row scanners is included. Noise-free projection data was generated on a 64-row detector with a fan-angle of ±30° and cone-angle of ±7.13°. Further experiments and details are found in [9].



(a) Phantom (b) Multi-slice, $\sigma = 0.134$



(c) PI-method, $\sigma = 0.050$ (d) PI-FAST, $\sigma = 0.034$

Fig. 6. Reconstruction results of the sphere clock phantom [9] together with the root mean square error. Radius $R = 2.0$ length units, pitch $P = 1.0$ length units, $N_\theta = 256$ projections per half turn, fan-angle $\pm 30°$, cone-angle $\pm 7.13°$ on a 64 row detector with 255 elements per row. Slice reconstructed on $256 \times 256 \times 1$ voxels of side $\frac{1}{128}$ length unit. Greyscale interval $[-0.05, 0.05]$.

REFERENCES

[1] Stefan Nilsson, *Application of Fast Backprojection Techniques for Some Inverse Problems of Integral Geometry*, Ph.D. thesis, Department of Mathematics, Linköping University, 1997, No. 499.

[2] Martin L. Brady, “A Fast Discrete Approximation Algorithm for the Radon Transform,” *SIAM Journal of Computing*, vol. 27, no. 1, pp. 107–119, 1998.

[3] Per-Erik Danielsson and Malin Ingerhed, “Backprojection in $O(N^2 \log N)$ time,” in *IEEE Medical Imaging, Nov 13–15, 1997, Alberquerque, New Mexico, USA*, 1997.

[4] Henrik Turbell and Per-Erik Danielsson, “Helical Cone-Beam Tomography,” *International Journal of Imaging Systems and Technology*, vol. 11, no. 1, pp. 91–100, 2000.

[5] M. Defrise, F. Noo, and H. Kudo, “A solution to the long object problem in helical CB tomography,” *Physics in Medicine and Biology*, vol. 45, pp. 623–643, 2000.

[6] G. L. Larson, C. C. Ruth, and C. R. Crawford, “Nutating slice CT image reconstruction,” Patent Application WO 98/44847, 1998.

[7] M. Kachelriess, S. Schaller, and W. Kalender, “Advanced Single-Slice Rebinning in Cone-Beam Spiral CT,” *Medical Physics*, vol. 27, no. 4, pp. 754–772, 2000.

[8] P. Edholm, R. M. Lewitt, and B. Lindholm, “Novel Properties of the Fourier Decomposition of the Sinogram,” in *Int. Workshop on Physics and Engineeing of Computerized Multidimensional Imaging and Processing, Proc. of the SPIE*, 1986, vol. 671, pp. 8–18.

[9] Henrik Turbell, *Cone-Beam Reconstruction Using Filtered Backprojection*, Ph.D. thesis, Linköping University, 2001, ISBN 91-7219-919-9.

[10] K Taguchi and H Aradate, “Algorithm for image reconstruction in multi-slice helical CT,” *Medical Physics*, vol. 25, no. 4, pp. 550–561, 1998, ISSN 0094-2405.

Why do Patients and Their Care-Givers Need 3D and 4D Imaging?
Thomas F. Budinger M.D., Ph.D.

Imagine you finally decide to learn the status and prognosis of your lower back pain. From your MRI study the local orthopedic or neuro-surgeon makes an 8 cm midline incision over your lower vertebrae. While looking at the X-ray view-box in the operating room, as well as what can be seen of your spinal cord and the disc after removing some bony bridges, the surgeon retracts your cord and chisels out part or all of your disc. He then adds some shims or screws to help support your vertebrae. This scenario could apply to cervical vertebrae or knee surgery. Having benefited from these repairs you later encounter an episode of robotic surgery of your clogged coronaries, or a blind biopsy of your prostate because your PSA is high or your breast because your mother has breast cancer and somebody saw something suspicious on a mammogram.

In all of these procedures few physicians would claim they had confidence that what they were able to see, allowed them to do their best for you. In fact, most would complain that the lack of anatomical landmarks and adequate surgical exposure limit diminished their effectiveness and therefore the safety and efficacy of their procedures.

Therefore, how can fully 3D imaging scientists practically change the practice of medicine? The answer lies in a few realistic examples from which one can extrapolate to many surgical and non-surgical procedures that promise to improve the quality and effectiveness of medical care.

Now imagine the surgeon has made perhaps three 1 cm incisions in your back and through a stereo projection system in the operating room, the surgeon can visualize your vertebrae in virtual reality from a previously obtained Spiral (helical) CT. The surgeon can also visualize the spatial position of the operating instruments relative to the bony architecture of your spine. At his oral command he can move your virtual reality anatomy to improve viewpoints and can call on additional MRI or even SPECT or PET bone blood flow information to aid in the surgical decisions. The surgeon's tools move around unbroken bony bridges and the remodeling operation is nearly bloodless and nearly atraumatic. A similar scenario can be portrayed for cardiac surgery.

In short, 3D imaging is a method of achieving surgical exposure without surgery. Why cannot this be done by in-situ endoscopy without prior imaging? Because the 3D relations of a wide field of view is needed for safe and accurate surgery.

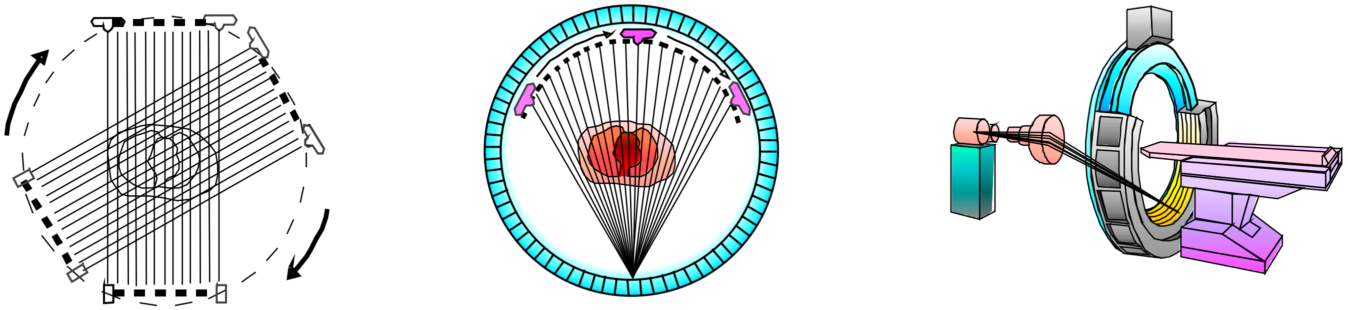
In addition to improved visualization of the spatial relationships, 3D data acquisition methods and reconstruction methods enable the realization of entirely new areas of medical imaging. One example is the recent discovery that multidetector helical CT can characterize coronary atheroma with clinically useful diagnostic resolution. The breakthrough here is a direct result of speed of volume coverage that in effect negates motion-based blurring. The volume coverage provides adequate data acquisition during a data collection sequence synchronized with the EKG window.

Other examples of 3D and 4D methods are dynamic MRI for studying the motion of bony joints and of the heart. Of particular interest is our proposed method for evaluating the motion of the human vertebral column. Even methods of in-vivo microscopy, as for example 50 μm resolution imaging, of the human cortex require 4D acquisition methods and motion compensation algorithms. As shown in the accompanying figures, the evolution of instruments continues to make demands on scientists who can create algorithms to optimize the data collection from these new geometries. Examples of these applications alert us to the importance for 4D algorithms that encompass the following applications:

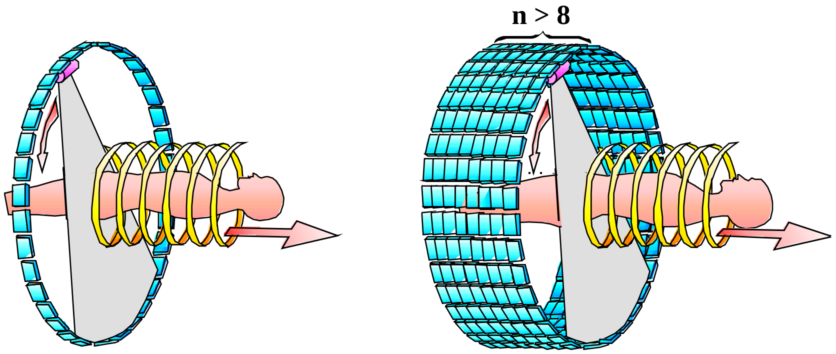
- Motion compensation and motion imaging in SPECT and PET
- Kinetic parameter extraction in dynamic SPECT
- Tensor tomography
- Diffusion tensor imaging
- Cardiac motion parametric imaging
- Motion parametric imaging
- Shoe fit optimization by dynamic MRI.

Examples of current medical applications and horizons will be given in this presentation.

Evolution of X-ray CT

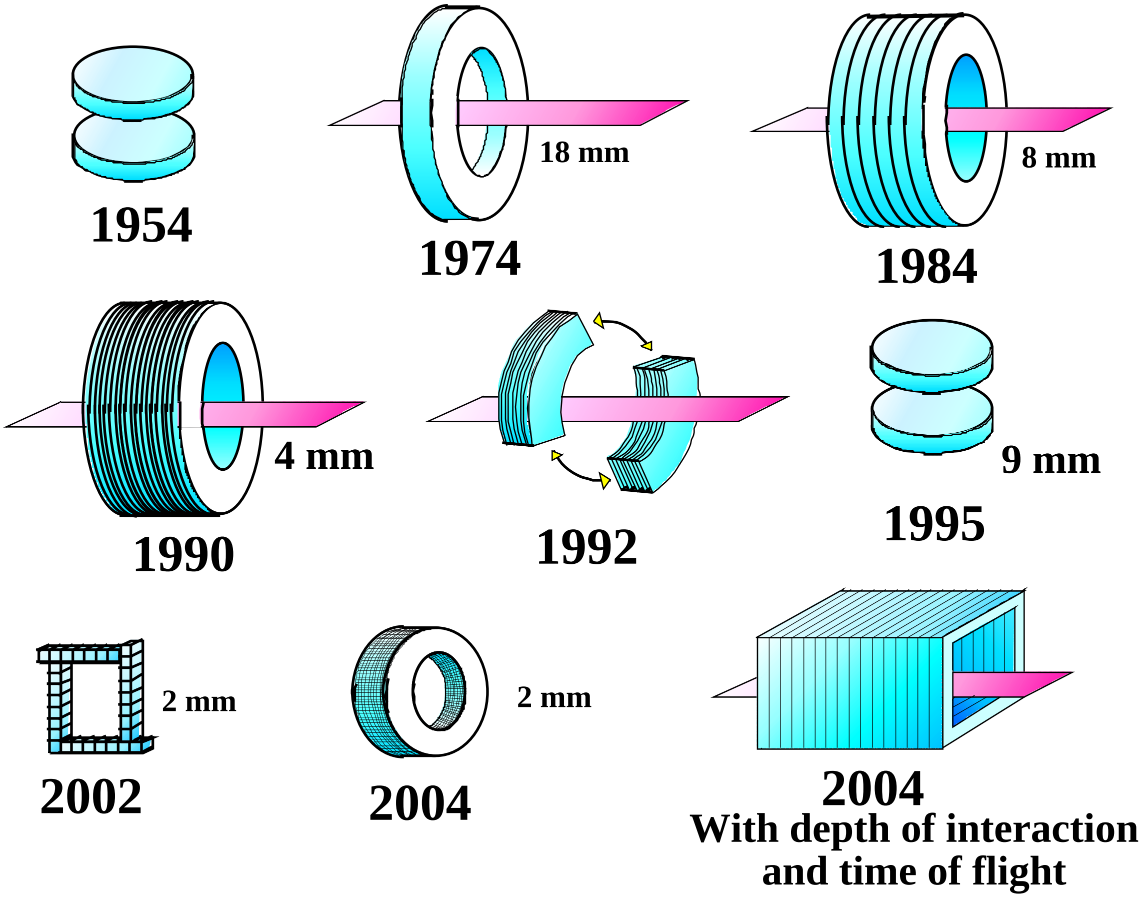


1st - generation CT **Later - generation CT** **Electron Beam CT**

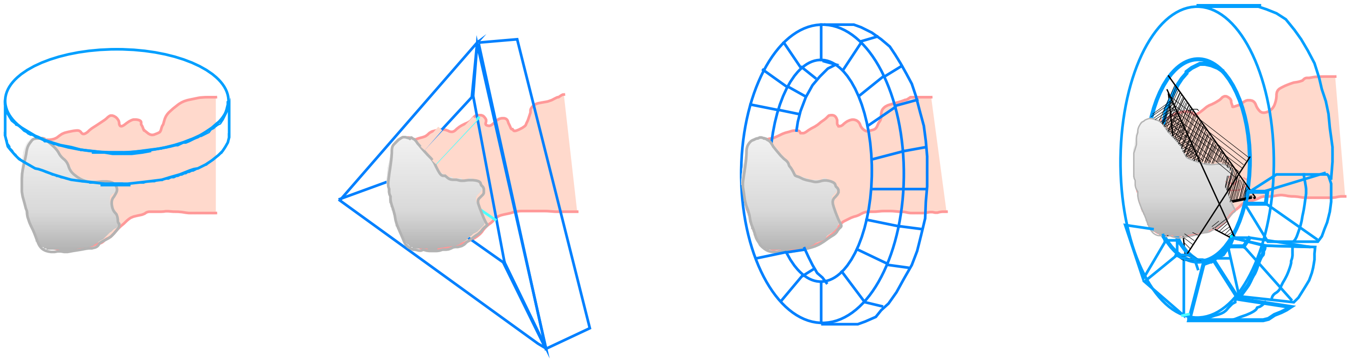


Spiral CT **Multidetector CT**

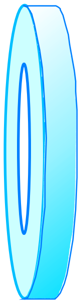
PET Developments



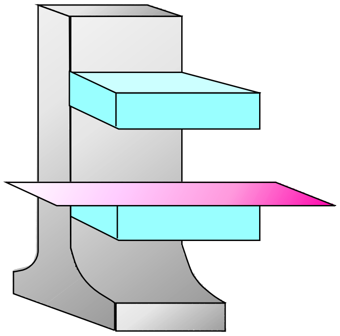
SPECT Developments



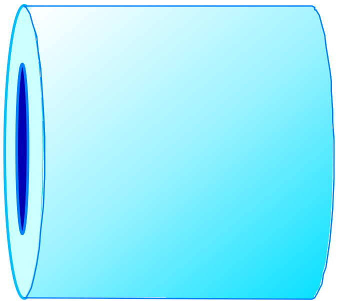
Magnetic Resonance Developments



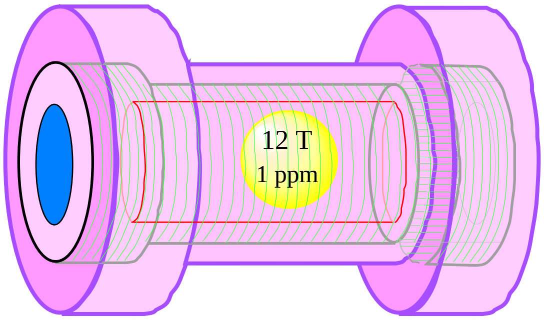
0.04 T



< 0.5 T



1.5 - 8 T



12 T

Image Reconstruction Algorithm for a SPECT System with a Convergent Rotating Slat Collimator

Gengsheng L. Zeng
University of Utah, Salt Lake City, UT 84108, USA
Daniel Gagnon
Marconi Medical Systems, Cleveland, OH 44143, USA

Abstract — In this paper we suggest the use of a convergent rotating slat collimator in a SPECT system that contains strip-shaped CdZnTe detectors. This imaging device is able to provide high spatial and energy resolution for small animal imaging. A novel design method and reconstruction technique are proposed for use in this system.

I. Introduction

A parallel, slat collimator shown in Figure 1 was proposed in 1975 by Keyes [1]. This collimator can be mounted on a gamma camera for SPECT (single photon emission computed tomography) imaging. This collimator does not consist of holes. Instead, it is made of parallel plates. One advantage of a slat collimator over a hole collimator is its better geometric efficiency [4]-[7]. However, this slat collimator can only measure a one-dimensional profile, and therefore cannot be used to directly measure a two dimensional image. A planar image can only be reconstructed using one-dimensional profiles. Three-dimensional imaging requires two motions: Collimator rotation and detector rotation as illustrated in Figure 1.

The idea of using a rotating collimator was later adopted in applications of semiconductor detectors [2][3] and NaI(Tl) Anger cameras [4]-[7].

In this paper we propose using a convergent rotating slat collimator for SPECT, and the development of a reconstruction algorithm.

This research is directed towards building a small animal SPECT imager with both high spatial resolution and excellent energy resolution. CdZnTe semiconductor chips have been used to build a gamma ray detector [8][9]. The detector is strip-shaped, like the detector depicted in Figure 1. However, the collimation slats are arranged so that they have a focal line in front of the detector, as illustrated in Figure 2. This design enables us to achieve a sub-millimeter spatial resolution. The distance between the focal line and the detector is referred to as the collimator's *focal length*. The focal length will be determined by the object size. A smaller object requires a shorter focal length.

II. Methods

A. Projection data

A common feature of slat collimator is that the measured projection is a weighted planar integral rather than a line integral.

Let's first assume that the detector is infinitely narrow. We have an approximate "point detector" between adjacent collimator slats. The planar integral that this point detector measures is weighted by a factor of $1/r$, where r is the distance between the point detector and the point-object.

If the detector is not narrow, we can treat it as rows of independent point detectors, each one having its own weighting factor of $1/r_i$. The integrated effect of each factor $1/r_i$ yields a distance dependent weighting factor in the planar integral. The measured projection value is the sum of all measurements from a row of individual "point-detectors." The goal behind this work is to develop a method to remove the $1/r_i$ weighting factor from each "point-detector."

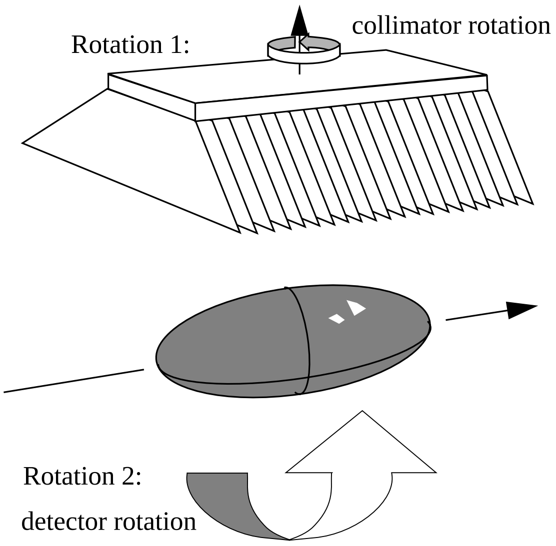


Figure 1. A Parallel rotating slat collimator

B. Introduction of a small tilt angle

In order to remove the $1/r_i$ weighting factor from the projection data, we modified the convergent slat collimator shown in Figure 2 by tilting every slat a small angle δ, as indicated by the broken lines in Figure 3.

C. Algorithm

Let us consider the local coordinate system shown in Figure 4. This local coordinate system is centered at an arbitrary “point-detector” which measures a weighted planar integral of the plane labeled **a**. Since the collimator slats are tilted by a small angle δ, after the collimator rotates 180° (Rotation 1 in Figure 1), another “point-detector” will be rotated to the location where the previous “point-detector” was located. The new “point-detector” measures a weighted planar integral of the plane labeled **a’**. If plane **a** is tilted *up* by a small angle δ, then plane **a’** is tilted *down* by a small angle δ.

The measurement at plane **a** is given as

$$g(\delta) = \iint \frac{1}{r} f(r, \alpha, r\tan\delta)\, r\,dr\,d\alpha \tag{1}$$

and the measurement at plane **a’** is given as

$$g(-\delta) = \iint \frac{1}{r} f(r, \alpha, -r\tan\delta)\, r\,dr\,d\alpha\,. \tag{2}$$

By subtracting Eq. (2) from Eq. (1) and multiplying the difference by a constant $(\cos^2\delta)/(2\delta)$, we have

$$\frac{g(\delta) - g(-\delta)}{2\delta}(\cos^2\delta) \tag{3}$$

$$= \frac{\iint \frac{1}{r} f(r, \alpha, r\tan\delta)\, r\,dr\,d\alpha - \iint \frac{1}{r} f(r, \alpha, -r\tan\delta)\, r\,dr\,d\alpha}{2\delta}.$$

Equation (3) can also be expressed as:

$$\frac{g(\delta) - g(-\delta)}{2\delta}(\cos^2\delta) \tag{4}$$

$$= \iint \frac{1}{r} \frac{f(r, \alpha, r\tan\delta) - f(r, \alpha, -r\tan\delta)}{2\delta}\, r\,dr\,d\alpha\,.$$

The $\frac{f(r, \alpha, r\tan\delta) - f(r, \alpha, -r\tan\delta)}{2\delta}$ portion of Eq. (4) is readily recognized as an approximation of the partial derivative of f, with respect to variable δ, and multiplied by a constant $\frac{r}{\cos^2\delta}$, because

$$\frac{d}{d\delta} r\tan\delta = \frac{r}{\cos^2\delta}\,. \tag{5}$$

Using the local coordinate system presented in Figure 4, and recognizing the fact that the r’*s* in $r/\cos^2\delta$ and $1/r$ cancel out, we have:

$$\iint \frac{\partial}{\partial w} f(r, \alpha, w)\, r\,dr\,d\alpha = \frac{\partial}{\partial w}\iint f(r, \alpha, w)\, r\,dr\,d\alpha \tag{6}$$

and

$$\frac{g(\delta) - g(-\delta)}{2\delta}(\cos^2\delta) \tag{7}$$

$$\approx \iint \frac{\partial}{\partial w} f(r, \alpha, w)\, r\,dr\,d\alpha$$

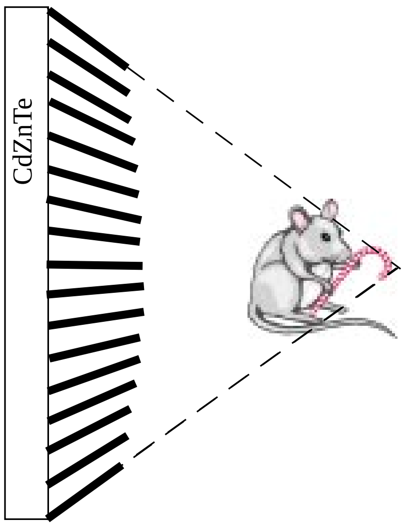


Figure 2. A convergent collimator.

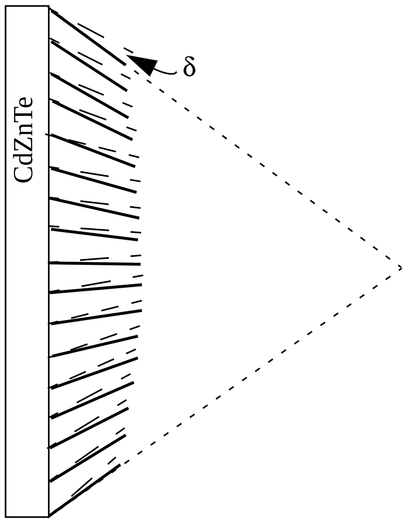


Figure 3. A convergent collimator with a small tilt-angle δ.

$$= \frac{\partial}{\partial w}\iint f(r, \alpha, w) r dr d\alpha .$$

Here, $\iint f(r, \alpha, w) r dr d\alpha$ is the Radon transform of object *f*.

The Radon inversion formula reconstructs the image f by backprojecting the second-order derivative of the Radon transform, i.e., by backprojecting $\frac{\partial^2}{\partial w^2}\iint f(r, \alpha, w) r dr d\alpha$.

The reconstruction procedure is as follows:

i) Measure data at plane **a**, see Eq. (1).

ii) Rotate the collimator 180˚ and measure data at plane **a’**, see Eq. (2).

iii) Take the difference of the above two measurements and normalize it by a constant $(\cos^2\delta)/(2\delta)$, See Eq. (7). [Note: This step could also be the very last step in the reconstruction, because $(\cos^2\delta)/(2\delta)$ is a constant.]

iv) Sort the data according to the detector orientation, then store the data in the Radon space.

v) Take the derivative of the data along the radial direction in the Radon space, obtaining the second-order derivative of the Radon transform.

vi) Backproject the data according to the Radon inversion formula:

$$f(\vec{x}) = -\frac{1}{8\pi}\int_0^{\pi}\int_0^{\pi} p''_{\theta\varphi}(\vec{x} \cdot \vec{\theta}) \sin\theta d\theta d\varphi \tag{8}$$

where $\vec{x}$ is a point in the 3D image space and $\vec{\theta}$ is the direction of the detector, which is the *w* direction in Figure 4. The angle θ is the angle of Rotation 1 indicated in Figure 1 and $\theta = 0°$ corresponds to the position where $\vec{\theta}$ is parallel to the axis of Rotation 2. The angle φ is determined by Rotation 2 and the starting angle is arbitrary. The Radon transform of the object *f* is denoted by *p*, and

$$p'' = \frac{\partial^2}{\partial w^2}\iint f(r, \alpha, w) r dr d\alpha . \tag{9}$$

D. Data sufficiency condition

In order to guarantee a sufficient measurement of the data, we require that every point $\vec{x}$ in the region-of-interest should see a backprojected plane from *ALL* orientations.

Let us define the *focal point* of the collimator to be the point directly in front of the detector center on the focal line. To obtain a sufficient data measurement, the focal point position must have a non-planar trajectory. Figure 5 illustrates two examples that can be used to acquire a complete data set.

III. Discussion

In this paper we outlined the development of a reconstruction method for SPECT imaging utilizing a convergent rotating slat collimator. The method introduced a small tilt-angle for the slats and a differential technique to remove the 1/*r* weighting factor

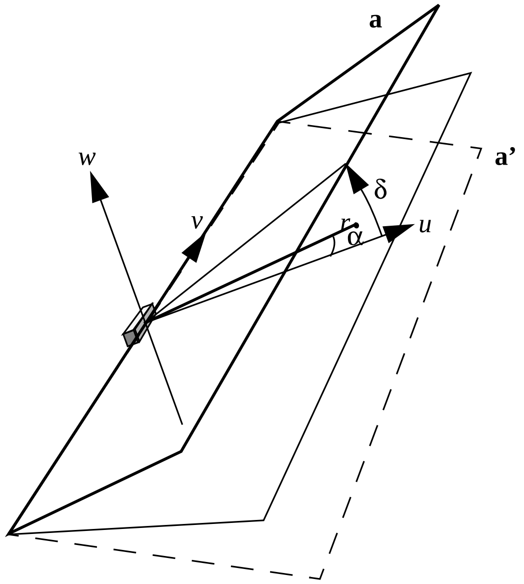


Figure 4. Local coordinate systems: The cylindrical system (r, α, w) and the Cartesian system (u, v, w).

in the projection data, which were then converted into Radon data. The Radon inversion formula was used to reconstruct the image. A data sufficiency condition for this imaging geometry was also discussed.

IV. References

1. Keyes W. I., “The fan-beam gamma camera,” Phys. Med. Biol. 20, 489-491 (1975).
2. Urie M. M., Mauderli W., Fitzgerald L. T., and Williams C. M., “Rotating laminar emission camera with GE-detector,” *IEEE Trans. Nucl. Sci.* 26, 552-558 (1979).
3. Mauderli W., Luthmann R. W., Fitzgerald L. T., Urie M. M., Williams C. M., Tosswill C. H., and Entine G., “A computerized rotating laminar radionuclide camera,” *J. Nucl. Med*. 20, 341-344 (1979).
4. Lodge M. A., Binnie D. M., Flower M. A., and Webb S., “The experimental evaluation of a prototype rotating slat collimator for planar gamma camera imaging,” *Phys. Med. Biol*. 40, 427-448 (1995).
5. Lodge M. A., Webb S., Flower M. A., and Binnie D. M., “A prototype rotating slat collimator for single photon emission computed tomography,” *IEEE Trans. Med. Imaging*, 15, 500-511 (1996).
6. Webb S., Binnie D. M., Flower M. A., and Ott R. J., “Monte Carlo modelling of the performance of a rotating slit-collimator for improved planar gamma-camera imaging,” *Phys. Med. Biol*. 37, 1095-1108 (1990).
7. Webb S., Flower M. A., and Ott R. J., “Geometric efficiency of a rotating slit-collimator for improved planar gamma-camera imaging,” *Phys. Med. Biol.* 38, 627-638 (1993).
8. Matherson K. J., Barber H. B., Barrett H. H., Eskin J. D., Dereniak E. L., Woolfenden J. M., Young E. T., and Augustine F. L., “Progress in the development of large-area modular 64 x 64 CdZnTe imaging arrays for nuclear medicine,” *IEEE Trans. Nucl. Sci*. 45, 354-358 (1988).
9. Butler J. F., Lingren C. L., Friesenhahn S. J., Doty F. P., Ashburn W. L., Conwell R. L., Augustine F. L., Apotovsky B., Pi B., Collins T., Zhao S., and Isaacson C., “CdZnTe solid-state gamma camera,” *IEEE Trans. Nucl. Sci*. 45, 359-363 (1998).

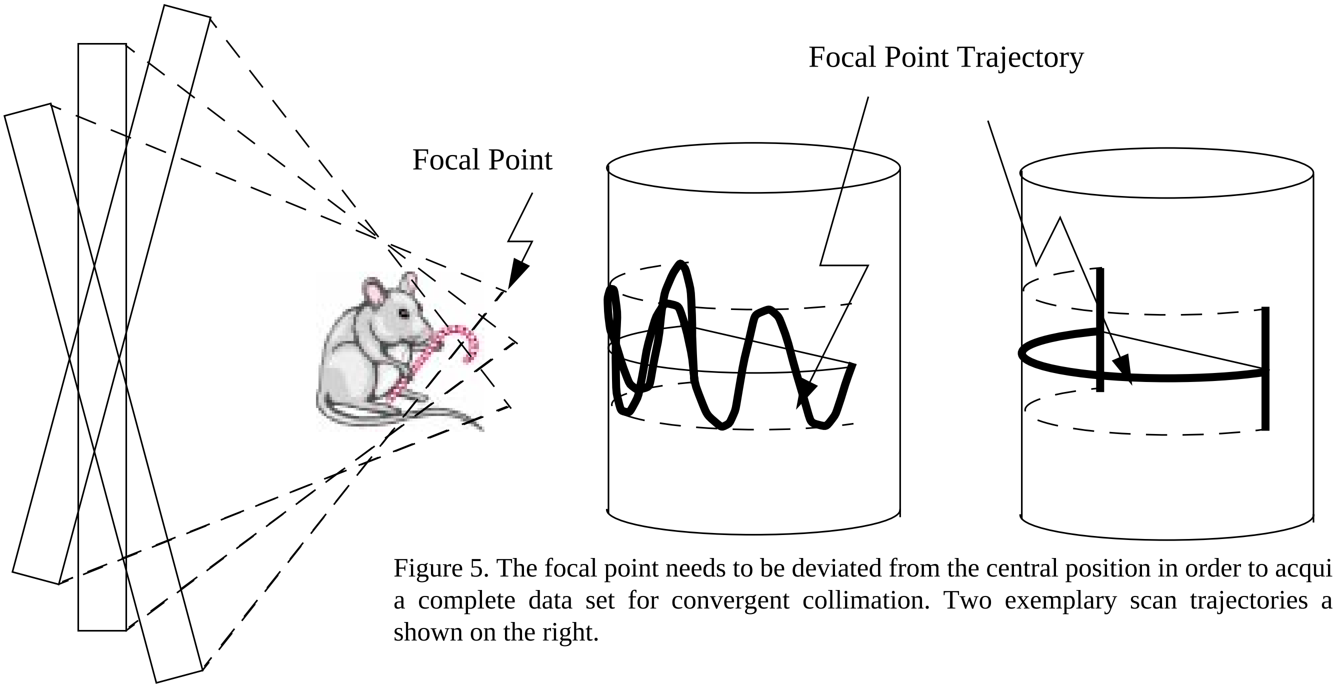


Figure 5. The focal point needs to be deviated from the central position in order to acquire a complete data set for convergent collimation. Two exemplary scan trajectories are shown on the right.

Image Reconstruction in 3D Short-Scan SPECT

Xiaochuan Pan, Chien-Min Kao, Charles Metz, and Alexander Kiselev*
Departments of Radiology and Mathematics*, The University of Chicago

Abstract

Physical factors such as photon attenuation degrade image quality and quantitative accuracy in single-photon emission computed tomography (SPECT). It is often considered (especially in situations with non-uniform attenuation and distance-dependent spatial resolution (DDSR)) that adequate compensation for the effects of these physical factors requires data acquired over 2π. However, as the analysis in this work suggests, one may need data acquired only over π to correct adequately for the effects of some of the physical factors. Reduction of the scanning angle in SPECT imaging is desirable because it can reduce the scanning time and thus minimize patient-motion and other artifacts, and because scans over less than 2π can allow the detector to be rotated at a fixed distance closer to the patient. Also, in certain cases (e.g., in cardiac SPECT), one can choose the scanning angular range for obtaining maximum numbers of photons. This work focuses on investigation of accurate image reconstruction in 3D SPECT from data that are acquired with parallel-beam collimation only over π and that contain the effects of photon attenuation (either uniform or non-uniform) and DDSR. For simplicity, we refer to such a scanning configuration as 3D short-scan SPECT. This work may have significant theoretical as well as practical implications for image reconstruction in SPECT.

I. BACKGROUND

Single-photon emission computed tomography (SPECT) is an important nuclear medicine imaging modality. Physical factors in SPECT such as photon attenuation and imperfect spatial resolution degrade image quality and quantitative accuracy [1, 2] and should be adequately corrected for. Because these physical factors are generally spatially variant, it is often considered (especially in the situation with non-uniform attenuation and distance-dependent spatial resolution (DDSR)) that adequate compensation for the effects of these physical factors requires data measured at projection angles over 2π. However, as the analysis below suggests, it appears possible that (at least under certain conditions) data acquired only over π can be used for adequately correcting for the effects of some physical factors such as photon attenuation. The reduction of the scanning angle in SPECT imaging is desirable because it can reduce scanning time and thereby minimize patient-motion and other artifacts. Also, in certain cases (e.g., in cardiac SPECT), one can choose the scanning angular range for obtaining maximum numbers of photons [3, 4]. This work focuses on investigation of accurate image reconstruction in 3D SPECT from data that are acquired with parallel-beam collimation only over π and that contain the effects of photon attenuation (either uniform or non-uniform) and DDSR. For simplicity, we refer to such a scanning configuration as the 3D short-scan SPECT.

II. MATHEMATICAL RATIONALES

A. *Redundant Information and Reduction of Scanning Angle*

In some tomographic imaging systems, the data measured over 2π contain redundant information. One example is measurement of the 2D Radon transform [5], $p(\xi, \phi)$, of a real function over 2π, where ξ is the detector bin index and ϕ is the measurement angle. Such measurements contain redundant information because, in the absence of noise and other inconsistencies, the measurements from conjugate views are mathematically identical, i.e.,

$$p(\xi, \phi) = p(-\xi, \phi + \pi). \tag{1}$$

It is well known that such information can be exploited for reducing the scanning angle from 2π to π because *the maximum difference between the values of the real angles on the two sides of Eq. (1) is* π.

In fan-beam computed tomography (CT), the quantity $q(t, \beta)$ is used to denote the measured transmission data, where t indicates the detector bin index and β the measurement angle. Although the fan-beam measurements (except for those with $t = 0$) from conjugate views are not mathematically identical, one can still show that the fan-beam data acquired over 2π contains redundant information, i.e.,

$$q(t, \beta) = q(-t, \beta + \pi + 2\alpha(t)), \tag{2}$$

where $\alpha(t)$ is a known and *real* function of t, and its explicit form depends upon the detector configurations [6, 7]. It can be shown that $\max\{2\alpha(t)\} = 2\alpha(t_{max})$, where t_{max} is a positive finite number and $|t| \leq t_{max}$. In most practical situations, $2\alpha(t_{max}) < \pi$. Therefore, *the maximum difference between the values of the real angles on the two sides of Eq. (2) is* $\pi + 2\alpha(t_{max})$, which is less than 2π. It is well known that such redundant information can be exploited for reducing the scanning angle from 2π to $\pi + 2\alpha(t_{max})$, which is referred to as the short-scan (or half-scan) fan-beam CT [8, 9].

In diffraction tomography (DT) [10, 11], one can derive a quantity $M(\nu, \phi)$ from the measured data at a particular angle ϕ, where ν can be interpreted as the spatial frequency of the measured data function. It has been shown that $M(\nu, \phi)$ measured at angles over 2π contains redundant information [12], i.e.,

$$M(\nu, \phi) = M(-\nu, \phi + \pi + 2\alpha(\nu)), \tag{3}$$

where $\alpha(\nu)$ is a known and *real* function of ν, and the explicit form of $\alpha(\nu)$ is determined by the DT scanning configurations. Also, it can be shown that $\max\{2\alpha(\nu)\} = 2\alpha(\nu_{max}) < \pi$, where ν_{max} is a positive finite number and $|\nu| \leq \nu_{max}$. Therefore, *the maximum difference between the values of real angles on the two sides of Eq. (3) is* $\pi + 2\alpha(\nu_{max})$, which is less than 2π. Again, such redundant information has been exploited for reducing the scanning angle from 2π to $\pi + 2\alpha(\nu_{max})$ in DT, which is referred to as the minimum-scan DT [13].

In summary, the data acquired over 2π in some tomographic imaging systems contain redundant information, which can be exploited for reducing the scanning angles and/or to reduce noise in the reconstructed image [12,14,15]. The above analysis suggests that, because the physical scanning angle must be real, the maximum scanning angles are determined by *the maximum differences between the values of the real angles on the two sides of the equations* that characterize the consistent conditions of (or, equivalently, the redundant information in) the data functions in these imaging systems.

B. Redundant Information and Possible Reduction of Scanning Angle in SPECT

Now we consider 3D SPECT with parallel beam projection. When the effects of photon attenuation and DDSR are considered, from the measured data one can obtain the so-called modified sinogram, which can be expressed readily as

$$m(\xi,\phi,z) = \int d\eta\, \mathrm{e}^{-\int_0^\eta d\eta'\, \mu(\xi\,\hat{\theta}(\phi)+\eta'\,\hat{\theta}^\perp(\phi),\, z)}$$

$$\times \iint d\xi'\, dz'\, a(\xi'\,\hat{\theta}(\phi) + \eta\,\hat{\theta}^\perp(\phi),\, z')\, h(\xi-\xi', z-z';\eta), \quad (4)$$

where ξ and z are the 2D detector indices, $a(x,y,z)$ is the 3D radioactivity-distribution function; $\hat{\theta}(\phi) = (\cos\phi, \sin\phi)$ and $\hat{\theta}^\perp(\phi) = (-\sin\phi, \cos\phi)$ are two orthogonal unit vectors; $x = \xi\cos\phi + \eta\sin\phi$; $y = -\xi\sin\phi + \eta\cos\phi$; $\mu(x,y,z)$ is the known 3D attenuation function; and $h(\xi,z;\eta)$ is the known distance-dependent spatial resolution function. The full-width-at-half-maximum (FWHM) of the latter function is dependent of the distance η. In this work, we consider $h(\xi,z;\eta)$ only as a shift-invariant resolution function. However, it can be a generally shift-variant resolution function.

1) Possible reduction of scanning angle in SPECT with uniform attenuation

Obviously, from inspection of Eq. (4) it is unclear whether the data acquired over $\phi \in [0, 2\pi)$ contain redundant information in 3D SPECT with the effects of non-uniform attenuation and a general DDSR. However, when only the effect of uniform attenuation is considered, the modified sinogram becomes the 3D exponential Radon transform (ERT) [16], and we have demonstrated previously that the 3D ERT acquired at $\phi \in [0, 2\pi)$ contain redundant information, i.e.,

$$M(\nu_m, \nu_z, \phi) = M(-\nu_m, \nu_z, \phi + \pi + \phi'(\nu_m)), \quad (5)$$

where $M(\nu_m, \nu_z, \phi)$ is the 2D Fourier transform of the modified sinogram $m(\xi, z, \phi)$ with respect to ξ and z, and where $\phi'(\nu_m)$ is a known and *purely imaginary* function of ν_m. We previously developed algorithms that exploit such redundant information for controlling noise in reconstructed SPECT images [14, 15, 17].

Can such redundant information in the ERT be exploited for reduction of the scanning angle? The analysis in Sec. II.A suggests that, because the physical scanning angle must be real, the maximum scanning angle is *the maximum difference between the values of real portions of the angles on the two sides of the equations* (e.g., see Eqs. (1-3) that characterize the redundant information.) In Eq. (5) of the ERT case, because $\phi'(\nu_m)$ is purely *imaginary*, we speculate that it may have no impact on the determination of the maximum physical scanning angle, which must be real, and thus that the ERT acquired only over $\phi \in [0, \pi)$ can be used for accurately reconstructing images in SPECT with uniform attenuation. Our numerical studies presented below seem to support our hypothesis strongly.

Furthermore, for SPECT with uniform attenuation and certain DDSR functions such as the Cauchy function [18, 19], one can show that the 2D Fourier transform of the modified sinogram satisfies Eq. (4). In this situation, $\phi'(\nu_m, \nu_z)$ becomes a function of ν_m and ν_z. However, most importantly, $\phi'(\nu_m, \nu_z)$ remains *purely imaginary*. Therefore, *the maximum difference between the values of the real angles on the two sides of Eq. (5) remains to be* π. This observation leads us to speculate that accurate images may be reconstructed from data acquired over π in SPECT with uniform attenuation and DDSR of certain forms. We investigated theoretically as well as numerically the image reconstruction from data acquired from only π in SPECT with uniform attenuation and DDSR. Indeed, our numerical investigation below suggests that the quality of images reconstructed from data acquired over π appear to be comparable to that of images reconstructed from data acquired over 2π.

2) Possible reduction of scanning angle in SPECT with non-uniform attenuation

So far, it has been generally believed that one needs data over 2π for accurate image reconstruction in SPECT with non-uniform attenuation. However, the seemingly promising numerical results below obtained for SPECT with uniform attenuation leads naturally to the question: Can accurate images be reconstructed from the data acquired only over π in SPECT with non-uniform attenuation? In an attempt to answer this question, we are currently conducting a theoretical investigation, which appears to be an exceedingly difficult task despite the fact that an analytic solution has recently been derived for full-scan SPECT with the effect of non-uniform attenuation [20, 21]. On the other hand, we also conducted numerical research on image reconstruction from data acquired over π in SPECT with non-uniform attenuation. These numerical results, which are presented in Sec. III below, indicate (at least for the cases that we studied) that the quality of images reconstructed from data acquired over π is comparable to that of images reconstructed from the data acquired over 2π in SPECT when attenuation is non-uniform.

We investigated numerically the image reconstruction from data acquired only at $\phi \in [0, \pi)$ in short-scan SPECT with the effects of both non-uniform attenuation and DDSR. Again, the results of these numerical studies, as shown below, seem to suggest that the quality of images reconstructed from data acquired over π appear to be comparable to that of images reconstructed from data acquired over 2π.

3) Reconstruction Algorithm

Although we speculate that one may need data acquired only over $\phi \in [0, \pi)$ for accurate image reconstruction in short-scan SPECT, it remains unclear whether “closed-form” algorithms can be derived to accomplish such reconstruction tasks. On the other hand, one may use iterative algorithms to reconstruct images (i.e., to obtain solutions $a(\vec{r})$ in Eq. (4)) from knowledge of the data $m(\xi, \phi, z)$ over $\phi \in [0, \pi)$. Additive iterative algorithms can be devised for obtaining the solution in Eq. (4). An important question is whether such algorithms converge and, if so, whether they converge uniquely to the correct solution. It may be possible to prove the convergence of such additive iterative algorithms in certain situations. However, for a data function that contains the effects of non-uniform attenuation and DDSR as shown in Eq. (4), it is generally difficult (if not impossible) to prove the convergence of additive iterative algorithms. More importantly, the additive iterative algorithms [22], in general, cannot guarantee the positivity of the solutions and thus can be susceptible to noise and other inconsistencies such as sample aliasing that always accompanies experimentally measured data.

For the purpose of simplicity, Eq. (4) can be rewritten symbolically as[1]

$$g(y) = \int_{D_x} dx\, h(x, y)\, f(x) \qquad \forall\, y \in D_y, \tag{6}$$

where the real and non-negative functions $g(y)$ and $f(x)$ denote the data and image functions with domains D_y and D_x, respectively, and $h(x, y)$ denotes the kernel of the imaging transformation, which, as shown in Eq. (4), is also non-negative. The task here is to find, from knowledge of the data $g(y)$ (or, equivalently, $m(\xi, \phi, z)$ over $\phi \in [0, \pi)$), a non-negative solution $f(x)$ (or, equivalently, $a(\vec{r})$) that satisfies Eq. (6) (or, equivalently, Eq. (4).) Because Eq. (6) (i.e., Eq. (4)) is an inherently non-negative integral equation, we propose to use the algorithm

$$f^{(n+1)}(x) = \frac{f^{(n)}(x)}{\int_{D_y} dy\, h(x, y)} \int_{D_y} dy\, \frac{h(x, y) g(y)}{\int_{D_x} dx\, h(x, y)\, f^{(n)}(x)} \tag{7}$$

to obtain the solution $f(x)$ (i.e., the image function $a(\vec{r})$), where n is the number of iterations.

The algorithm in Eq. (7) is related to the expectation maximization (EM) algorithm [23–27], which has been shown to yield the maximization-likelihood solution when the data function $g(y)$ contain Poisson noise [24]. Even in the absence of Poisson noise, from the perspective of solving the positive integral equation, it can be shown [28] that *the algorithm in Eq. (7) converges* in the sense that the Kullback-Leibler discrepancy between $f^{(n+1)}(x)$ and $f^{(n)}(x)$ approaches zero monotonically at a rate faster than $(\frac{1}{n})$. (Recall that it is generally difficult to prove the convergence of additive algorithms for Eq. (4).) If a unique solution to Eq. (7) exists, *one can show that $f^{(n+1)}(x)$ converges to that unique solution.*[2] Additionally, from a practical point of view, the algorithm in Eq. (7) is easy to implement because it involves only forward and backward transformations and, more importantly, guarantees the positivity of the solution. Numerical investigation suggests that the EM algorithm is generally less susceptible than additive algorithms to the unavoidable noise and inconsistencies contained in experimentally acquired SPECT data [22].

Fig. 1: Representative slices of the activity map, superimposed on the attenuation maps.

III. NUMERICAL RESULTS

We conducted computer simulation to evaluate the possibility of short-scan SPECT reconstruction, as we stipulated above. We considered a SPECT cardiac study. The 3D radioactivity map has a uniform concentration inside the myocardium and is zero elsewhere. The 3D attenuation map is non-uniform, with a value of $0.15\,\mathrm{cm}^{-1}$ inside the chest, except for the lungs, which have no attenuation. Representative slices of the radioactivity map, superimposed on the attenuation map, are shown in Fig. 1. For purposes of comparison, we also considered a uniform attenuation map that is obtained from the non-uniform one by removing the lungs. We used a Gaussian DDSR of the form $h(\xi, z; \eta) = \frac{1}{2\pi^2\sigma_\eta^2} \exp\left\{-\frac{\xi^2+z^2}{2\sigma_\eta^2}\right\}$ to model the blurring effect in the data, where the standard deviation $\sigma(\eta) = \sigma_0 + \sigma_1\eta$ depends linearly upon the distance η. Noiseless SPECT projection data, covering the full 2π view, was generated with Poisson noise subsequently to produce noisy data. The simulated projection data consist of 120 views of 128 $(\mathrm{radial}) \times 32\,(\mathrm{axial})$ sinograms with a bin size of $0.25 \times 0.25\,\mathrm{cm}^2$. The simulated noisy 2π data have about ten thousand counts per slice.

In reconstruction, all the simulated views were used for producing “full-scan” reconstruction, whereas only half of the views, ranging from $\pi/2$ to $3\pi/2$, are used for generating “short-scan” reconstruction. Both modes of reconstruction generated $128 \times 128 \times 32$ images with a $0.25 \times 0.25 \times 0.25\,\mathrm{cm}^3$ voxel size. In the results shown below, we considered three types of reconstructions: (a) EM reconstruction from the short-scan data over π; (b) EM reconstruction from full-scan data over

---

[1] The symbols x and y here denote general 3D spatial coordinates in the image and data spaces, respectively, and should not be confused with the Cartesian coordinates x and y in Eq. (4).

[2] We are currently investigating the existence of the unique solution to Eq. (4) and the conditions under which the unique solution may exist.

2π; and (c) FBP reconstruction from short-scan data over π.

Figure 2 shows representative slices of the images reconstructed from noiseless and noisy data, which contain only the effect of uniform attenuation. These results demonstrate qualitatively that, for both noiseless and noisy data, images reconstructed from the short-scan and full-scan data are of essentially similar quality. These reconstructions have effectively removed the effects of attenuation that can be clearly observed in the FBP reconstructions (in the form of reduced image brightness toward the center). Notice here that the total number of counts in the short-scan reconstruction is only half of that in the full-scan reconstruction. However, for a fixed scanning time, one would expect the short-scan data to have approximately the same number of counts as in the full-scan data. When this factor is taken into consideration, it would be interesting to compare quantitatively the noise properties between reconstructions from both short-scan and full-scan data, where the former have high signal-to-noise ratio than do the latter.

Figure 3 shows representative slices of the images reconstructed from noiseless and noisy data, which contain only the effect of non-uniform attenuation. Again, these results show that, for both noiseless and noisy data, images reconstructed from the short-scan and full-scan data are of comparable quality. These reconstructions have compensated effectively for the effects of attenuation that can be clearly observed in the FBP reconstructions. It is also interesting to note the differences in the attenuation artifacts in the FBP images between this and the above case.

We subsequently introduced the effect of a Gaussian DDSR, with $\sigma_0 = 0.2\,\text{cm}$ and $\sigma_1 = 0.02$, into the simulated data and repeated the studies above. Figure 4 displays representative slices of the images reconstructed from noiseless and noisy data, which contain the effects of both uniform attenuation and Gaussian DDSR. Again, these results suggest that, for both noiseless and noisy cases, images reconstructed from the short-scan and full-scan data are of comparable quality. These reconstructions have compensated effectively for the effects of uniform attenuation and DDSR, which can be clearly observed in the FBP reconstructions. Finally, Figure 5 displays representative slices of the images reconstructed from noiseless and noisy data, which contain the effects of both non-uniform attenuation and Gaussian DDSR. Again, these results indicate that, for both noiseless and noisy cases, images reconstructed from the short-scan and full-scan data have comparable quality and that the effects of non-uniform attenuation and DDSR have been compensated effectively for.

IV. CONCLUSIONS AND DISCUSSION

It has been observed that the redundant information contained in some tomographic imaging systems can be exploited for devising short-scan configurations in these imaging systems [8, 9, 13]. We have shown previously [14, 15, 17] that redundant information exists in data acquired over 2π in SPECT with uniform attenuation and DDSR of certain forms. These observations led us to hypothesize or speculate that one may need data acquired over π instead of 2π for accurate image reconstruction in 3D SPECT, and thus to suggest the concept of short-scan SPECT. We propose the use of a non-linear EM algorithm to reconstruct images in short-scan SPECT. It can be shown that this EM algorithm converges, and that it converges to the unique solution if such a unique solution exists. We also performed a numerical investigation to verify and evaluate accurate image reconstruction in short-scan SPECT. These results indicate (at least for the examples studied) that the quality of the reconstructed images in short-scan 3D SPECT appears to be essentially similar to that of reconstructed images in full-scan 3D SPECT. We are currently performing a detailed quantitative evaluation of image quality and accuracy and will report its results in the near future.

We are also conducting a theoretical investigation on the possibility of obtaining analytical solutions in 3D short-scan SPECT when the effects of attenuation and DDSR are considered. Recent work [20, 21] on the analytical solutions for 2D full-scan SPECT with only the effect of non-uniform attenuation may provide useful insights into our investigation of short-scan SPECT. Investigations on such analytic solutions can be theoretically important in understanding the reconstruction problems in short-scan SPECT. For example, the existence of such an analytical solution[3] will imply that the solution of Eq. (4) is unique, and consequently, that the non-linear iterative algorithm in Eq. (7) converges to the unique solution.

In this work, we discuss only the image reconstruction from data acquired at angles from 0 to π. In fact, this can be considered as a special case of the so-called *π-scheme short-scan SPECT* that we have proposed. Basically, in the π-scheme short-scan SPECT, the data can be acquired over disjointed angular intervals. We speculate that, as long as the summation of these intervals without conjugate views is larger than or equal to π, images with quality comparable to that of images in full-scan SPECT may be reconstructed. Our preliminary numerical investigation has confirmed this observation and suggests that reconstruction from data acquired over disjointed angular internals converges even faster than that from data acquired from 0 to π in short-scan SPECT.[4]

This work is theoretically intriguing because it poses several theoretically interesting and challenging questions. Does a unique solution exist in 3D short-scan SPECT with the effects of attenuation and DDSR? If so, under what conditions does such a solution exist? Research intended to provide answers to these questions is currently under way. Also, the practical implications of this work seem to be significant because the proposed π-scheme short-scan allows data acquisitions at desired projection views at which the emitted gamma-rays may

---

[3] For instance, the Tretiak-Metz method is an analytic solution to the reconstruction problem in 2D full-scan SPECT with uniform attenuation. It is highly susceptible to data noise and inconsistencies. However, its existence guarantees that there is a unique solution to the inversion of the ERT from its knowledge over 2π.

[4] This is understandable because, from the numerical perspective, the linear transformation associated with data acquired over disjointed angular internals generally is better conditioned than is that for data acquired from 0 to π in π-scheme short-scan SPECT.

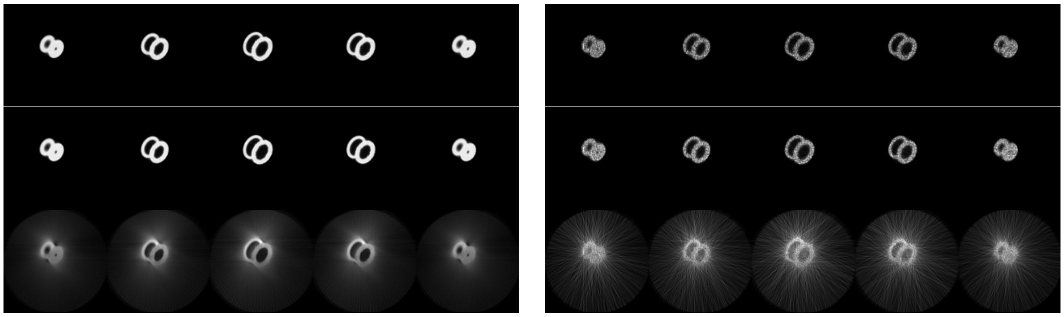


Fig. 2: Images reconstructed from simulated data with the only effect of uniform attenuation. Representative slices in the images reconstructed from noiseless (left panel) and noisy (right panel) short-scan data (1st row), the full-scan data (2nd row), and FBP reconstruction from short-scan data (3rd row). 50 and 20 iterations were used in the case of noiseless and noisy data, respectively, so that their total computation cost is identical.

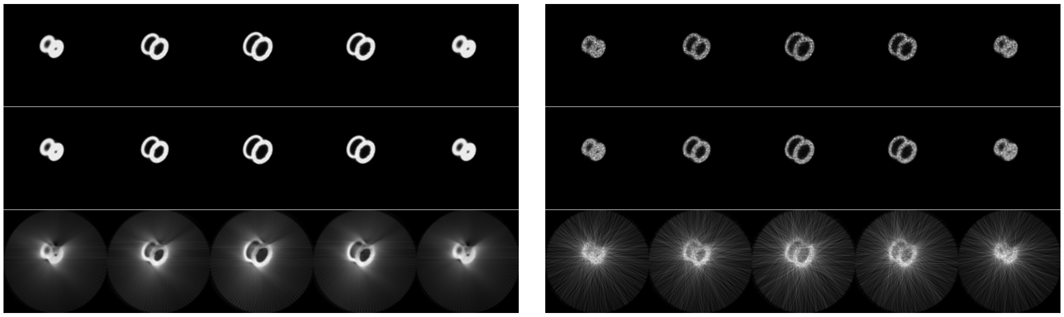


Fig. 3: Images reconstructed from simulated data with the only effect of non-uniform attenuation. Representative slices in the images reconstructed from noiseless (left panel) and noisy (right panel) short-scan data (1st row), the full-scan data (2nd row), and FBP reconstruction from short-scan data (3rd row). 50 and 20 iterations were used in the case of noiseless and noisy data, respectively.

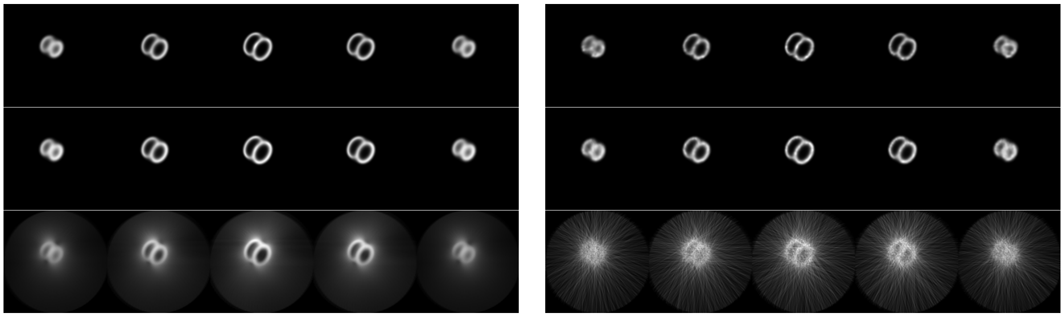


Fig. 4: Images reconstructed from simulated data with the effects of both uniform attenuation and DDSR. Representative slices in the images reconstructed from noiseless (left panel) and noisy (right panel) short-scan data (1st row), the full-scan data (2nd row), and FBP reconstruction from short-scan data (3rd row). 50 and 20 iterations were used in the case of noiseless and noisy data, respectively.

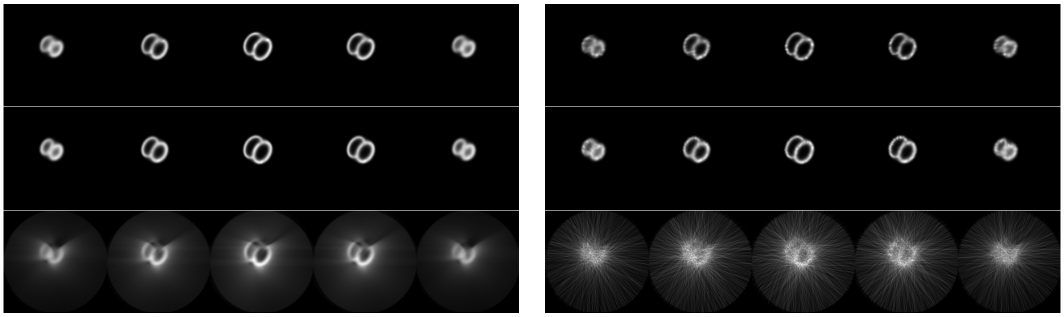


Fig. 5: Images reconstructed from simulated data with the effects of both non-uniform attenuation and DDSR. Representative slices in the images reconstructed from noiseless (left panel) and noisy (right panel) short-scan data (1st row), the full-scan data (2nd row), and FBP reconstruction from short-scan data (3rd row). 50 and 20 iterations were used in the case of noiseless and noisy data, respectively.

undergo the least attenuation and blurring, thus providing the freedom for significantly reducing the scanning time and for obtaining data with a high signal-to-noise ratio. One clinical study that may benefit from such a π-scheme scan is cardiac imaging with SPECT.

V. ACKNOWLEDGMENTS

This work was supported in part by National Cancer Institute Grants CA-70449 and CA-85593. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute.

VI. REFERENCES

[1] T. F. Budinger, G. T. Gullberg, and R. H. Huesman. Emission computed tomography. In G. T. Herman, editor, *Image Reconstruction from Projections: Implementation and Application*, pages 147–246, Springer, New York, 1979.

[2] R. J. Jaszczak, R. E. Coleman, and F. R. Whitehead. Physical factors affecting quantitative measurements using camera-based single-photon emission computed tomography (SPECT). *IEEE Trans. Nucl. Sci.*, 28:69–80, 1981.

[3] R. T. Go, W. J. MacIntyre, T. S. Houser, M. Pantoja, D. H. F. J. K. O'Donnell, B. J. Sufka, D. A. Underwood, and T. F. Meaney. Clinical evaluation of 360° and 180° data sampling techniques for transaxial SPECT thallium-201 myocardial perfusion imaging. *J. Nucl. Med.*, 26:695–706, 1985.

[4] K. Knesaurek, M. A. King, S. J. Glick, and B. C. Penney. Investigation of causes of geometrical distortion in 180° and 360° angular sampling in SPECT. *J. Nucl. Med.*, 30:1666–1675, 1989.

[5] F. Natterer. *The Radon Transform*. Wiley and Sons, New York, 1986.

[6] G. T. Herman. *Image Reconstruction from Projections*. Academic Press, New York, 1980.

[7] A. Rosenfeld and A. C. Kak. *Digital Picture Processing*, volume 1. Academic Press, New York, 1982.

[8] D. L. Parker. Optimal short scan convolution reconstruction for fan-beam CT. *Med. Phys.*, 9:245–257, 1982.

[9] C. R. Crawford and K. King. Computed tomography scanning with simultaneous patient translations. *Med. Phys.*, 17:967–982, 1990.

[10] R. Mueller, M. Kaveh, and G. Wade. Reconstructive tomography and applications to ultrasonics. *Proceedings of the IEEE*, 67:567–587, 1979.

[11] S. Pan and A. Kak. A computational study of reconstruction algorithms for diffraction tomography: Interpolation versus filtered backpropagation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31:1262–1275, 1983.

[12] X. Pan. A unified reconstruction theory for diffraction tomography with considerations of noise control. *J. Opt. Soc. Am.*, 15:2312–2326, 1998.

[13] X. Pan and M. Anastasio. Minimum-scan filtered backpropagation algorithms in diffraction tomography. *J. Opt. Soc. Am.*, 16:2896–2903, 1999.

[14] C. E. Metz and X. Pan. A unified analysis of exact methods of inverting the 2-D exponential Radon transform, with implications for noise control in SPECT. *IEEE Trans. Med. Imaging*, 14:643–658, 1995.

[15] X. Pan and C. E. Metz. Analysis of noise properties of a class of exact methods of inverting the 2-D exponential Radon transform. *IEEE Trans. Med. Imaging*, 14:659–668, 1995.

[16] O. J. Tretiak and C. E. Metz. The exponential Radon transform. *SIAM J. Appl. Math.*, 39:341–354, 1980.

[17] X. Pan and C. E. Metz. Analytical approaches for image reconstruction in 3D SPECT. In P. Grangeat and J. Amans, editors, *3D Image Reconstruction in Radiology and Nuclear Medicine*, pages 103–116, Kluwer Academic Publishers, New York, 1996.

[18] C. P. Appledorn. An analytical solution to the non-stationary reconstruction problem in single photon emission computed tomography. In D. A. Ortenhdahl and J. Llacer, editors, *Information Processing in Medical Imaging*, pages 69–79, Wiley-Liss, New York, 1990.

[19] X. Pan. Analysis of 3D SPECT image reconstruction and its extension to ultrasonic diffraction tomography. *IEEE Trans. Nucl. Sci.*, 45:1308–1316, 1998.

[20] R. G. Novikov. An inversion formula for the attenuated X-ray transformations. (preprint), 2000.

[21] F. Natterer. Inversion of the attenuated Radon transform. (preprint), 2000.

[22] B. M. W. Tsui, G. T. Gullberg, E. R. Edgerton, J. G. Ballard, J. R. Perry, W. H. McCartney, and J. Berg. Correction of non-uniform attenuation in cardiac SPECT imaging. *J. Nucl. Med.*, 30:497–507, 1989.

[23] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *JRSS*, 39:1–38, 1977.

[24] L. A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imaging*, 1:113–122, 1982.

[25] T. M. Cover. An algorithm for maximizing expected log investment return. *IEEE Trans. Information Theory*, 30:369–373, 1984.

[26] I. Csiszár and G. Tusnády. Information geometry and alternating minimization procedures. *Statistics and Desisions*, Supplement issue 2:205–237, 1984.

[27] K. Lang and R. Carson. EM reconstruction algorithms for emission and transmission tomography. *J. Comput. Assist. Tomogr.*, 8:306–318, 1984.

[28] W. H. Wong. On Vardi's algorithm for positive integral equations. *Technical Report*, 334:1–12, Department of Statistics, The University of Chicago, 1991.

4D Processing of Gated SPECT Images Using Deformable Mesh Modeling[1]

Jovan G. Brankov, Yongyi Yang, and Miles N. Wernick

Illinois Institute of Technology
Department of Electrical and Computer Engineering
3301 S. Dearborn St., Chicago, IL 60616, USA

Abstract

In this work we present a new 4D approach for reducing noise in gated SPECT perfusion images while preserving accurate cardiac motion. The method is based on motion-compensated temporal smoothing using a deformable content-adaptive mesh to model cardiac motion. We use a new, fast method for initial mesh generation. This mesh is then deformed to track cardiac motion and smoothing is performed along motion tractories through the space-time coordinate system.

I. INTRODUCTION

The quality of SPECT images is adversely affected by noise caused by low photon counts. The problem of noise is especially serious in gated studies, where the counts are divided into a number of time intervals to obtain an image sequence [1]. Because of their relatively high noise level, gated images can potentially benefit most from appropriate image processing.

In this paper we propose a new spatial-temporal processing method for gated images that uses motion tracking based on deformable mesh modeling of the images. In nuclear medicine, spatial-temporal processing has become popularly known as four-dimensional (4D) processing to reflect the use of three spatial dimensions plus time. Therefore, we will use the terms "4D" and "spatial-temporal" interchangeably, although our preliminary studies are based on a single slice of a gated image sequence, so that we only have two spatial dimensions plus time.

4D processing is an example of multichannel image recovery, which we reviewed in [2]. The basic idea of this approach is to exploit the statistical correlations between the desired signal components of different image frames in a sequence or other collection of related images.

Methods of 4D processing have received increasing interest lately. Our group has proposed several 4D methods designed for reconstruction of motion-free images, such as those obtained in dynamic PET studies [3]. Lalush and Tsui [4] applied 4D image reconstruction to cardiac SPECT images, but did not incorporate motion estimation explicitly in their techniques. In the broader image-processing field, motion-compensated processing is a well-known approach to reduce the noise in an image sequence [5]. In the nuclear medicine field, Klein *et al.* [6] developed a motion-compensated summing method using motion estimation, based on the optical-flow method [7,8], for obtaining a single image from a gated PET study.

In this paper we propose the following method. We represent the images and account for motion in a gated SPECT sequence by way of a content-adaptive mesh model (CAMM) (Fig. 1), which is allowed to deform over time. We apply temporal smoothing along the trajectories that the nodes of this mesh traverse through the space-time coordinate system (see Fig. 2). This approach aims to reduce noise, while avoiding potential distortions of the cardiac motion.

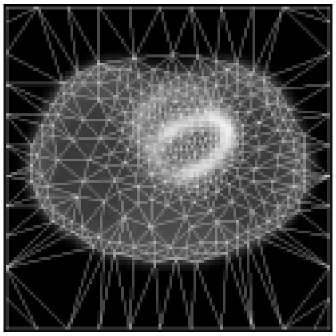


Figure 1. Mesh structure used in our experiment.

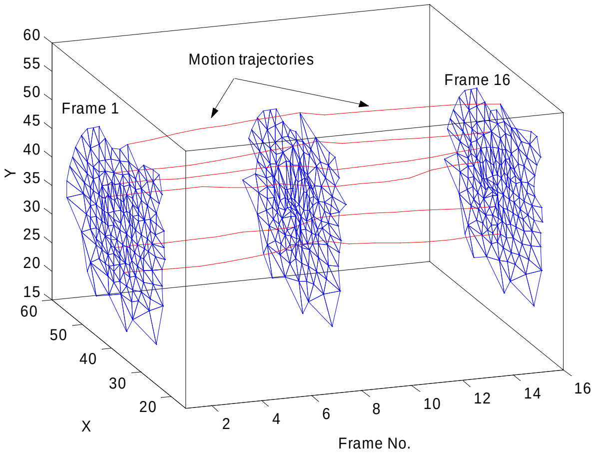


Figure 2. Deformable mesh (shown for frames 1, 8, and 16), and motion trajectories for some selected mesh nodes throughout the sequence.

---

[1] This research was supported by the National Institutes of Health under grant HL65425 and by the Whitaker Foundation.

II. METHODS

A. *Motion Field Modeling*

In a CAMM, the image domain is subdivided into a number of mesh elements, the vertices of which are called *nodes*. By deformation of the individual mesh elements, a deformable CAMM can be used to describe the image motion through the inter-frame displacement of the nodes [9]. Such a deformable CAMM is well-suited for modeling complex, non-rigid motion, such as that of the heart.

The image domain D is partitioned into M non-overlapping mesh elements, denoted by D_m, $m = 1, 2, \ldots, M$, defined by the nodes. The image motion is then derived from the inter-frame displacement of these nodes as follows. Over a particular element D_m, the motion field is described as

$$\mathbf{d}(\mathbf{x}) = \sum_{n=1}^{N} \varphi_n(\mathbf{x})\mathbf{d}_n, \tag{1}$$

where $\mathbf{d}_n$ and $\varphi_n(\mathbf{x})$ are the displacement vector and interpolation basis function associated with node n, respectively, and N is the total number of mesh nodes. Note that the support of each basis function $\varphi_n(\mathbf{x})$ is limited only to those elements D_m associated with node n .

In practice the nodal vectors $\mathbf{d}_n$ in the motion model in (1) are unknown, and must be determined from the observed data. To track the motion between image frames, a natural approach is to displace the mesh nodes so that the corresponding mesh elements in the two frames achieve the best match in terms of their image values.

Let $f_r(\mathbf{x})$ and $f_t(\mathbf{x})$ denote, respectively, the image functions of a reference frame and a neighboring frame in the sequence, known as the target frame. As a matching criterion the following objective function is adapted for our application from [9]:

$$J = \frac{1}{2} W_m \sum_{m=1}^{M} \left[\int_{D_m} \left(f_t(\mathbf{x} + \mathbf{d}(\mathbf{x})) - f_r(\mathbf{x}) \right)^2 d\mathbf{x} \right] + \frac{1}{2}(1 - W_m) E_d \ , \tag{2}$$

where the first term is the matching error accumulated over all M mesh elements between the two frames, the second term E_d is a measure of mesh regularity (to be defined below), and W_m is a constant chosen for trade-off between mesh matching accuracy and mesh regularity.

The mesh regularity measure E_d in (2) is defined as:

$$E_d = \frac{1}{2} \sum_{n=1}^{N} \left\| \sum_{l \in \Im_n} (\mathbf{x}_n - \mathbf{x}_l) \right\|^2 , \tag{3}$$

where N is the total number of mesh nodes in the image, and $\Im_n$ is the set of immediate neighboring mesh nodes that are connected to node n.

The nodal vectors $\mathbf{d}_n$ are then solved numerically by minimizing the objective function in (2) with a gradient descent algorithm. More details of the implementation can be found in [9].

B. *Spatial-Temporal Processing*

Let $\mathbf{f}_k$ and $\hat{\mathbf{f}}_k$ denote image frame k (in vector form) before and after processing, respectively, with k = 1,...,K. Further, let $\mathbf{f} = \begin{bmatrix} \mathbf{f}_1^T & \cdots & \mathbf{f}_K^T \end{bmatrix}^T$ and $\hat{\mathbf{f}} = \begin{bmatrix} \hat{\mathbf{f}}_1^T & \cdots & \hat{\mathbf{f}}_K^T \end{bmatrix}^T$ represent the entire image sequence before and after processing, respectively. Then the proposed 4D processing framework can be described as a separable operation of the following form:

$$\hat{\mathbf{f}} = (\mathbf{H}_s \cdot \mathbf{H}_t)\mathbf{f} \tag{4}$$

where $\mathbf{H}_s$ and $\mathbf{H}_t$ represent the spatial and temporal processing operators.

In this paper, a spatial low-pass Butterworth filter (described in the next section) is used for the operator $\mathbf{H}_s$ in Eq. (4).

The temporal processing operator $\mathbf{H}_t$ is implemented as a finite impulse response (FIR) filter along the motion trajectories. For each voxel $\mathbf{x}$ in frame k the image value is processed by $\mathbf{H}_t$ according to the following equation:

$$\hat{f}_k(\mathbf{x}) = \sum_{l=-K/2}^{K/2} h_t(l) f_{k+l}\left(\mathbf{x} - \mathbf{d}_{k,l}(\mathbf{x})\right) \tag{5}$$

where $\mathrm{d}_{k,l}(\mathrm{x})$ denotes the relative motion between voxel $\mathbf{x}$ in frame k and its corresponding voxel in frame $k+l$. The filter coefficients $h_t(l)$ are defined as:

$$h_t(l) = \frac{1}{C}\left(1 - \frac{2|l|}{K}\right)^{\gamma}, l = -K/2, \cdots, K/2, \tag{6}$$

where γ is a parameter used to control the degree of temporal smoothing, and C is a normalization constant defined so that the filter has unity DC response. Filters that are more optimal will be considered in future studies.

III. EXPERIMENTS

A. *Evaluation Data*

The proposed spatial-temporal processing algorithms were tested using the 4D gated mathematical cardiac-torso (gMCAT) D1.01 phantom [10]. The field of view is 36 cm; the pixel size is 5.625mm. Poisson noise, at a level of 4 million total counts for the entire sequence was introduced to represent a clinical Tc^{99m} study. In this preliminary study, a single slice (No.70) was used. This slice had approximately 5.5×10^4 counts per frame (a total of 16 frames). No attenuation correction was used.

B. *Mesh generation*

The mesh structure was constructed using a new method we have proposed [11]. A total of 389 mesh nodes were used in the mesh shown in Figure 1, which is only about one-tenth the number of pixels. Note that the algorithm automatically places the mesh nodes densely in the important heart region, and sparsely in the background.

C. *Motion Field Estimation*

The noisy projection data were first reconstructed by using the maximum-likelihood expectation-maximization (MLEM) algorithm [12]. In this step image frames were reconstructed in an independent, frame-by-frame fashion. To help suppress the noise level in the reconstructed images, individual frames were smoothed spatially with an order-5 Butterworth filter with a cutoff frequency of 0.3 cycles/pixel. Afterward, level equalization was applied to enhance the image features in the relatively weak right-ventricular region.

The resulting sequence was then used for motion estimation based on (2), where the parameter W_m was set to 0.95. The mesh structure in Fig. 1 was used as the initial mesh. In our experiment the nodal positions were updated only for nodes belonging to a circular region of interest containing the heart. This served to reduce the computational burden.

In Figure 2 we show the deformable mesh obtained by the procedure described above. For illustration purposes, mesh structures are shown for frames 1, 8, and 16. In addition, the motion trajectories of some selected mesh nodes are also shown throughout the sequence.

D. *Results*

In this section we present results obtained from processing of images reconstructed using the MLEM algorithm. For comparison, the following processing methods were considered: (1) spatial-only filtering (“Spatial”), in which an order-5 Butterworth filter with a cutoff frequency of 0.3 cycles/pixel was applied to the reconstructed images; (2) the proposed 4D processing method (“ST-DM”), applied to the MLEM reconstructed images; and (3) the same smoothing filters as in (2), except that motion compensation was omitted (“ST-NM”). The purpose of evaluating the third method is to demonstrate that, while temporal smoothing without motion compensation can reduce noise, it yields a significant degradation of the representation of cardiac motion.

In Figure 3 we present some reconstructed images for visual evaluation. Note that “Original” is the phantom degraded by the system blur to represent an approximate best case image for comparison. The image results suggest that both ST-NM and ST-DM can significantly reduce the noise level in the reconstructed images. However, the images from the ST-NM method suffer from significant motion distortion. This is evident when viewing the images as a cine loop (movie),[2] but it can also be measured quantitatively. Failure to compensate for motion (in the ST-NM method) also reduces the frame-to-frame variation in the left ventriclar volume (Figure 3), which we expect will distort measurements of ejection fraction.

To quantify these observations, we plot in Figure 4 the time activity curves (TAC) for a small region in the left ventricular wall vs. the frame number, computed for images obtained by the three methods. The total squared errors between the original TAC and TACs obtained by the Spatial, ST-NM, and ST-DM methods were 3.45, 0.50 and 0.28, respectively. Again, the best performance was achieved by the proposed ST-DM method. In future studies we will evaluate quantitatively the effect of the algorithms on ejection fraction measurements, perfusion-defect detection, and apparent wall motion.

IV. DISCUSSION

In this paper we demonstrated that one can improve the quality of the reconstructed images in gated SPECT by use of spatial-temporal processing with deformable content-adaptive mesh modeling. Such an approach can effectively suppress the noise in the images without distorting cardiac motion. By the time of the conference, we hope to extend our implementation to 3D volumetric reconstruction of gated image sequences.

V. REFERENCES

[1] K. Nichols and E. G. Depuey, "Regional and global ventricular function analysis with SPECT perfusion imaging," in *Nuclear Cardiology: State of the Art and Future Directions*, 2nd ed., pp. 137-187, St. Louis: Mosby, 1999.

[2] N.P. Galatsanos, M.N. Wernick, and A.K. Katsaggelos, “Multichannel Image Recovery,” in *Handbook of Image and Video Proc.*, A. Bovik, ed. San Diego: Academic Press, 2000, pp. 155-168.

[3] M. N. Wernick, E. J. Infusino, and M. Milosevic, “Fast spatio-temporal image reconstruction for dynamic PET,” *IEEE Trans. Med. Imag*., vol. 18, pp. 185-195, 1999.

[4] D. S. Lalush and B. M. W. Tsui, “Block-iterative techniques for fast 4D reconstruction using *a priori* motion models,” *Phys. Med. Biol*., vol. 43, pp. 875-886, 1998.

[5] A. M. Teklap, *Digital Video Processing*. NJ: Prentice-Hall, 1995.

[6] G. J. Klein, B. W. Reutter, and R. H. Huesman, “Non-rigid summing of gated PET via optical flow,” *IEEE Trans. Nucl. Sci.*, vol. 44, pp. 1509 -1512, 1997.

[7] J. F. Vega-Riveros and K. Jabbour, “Review of motion analysis techniques,” *IEEE Comm. Speech and Vision*, vol. 136, pp. 397-404, Dec. 1989.

[8] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, “Performance of optical flow techniques,” presented at *Inter. Jour. of Comp. Vision*, 1994.

[9] Y. Wang and O. Lee, “Active Mesh-A Feature Seeking an Tracking Image Sequence Representation Scheme,” *IEEE Trans. Image Processing*, vol. 3, pp. 610-624, 1994.

[10] P. H. Pretorius, W. Xia, M. A. King, B. M. W. Tsui, T. S. Pan, and B. J. Villegas, “Evaluation of right and left ventricular volume and ejection fraction using a mathematical cardiac torso phantom for gated pool SPECT,” *J. Nucl. Med.*, vol. 38, pp. 1528-1534, 1997.

[11] J. G. Brankov, Y. Yang, and M. N. Wernick, “Content-Adaptive Mesh Modeling for Tomographic Image Reconstruction,” submitted to this conference.

---

[2] Image sequences available for cine viewing at *http://www.iit.edu/~branjov/3D01.htm*

[12] K. Lange and R. E. Carson, “EM reconstruction algorithms for emission and transmission tomography,” *J. Comput. Assist. Tomog*r., vol. 8, no.2, pp. 306–316, 1984.

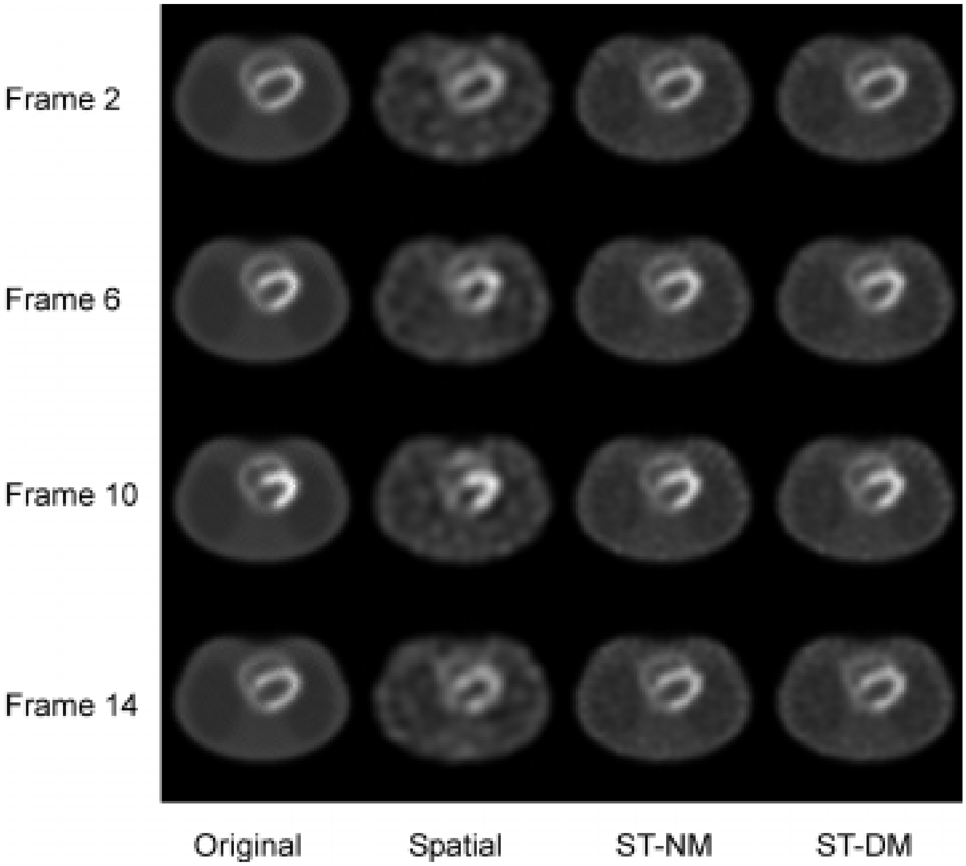


Figure 3. Results obtained by maximum-likelihood expectation-maximization. “Spatial” denotes spatial smoothing only. “ST-NM” denotes spatial-temporal smoothing without motion compensation. “ST-DM” denotes the proposed spatial-temporal smoothing with motion compensation achieved using a deformable mesh. “Original” denotes the phantom degraded by the system blur to represent an approximate best-case image for comparison.

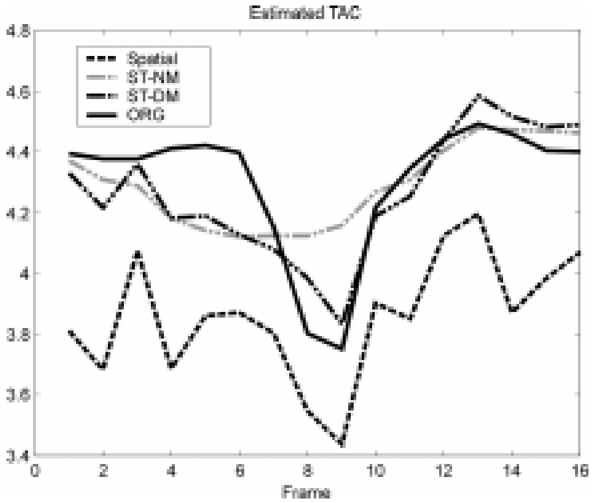


Figure 4. Time activity curves (TAC) for a small region in the left ventricular wall vs. the frame number. Note the failure of ST-NM to capture the motion at frame 8 and 9.

A Fully 4D Expectation Maximization Algorithm Using Gaussian Diffusion Based Detector Response for Slow Camera Rotation Dynamic SPECT [1]

T. H. Farncombe†, M. A. King†, A. M. Celler‡, S. Blinder‡

† Department of Nuclear Medicine, University of Massachusetts Medical School, Worcester, MA, USA

‡ Division of Nuclear Medicine,Vancouver Hospital and Health Sciences Centre, Vancouver, B.C., Canada

Abstract

The dynamic expectation maximization algorithm (dEM) has been developed to determine the kinetic information of various metabolic processes from SPECT projection data acquired with a single slow camera rotation. Depth-dependent detector response compensation (DRC) has been tested as a means to improve the accuracy of these reconstructions. As well, by limiting the number of temporal frames in a dSPECT reconstruction, it is hoped that improved image quality will result. With this reduction in temporal frames, it becomes possible to use block-iterative type reconstruction methods to decrease reconstruction time.

To test the effect of DRC in the dEM algorithm, a series of computer simulations have been performed using the dMCAT phantom modified to model the kinetic response to Tc-99m Teboroxime. Reconstructions indicate that the inclusion of three dimensional DRC with the dEM algorithm improves image quality compared to no DRC. Furthermore, dynamic reconstructions are able to provide additional information, namely kinetic parameters, not attainable with static reconstruction methods. A reduction in the number of temporal frames reconstructed resulted in slightly increased image quality and time activity curve accuracy, but at the expense of decreased temporal resolution. Such a reduction may be acceptable if the temporal changes present are not great compared to the data acquisition time.

I. INTRODUCTION

In a conventional SPECT study, it is possible for the distribution of the radiotracer within the body to change over the acquisition time. If such a temporal change is significant, inconsistent projection measurements will result when acquired with a conventional slow camera rotation [1]. Often such a temporal change may be the result of physiologcal changes in the body related to organ function and as such, may provide information useful for diagnostic purposes. In order to obtain estimates of these dynamic processes, dynamic SPECT using multiple fast rotations of a SPECT camera has often been used in the past [2]. However, such data collection often results in a low signal to noise ratio, thus resulting in poor reconstructed images. Alternative methods of dynamic SPECT have been proposed that produce tomographic kinetic information, while maintaining a conventional slow camera rotation. We will denote such a data acquisition technique as dSPECT (dynamic SPECT) and will present in this paper, improvements to the dynamic expectation maximization algorithm (dEM) for determining the temporal behaviour of activity within an object from data acquired in such a manner.

In our method of dSPECT, temporal changes in the radiotracer distribution within an object are represented in terms of linear inequality constraints over time [3]. Once transformed into this representation, it is possible to use a modified version of the expectation maximization algorithm in order to reconstruct the changes in activity that occur over the data acquisition time.

In conventional SPECT, it has been determined that in most cases, improvements in reconstruction accuracy and lesion detectability result when the 3D depth-dependent nature of SPECT spatial resolution is taken into account in the reconstruction process [4]. This aids the reconstruction by making the projection data more consistent as each object voxel is sampled over a number of detector elements at each projection angle. This is similarly the case in dSPECT and it has been shown that reconstruction accuracy improves when multiple projection angles are acquired simultaneously at the same time point [3]. Thus, the inclusion of 3D detector response compensation (DRC) into the dEM algorithm should result in improved dSPECT reconstruction quality, and hence more accurate kinetic estimates.

Because of the rather large computational demand involved in reconstructing a dynamic SPECT data set, previous implementations of the dEM algorithm have focused on reconstructing transverse slices with depth-dependent detector-response within each slice only. As well, at each projection angle, a separate radiotracer distribution was reconstructed. While such a method results in high temporal resolution, this quickly produces an overabundance of data when reconstructed with fully 4D methods (eg., $64 - 128 \times 128 \times 128$ images $\approx$ 500MB of data). Because of this, when a high degree of temporal resolution is not required, it would be advantageous to reduce the number of temporal frames reconstructed. As well as reducing computer requirements, this should provide improved image quality as the data from each temporal frame is better supported across a subset of projection angles. Computation time can also be reduced in this scenerio as it allows for block-iterative reconstruction methods to be used as we shall see.

II. DYNAMIC EXPECTATION MAXIMIZATION

The dEM algorithm [5] reconstructs a series of three dimensional spatial images of the tracer distribution within the object. For example, from dSPECT projection data acquired of

---

[1] This work was supported in part by grant number HL 50349 from the National Heart, Lung and Blood Institute

a $128 \times 128 \times 128$ object over 64 projections, the current dEM algorithm reconstructs 64 different $128 \times 128 \times 128$ images, where each image corresponds to one of the 64 time frames when a projection was acquired.

A. Linear Inequality Temporal Constraints

For a given object voxel, any change in activity from one time frame to the next can be represented in terms of the activity difference over the time frames. Two possibilities exist for how the activity can vary over two consecutive time frames. These are:

i) Decreasing activity from one frame to the next.

$$x_0 \geq x_1 \geq 0 \quad or \quad x_0 - x_1 \geq 0 \tag{1}$$

ii) Increasing activity from one frame to the next.

$$0 \leq x_0 \leq x_1 \quad or \quad x_1 - x_0 \geq 0 \tag{2}$$

In both cases, the differences have been written so as to provide a positive quantity. As well, when the difference between two consecutive time frames is equal to zero, we have $x_0 = x_1$ (ie., static behaviour). Over all J time frames, we can denote the differences in activity within the kth voxel as the vector $\tilde{\mathbf{x}}_\mathbf{k}$, and relate it to the voxel activity $\mathbf{x}_\mathbf{k}$ by the matrix-vector product,

$$\tilde{\mathbf{x}}_\mathbf{k} = A_k \mathbf{x}_\mathbf{k} \tag{3}$$

where A_k is a matrix of size $(J \times J)$ and $\mathbf{x}_\mathbf{k}$ and $\tilde{\mathbf{x}}_\mathbf{k}$ are vectors of length J, representing the activity and the activity difference respectively, within the kth object voxel over the J time frames. It should be pointed out that each object voxel will have a corresponding difference matrix A_k, which can then be combined together in a large matrix A for the entire object.

The simple temporal constraints of (1) and (2) are sufficient to describe a variety of clinical circumstances (eg., washout from the kidneys, accumulation in the liver or bladder, etc.) [6], but they are not able to describe other cases comprised of a combination of both increasing and decreasing behaviour. For such cases, it is possible to formulate a difference matrix by using a combination of (1) and (2) that, when operating on the activity vector, will produce positive activity differences at each time and with the desired overall temporal behaviour. With a temporal link established through the use of the linear inequality temporal constraints, we can now turn our attention to determining the activity difference vector $\tilde{\mathbf{x}}_\mathbf{k}$. For this, we will make use of iterative reconstruction methods, particularly the expectation maximization algorithm [7], although it will be seen that any non block-iterative reconstruction method (eg., ART, MART, etc) will suffice at this point.

B. Dynamic Projection Operator

In the conventional expectation maximization (EM) algorithm, a forward projection operator is applied to the estimated activity distribution within the object at each projection angle. This estimated projection is then compared to the actual collected data for the same angle and then backprojected into the object space in order to arrive at a scaling factor for each object voxel. Mathematically, the projection and backprojection operations can be written as matrix-vector products so that the EM algorithm for SPECT can be written,

$$x_k^{n+1} = \frac{x_k^n}{\sum_{i,j=1}^{I,J} C_{ijk}} \sum_{i,j=1}^{I,J} \frac{C_{ijk} y_{ij}}{\sum_{k'=1}^{K} C_{ijk'} x_{k'}^n} \tag{4}$$

where x_k represents the estimated activity in the kth object voxel, C_{ijk} is the projection matrix that maps the voxelized three dimensional activity distribution into the two dimensional camera space, and y_{ij} represents the actual measured values in the ith detector element at the jth projection stop. Here the object activity distribution x_k is assumed to remain fixed over all the projection measurements acquired and so the size of the projection operator is $(I * J \times K)$ where I is the number of detector elements, J is the number of projection stops, and K is the number of object voxels in the reconstructed three dimensional object space. The projection operator, C_{ijk} can be written as a matrix comprised of the following:

$$C = \begin{pmatrix} (C_{i1k}) \\ (C_{i2k}) \\ \vdots \\ (C_{iJk}) \end{pmatrix} \tag{5}$$

where the (C_{ijk})'s represent the projection of the kth object voxel into the detector element i at the jth projection angle.

In the dynamic case however, the projection operator is no longer an $(I * J \times K)$ matrix since the object distribution is different at each projection angle, but rather is a $(I * J \times K * J)$ matrix. In the dynamic case, the dynamic projection operator (for a single detector head) can therefore be written as,

$$\tilde{C} = \begin{pmatrix} (C_{i1k}) & 0 & 0 & 0 \\ 0 & (C_{i2k}) & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & (C_{iJk}) \end{pmatrix} \tag{6}$$

The use of a Gaussian diffusion projection matrix for modelling the three dimensional depth dependent camera response has been shown to be both computationally efficient and accurate in static SPECT reconstruction algorithms [4]. For this reason, it was chosen to use this method in order to perform the projection and backprojection steps in the dEM algorithm. Due to space limitations, the reader is directed to [4, 8, 9] for more information on Gaussian diffusion DRC.

C. The Dynamic EM Algorithm

As mentioned previously, for any temporal behaviour, an appropriate linear difference matrix can be obtained so that the difference in activity between any two consecutive time frames will always be a positive quantity. As positivity constraints in the unknown variable are inherent in the expectation maximization algorithm, it is natural to proceed to use this algorithm in order to solve the dynamic SPECT problem.

Following along the lines of the static EM algorithm (4), a dynamic version can be obtained using the activity differences. Thus, the dEM algorithm can be written as,

$$\tilde{\mathbf{x}}_k^{n+1} = \frac{\tilde{\mathbf{x}}_k^n}{\sum_{i,j=1}^{I,J} (\tilde{C}A^{-1})_{ijk}} \sum_{i,j=1}^{I,J} \frac{(\tilde{C}A^{-1})_{ijk} y_{ij}}{\sum_{k'=1}^{K} (\tilde{C}A^{-1})_{ijk'} \tilde{\mathbf{x}}_{k'}^n} \quad (7)$$

where the vector $\tilde{\mathbf{x}}_k$ is the activity difference vector for the kth voxel and is related to the activity by equation (3). Notice that the projection step is actually a projection of the estimated activity distribution within the object at the appropriate time interval. Following reconstruction, the activity difference vector is converted into the activity vector by $\mathbf{x}_k = A_k^{-1} \tilde{\mathbf{x}}_k$.

Given the amount of data produced in a typical dSPECT scan, it is often desireable to reduce the number of unknown variables in order to reduce computational demands. In order to do this, when a high degree of temporal resolution is not required, the number of frames reconstructed can be reduced by altering the dynamic projection matrix such that the tracer distribution remains constant over a subset of angles. Thus, if the number of projection angles is J, and N is the desired number of temporal frames ($N \leq J$), then over each J/N angles, the object activity vector $\mathbf{x}$ can be held constant. In such a case, the projection matrix can be written,

$$\tilde{C} = \begin{pmatrix} (C_{i1k}) & 0 & 0 & 0 \\ \vdots & 0 & 0 & 0 \\ 0 & (C_{i(J/N+1)k}) & & \\ 0 & \vdots & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \vdots \\ 0 & 0 & 0 & (C_{iJk}) \end{pmatrix} \quad (8)$$

Similarly, the sizes of the difference matrix A_k and the activity vector $\mathbf{x_k}$ shrink to $(N \times N)$ and $(N \times 1)$ respectively. As each temporal frame is reconstructed based upon the projection data of a few angles, it is possible to apply block-iterative reconstruction methods at this point in order to decrease reconstruction time, for example, grouping the subsets as $\{(C_{i1k}), (C_{i(J/N+1)k}), ...\}$,$\{(C_{i2k}), (C_{i(J/N+2)k}), ...\}$, etc.

III. SIMULATIONS

To test the accuracy of the fully 4D dEM algorithm, simulations were performed using a version of the dMCAT [10] model modified to model the extraction and washout of Tc-99m Teboroxime [11]. Kinetic parameters used in this model were based on those found from compartmental modelling of Teboroxime within canine myocardium [2].

Two scenerios were simulated with the dMCAT consisting of a healthy myocardium and the same myocardium with an anterior wall defect. Projection data starting at 1 min post-injection was created using an analytic projector model of a dual head SPECT camera. Detector heads were placed in the 90° configuration and acquisition proceeded from LAO and RAO to LPO and LAO respectively for each head (ie., 90° rotation per head) in a circular orbit of radius 30 cm. A total of 32 projections per head (20 s per projection) were acquired into 64×64 pixel matrices with a pixel size of 6.25 cm. For both heads, a low energy, high resolution collimator was simulated with a FWHM of 1.59 cm at 30 cm. Noiseless projection data was generated and scaled to 5 million total counts and randomly generated Poisson noise added.

Reconstructions were carried out using the static MLEM and the dEM algorithms both with and without the inclusion of 3D detector response. dEM was used with high temporal resolution using all 32 time frames, as well as with 8 and 16 frames. In all cases, a total of 150 iterations were performed.

IV. RESULTS AND DISCUSSION

In Figure 1, the impact of using 3D DRC can be seen on a single short axis slice shown at t = 2 min. Reconstructions were performed both with and without 3D DRC. It is apparent from these reconstructions, that when DRC is applied to a dEM reconstruction, improved images result. In this time frame, it appears that the inferior wall of the myocardium is better distinguished in the DRC reconstruction. Additionally, contamination from the liver into the myocardium also appears to be lessened in the 3D DRC image compared to the no DRC image. In both cases, reconstructed time activity (TA) curves disagree to some extent with true TA behaviours, although the general trend for each organ is obtainable and in fact, the initial fast uptake of tracer into the myocardium is distinguishable in the first two time frames.

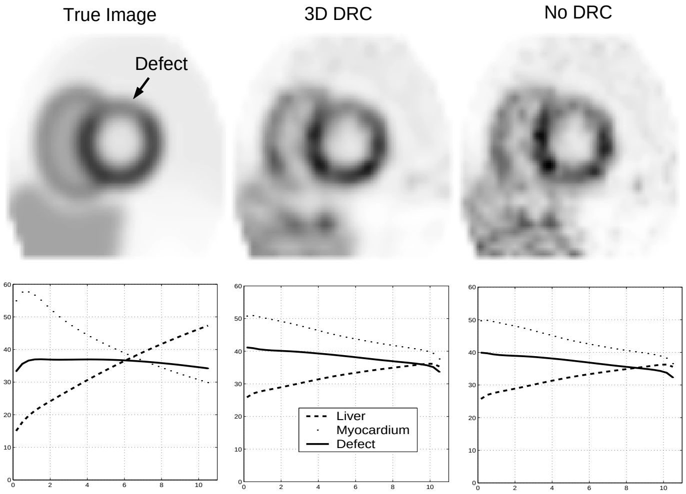


Figure 1: Short axis slice of dMCAT phantom at t = 2 min obtained from dEM reconstruction with and without DRC.

Figure 2 shows sample images of the same short axis slice reconstructed with differing numbers of temporal frames. Images produced from 16 or 8 frames appear very similar, but slight differences can be seen between the dEM and static MLEM reconstructions. At the time frame shown, contrast between the myocardium and the defect appears slightly greater in the dEM reconstructions compared to the MLEM result. Additionally, the effect of activity uptake into the liver is reduced in dEM vs static reconstructions as the dynamic reconstruction is able to account for the increasing liver activity over time, thus minimizing streak artifacts in the inferior heart wall. Regional TA curves appear very similar in all dEM reconstructions, although the initial fast myocardial uptake is no longer apparent when decreased to 16 temporal frames.

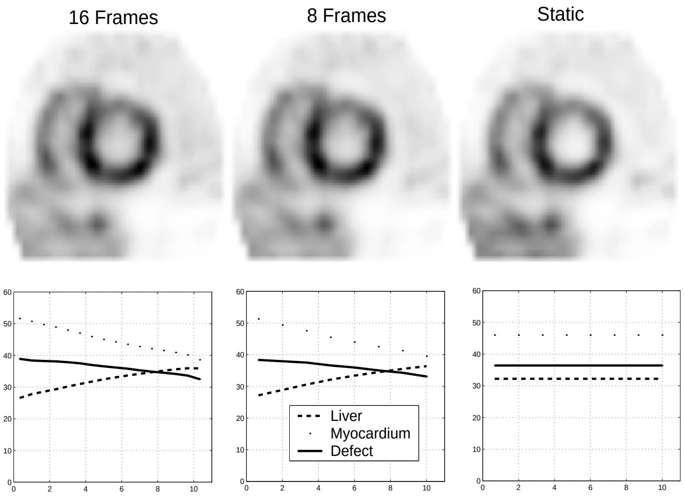


Figure 2: Short axis slice of dMCAT phantom at t = 2 min obtained from dEM reconstruction with fewer temporal frames and with DRC.

It was seen in this simulation that the location of the myocardial defect was able to be determined through both dynamic and static reconstructions. With static methods, images depict an average activity within each object voxel. In this simulation, the data acquisition was started soon enough and was short enough so that the average activity within the myocardium was higher than that in the defect. However, if the acquisition is started later, or if the acquisition time is longer, a static reconstruction may depict the defect with a greater activity, thus giving the appearance of a healthy myocardium. If reconstructed with a dynamic method such a dEM, this effect will not occur as the defect and healthy myocardium can be distinguished based on their dynamic parameters. However, as one decreases the number of temporal frames in a dEM reconstruction, this effect may become more important.

V. CONCLUSION

A fully 4D dynamic expectation maximization algorithm has been presented for use in dynamic SPECT imaging using slow acquisitions. Depth-dependent spatial resolution is modelled through an incremental Gaussian diffusion, while temporal constraints are enforced in the object at each time frame. These constraints can model a wide range of temporal behaviours, and can be reduced in number in order to decrease reconstruction time and computer requirements. Additionally, reducing temporal frames allows for the possibility of performing an adaptive framing method whereby temporal frames are finely sampled when rapid changes in activity occur and are more coarsely sampled during slow changes.

In computer simulations, the inclusion of DRC in this algorithm has resulted in increased reconstruction accuracy compared with no DRC. As well, by limiting the number of temporal frames in the reconstruction, slightly improved reconstructions result as each reconstructed frame is supported by multiple projection angles. However, one must be cautious of reducing the number of temporal frames too much as a resultant decrease in temporal resolution follows.

VI. REFERENCES

[1] B. Bok, A. Bice, M. Clausen, D. Wong, and H. Wagner, “Artifacts in camera based single photon emission tomography due to time activity variation,” *Eur. J. Nucl. Med.*, vol. 13, pp. 439–442, 1987.

[2] A. Smith, G. Gullberg, P. Christian, and F. Datz, “Kinetic modeling of teboroxime using dynamic SPECT imaging of a canine model,” *J. Nuc. Med.*, vol. 35, pp. 484–495, 1994.

[3] T. Farncombe, *Functional Dynamic SPECT Imaging Using a Single Slow Camera Rotation*. PhD thesis, University of British Columbia, 2000.

[4] H. C. Gifford, M. A. King, R. G. Wells, W. G. Hawkins, M. V. Narayanan, and P. H. Pretorius, “LROC analysis of detector-response compensation in SPECT,” *IEEE Trans. Med. Imag.*, vol. 19, pp. 463–473, 2000.

[5] T. Farncombe, S. Blinder, A. Celler, D. Noll, J. Maeght, and R. Harrop, “A dynamic expectation maximization algorithm for single camera rotation dynamic SPECT (dSPECT),” *IEEE Nuclear Science Symposium Conference Record*, 2000.

[6] A. Celler, T. Farncombe, C. Bever, D. Noll, J. Maeght, R. Harrop, and D. Lyster, “Performance of the dynamic single photon emission computed tomography (dSPECT) method for decreasing or increasing activity changes,” *Phys. Med. Biol.*, vol. 45, pp. 3525–3543, 2000.

[7] L. Shepp and Y. Vardi, “Maximum liklihood reconstruction for emission tomography,” *IEEE Trans. Med. Imag.*, vol. MI-1, pp. 113–122, 1982.

[8] A. W. McCarthy and M. I. Miller, “Maximum likelihood SPECT in clinical computation times using mesh-connected parallel computers,” *IEEE Trans. Med. Imag.*, vol. 10, pp. 426–436, 1991.

[9] M. A. King, T.-S. Pan, and D.-S. Luo, “An investigation of aliasing with Gaussian-diffusion modelling of SPECT spatial resolution,” *Trans. Nucl. Sci.*, vol. 44, pp. 1375–1380, 1997.

[10] P. Pretorius, W. Xia, M. King, B. Tsui, and K. Lacroix, “A mathematical model of motion of the heart for use in generating source and attenuation maps for simulating emission imaging,” *Med. Phys.*, vol. 26, pp. 2323–2332, 1999.

[11] A. Celler, S. Blinder, D. Noll, T. Tyler, F. Duclercq, and R. Harrop, “Investigation of the dynamic SPECT (dSPECT) method for teboroxime using a 4-D kinetic thorax model dMCAT,” in *Sixth International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, submitted*, 2001.

Transfer Function Restoration in 3D Electron Microscopy via Iterative Data Refinement

C.O.S. Sorzano(1), J.J. Fernández(2), R. Marabini(3), G.T. Herman(3),Y. Censor(4)
and J.M. Carazo(1)

5th April 2001

(1)Centro Nacional de Biotecnología, Campus Universidad Autónoma s/n, 28049 Cantoblanco, Madrid, Spain.

(2)Dept. Arquitectura de Computadores y Electrónica. Universidad de Almería. 04120 Almería, Spain.

(3)Center for Computer Science and Applied Mathematics, Temple University, Philadelphia, PA, 19122-6094, USA.

(4)Dept. of Mathematics, University of Haifa, Mt. Carmel, Haifa 31905, Israel.

Abstract

Three-dimensional electron microscopy (3D-EM) aims at obtaining structural information of macromolecular complexes within a typical resolution range of between 2 to 0.3 nm from the projection images produced by an electron microscope. As any other imaging device, the electron microscope introduces a transfer function (called in this field Contrast Transfer Function, CTF) into the image acquisition process which modulates the different frequencies of the projection signal. Thus, the 3D reconstructions performed with these CTF-affected projections is also affected by an implicit 3D transfer function. Depending on the preparation procedure, the effect of this CTF is quite dramatic limiting severely the achievable resolution. In this work we make use of the *Iterative Data Refinement* technique to obtain CTF-free reconstructions. It is shown that the approach can be successfully applied to noiseless as well as to noisy data.

1 Introduction

The structural information of biological complexes, i.e., their shape and spatial conformation, is vital in molecular biology as complementary information to biochemical studies. This knowledge can help, for example, in new drug development and in the understanding of many diseases. Nowadays, the 3D structure of a protein can be addressed using two different approaches: one predicts the conformational state of the complex based on its biochemical properties and the possible similarity with other proteins whose structure is known; and the other applies 3D reconstruction algorithms to data collected by some experimental technique. There exist several possible biophysical techniques for visualizing a protein, such as X-rays, NMR or electron microscopy (EM). The main drawbacks of the first two are that they are very restrictive with respect to the range of proteins that can be studied. 3D-EM is a technique which provides only medium-low resolution structural information of macromolecules, however, it does not need special conditions of the specimens and it has been established as a useful technique in the field of structural biology.

One of the limiting factors of 3D electron microscopy is that it is difficult to obtain high resolution structural details due to the strong effect of the microscope transfer function on the experimental projections. Particularly, as will be shown later, the CTF introduces severe phase shifts and eliminates all information at certain frequencies. One way of obtaining higher resolution results is by compensating for the effect of the microscope transfer function. Several such methods have been proposed

[1, 6, 7, 12, 14, 16], but they usually apply some kind of Wiener-like division by the transfer function in the Fourier space, and also amplify noise at those frequencies where the transfer function has small magnitude.

In this work we apply the Iterative Data Refinement (IDR) technique introduced by Censor, Elfving and Herman [2] and further studied by Herman and co-workers [8, 9, 11] to remove the effect of the microscope imperfection and, thus, obtain high resolution structural information about the macromolecules under study. Although in this work only simulations with phantoms have been carried out, the results are encouraging enough to justify future tests on experimental data sets.

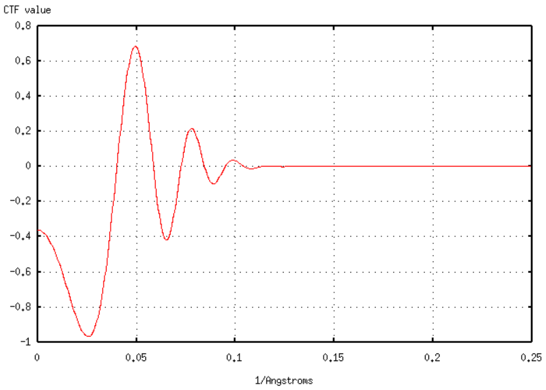


Figure 1: CTF used in the cryomicroscopy simulations.

2 Materials and methods

Contrast Transfer Function

Image formation by an the electron microscope is due to two different physical processes during the electron interaction with the specimen: first, a shift in the electrons phase, and second, an electron direction change. Both effects are combined to produce a single modulation transfer function called Contrast Transfer Function [5, Chapter 2.II]. A cross-sectional plot of such a modulating function, typical for cryomicroscopy, can be seen in Figure 1. Notice that the sign changes and the crossings of 0 are responsible for a contrast inversion in the projection image and for the complete elimination of the information at certain frequencies. A parametric model of this transfer function is given in [15], and it is used in the simulations presented in this work.

Iterative Data Refinement

This technique was first introduced in Censor [2] in the context of 2D medical tomography, see Section 10.5 in [3]. The underlying idea is to change iteratively the projection images so that the process converges to the ideal (without transfer function) projections. Calling g_i^k the projection in the direction i at the iterative step k, g^k the set of projections at step k, P_i the projection operator in the direction i, C_i the contrast transfer function operator for that image, and R the reconstruction operator, the IDR algorithm is formulated as $g_i^{k+1} = \mu^{k+1} g_i^0 + (P_i - \mu^{k+1} C_i P_i) R g^k$ where μ^{k+1} is a relaxation parameter applied at iteration $k+1$ and g^0 represents the set of experimental images. Basically, the procedure proceeds as follows: a first reconstruction is done with the experimental images g^0, then every projection is modified following a mixture of the original projections, the reprojection of the just reconstructed volume and this same reprojection after the CTF. The new set of images is again used for reconstruction and so on. ART+blobs [10] has been used as the reconstruction operator R.

Simulations

For the sake of objective assesment of quality, simulations with a phantom have been run following the Figures of Merit approach described in Sorzano [13]. The phantom corresponds to bacteriorhodopsin, whose structure is known at atomic resolution (3.5Å) [4]. The surface rendering of this structure, calculated from a volume sampled at 2Å/pixel, can be seen on Figure 2.

Computational projections have been done simulating the cryomicroscopy conditions. In cryomicroscopy the specimens are embedded in ice, and are then imaged at very low electron doses to preserve structural details**.** As a consequence, the projection images that are obtained with this technique are extremely noisy, with signal-to-noise ratios lower than 1.0. The CTF that has been used for the projections is shown in Figure 1. 2000 images were taken all over the projection space, a selection of these projections is shown in Figure 3. Angular Gaus-

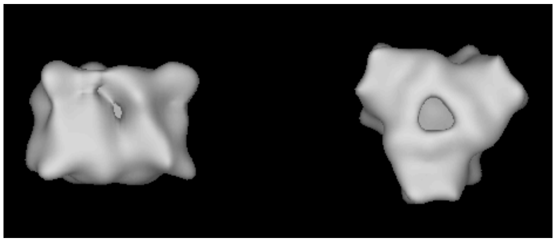


Figure 2: Bacteriorhodopsin phantom filtered at 17Å (maximum resolution permitted by the simulated CTF).

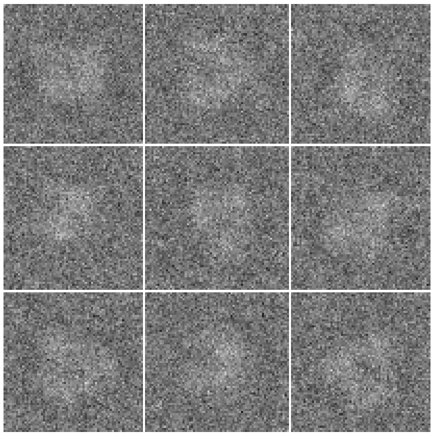


Figure 3: A selection of projections simulating cryomicroscopy images from the bacteriorhodopsin phantom.

sian noise (N(0,5)) and shift Gaussian noise (N(0,2)) have been added to simulate the positional uncertainty in 3D electron microscopy.

3 Results

Three different reconstructions have been performed to test the efficacy of the IDR algorithm: the first one without any CTF correction, the second one with phase correction, and the third with amplitude and phase correction via IDR. Figure 4 shows the results for each case while Table 1 shows the L2 and L1 measures of the error [13] between each reconstruction and the original phantom and the maximum resolution [5, Chapter 5.V] achieved with each method. Table 1, as well as the 3D reconstructions performed with the different degrees of CTF correction, show that the IDR technique achieves better reconstructions with higher resolution.

Corrective action	L2error	L1 error	Resolution
No action	0.995844	0.961378	28Å
Phase corrected	0.996125	0.963414	20Å
IDR	0.997908	0.967625	17Å

Table 1: Reconstruction similarity measure after different corrective actions.

4 Conclusions

The IDR's ability of removing the transfer function effect with noiseless data has been proved in [9]. In this work we have extended those results to extremely noisy data. At the same time, we have shown the importance of applying CTF correction to 3D electron microscopy in order to obtain high resolution reconstructions. Further work must be done to tune the IDR free parameters to the specific cases of cryomicroscopy and negative staining. However, these preliminary results encourage us to proceed further on experimental data sets with this CTF correction technique which allows the application of particular transfer functions to each projection and which avoids the noise amplification effect caused by most other correction methods used so far.

Acknowledgements

Dr. Marabini is partially supported by a grant from the Spanish "Ministerio de Educación y Ciencia". Partial support is acknowledged to the "Comisión Interministerial de Ciencia y Tecnología" of Spain through project BIO980761 and TIC990361. The work of Y. Censor was supported by Research Grant 592/00 from the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities.

References

[1] N. Boisset and F. Mouche. *Sepia officinalis Hemocyanin*: A refined 3D structure from Field Emission Gun microscopy. *J. Mol. Biol.*, 296:459–472, 2000.

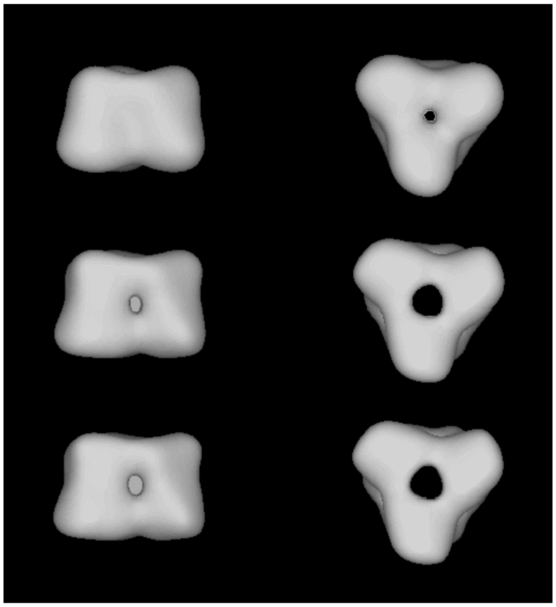


Figure 4: From top to bottom: Reconstruction without CTF correction, reconstruction with phase correction, reconstruction with amplitude and phase correction.

[2] Y. Censor, T. Elfving, and G.T. Herman. A method of iterative data refinement and is applications. *Math. Meth. Appl. Sci.*, 7:108–123, 1985.

[3] Y. Censor and S.A Zenios. *Parallel optimization: theory, algorithms and applications*. Oxford University Press, New York, NY, USA, 1997.

[4] T.A. Ceska, R. Henderson, J.M. Baldwin, F. Zemlin, E. Beckmann, and K. Downing. An atomic model for the structure of bacteriorhodopsin, a seven-helix membrane protein. *Acta Physiol. Scand. Suppl.*, 607:31–40, 1992.

[5] J. Frank. *Three Dimensional Electron Microscopy of Macromolecular Assemblies*. Academic Press, San Diego, 1996.

[6] J. Frank and P. Penczek. On the correction of the contrast transfer function in biological electron microscopy. *Optik*, 98:125–9, 1995.

[7] N. Grigorieff. Three-dimensional structure of bovine NADH : Ubiquinone oxidoreductase (complex I) at 22 Å in ice. *J. Mol. Biol.*, 277:1033–1046, 1998.

[8] G.T. Herman. Standard image recovery methods in the Iterative Data Refinement framework. In *Proc. SPIE*, number 1199, pages 1501–1511, 1989.

[9] G.T. Herman and D.W. Ro. Image recovery using iterative data refinement with relaxation. *Opt. Eng.*, 29:513–523, 1990.

[10] R. Marabini, G.T. Herman, and J.M. Carazo. 3D reconstruction in electron microscopy using ART with smooth spherically symmetric volume elements (blobs). *Ultramicroscopy*, 72:53–65, 1998.

[11] D.W. Ro, G.T. Herman, and P.M. Joshep. Resolution enhancement of magnetic resonance images using an iterative data refinement technique. In *Proc SPIE*, volume 1199, pages 952–962, 1989.

[12] U. Skoglund, L.G. Ofverstedt, R. Burnett, and G. Bricogne. Maximum-entropy three-dimensional reconstruction with deconvolution of the contrast transfer function: A test application with adenovirus. *J. Struct. Biol.*, 117:173–188, 1996.

[13] C.O.S. Sorzano, R. Marabini, N. Boisset, E. Rietzel, R. Schröder, G.T. Herman, and J.M. Carazo. The effect of overabundant projection directions on 3D reconstruction algorithms. *J. Struct. Biol.*, 2001. (in press).

[14] H. Stark, EV. Orlova, J. Rinke-Appel, N. Junke, F. Mueller, M. Rodnina, W. Wintermeyer, R. Brimacombe, and M. van Heel. Arrangement of tRNAs in pre- and posttranslocational ribosomes revealed by electron cryomicroscopy. *Cell*, 88:19–29, 1997.

[15] Z.H. Zhou, S. Hardt, B. Wang, M.B. Sherman, J. Jakana, and W. Chiu. CTF determination of images of ice-embedded single particles using a graphics interface. *J. Struct. Biol.*, 116:216–222, 1996.

[16] J. Zhu, P.A. Penczek, R. Schröder, and J. Frank. Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: Procedure and application to the 70S *Escherichia coli* ribosome. *J. Struc. Biol.*, 118:197–219, 1997.

Implementation of Continuous 3D Whole Body PET Scanning Using On-the-fly Fourier Rebinning

Keishi Kitamura, Kazumi Tanaka, and Tomohiko Sato
Shimadzu Corporation, 1 Nishinokyo-kuwabara-cho, Nakagyo-ku, Kyoto, 604-8511 Japan

Abstract

Continuous scanning mode in 3D whole body PET studies has the advantage of axial sensitivity uniformity over the majority of the axial FOV. However this scan mode requires large data handling compared to conventional discrete scans. In this work, we have implemented and evaluated a new continuous 3D scan method using 'on-the-fly' Fourier rebinnig. In this method, sinograms for the pair of rings are added in real-time into the sinograms of the incremented ring pairs by moving the bed axially one detector width at a time. For a *N*-ring scanner, *2N-1* sinograms are transferred to a host computer at each bed position and rebinned into direct 2D sinograms using Fourier rebinning. Phantom and human studies showed that the axial image uniformity is achieved. This method can minimize the time for offline data processing and makes the continuous 3D scan more practical in clinical whole body studies.

I. INTRODUCTION

3D whole body PET scan offers the advantage of high sensitivity and can reduce patient dose while maintaining high signal to noise ratio (S/N). In conventional whole body scan, acquiring at distinct bed positions overlapping several slices results in an axially varying sensitivity profile due to the cylindrical geometry. In order to minimize the axial non-uniformity, it is necessary to optimize the number of overlapping end slices [1] and maximum ring differences being incorporated in the final images.

Several investigators have shown that the continuos axial sampling method is effective for the elimination of the axial S/N non-uniformity in the reconstructed images [2, 3, 4, 5]. This sampling method can be implement using conventional histogram mode acquisitions by moving the bed in a small discrete step or using list mode acquisitions by moving the bed in a true continuous motion. However, due to the vast amount of data produced by this scan mode, the total axial distance that can be covered is limited by the available memory and disk space. And the computation time for data to be re-sorted and added into sinograms may decrease the throughput of the continuous 3D scans.

In this work, we report on a new implementation of the 3D continuous scan mode using 'on-the-fly' Fourier rebinning, in which the bed is moved by axial width of detectors and synthesized sinogram data sets are transferred and rebinned into 2D sinograms at each bed position. This method does not require large memory and disk space and can improve data processing efficiency.

II. MATERIAL AND METHODS

A. Data Acquisition

All data were acquired on a Shimadzu SET-2400W PET scanner [6], which is a 32-ring system with an axial detector width of 6.25 mm and an axial FOV of 200 mm. In 3D mode, all possible coincidence pairs of 1024 sinograms are acquired. The front-end data acquisition system consists of large-scale acquisition memories (1GB) and a microprocessor that controls data collection in histogram mode, real-time corrections for dead time and decay of radioisotopes, and data transfer to a UNIX host computer with a storage disk. The acquisition software was modified for the continuous scan mode to move the bed by the axial width of detector, shift the memory address for sinogram matrices at each bed position, and write the data acquired to the disk sequentially, as described below.

For a *N*-ring scanner, N^2 sinograms $q_k(s, \phi, r_1, r_2)$ are acquired for the pair of rings (r_i, r_j) at the *k*th bed position. In the continuous scan mode, $q_k(s, \phi, r_1, r_2)$ are added into *2N-1* sinogram data set $p_h(s, \phi, \Delta r)$ in real-time as follows:

$$p_h(s, \phi, \Delta r) = \sum_{k=k_{\min}}^{k_{\max}} q_k(s, \phi, r_1 + k, r_2 + k) \qquad (1)$$

Where

$h = 0, 1, 2, \cdots, N + n - 1$,

$\Delta r = (r_1 - r_2) = 0, \pm 1, \pm 2, \cdots, \pm \Delta r_{\max}$,

$k_{\min} = \max\{0, h - n + 1\}$,

$k_{\max} = \min\{h, N - |\Delta r| - 1\}$.

In this work, we used $\Delta r_{\max} = N - 1$. With *n* bed positions, *N+n–1* data sets are acquired and each data set $p_h(s, \phi, \Delta r)$ consists of *2N-1* sinograms except for last *N-1* data sets. An example of the data sets is shown in Figure 1.

Lines of response (LOR's) corresponding to larger ring differences (larger azimuth angle) are measured less frequently than those with smaller ring differences. In general, LOR's with ring differences of Δr are measured $N-|\Delta r|$ times. In our implementation, the sets of ring difference data remain unscaled relative to each other. This means the sinograms being weighted by the number of times they are measured, which is able to improve S/N ratio in the final reconstructed image [3].

After the acquisition at the *k*th bed position, $p_h(s, \phi, \Delta r)$ is transferred to the host computer. After the data transfer is completed, the memory buffer can be reuse for a new data set. As a result, the total size of sinogram data sets for a 3D continuous scan is given by $(2N\text{-}1)\times(N+1)\times N_b$, not depending on the number of bed movements, where N_b is the number of bytes of memory buffer per sinogram.

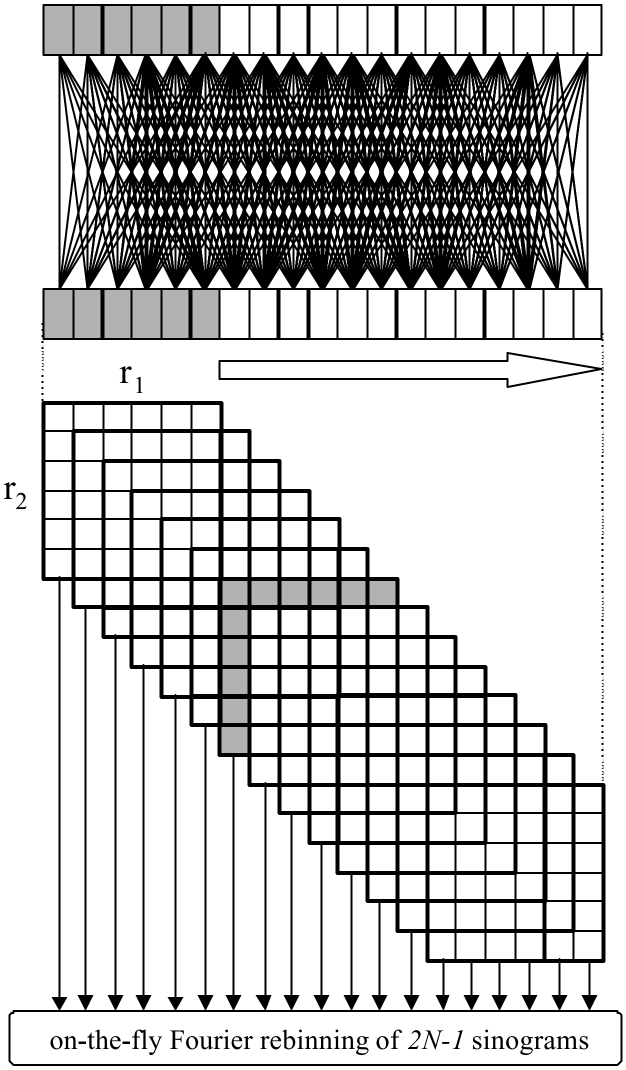


Fig. 1: Schematic representation of the lines of response and 3D data matrices for the continuous 3D whole body scan with a N = 6-ring scanner, which is indicated by the shaded detectors. In the 3D data space, each square corresponds to one oblique sinogram for the pair of rings (r_1, r_2), and each 6x6 matrix corresponds to sinogram data set at each bed position. The shaded area in the sinogram matrices is an example of synthesized sinogram data set, which is transferred and processed at each bed position.

B. Data Processing

To correct the transferred *2N-1* sinograms for detector efficiency variations, normalization correction factors are calculated from the conventional 3D normalization data by averaging normalize factors of the same ring differences in a similar way as (1). Attenuation correction factors are calculated from attenuation maps, which are reconstructed and assembled from 2D whole body transmission scans with conventional discrete bed motions.

The corrected sinograms are rebinned into 2D direct sinograms using Fourier rebinning [7]. In a given oblique sinogram, the azimuth angle is approximately constant, so a 1D interoperation across different oblique sinograms is not required. As a result, one data set transferred at each bed position can be processed independently, because it includes pairs of oblique sinograms with opposite values of Δr, which are merged to calculate discrete 2D Fourier transform as equation (6) in [7]. The resulting 2D data sets are reconstructed using 2D filtered backprojection or ordered subset EM (OSEM) algorithm [8].

C. Phantom and Human Studies

Phantom studies were performed in order to determine the axial uniformity in S/N using a 3D continuous data collection protocol compared with the conventional 3D whole-body scan protocol in the same total scan time. A 15cm diameter, 60 cm high cylindrical phantom containing about 2 mCi of F-18 were acquired. In the conventional discrete scans, data were acquired with 200 sec x 3 bed positions overlapping end slices of 5 or 10, and maximum ring difference of 22 was used for reconstruction. In the continuous scan, data were acquired with 10 sec x 60 bed positions. The percent standard deviation of the reconstructed images were calculated by drawing a 8 cm diameter ROI at the center of each image plane.

3D Whole body scans of a normal subject were also acquired in the conventional discrete scan and the continuous scan. In the discrete scan mode, emission data were acquired with 4 min x 3 bed positions overlapping 5 slices, 90 min after injection of 4 mCi ^{18}FDG. Transmission data were extracted from a 2D simultaneous emission and transmission scan acquired with 2 min x 3 bed positions. Attenuation correction factors were calculated from the reconstructed attenuation map processed with the non-linear Gausssian filters [9]. In the continuous scan mode, emission data were acquired with 12 sec x 60 bed positions. Emission images were in all cases reconstructed using OSEM reconstruction with 1 iteration and 24 subsets [6]. A post-reconstruction 2D Butterworth filter with 16 mm cut-off was applied to the images. No axial smoothing and weighted summation of overlapped images were performed.

III. RESULTS

Figure 2 shows the axial variation in standard deviation of the uniform cylinder for the discrete scan (solid line) with different overlap and for the continuous scan (circles). These plots show that the continuous scan provide a better noise uniformity, while noise is amplified considerably near the gaps between bed positions in the discrete scan. Using 10 slice overlap, axial nonuniformity is slightly improved while the axial coverage is shortened.

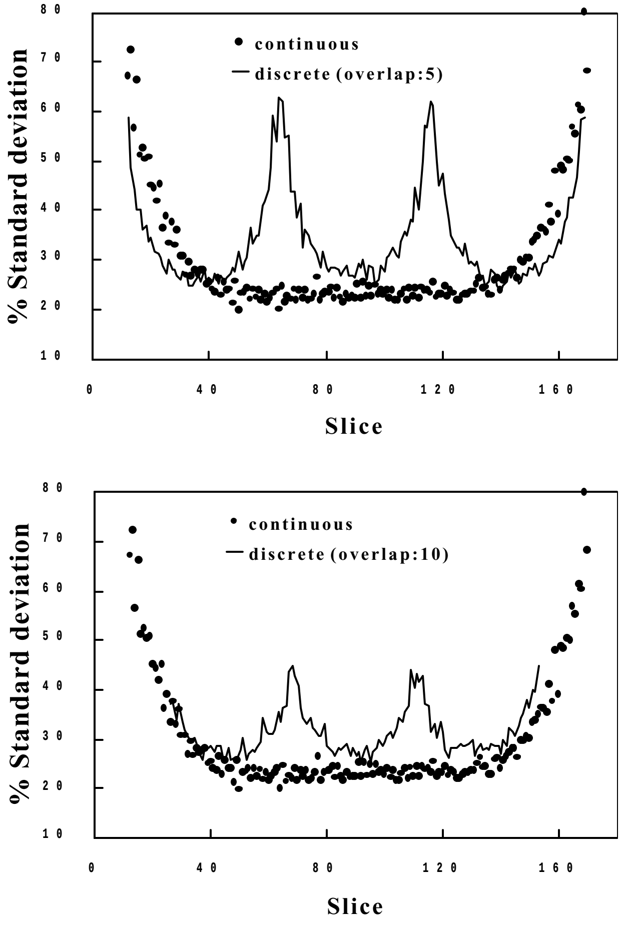


Fig. 2: Standard deviation/Mean of the 3D cylinder in a 10 cm diameter ROI placed at the center of the reconstructed image. The data were acquired using the discrete scan (solid lines) with overlapped slices of 5 (top) and 10 (bottom) and using the continuous scan (circles).

Figures 3 and 4 show coronal and sagittal cross sections from the normal subject scan using the conventional discrete scan (left) and the continuous scan (right). The images in the top rows have not been attenuation corrected and the images in the bottom rows have been attenuation corrected using the same attenuation map for the both scan modes.

As shown in these figures, in the discrete scan, there is an increase in noise and visual artifacts at the end of the slices of each axial FOV indicated by the arrows. In the continuous scan, these noise amplifications were eliminated and the spine appears much more clearly, indicating the overall improvement in S/N.

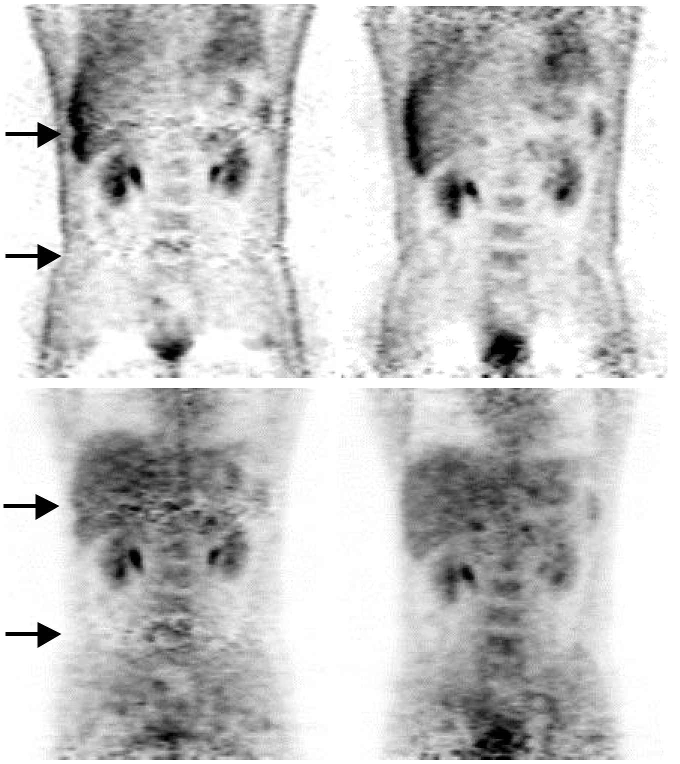


Fig. 3: Coronal cross sections through a normal subject using the discrete scan (left) and the continuous scan (right). The bottom images have been attenuation corrected.

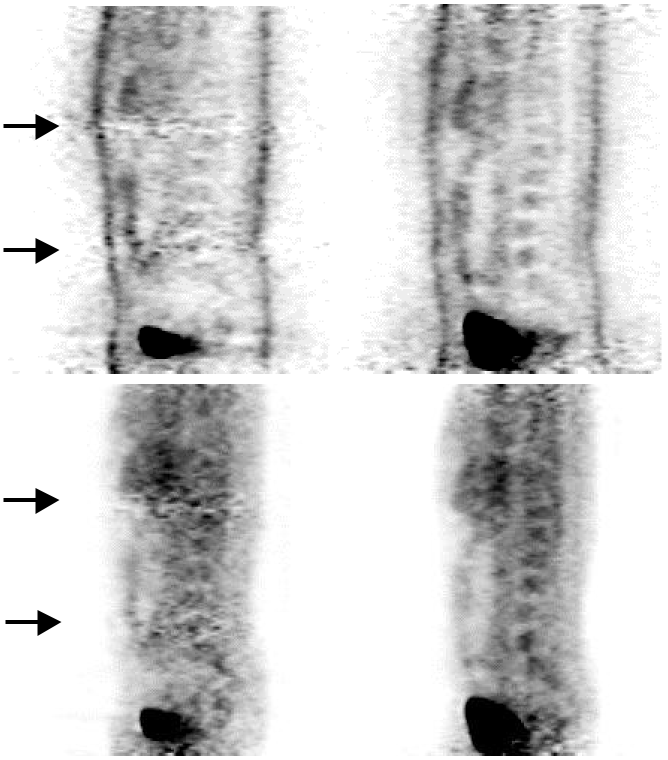


Fig. 4: Sagittal cross sections through a normal subject using the discrete scan (left) and the continuous scan (right). The bottom images have been attenuation corrected.

IV. DISCUSSION

The main advantage of the continuous scan is the axial S/N uniformity as shown in Fig. 2. In addition, the overall improvements in S/N of the whole body images can be seen in Fig. 3 and Fig. 4. In the continuous scan, all the data points in the central slices are sampled with all the detectors in the axial direction hence the detector efficiencies in the axial direction are averaged. As a result, S/N of the data corrected for detector efficiencies is slightly improved [10].

In this work, we have implemented continuous 3D whole body PET scanning using on-the-fly Fourier rebinning. This implementation has only required a slight modification to the acquisition software and a small size of acquisition memory and storage disk not depending on the axial coverage. Compared to the complete 3D data set, the size of sinogram data set that is transferred and processed at each bed position is small enough to be easy to handle. The program can be easily modified to improving axial sampling by decreasing the bed motion step and increasing the number of data set. But an axial sub-sampling is usually unnecessary in clinical whole body studies, because the axial resolution improvement is not significant [5].

The one drawback of the small step bed motion is that there is a dead time introduced between each bed motions. During the patient scan, it was found that this dead time loss is less than 8% of total acquisition time, which is not critical in clinical studies. It may be possible to modify the acquisition software for the microprocessor to move the bed in a true continuos motion and to be triggered at sampling time intervals to synthesize continuos sinograms.

V. SUMMARY

We have implemented the continuous scan for 3D whole body PET studies using the on-the-fly Fourier rebinning, in which sinograms are synthesized in the acquisition memory and transferred and processed after the acquisition at each bed position. Improved data processing efficiency can make the continuous 3D whole body scan more practical in routine clinical studies, while maintaining the axial S/N uniformity

VI. ACKNOWLEDGEMENTS

The authors wish to thank Mr. Masahiko Komatsu from Shimadzu SD Corporation for providing assistance in the implementation of the acquisition software, and staff of Yokohama City University Hospital for the acquisition of data.

VII. REFERENCES

[1] P. D. Cutler and M. Xu, "Strategies to improve 3D whole-body PET image reconstruction," Phys. Med. Biol. vol.41, pp.1453-1467, 1996

[2] M. Dahlbom, D.-C. Yu, S. Cherry, A. Chatziioannou, and E. J. Hoffman, "Methods for improving image quality in whole body imaging," IEEE Trans. Nucl. Sci., vol. NS-39, pp. 1079-1083, 1992.

[3] S. R. Cherry, M. Dahlbom, and E. J. Hoffman, "High sensitivity, total body PET scanning using 3D data acquisition and reconstruction," IEEE Trans. Nucl. Sci., vol.39, no.4, pp.1088-1092, 1992

[4] M. Dahlbom, P. Cutler, W. Digby, P. Luk, and J. Reed, "Characterization of sampling schemes for whole body PET imaging," IEEE Trans. Nucl. Sci., vol. NS-41, pp. 1571-1576, 1994.

[5] M. Dahlbom, J. Reed, and J. Young, "Implementation of true continuous 2D/3D whole body PET Scanning," Conference Record of the IEEE Nuclear Science Symposium and Medical Imaging Conference, Lyon, 2000

[6] T.Fujiwara, S Watanuki, S Yamamoto, M Miyake, S Seo, M Itho, K Ishii, H Orihara, H Fukuda, T Satoh, K Kitamura, K Tanaka, and S Takahashi, "Performance evaluation of a large axial field-of-view PET scanner:SET-2400W," Annals of Nuclear Medicine, Vol.11, No.4, pp.307-313, 1997

[7] M. Defrise, P. E. Kinahan, and D. W. Townsend, "Exact and approximate rebinnig algorithm for 3-D PET data," IEEE Trans. Med. Imag., vo.16, no.2, pp.145-157 ,1997

[8] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," IEEE Trans. Med. Imag., vol. MI-13, pp. 601-609, 1994.

[9] K. Kitamura, H. Iida, M. Shidahara, S. Miura, and I. Kanno, "Noise reduction in PET attenuation correction using non-linear Gaussian filters," IEEE Trans. Nucl. Sci., vol.47, no.3, pp.994-999, 2000

[10] A Chatziioannou and M Dahlbom, "Study on the effects of whole body PET spatial sampling schemes on data SNR", IEEE Nuclear Science Symposium and Medical Imaging Conference Record, pp. 1295-1299, vol.2, 1996

Scatter Correction for Positron Emission Mammography

Jinyi Qi and Ronald H Huesman

Abstract—**We have previously presented a regularized list mode maximum likelihood reconstruction algorithm for the positron emission tomograph that is being developed at our laboratory. Here we will present a scatter correction method for this algorithm. The mean scatter sinogram is estimated using a Monte Carlo simulation program. It is then incorporated into the forward model of the reconstruction algorithm. With the assumption that the background activity is nearly uniform, the Monte Carlo scatter simulation need only run once for each PEM configuration. This saves computational time and makes the Monte Carlo scatter correction viable in clinical situations. The propagation of the noise from the estimated scatter sinogram into the reconstruction is theoretically analyzed. The results provide an easy way to calculate the required number of events in the Monte Carlo scatter simulation for a given noise level in the image. The analysis is also applicable to other scatter estimation methods provided that the covariance of the estimated scatter sinogram is available.**

I. INTRODUCTION

A rectangular positron emission tomograph (Fig. 1), dedicated to imaging the human breast, is under development at our Laboratory [1]. The tomograph consists of four banks of detector modules (two banks of 3×3 modules left and right and two banks of 3×4 modules top and bottom). Each module consists of an 8×8 array of 3mm×3mm×30mm lutetium oxyorthosilicate (LSO) crystals. The maximum field of view (FOV) of the system is 96×72×72 mm^3. For each crystal, the system digitizes the depth of interaction (DOI) of the photon with three bits. Each detector is placed in coincidence with all detectors in the other three banks, giving rise to 172 million possible lines of response (LORs). The system operates exclusively in fully 3D mode.

The data from the new tomograph is stored in list mode format because the total number of detections will generally be far less than the total number of LORs. We have developed a list mode likelihood reconstruction algorithm for the tomograph [2]. The DOI information was explicitly modeled in the forward projection for each LOR.

Here we present a scatter correction method for the list mode likelihood reconstruction algorithm. The scatter sinogram is estimated using a Monte Carlo simulation program. The results are then incorporated in the forward model of the reconstruction algorithm. For breast imaging with F-18-labeled deoxyglucose (FDG), we can assume that the FOV is filled with uniform activity and that features such as cancerous lesions account for a small fraction of the radioactivity, so we only need to run the Monte Carlo scatter simulation once for each system configuration. For each data set, we can scale the scatter sinogram by the ratio of the total events between the Monte Carlo simulation and the data set, assuming the scatter profiles are the same. This saves a large amount of computation time. We also theoretically analyzed the noise propagation from the Monte Carlo scatter sinogram into the reconstructed image. Such analysis is also useful in determining the total number of events that is required for the Monte Carlo scatter simulation. Using Monte Carlo simulation to estimate scatter sinogram is not a new idea. It is, however, the application of the method to the PEM reconstruction and the theoretical analysis of the noise propagation that make this paper novel.

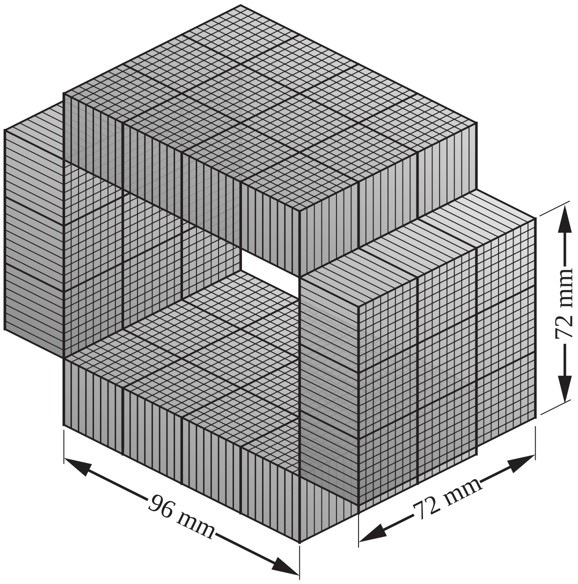


Fig. 1. PEM Geometry.

II. THEORY

A. List Mode Likelihood Reconstruction with Scatter Correction

Histogrammed PET data are generally modeled as a collection of independent Poisson random variables. By treating the detections in each LOR separately, we can derive the appropriate log-likelihood function for list mode data [2]:

$$L(\boldsymbol{x}) = \sum_{k=1}^{K} \log \left[\sum_{j=1}^{N} p(i_k, j) x_j + s_{i_k} \right] - \sum_{j=1}^{N} \varepsilon_j x_j, \qquad (1)$$

This work was supported by the U.S. Department of Health and Human Services under grant P01 HL25840, by the National Cancer Institute under grant R01 CA 59794, and by the Director, Office of Science, Office of Biological and Environmental Research, Medical Sciences Division of the US Department of Energy under contract DE-AC03-76SF00098.

J. Qi and R.H. Huesman are with the Center for Functional Imaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA (telephone: 510-486-4695, e-mail: {jqi,rhhuesman}@lbl.gov).

where x_j is mean activity inside the j^{th} voxel of the unknown image, $p(i,j)$ is the probability of detecting an event from the j^{th} voxel in the i^{th} LOR, s_i is the mean scatter in the i^{th} LOR, i_k is the index of the LOR of the k^{th} detection, $\varepsilon_j \equiv \sum_i p(i,j)$, K is the total number of detections, and N is the total number of image voxels.

The maximum likelihood (ML) estimate can be found by maximizing (1). A popular ML algorithm for PET reconstruction is the expectation maximization (EM) algorithm [3], [4], [5]. However, the ML solution is unstable (i.e. noisy) because of the ill-conditioness of tomography systems. Hence some form of regularization (or prior function) is needed to reconstruct a reasonable image. The prior function used in [2] is a Gaussian prior whose logarithm is of the form

$$\beta U(\boldsymbol{x}) = \frac{\beta}{2}(\boldsymbol{x}-\boldsymbol{m})'\boldsymbol{R}(\boldsymbol{x}-\boldsymbol{m}), \tag{2}$$

where β is the smoothing parameter, $\boldsymbol{m}$ is the estimated mean of the unknown image, $\boldsymbol{R}$ is a positive definite (or semidefinite) matrix.

Combining the likelihood function (1) and the image prior (2), the reconstruction is found as:

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}\geq 0}\left[L(\boldsymbol{x}) - \beta U(\boldsymbol{x})\right]. \tag{3}$$

For further simplification, $\boldsymbol{R}$ is chosen to be a diagonal matrix, so the EM algorithm can be used to solve (3). The EM update equation is [2]

$$\hat{x}_j^{n+1} = \frac{1}{2}\left(m_j - \frac{\varepsilon_j}{\beta r_{jj}}\right) + \sqrt{\frac{1}{4}\left(m_j - \frac{\varepsilon_j}{\beta r_{jj}}\right)^2 + \frac{\hat{x}_j^n}{\beta r_{jj}}\sum_{k=1}^{K}\frac{p(i_k,j)}{\sum_{l=1}^{N} p(i_k,l)\hat{x}_l^n + s_{i_k}}},$$

where r_{jj} is the (j,j)th element of $\boldsymbol{R}$.

B. *Estimate Mean Scatter Sinogram using Monte Carlo Simulation*

The scatter correction method described in the previous section requires the mean scatter sinogram being known before the reconstruction starts. For conventional PET systems, a scatter sinogram can be estimated by single scatter simulation, deconvolution of emission sinogram, dual energy windows, or Monte Carlo methods. Most analytical scatter estimation methods require fitting an computed scatter sinogram to the tails of the emission sinogram that consist of pure scatter events. This is not practical for PEM geometry as the whole FOV is filled with activity. Therefore, we adopt the Monte Carlo method here.

The Monte Carlo scatter simulation requires both emission and attenuation maps. Generally they are obtained from an initial reconstruction without scatter correction. One advantage of breast imaging with FDG is that the background is quite uniform. If we assume that whole FOV is filled with uniform activity and that features such as cancerous lesions account for a small fraction of the radioactivity, then we only need to run the Monte Carlo simulation once for each system configuration. This will saves a large amount of computational time.

The Monte Carlo simulation program estimates the scatter sinogram by tracing all 511 keV photon pairs randomly generated inside the FOV. For each photon, it first computes the interaction point based on the attenuation length. Then, it determines whether it is a photo-electric or Compton interaction. If it is photo-electric, it dumps all of its current energy; if it is Compton, it computes the energy deposited and the new direction of the photon using the Klein-Nishima formula, and continues to trace the photon until the photon has dumped all its energy or traveled outside of the system. A photon is detected when the energy deposited at one detector is greater than a preselected threshold. A coincidence event is recorded if both photons are detected. The Monte Carlo simulation program histograms separately the scattered and unscattered (true) events.

For each individual data set, the scatter sinogram is then estimated by

$$\hat{s}_i = \frac{\text{total events in the data}}{\text{total events in Monte Carlo Simulation}} s_i^{MC}, \tag{4}$$

where s_i^{MC} is the number of scatter events in the i^{th} LOR in the Monte Carlo simulation. This assumes the scatter fraction and profile in each data set are the same, as most scatters are generated from the uniform background.

C. *Noise Propagation*

Noise is inevitable in the Monte Carlo scatter sinogram. Because the Poisson nature of the counting process, the longer it runs, the less the noise is. This presents a trade off between time and accuracy. In this section, we analyze how the noise in the scatter sinogram propagates into the reconstruction.

We denote the MAP reconstruction in (3) as $\hat{\boldsymbol{x}}(\boldsymbol{y},\hat{\boldsymbol{s}})$ to indicate that $\hat{\boldsymbol{x}}$ is dependent on estimated scatter sinogram $\hat{\boldsymbol{s}}$. Since $\boldsymbol{y}$ and $\hat{\boldsymbol{s}}$ are independent, we first focus on noise from $\hat{\boldsymbol{s}}$ and assume $\boldsymbol{y}$ noisefree (i.e. $\boldsymbol{y} = \bar{\boldsymbol{y}} = \boldsymbol{P}\boldsymbol{x} + \boldsymbol{s}$). We can approximate $\hat{\boldsymbol{x}}(\boldsymbol{y},\hat{\boldsymbol{s}})$ using a first order Taylor series expansion at the point $\hat{\boldsymbol{s}} = \boldsymbol{s}$:

$$\hat{\boldsymbol{x}}(\boldsymbol{y},\hat{\boldsymbol{s}}) \approx \hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s}) + \nabla_s\hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s})(\hat{\boldsymbol{s}}-\boldsymbol{s}). \tag{5}$$

This approximation is similar to that presented in [6]. From (5), we have the following expression for the covariance of noise in the reconstruction caused by the noise in the estimated scatter sinogram

$$\Sigma(\hat{\boldsymbol{x}}) \approx \nabla_s\hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s})\Sigma(\hat{\boldsymbol{s}})[\nabla_s\hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s})]' \tag{6}$$

where $\Sigma(\hat{\boldsymbol{s}})$ is the covariance matrix of the estimated scatter sinogram.

To compute $\nabla_s\hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s})$, we follow the idea presented in [6]. We restrict our attention to the situations where the solution of (3) satisfies

$$0 = \frac{\partial}{\partial x_j}\left[L(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{s}) - \beta U(\boldsymbol{x})\right]\Big|_{\boldsymbol{x}=\hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s})}, \quad j = 1,\ldots,M. \tag{7}$$

While this assumption precludes inequality constraints, it should work fine here because of the uniform background. Differentiating (7) with respect to s_i by applying the chain rule and solving the resulting equation, we get [6]

$$\begin{aligned}\nabla_s \hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s}) &= \left\{-\nabla_{xx}\left[L(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{s}) - \beta U(\boldsymbol{x})\right]|_{\boldsymbol{x}=\hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s})}\right\}^{-1} \\ &\quad \nabla_{xs}\left[L(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{s}) - \beta U(\boldsymbol{x})\right]|_{\boldsymbol{x}=\hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s})}\end{aligned} \tag{8}$$

where the $(j,k)^{th}$ element of the operator ∇_{xx} is $\frac{\partial^2}{\partial x_j \partial x_k}$, and the $(j,l)^{th}$ element of the operator ∇_{xs} is $\frac{\partial^2}{\partial x_j \partial s_l}$.

From (1) and (2), we can derive

$$\begin{aligned}&\nabla_{xx}\left[L(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{s}) - \beta U(\boldsymbol{x})\right] \\ &\qquad = \boldsymbol{P}' \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s}+\boldsymbol{r})_i^2}\right]\boldsymbol{P} + \beta\boldsymbol{R}\end{aligned}$$

and

$$\nabla_{xs}\left[L(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{s}) - \beta U(\boldsymbol{x})\right] = \boldsymbol{P}' \operatorname{diag}\left[\frac{-y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s})_i^2}\right].$$

Then

$$\begin{aligned}\nabla_s \hat{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{s}) &= \left\{\boldsymbol{P}' \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s}+\boldsymbol{r})_i^2}\right]\boldsymbol{P} + \beta\boldsymbol{R}\right\}^{-1} \\ &\quad \boldsymbol{P}' \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s})_i^2}\right].\end{aligned} \tag{9}$$

Substituting (9) into (6) results in

$$\begin{aligned}\Sigma(\hat{\boldsymbol{x}}) &\approx \left\{\boldsymbol{P}' \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s}+\boldsymbol{r})_i^2}\right]\boldsymbol{P} + \beta\boldsymbol{R}\right\}^{-1} \\ &\quad \boldsymbol{P}' \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s})_i^2}\right]\Sigma_{\hat{\boldsymbol{s}}} \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s})_i^2}\right]\boldsymbol{P} \\ &\quad \left\{\boldsymbol{P}' \operatorname{diag}\left[\frac{y_i}{(\boldsymbol{P}\hat{\boldsymbol{x}}+\boldsymbol{s}+\boldsymbol{r})_i^2}\right]\boldsymbol{P} + \beta\boldsymbol{R}\right\}^{-1}\end{aligned} \tag{10}$$

In general, $\hat{\boldsymbol{x}}$ is a slightly blurred version of $\boldsymbol{x}$, so the projection $\boldsymbol{P}\hat{\boldsymbol{x}} + \boldsymbol{s}$ is approximately equal to the mean of the data, $\bar{\boldsymbol{y}}$. Therefore, we can simplify the above expression to

$$\Sigma(\hat{\boldsymbol{x}}) \approx [\boldsymbol{F}+\beta\boldsymbol{R}]^{-1}\boldsymbol{P}' \operatorname{diag}\left[\frac{\sigma_{\hat{s}_i}^2}{\bar{y}_i^2}\right]\boldsymbol{P}[\boldsymbol{F}+\beta\boldsymbol{R}]^{-1}, \tag{11}$$

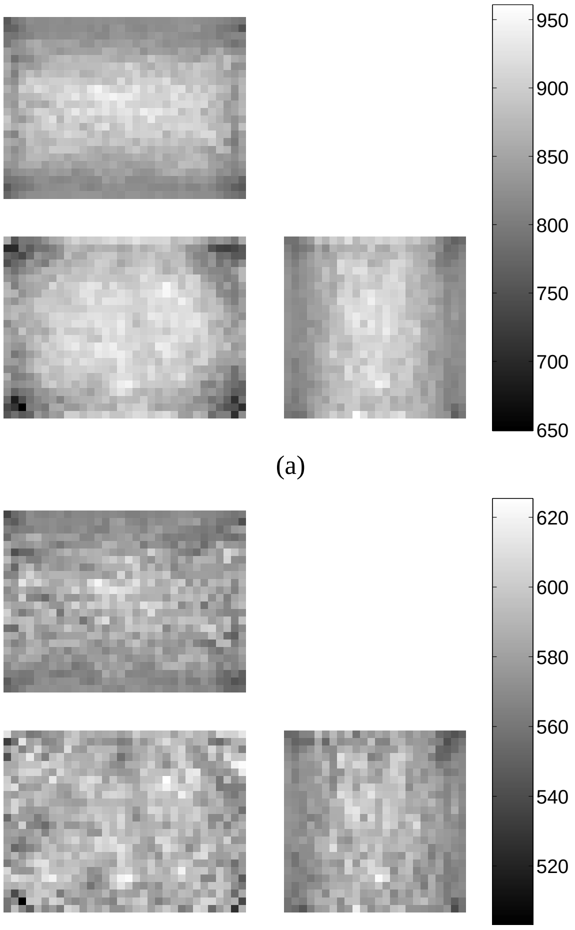
where $\boldsymbol{F} = \boldsymbol{P}' \operatorname{diag}\left[\frac{1}{\bar{y}_i}\right]\boldsymbol{P}$ is the Fisher information matrix and $\sigma_{\hat{s}_i}$ is the variance of $\hat{s}_i$. Eq. (11) is the covariance matrix of the noise in the reconstruction that is propagated from the estimated scatter sinogram.

The covariance of noise in the reconstruction caused by the Poisson noise in the data is [6]

$$\begin{aligned}\Sigma_{\text{Poisson}}(\hat{\boldsymbol{x}}) &\approx \nabla_y \hat{\boldsymbol{x}}(\bar{\boldsymbol{y}})\text{Cov}(\boldsymbol{y})[\nabla_{\text{y}}\hat{\boldsymbol{x}}(\bar{\boldsymbol{y}})]' \\ &\approx [\boldsymbol{F}+\beta\boldsymbol{R}]^{-1}\boldsymbol{F}[\boldsymbol{F}+\beta\boldsymbol{R}]^{-1}.\end{aligned} \tag{12}$$

Adding (11) and (12), we get the covariance of the total noise

$$\Sigma_{\text{total}}(\hat{\boldsymbol{x}}) = [\boldsymbol{F}+\beta\boldsymbol{R}]^{-1}\boldsymbol{P}' \operatorname{diag}\left[\frac{\sigma_{\hat{s}_i}^2}{\bar{y}_i^2} + \frac{1}{\bar{y}_i}\right]\boldsymbol{P}[\boldsymbol{F}+\beta\boldsymbol{R}]^{-1} \tag{13}$$



(a)

(b)

Fig. 2. (a) Reconstruction without scatter correction. (b) Reconstruction with scatter correction. The images are top view slice (upper left), front view slice (lower left), and side view slice (lower right) through the center voxel. Images in (a) and (b) are individually scaled as indicated by the gray level bars.

It shows that the noise from the scatter sinogram is equivalent to an increase of the noise in the data by a factor of $1 + \sigma_{\hat{s}_i}^2/\bar{y}_i$. $\sigma_{\hat{s}_i}^2/\bar{y}_i$ is equal to the scatter fraction of the i^{th} LOR divided by the ratio between total number of detection in the Monte Carlo simulation and the total number of detection in the data. For example, if the average scatter fraction is 30%, and the Monte Carlo simulation has 30 times as many events as the data, then the noise variance increase in reconstruction cause by the scatter sinogram will be about 1%. Equation (13) can be used to determine the number of events required in the Monte Carlo simulation and to design better simulation strategy.

III. SIMULATION RESULTS

In simulation we assume a subject weighing 70 kg and an injection of 1 mCi of FDG, which is uniformly distributed within the body. This activity density within the 72 × 72 × 96 mm^3 field of view and an imaging time of 60 s gives about 16 million disintegrations within the imaging volume.

The Monte Carlo simulation uses the appropriate energy-dependent cross sections for the interaction of photons in water (in the field of view) and in LSO detector. The average detection efficiency is about 13% for an energy threshold of 270 keV. Of all the detected events, there are about 35% unscattered events, 32% events scattered in the FOV, and 33% events scattered in the detector (not scattered in the FOV).

Fig. 2 shows some example reconstructions of a simulated flood source with and without scatter correction. The images shown are three orthogonal slices through the center voxel: top view slice (upper left), front view slice (lower left), and side view slice (lower right). The reconstructed image without scatter correction (Fig. 2a) shows brighter at the center of FOV and darker at the corners, especially in the front view slice. The scatter corrected image (Fig. 2b) shows more uniform activity distribution. Note the gray level maps in Fig. 2a and Fig. 2b are different. Here we have only corrected for the scatters in the FOV. We are looking for a better way to deconvolve the scatters in the detector as they are more localized.

IV. CONCLUSION

We have implemented the Monte Carlo scatter correction method for the list mode likelihood reconstruction algorithm for the PEM and shown some results based on computer simulations. For breast imaging with FDG, we can assume that the FOV is filled with uniform activity and that features such as cancerous lesions account for a small fraction of the radioactivity. This specific application of PEM allows us to run the Monte Carlo scatter simulation only once for each scanner configuration. The scatter sinogram for each individual data set can then be estimated using the total number of detections. This saves a large amount of computation time.

We also theoretically analyzed the noise propagation from the estimate scatter sinogram into the final reconstructed image. The results show that the noise propagated from the estimated scatter sinogram is equivalent to increasing the noise variance in each LOR by a factor of $1 + \sigma^2_{\hat{s}_i}/\bar{y}_i$. If we assume the scatter fraction for each LOR is the same, then this factor is a constant for all LORs and it provides a easy way to calculate the required number of events in the Monte Carlo scatter simulation for a given noise level in reconstruction. This noise analysis is applicable to other scatter estimation methods provided that an estimate of the covariance of the estimated scatter sinogram is available.

REFERENCES

[1] P. Virador, W. Moses, and R. Huesman, “Reconstruction in PET cameras with irregular sampling and depth of interaction capability,” *IEEE Transactions on Nuclear Science*, vol. 45, pp. 1225–1230, 1998.

[2] R. Huesman, G. Klein, W. Moses, J. Qi, B. Reutter, and P. Virador, “List mode maximum likelihood reconstruction applied to positron emission mammography with irregular sampling,” *IEEE Transactions on Medical Imaging*, vol. 19, pp. 532–537, 2000.

[3] A. Dempster, N. Laird, and D. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of Royal Statistical Society, Series B*, vol. 39, pp. 1–38, 1977.

[4] L. Shepp and Y. Vardi, “Maximum likelihood reconstruction for emission tomography,” *IEEE Transactions on Medical Imaging*, vol. 1, pp. 113–122, 1982.

[5] K. Lange and R. Carson, “EM reconstruction algorithms for emission and transmission tomography,” *Journal of Computer Assisted Tomography*, vol. 8, pp. 306–316, 1984.

[6] J. Fessler, “Mean and variance of implicitely defined biased estimators (such as penalized maximum likelihood): Applications to tomography,” *IEEE Transactions on Image Processing*, vol. 5, pp. 493–506, 1996.

Internet-based 3D PET Image Reconstruction using A Beowulf PC Cluster

‡D Shattuck, ‡J. Rapela, ‡E. Asma, ⋆A. Chatziioannou, †J. Qi, ‡R. Leahy
‡Signal and Image Processing Institute, Univ. of Southern California, Los Angeles, CA 90089
⋆Crump Institute for Molecular Imaging, University of California, Los Angeles CA90095
†Center for Functional Imaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

Abstract— **We describe an approach to fast iterative reconstruction from fully 3D PET data using a network of PentiumIII PCs configured as a Beowulf cluster. To facilitate the use of this system, we have developed a browser-based interface using Java. The system compresses PET data on the user's machine, sends this data over a network, and instructs the PC cluster to reconstruct the image. The cluster implements a parallelized version of our preconditioned conjugate gradient method for fully 3D MAP image reconstruction. We report on the speed-up factors using the Beowulf approach and the impacts of communication latencies in the local cluster network and the network connection between the user's machine and our PC cluster.**

Keywords— **3D PET, Beowulf cluster, distributed computing, iterative reconstruction**

I. INTRODUCTION

Iterative reconstruction of clinical PET images using statistically optimal algorithms can require an hour or more of computation on a single-processor computer for fully 3D data sets. Dramatic reductions in computation time have been achieved by converting fully 3D data sets to 2D using rebinning algorithms and then using iterative 2D reconstruction methods [1]. Further reductions have been achieved using the ordered-subsets EM (OSEM) algorithm which can achieve acceptable results in just a few passes through the data. However, these speed-ups are achieved at a price: the rebinning methods, even if exact for true line integrals, are unable to accurately model the true physical response of the scanner. Similarly, the OSEM method never optimizes the likelihood objective function and the results can be highly dependent on the number of subsets and number of iterations that are used. In our work [3], [2] we have concentrated on using convergent algorithms to compute maximum a posteriori (MAP) or equivalently penalized-ML solutions to the PET reconstruction problem. The more accurate models that we use with fully 3D data sets have been shown to improve image resolution [3] but inevitably lead to longer computation times.

Our approach to reducing reconstruction time is to use rapidly converging methods such as the preconditioned conjugate gradient method. Further reductions can be obtained using multiprocessor computing. Previously we have used multithreading methods to parallelize the code across multiple CPUs in a single symmetric multiprocessor (SMP) server. This arrangement is attractive since the servers typically have shared memory and hence there is minimal overhead incurred in distributing data across the processors. In tests with a four processor server, we were able to achieve speed-up factors of approximately 3.4 relative to a single processor. Unfortunately, the number of processors in standard servers is usually limited to four and the cost is high relative to single or dual processor systems. For this reason we have recently investigated the use of a Beowulf PC cluster that allows us to use a large number of low cost systems to achieve substantial speed-up relative to a single computer. Vollmar et al [5] recently reported the use of a PC cluster for 3D PET reconstruction. Their approach differs from that described here in that the forward and backprojection were based on on-the-fly computation rather than a precalculated system matrix. Similarly Labbe et al [6] present a set of forward and backprojection operators suitable for cluster and parallel computing but again these are based on on-the-fly computation.

The Beowulf cluster is simply a network of Unix or Linux workstations. For the purposes of code parallelization, the cluster is configured with a head-node that controls the program and a set of worker-nodes that handle processes spawned by the head node. The difference between the Beowulf cluster and a multiple CPU server is that the former do not have shared memory, and data must be transferred via a local ethernet between processors. This is often the bottleneck in performance of these clusters and of particular importance in PET image reconstruction where the data sets and image volumes are large. Here we report on our progress using a combination of multithreading and distributed computing on a Beowulf cluster consisting of 9 dual 933MHz PentiumIII computers connected via a 100mb/s switched ethernet.

A second goal of our work was to decouple the computer used for reconstruction from that used to acquire data. To do this we have developed a web-browser based interface to our distributed computing code using Java. Thus data can be processed using the cluster from any computer connected to the Internet. While data transfer may be slow for standard Internet connections, the availability of Internet2 connections at many research facilities make this approach viable. We report on an experiment we have performed by reconstructing data residing on the PET system computers in the Nuclear Medicine clinic at UCLA using the cluster at the Signal and Image Processing Institute at USC.

This work was supported by the National Cancer Institute under Grant No. R01 CA59794 and the National Foundation for Functional Brain Imaging.

II. Methods

A. MAP Image Reconstruction

We use a MAP estimation algorithm to reconstruct PET images [3]. In this approach, the data are modeled as:

$$\overline{\mathbf{y}} = \mathbf{P}\mathbf{x} + \overline{\mathbf{r}} + \overline{\mathbf{s}} \tag{1}$$

where $\overline{\mathbf{y}}$ is the mean of the data, $\mathbf{x}$ is the source distribution, $\overline{\mathbf{r}}$ is the mean of the randoms, and $\overline{\mathbf{s}}$ is the mean of the scattered events. $\mathbf{P}$ is the system matrix describing the probability that an event is detected, which we factor as:

$$\mathbf{P} = \mathbf{P}_{\mathbf{norm}}\mathbf{P}_{\mathbf{blur}}\mathbf{P}_{\mathbf{attn}}\mathbf{P}_{\mathbf{geom}} \tag{2}$$

where $\mathbf{P}_{\mathbf{geom}}$ is the geometric projection matrix describing the probability that a photon pair reaches the front faces of detector pair in the absence of attenuation and assuming perfect photon pair colinearity, $\mathbf{P}_{\mathbf{blur}}$ models photon pair non-colinearity, inter-crystal scatter and crystal penetration, $\mathbf{P}_{\mathbf{attn}}$ contains attenuation correction factors for each detector pair, and $\mathbf{P}_{\mathbf{norm}}$ is a diagonal matrix containing the normalization factors.

Reconstructions are computed as the maximizer of a posterior probability equal to the sum of the log-likelihood of the data, $\mathbf{y}$, conditioned on the image, $\mathbf{x}$, and the log-prior, which has the form of a Gibbs energy function:

$$\begin{aligned} \ln\ p(\mathbf{y}|\mathbf{x}) \ &= \ \textstyle\sum_i \{\overline{y}_i + y_i \ln(\overline{y}_i)\} \\ &- \ \textstyle\sum_j \sum_{\substack{k\in\mathcal{N}_j \\ k>j}} \kappa_{jk} V(x_j - x_k) \end{aligned} \tag{3}$$

where $V(x)$ is the potential function.

As in our previous work on PET image reconstruction, a preconditioned conjugate-gradient algorithm was used for optimization. In particular, the following preconditioned Polak-Ribiere form of conjugate gradient method was used.

$$\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} + \alpha^{(n)}\mathbf{s}^{(n)} \tag{4}$$

$$\mathbf{s}^{(n)} = \mathbf{d}^{(n)} + \beta^{(n-1)}\mathbf{s}^{(n-1)} \tag{5}$$

$$\mathbf{d}^{(n)} = \mathbf{C}^{(n)}\mathbf{g}^{(n)} \tag{6}$$

$$\beta^{(n-1)} = \frac{(\mathbf{g}^{(n)} - \mathbf{g}^{(n-1)})'\mathbf{d}^{(n)}}{\mathbf{g}^{(n-1)'}\mathbf{d}^{(n-1)}} \tag{7}$$

The PCG algorithm is initialized with $s^{(0)} = d^{(0)}$ and iteratively computes the conjugate directions. It it necessary to check that $s^{(0)}$ is an ascent direction. In the case that $s^{(n)'}g^{(n)} < 0$, $s^{(n)}$ is a descent direction and the algorithm is reinitialized with $s^{(n)} = d^{(n)}$. The step size, $\alpha^{(n)}$, is computed at each iteration using a Newton-Raphson line search to maximize the objective function. We incorporate a positivity constraint by using a bent-line search as we describe in [3].

Here we report on application of this algorithm, using the PC cluster, to data collected in 3D mode using the CTI ECAT HR+ scanner. The data from the HR+ scanner was a standard 3D dataset rebinned with a span of 9 and a maximum ring difference of 22. There were 239 sinograms each of size 288 (elements) by 144 (angles) giving a total emission sinogram size of 40 MB. Attenuation correction requires a second sinogram of the same size, thus the total data for a single frame is on the order of 80MB in size. These file sizes are important when considering the impact of reconstruction via a browser over an internet connection. The other files that are required are either small (such as the factored normalization file) or can be stored on the cluster (such as the forward projection matrix, $\mathbf{P}$). Voxel sizes used in our reconstructions were $2.25mm \times 2.25mm \times 2.42mm$ for the HR+ and images were of size $128 \times 128 \times 63$ so that the image is of size 4MB when saved as 4-byte real values. Thus, transfer times for the reconstructed images back to a remote user are small compared to those for sending the data to the cluster.

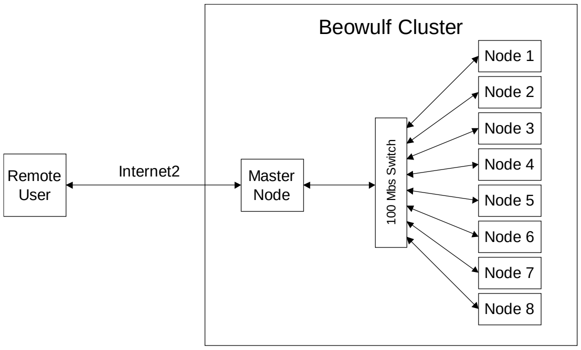


Fig. 1. Architecture of the PC cluster.

B. Beowulf Clusters and Code Parallelization

B.1 System setup

We built a Beowulf cluster consisting of one master node and eight worker nodes. Each worker node is a rackmounted dual processor Intel Pentium III 933 MHz system with 512MB of RAM and 20GB of disk space. The master node is also a rackmounted dual processor Intel Pentium III 933 MHz system, but has 1GB of RAM and a 36GB hard drive. The master node has dual network interface cards, allowing the cluster to have a private network but still be accessible from the Internet. The configuration of the cluster is shown in Fig 1. We configured the system with the Linux operating system (RedHat v. 7.0; Linux kernel version 2.4). We also installed the Local Area Multicomputing (LAM) 6.5.1 version of the Message Passing Interface (MPI) software onto each node. MPI is an open standard for communicating data between computer processes; LAM is an implementation for use on clustered computers and provides a programming environment that is portable to other architectures.

B.2 Code parallelization

Analysis of our algorithm's performance on a single computer revealed that two operations dominated computation: the back projection of the sinograms into image space and the forward projection of the image into sinogram space. Computation of the gradient in (6) requires a forward and back projection; a second forward projection is required prior to doing the line search to compute $\alpha^{(n)}$ in (4). We distributed the processing of these key operations across the cluster. Ignoring communication costs, we are able to achieve roughly a factor-$0.75 \cdot N$ speed-up on the forward and backward projections using N nodes of the cluster. This number is less than N because the workload is not perfectly balanced across the processors, but still represents a dramatic improvement in computing time. Were we to communicate the sinogram data during the iterations of the algorithm, the high cost of passing these results among the nodes would rapidly consume the gain in performance. Fortunately, we can decompose our problem such that the slave nodes never need to receive or transmit sinogram data once the iterations have begun.

The forward and back projection operators are both linear transformations, and represented as a factored system matrix as described above. During back projection, each element of the image may be a function of several elements of the sinograms. Each sinogram is transformed by the system matrix to contribute to the reconstructed image. We can partition this transformation based on arbitrary sets of sinograms, apply the system matrix separately to these sinograms to obtain their contribution to the reconstructed image, and then sum these partial results to obtain the entire transformation. In our distributed implementation, we assign each node a range of sinograms for which it is responsible. The node keeps updated versions of the data for these sinograms, and backprojects them into imagespace when requested by the head node. These results are sent back to the head node, where they are combined into a single image.

The forward projection problem can also be decomposed into functions producing individual sinograms; however, each operation will still require the full image that is being forward projected. Fortunately, the communication cost of transmitting images to each node is relatively small compared to the cost of transmitting sinogram data or performing the reconstruction computation. During forward projection, the head node broadcasts the image to each node; the nodes are responsible for producing the same sinograms that they will use during backprojection. During the iterations, the nodes generate any sinogram data they will need, and thus do not need to communicate their sinogram data to other nodes of the cluster. We distributed some additional computation to the nodes to eliminate the need to send any sinogram data to any other nodes; the effects of this distributed processing are small compared to the gains from distributing the projection operations.

A second layer of parallelization is used on each node, as each has dual processors. The projection problems are again decomposed based on sinograms, and two threads are spawned on each node to handle the projections. In this case, we achieve better load balancing as the node can dynamically assign sinograms to the threads as soon as they have finished processing. The computation of the image prior (3) is also multi-threaded on the head node. This operation may be distributed to the cluster in future work; however, this would require broadcasting of image vectors and the cost of communication may outweigh the benefit of distributed processing.

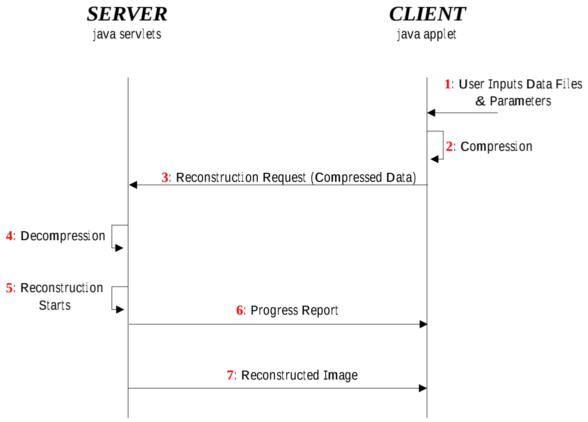


Fig. 2. Architecture of the web interface to the 3D PET reconstruction program. The user supplies the data files and parameters to the java applet. The applet compresses the data files and submits a reconstruction request to the server. The server decompresses the data files, starts the reconstruction and sends the reconstruction progress to the applet, which displays it in the client browser. When the reconstruction finishes the server sends the reconstructed image to the applet which displays the results in the client browser.

The projection routines are also used during initialization, so improvements to them will reduce start-up costs. Additionally, the geometry matrices used in projection can be hundreds of megabytes in size, and are needed on each node. These matrices are used repeatedly for a particular scanner and voxel size, thus we store copies of these files on the local hard drive. This reduces the network burden further.

C. *Java Browser-based Interface*

We developed a Java based interface to the 3D MAP reconstruction program that allows users across the Internet to run reconstructions on our Beowulf PC cluster. The interface consists of two components, a client module and a server module. Fig. 2 illustrates the architecture of the interface.

The client module was implemented as a Java applet and can run on standard web browsers. The user supplies the data files (emission file, normalization file, etc) and parameters (number of bed positions, number of frames, etc) for the MAP reconstruction and submits a reconstruction request. The client module collects the parameters and data files for the reconstruction and sends them to the server. The server module was implemented as a Java servlet. It receives the data from the client and starts the reconstruction. Text messages describing the reconstruction progress are sent to the client. The applet displays the reconstruction progress in the browser. When the reconstruction finishes the applet receives the reconstructed images and statistics about the reconstruction process.

The data sizes used in 3D PET reconstructions are large. The data size we used for HR+ reconstruction was 80MB per frame. To reduce transfer time the applet compresses

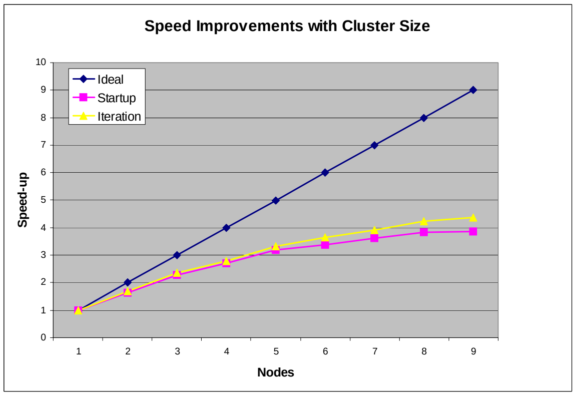


Fig. 3. Speed increase to the iterative portion of reconstruction code for different sized clusters.

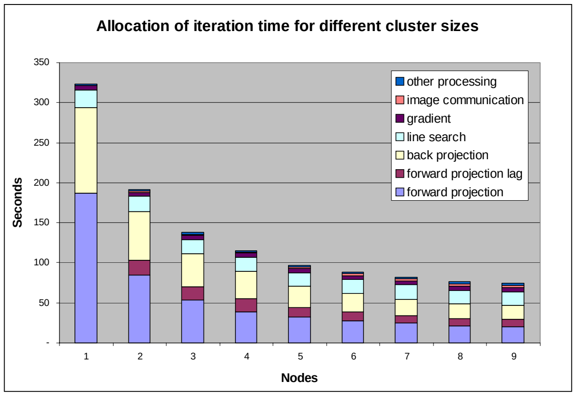


Fig. 4. Usage of computing time in a single iteration of reconstruction. Forward projection lagtime is the difference between when the head node finishes its portion of the forward projection and when all nodes are finished; some of this time may be used by the head node to perform additional computation.

the data files before sending the reconstruction request to the server side. The compressed data is transferred over the network and received by a servlet on the cluster. The servlet decompresses the data before starting the MAP reconstruction.

D. Network Connections

To evaluate performance over Internet2, the PC-cluster server was connected through a 100mb/s network to the University of Southern California backbone to Internet2. The client computer containing the data was similarly connected through the UCLA computer network. The link between USC and UCLA is part of the California Research and Education Network-2 (CalREN-2). CalREN-2 is a high-performance advanced-services network with a minimum communication bandwidth of 622Mbs.

III. RESULTS

We performed reconstructions on our cluster using different numbers of nodes to assess the benefit of using distributed processing. Figure 3 shows the performance gains achieved on both the initialization and the main loop of the program for reconstruction of HR+ 3D data. The chart show that we achieved better than $N/2$ increases in processing speed, where N is the number of nodes in the cluster, for up to 8 nodes. Our performance begins to flatten with the 9th node, which reduced the iteration time of the reconstructions by a factor of 4.36. This represents a significant performance gain over using a single dual-processor machine. Figure 4 shows the key components of a reconstruction iteration for the HR+ data as computed with different sized clusters. Forward and back projection clearly dominate the computation time when the algorithm is performed on a single node. As the number of nodes increases, the times for these operations are greatly reduced. When the ninth node is added to the cluster, the line search requires almost as much time as the forward and back projection. This figure indicates that to achieve further gains by adding more nodes we must either distribute additional processing or perform better load balancing.

Additional overhead for data transfers over Internet2 were minimal. For example transfer of the combined transmission and attenuation correction file (80MB) in uncompressed format took 54 seconds. from UCLA to USC. Typically we can achieve 50-75% compression using the compression applet which will reduce the transfer time to between 27 and 13.5 seconds. However, the time taken to compress the two files is approximately 2 minutes on a 450MHz UltraSPARC workstation, which exceeds the transfer time required for the uncompressed file. For systems with slower Internet connections, the trade-off between compression and transfer times will be different and use of compression will be appropriate. This preliminary study demonstrates the feasibility of using remote PC-clusters for image reconstruction , particularly for PET sites with access to fast networks. Furthermore, the cluster presents a relatively low-cost approach to achieving practical reconstruction times in 3D iterative PET reconstruction.

REFERENCES

[1] P. Kinahan, C. Michel, M. Defrise, D. Townsend, M. Sibomana, M. Lonneux, D. Newport, and J. Luketich: Fast iterative image reconstruction of 3D PET data, Proc. IEEE Nuclear Science Symposium and Medical Imaging, pp. 1918–1922, 1996.

[2] Mumcuoglu E.U., Leahy R. M.,Cherry S.R. and Zhou Z.: Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images. IEEE Trans. Med. Imag. **13**, (1994), 687–701

[3] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou and T. H. Farquhar: High resolution 3D Bayesian image reconstruction using the microPET small-animal scanner. Phys. in Med. Biol. **43**, 1998, 1001-1013

[4] D. Becker and T. Sterling and D. Savarese and B. Fryxelland K. Olson: Communication Overhead for Space Science Applications on the Beowulf Parallel Workstation, Proc. High Performance and Distributed Computing, 1995.

[5] St. Vollmar, M. Lercher, C. Knvss, C. Michel, K. Wienhard and W.D. Heiss: BeeHive: Cluster Reconstruction of 3-D PET Data in a Windows NT network using FORE, Proc. IEEE Med Imag Conf, Lyons, Oct, 2000.

[6] C. Labbe, H. Zaidi, C. Morel, K. Thielemans: An object-oriented library incorporating efficient projection/backprojection operators for volume reconstruction in 3D PET, Proc. 3D99, Egmond aan Zee, Netherlands, pp 137-140, 1999.

Object Dependency of Resolution and Convergence Rate in OSEM with Filtering

S. Mustafovic, K.Thielemans, D. Hogg and P. Bloomfield

***Abstract*--. Convergence properties of the Maximum Likelihood Expectation Maximization (MLEM) algorithm depending on the activity distribution in the field of view is extended to MLEM/Ordered Subsets EM (OSEM) where different types of regularization are applied.**

It will be shown that although different parts of the image converge at different rates, pure and post filtered MLEM/OSEM achieves reasonably uniform resolution.

By contrast, inter iteration filtering (IF OSEM) with smoothing filters, such as Gaussian, renders images with varying spatial resolution that is dependent on the surrounding activity. Furthermore, a similar effect is noticed on images reconstructed with MAP using a Gaussian root prior.

We conclude that resolution non-uniformity is entirely due to the filtering.

***Index Terms*—iterative reconstruction, OSEM, convergence, resolution**

I. INTRODUCTION

The major advantage of iterative over analytical algorithms is the option of emission and detection process to be accurately modelled [11]. Furthermore, iterative algorithms allow statistical noise models to be included as well as incorporation of prior knowledge. Also, provided that some kind of regularisation is used, images obtained with iterative algorithms are more acceptable.

On the other hand, filter-back projection (FBP) as a linear algorithm produces images which have nearly spatially invariant, object independent resolution.

Pure MLEM/OSEM produces images which possess unacceptable noise properties as the iteration number increases. That is why regularisation is needed. Different types of regularisation can be used such as early stopping, based on different criteria [9], inter-iteration filtering (IF) where filtering is used in between the iterations [6,10], post filtering where one keeps on iterating up to the convergence after which post filtering is applied [12] as well as incorporating penalty term in the objective function leading to MAP algorithms [1,7,8].

Defining criteria for stopping the iteration before the image becomes too noisy is very difficult since different parts of the image converge at different rates leading to non-uniform and object dependent resolution [1]. To avoid this problem post filtering can be used after the convergence is achieved, but this leads to long reconstruction times.

Previously it was shown that standard regularisation penalties produce non-uniform resolution even for space-invariant tomographs [2] and a modified penalty was proposed that improved resolution properties.

This abstract concentrates on the non-uniform convergence properties and the influence of the activity distribution present in the image.

Even though some people desire and expect non-uniform resolution based on the idea that high-counts regions provide higher resolution as mentioned in [2], there are applications where uniform resolution or at least object independent resolution is of crucial importance. One obvious application would be to dynamic PET studies where there are different activity distributions in different frames and hence different resolution properties. Also, for cross-patient studies or single patient studies taken over a period of time, the same resolution properties across the image are desirable [3].

Post and inter–filtering as methods of regularising an image were compared previously [6] and it was concluded that the coefficient of variation (CV) and contrast of the inter-filtering was slightly better. However, that analysis did not account for the distribution of the radioactivity in the image itself. Our experiments show that smoothing filters, such as Metz and Gaussian, incorporated directly into reconstruction either as the part of the image updating process (inter-filtering) or as a prior (MAP with Gaussian root prior) influence the resolution properties of the object surrounded with a large activity. Images obtained in this way have spatially varying resolution.

Section I reviews the algorithms used and the methods are presented in Section II. The experimental results are

Manuscript received April 6, 2001.

S. Mustafovic is with Imperial College of Science, Technology and Medicine and PET Methods Group, Hammersmith Hospital, Du Cane Road, London W12 0NN,UK telephone: + 4420 8383 3731, e-mail: sanida@ic.ac.uk.

Dr. K.Thielemans is with Imaging Research Solutions Limited, PET Methods Group, Hammersmith Hospital, Du Cane Road, London W12 0NN,UK telephone: +44020 8383 3731, e-mail: kris.thielemans@csc.ac.uk

D.Hogg is with Surrey University and PET Methods Group, Hammersmith Hospital, Du Cane Road, London W12 0NN,UK telephone: + 4420 8383 3731, e-mail: darren.hogg@csc.mrc.ac.uk.

P.Bloomfield is with Imaging Research Solutions Limited, PET Methods Group, Hammersmith Hospital, Du Cane Road, London W12 0NN,UK peter.bloomfield@csc.mrc.ac.uk

presented in Section III. Finally, we conclude with a discussion.

I.ALGORITHM DESCRIPTION

Images for this study were reconstructed with OSEM through various schemes. OSEM requires projections to be organised into a number of subsets. The subsets were chosen as shown in [6]. It is worth mentioning that one pass through all the subsets is considered one full iteration. The number of subsets used was fixed to 10 subsets for all reconstructions.

A uniform image as a starting point was used in all experiments. In the case of post-filtering plain OSEM needed to be run to the convergence after which images were filtered. The convergence point was determined by the stability of resolution/coefficient of variation.

Inter-iteration filtering incorporated filtering directly in the reconstruction process where the filtering process itself was applied at different intervals i.e. filtered every 5 subsets, every 10 and every 20.

Additionally, maximum a posteriori (MAP) with a Gaussian root prior was implemented where it was assumed that images are locally smooth and as such pass unaltered by Gaussian filter. Gaussian outputs a “weighted average” of each pixel's neighbourhood, with the average weighted more towards the value of the central pixels and therefore providing gentler smoothing than the one obtained with simple averaging. The prior will only be applied if the pixel values possess huge variations. The approach is similar to median root prior but differs in that instead of smoothness, privilege is given to those solutions, which are roots of the median [8].

The reconstructions were implemented using the object-oriented software library PARAPET [15,16].

II. METHODS

In the experimental study we have used simulated and real data. The real data were obtained from a HiDAC, 3D small animal PET scanner, which has about 1mm resolution.

The camera consists of four planar, rectangular detector banks each consisting of 8 HiDAC modules, rotating backwards and forwards every 6sec over 180°.

The data acquired in list mode was rebinned into 0.5mm bins where the axial field of view was set to 100mm and diameter to 60mm. The maximum acceptance angle was 59.03° discretized in 15 steps. We have used 160 views for the rebinning, resulting in the projection data of size $15 \times 160 \times 161$. Data were reconstructed on a grid $161 \times 161 \times 244$ cubic voxels of side 0.5mm.

The simulated data was forward projected and the resulting sinograms were used for reconstruction purposes. The sizes and angles were the same as for the experimental data. The effects of attenuation, scatter or noise were not simulated so that only resolution effects could be examined.

Furthermore, we have simulated two cylinders and two line sources placed such that one lied in between the cylinders and the other one is placed a bit further apart. The line sources were longer than the cylinders.

The real data consisted of a 1h scan of a germanium cylinder (external length 7.2cm and source length 6.2 cm with external diameter 3.6cm and source diameter 3cm) with a 10cm long aluminum oxide (Al_2O_3) filament line source aligned to it.

These configurations are similar to the 2D case in [1], to illustrate the effect of convergence of FWHM by surrounding activity (Fig. 1a).

Resolution was measured by FWHM of the line source where two values were recorded, one obtained from the part of the line source surrounded with the activity and the other one from the opposing end where there was no surrounding activity present.

A Gaussian smoothing filter with FWHM = 2mm was used in all experiments. This filter was incorporated directly into the reconstruction either after the normal OSEM image update (inter filtering) or as a prior (MAP with Gaussian root prior) or applied to the converged images reconstructed with pure OSEM (post filtering).

III. EXPERIMENTAL RESULTS

The analysis of the inter filtered case showed that the part of the image sandwiched between the two cylinders failed to achieve the same resolution as the one obtained in the post filtered case (Fig.1b and 1c). This indicates that application of inter filtering with smoothing filters renders images with spatially varying resolution. The more frequent the filter is applied the bigger the difference in the resolution properties between the two parts of the line sources (Fig. 2a). As the frequency of filtering is decreased the effect is less pronounced and the resolution properties approximate to the nonregularized case (Fig. 2b).

Furthermore, images reconstructed with MAP with Gaussian root prior showed that the same effect, i.e. non-uniform resolution, is present (Fig. 1d)

We conclude that these resolution non–uniformities are due exclusively to the filtering with smoothing filters (IF OSEM) or the interaction between likelihood function and the prior for MAP. Moreover, resolution properties depend on the object in the case of inter-iteration filtering whereas this effect is not present in the case of post-filtering. .

Similarly, simulations of the line source and the two cylinders placed this time further apart did not show this effect. Hence showing the object dependency once again.

Furthermore, the real data was reconstructed using 3D pure and post-filtered OSEM, IF OSEM where filtering was performed as explained in section II. Once again it was confirmed that smoothing filters incorporated in the reconstruction produce images with spatially varying resolution (Fig. 3).

IV. DISCUSSION

We have investigated object dependency and the convergence rate in the pure and regularized OSEM (post/inter filtering) as well as MAP with the Gaussian root prior. It was found that once a smoothing filter is applied either as a part of image updating process (IF OSEM) or in MAP the obtained images have spatially varying resolution depending on the activity distribution in the image.

Figures 2a, 2b give a clear intuitive explanation for this behavior in the interfiltering case. There are two competing effects on the FWHM. The normal OSEM update decreases the FWHM, while the filtering step increases it. At the convergence, the balance between these two updates is influenced by two factors: the frequency of filtering (higher the frequency, higher FWHM) and the convergence rate of pure OSEM in that point (slower convergence, higher FWHM). This latter factor gives an interesting connection between the object dependency of the local convergence rate and the resolution obtained in an inter filtering case. It is clear that this connection will also exists when filtering occurs after every subitertaion, which in the case of 1 subset is the original EMS algorithm. It is also clear that exactly the same behavior will happen in any algorithm, i.e. with different update steps than (OS) EM, that uses interfilterinng. This is because any algorithm will have different local convergence rates depending on surrounding activity. So we generally conclude: for any iterative algorithm, interfiltering with spatially invariant smoothing filtering will lead to object dependent resolution.

This situation is very similar to the case analyzed in [4]. Instead of intefiltering, there, a (smoothing) filtering term is added to the likelihood. It was found that having a spatially invariant penalty term leads to object dependent (and hence non-uniform) resolution. In this case, no obvious connection with the convergence rate is present. Indeed, the analysis in [4] is independent of the algorithm used to find the MAP maximum. In contrast, for interfiltering, the object dependency will vary if a different algorithm is used, as the local convergence rate would be different.

It is likely that, similar to the MAP case [3,4,5], uniform resolution could be obtained in an interfiltering case by adapting the filter locally. We plan to investigate this in the future. To do this, the fixed-point equations for EMS (and its variations) will have to be analyzed. This is worthwhile as we found that when no surrounding activity is present, interfiltering gives a better resolution vs. noise (measured as CV in a uniform region) trade-off compared to postfiltering (data not shown but see also [6]).

At the moment, it is unclear if a filtering approach is better than using a penalty term. To investigate this, we included some results on MAP with a Gaussian root prior. Although this algorithm does not correspond to maximizing an objective function (there is generally no penalty term associated with a root prior), we used this algorithm because a clear connection exists with the filter. We were surprised to see (Fig. 1d) that the FWHM obtained by this algorithm is a lot smaller than in the case of postfiltering (and hence interfiltering). Obviously, this will depend on the choice of beta, and one would have to look at the corresponding noise properties. We leave this for future work.

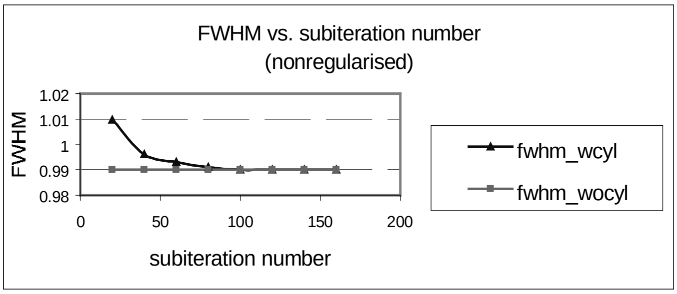


Figure 1a. Resolution vs. subiteration number (nonregularised OSEM) for two parts of the line source (part of the line source aligned with a cylinder – triangles, and part of the line source on its own – squares)

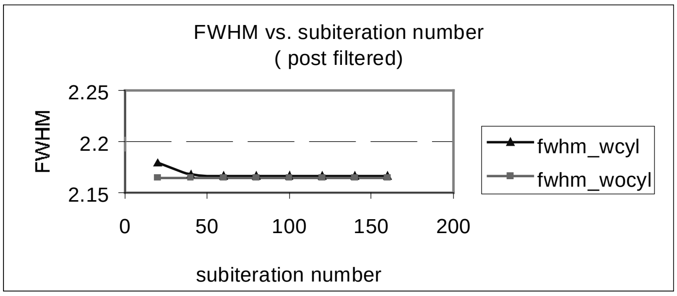


Figure 1b. Resolution vs. subiteration number (post filtered OSEM) for two parts of the line source (part of the line source aligned with a cylinder –triangles, and part of the line source on its own – squares)

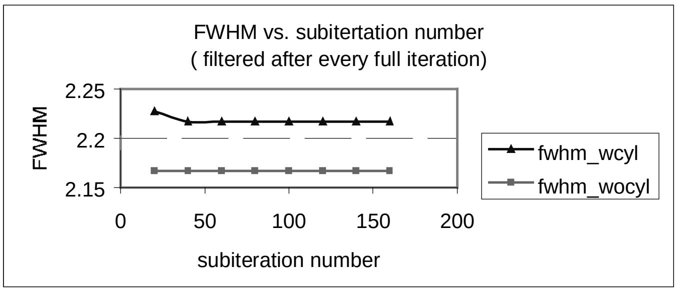


Figure 1c Resolution vs. subiteration number (inter filtered OSEM) for two parts of the line source (part of the line source aligned with a cylinder –triangles, and part of the line source on its own – squares)

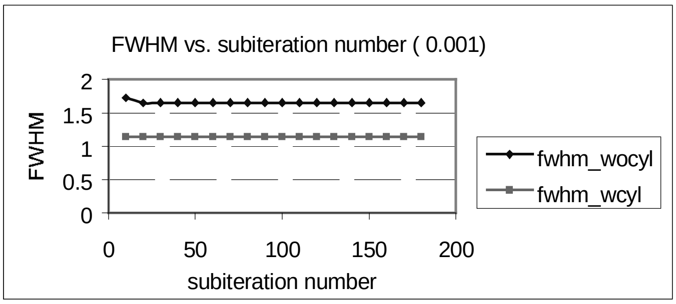


Figure 1d. Resolution vs. subiteration number (MAP with Gaussian root prior) for two parts of the line source (part of the line source aligned with a cylinder – triangles, and part of the line source on its own – squares)

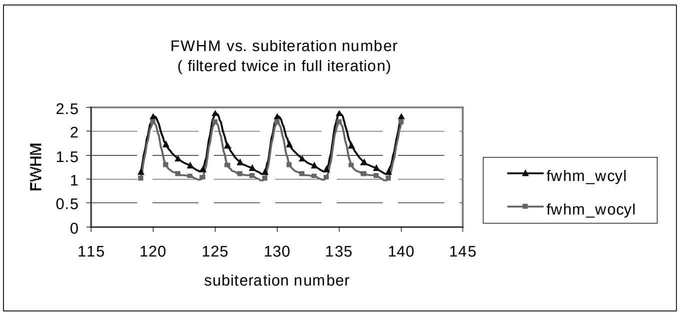


Figure 2a. Resolution vs. subiteration number for inter-filtered OSEM where filtering was applied twice in every full iteration

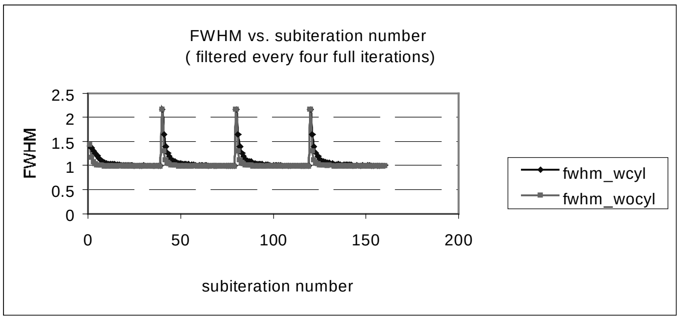


Figure 2b. Resolution vs. subiteration number for inter-filtered OSEM where filtering was applied every four full iteration

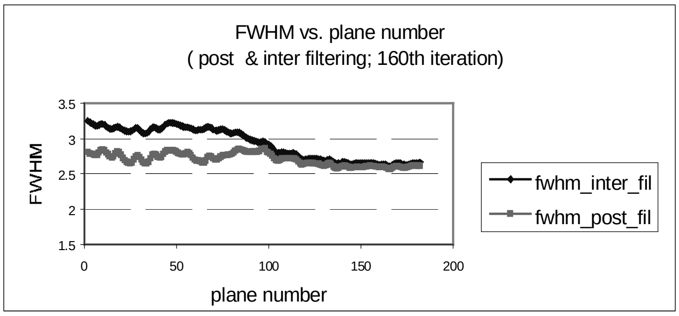


Figure 3 Comparison of resolution vs. plane number for real data of inter and post filtering

REFERENCES

1. J.Nuyts, "ML and MAP Reconstruction for Emission and Transmission Tomography", Nuclear Science Symposium and Medical Imaging Conference, Lyon, 2000
2. J. A. Fessler and W.L. Rogers, "Spatial resolution properties of the penalized-likelihood image reconstruction methods: Space invariant tomographs", IEEE Tr. Im. Proc., 5(9) ,Sep. 1996, pp. 1346-58,
3. J.W. Stayman and J. Fessler, " Regularization for uniform spatial resolution properties in penalized-likelihood image reconstruction", IEEE Trans. on Med. Imaging, March 8, 2000
4. J.A. Fessler and W.L.Rogers. " Uniform quadratic penalties cause nonuniform image resolution (and sometimes vice versa)", in Conf. Rec. of the IEEE Nuc. Sci.Symp. Med. Im.Conf., Vol.4., pp. 1915-1919, 1994
5. Qi J., R. Leahy, " Resolution and noise properties of MAP reconstruction for fully 3-D PET", Vol. 19, No5, May 2000, 493-506
6. M. Jacobson, R. Levkovitz, A. Ben-Tal, K.Thielemans, T.Spinks, D.Belluzzo, E.Pagani, V. Bettinardi, M. Gilardi, A.Zverovich and G.Mitra, "Enhanced 3D PET OSEM reconstruction using inter-update Metz filtering", Phys. Med.Biol. 45 (2000) 2417-2439
7. Green P.J "Bayesian reconstruction from emission tomography data using a modified EM algorithm", IEEE Tr. Med. Im., Vol.9, No 1, March 1990, pp. 84-93.
8. S. Alenius, " On noise reduction in iterative image reconstruction algorithms for emission tomography : median root prior", Ph.D. Thesis, 1999
9. E. Veklerov and J. Llacer "Stopping rule for the MLE algorithm based on statistical hypothesis testing," IEEE Tr. Med. Im. Vol.6, No. 4, pp.313-319, December 1987
10. S.Silverman, M.C. Jones, J.D. Wilson, D.W. Nychka " A smooth EM approach to indirect estimation problems, with particular reference to stereology and emission tomography" , Journal Royal Statistical Society, 1990, 271-324.
11. B.F. Hutton, H.M. Hudson, F.J. Beekman " A clinical perspective of accelerated statistical reconstruction", European Journal of Nuclear Medicine Vol. 24.No. 7, July 1997
12. Beekman F.J., Slijpen T.P.E., Niessen J.W., "Selection of task-dependent diffusion filters for the post-processing of SPECT images", Phys. Med. Biol. 43, Feb. 1998, pp. 1713-1730
13. A.P. Jeavons, R.A. Chandler, C.A. Dettmar, " A 3D HiDAC PET camera with sub-millimeter resolution for imaging small animals ", IEEE NSS and MIC Conference, Toronto Canada, 1998
14. Y. Vardi, L.A.Shepp, L. Kaufman " A statistical model for positron emission tomography", Journal of the American Statistical Association, Vol. 80, No. 389, March 1985
15. Labbé C., Thielemans K., Belluzzo D., Bettinardi V., Giardi M.C., Hague D.S., Jacobson M., Kaiser S, Levkovitz R, Margalit T, Mitra G., Morel C., Spinks T.J., Valente P., Zaidi H., Zverovich A. " An object –oriented library for 3D PET reconstruction using parallel computing ", 1999
16. Labbé C., Thielemans K., Zaidi H., Morel C., " An object-oriented library incorporating efficient projector/back projection operators for volume reconstruction in 3D PET", Proc. 1999 Int. meeting on fully three-dimensional image reconstruction in radiology and nuclear medicine, pp. 137-140,1999

Performance Evaluation of Exact Cone-Beam Algorithms for the Long-Object Problem in Spiral Computed Tomography

K. Sourbelle[1], H. Kudo[2], G. Lauritsch[3], K.C. Tam[4], M. Defrise[5] and F. Noo[6]
[1]Institute of Medical Physics, University of Erlangen, 91054 Erlangen, Germany
[2]Institute of Information Sciences and Electronics, University of Tsukuba, Tsukuba 305-8573, Japan
[3]Siemens Medical Solutions, P.O. Box 3260, 91050 Erlangen, Germany
[4]Siemens Corporate Research, Inc., Princeton, NJ 08540, USA
[5]Division of Nuclear Medicine, AZ-VUB, Free University of Brussels, 1090 Brussels, Belgium
[6]Institute of Electricity Montefiore, University of Liège, 4000 Liège, Belgium

I. INTRODUCTION

Spiral Computed Tomography (CT) with area detectors is of increasing interest for scanning large volumes in a short time and achieving isotropic resolution. When the cone-angle is large, exact reconstruction algorithms are needed which go beyond approximate solutions.

A major problem of spiral cone-beam CT is the so-called long-object problem, which concerns the exact reconstruction of a region-of-interest (ROI) of a long object using projection data of a pure spiral scan covering the ROI and its immediate vicinity only. This problem has been recently solved with different approaches: the Virtual Circle method [1], the Zero Boundary method [2], and the Local ROI technique [3]. Although all these methods are theoretically justified, they have different approaches and therefore some differences in numerical behavior and in effectiveness are expected.

The purpose of this study is to evaluate the performances of these algorithms regarding image quality and practicability. Here, we concentrate on Filtered-Backprojection-type (FBP) algorithms. A comparison between Radon- and FBP-based algorithms for the Local ROI method can be found in [4]. Image quality is determined on the basis of image artifacts, spatial resolution and noise properties. For practicability, the computing time as well as the overscan range required for the ROI reconstruction (i.e. the z-extension of the spiral below and above the ROI) are compared.

II. DESCRIPTION OF THE METHODS

The reconstruction of a short-object from axially truncated projections can be performed with an exact (or quasi-exact) 3D Filtered Backprojection (FBP) approach, as proposed in [5][6]. The method is based on the generalization of Grangeat's formula for calculating the Radon data and requires the cone-beam projections to be measured in some region B on the detector which is bounded by the projection of the upper and of the lower turn of the spiral. A straightforward implementation of the filtering step is the 4-step algorithm, which consists in performing some line integral derivatives on the cone-beam projection, backprojecting them on a virtual detector (which is large enough to cover the cone-beam projection of the whole object) and applying a final derivative filter. Alternatively, the filtering step can be carried out by convolution with an 1D ramp filter plus some boundary correction term involving only data at the boundaries of the region B [5][6] (these boundaries will be referred below as the mask boundaries). The 4-step algorithm is the basis of the Local ROI method whereas the 1D ramp filter approach is the basis of the Virtual Circle and of the Zero Boundary methods.

A. *The Virtual Circle Method (VC)*

One solution to the long-object problem is to add two supplementary circle scans at the bottom and at the top of the ROI [7]. However, this is not practical for medical imaging. It can be shown that the particular projection data of the circle scans needed for the reconstruction of the ROI can be synthesized from the data of the spiral scan. Therefore a real measurement of the circle scans is not necessary and the synthesized data are regarded as virtual data.

B. *The Zero Boundary Method (ZB)*

Since only the data on the boundaries of the region B cause troubles in the long-object problem, the Zero Boundary method treats them separately. The image to be reconstructed is expressed as the sum of two partial images. The first image can be reconstructed such that its cone-beam projections are equal to the data on the mask boundaries, exploiting the property that each point in the field-of-view belongs to one and only one PI-line. The second image can be reconstructed from the projections with zeroed mask boundaries, using a standard 3D FBP with ramp filtering and cone-beam backprojection.

C. *The Local ROI Method (LR)*

In the Local ROI technique [3], the Radon derivative data are grouped on ϕ-planes (meridian planes) containing the z-axis. The Radon derivative data on different ϕ-planes are computed for different portions of the object, which are called local ROI's. Each local ROI is defined by the parallel projection of the spiral scan path onto the corresponding ϕ-plane, so that the contributing cone beams are not contaminated by object information outside the local ROI. A (global) ROI consisting of the intersection of all local ROI's can be reconstructed without interference from the parts of the object outside the ROI. The Local ROI method was formulated as a Radon-based algorithm [8], as well as an FBP-based 4-step algorithm [9][10].

III. SIMULATION EXPERIMENTS

Image quality is evaluated with simulated projection data. The scanner geometry is based on standard parameters of medical scanners: the radius of the focus path is 57 cm and the detector is at the distance 43.5 cm from the z-axis. We take 1056 projections per rotation. We choose a table feed value of 12.8 cm per turn, which corresponds to a full cone-angle of 7.2°. For this large angle, approximate algorithms usually do not perform well any more.

Different phantoms are used for this study: a 3D anthropomorphic Head Phantom [11] for the investigation of artifacts, simulated ideal delta points for the evaluation of spatial resolution, and a homogeneous water sphere phantom of 20 cm diameter for the study of noise properties. The simulations are done for a flat detector with square pixels of dimensions $(0.88\text{ mm})^2$, which corresponds to a square pixel size of $(0.5\text{ mm})^2$ at the center of rotation. For the Head phantom and the water sphere, the simulated detector has 512×256 pixels.

The VC and the ZB methods have been implemented as originally published in [1] and [2], respectively. The filtering of the 4-step in the LR method has been implemented with the help of the linogram method [12], yielding an improvement in efficiency and accuracy, compared to the original published algorithm [10].

IV. EVALUATION OF IMAGE QUALITY

A. Visual Image Impression and Artifacts

The reconstructions of the Head Phantom for the three methods are shown in Fig. 1. The results of the reconstruction of the whole object are compared with a ROI reconstruction. For the whole object reconstruction, the spiral consists of 2.97 turns covering a z-range of [-19 cm, 19 cm]; the ROI reconstruction is obtained from projections of 2 turns covering the range [-12.8 cm, 12.8 cm]. The reconstructed volumes consist of 512^3 voxels of dimensions $(0.5\text{ mm})^3$.

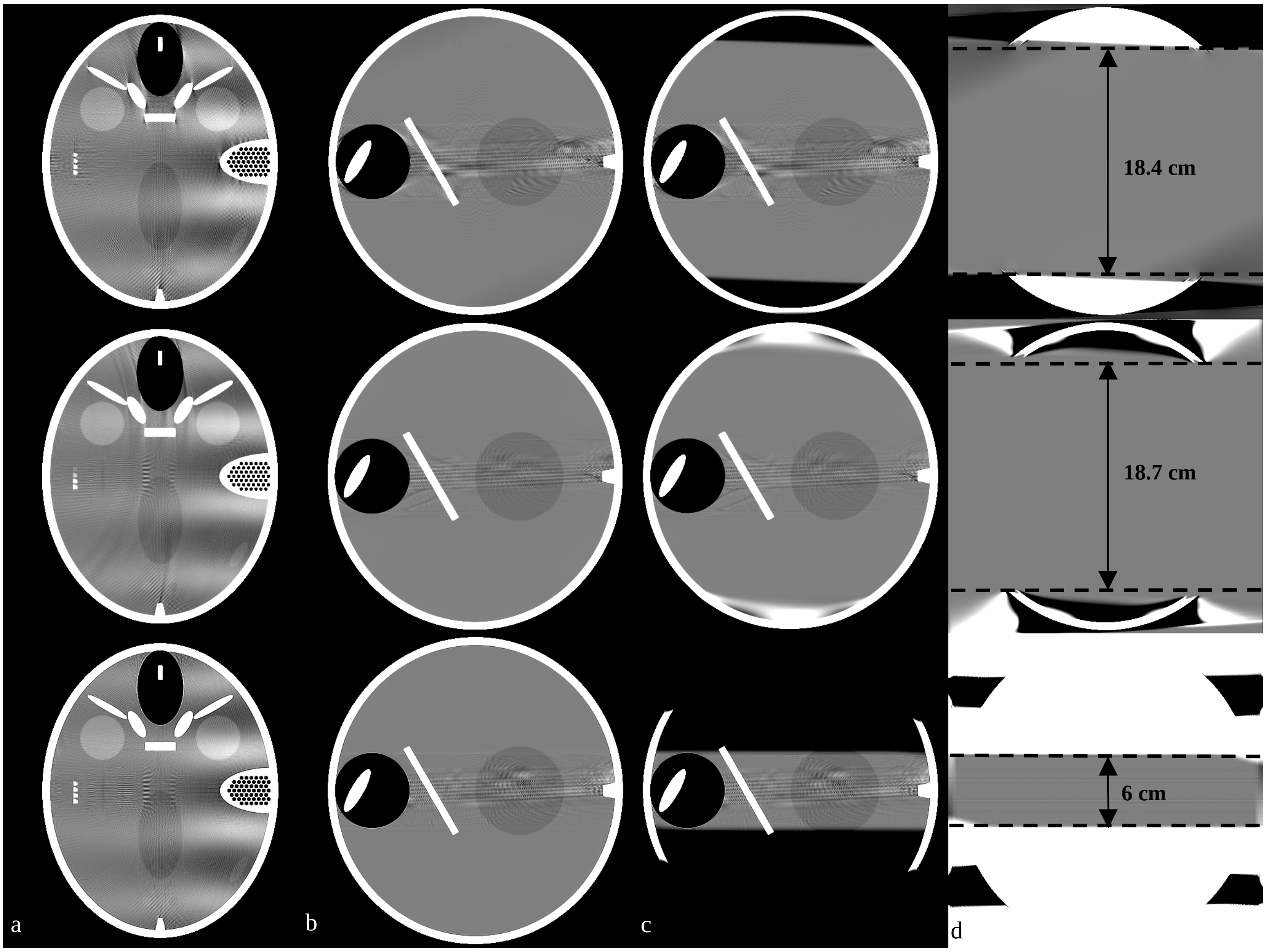


Fig. 1: Reconstruction of the 3D Head phantom for the three different algorithms. First row: VC. Second row: ZB. Third row: LR. a) xy-plane at z=0 b) yz-plane at x=0 with reconstruction of the whole object, c) yz-plane at x=0 with ROI reconstruction, d) difference between b) and c) - Window settings: a) to c) (C50, W80) d) (C0,W20)

The images are of good quality for all algorithms. The different objects of the phantom are well represented despite the relatively large cone-angle. The ROI reconstruction does not bring additional artifacts, which proves that the methods for short- and long-object reconstructions are equivalent.

A noticeable artifact is the wave in the axial slice, which comes from the small air bubbles of the right inner ear and is present for all methods. At the moment, we do not understand the origin of this artifact.

The ZB method presents some small artifacts in the region of the frontal sinus, which is probably due to a resolution mismatch between the 2 partial images. Some small artifacts are also to be seen in the VC method. None of these artifacts are present for the LR method, but we can notice undershooting around some of the high contrast objects. This undershooting is probably related to the calculation of the derivative filters in the 4-step algorithm.

B. Spatial Resolution

We simulate an ideal delta point at the origin of the coordinate system (x=y=z=0) and reconstruct the 3D Point Spread Function (PSF) for each algorithm. The reconstruction is performed with a voxel size of (0.02 mm)^3.

The in-plane Modulation Transfer Function (MTF) is obtained from the 3D PSF in the slice z=0 by averaging the radial profiles over all angles and performing a subsequent Hankel transform (see Fig. 2).

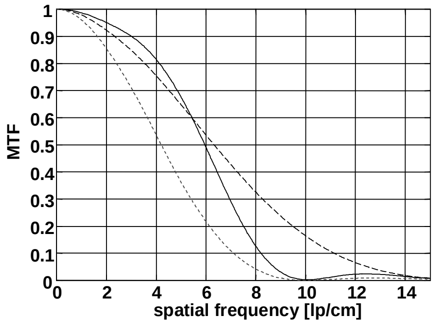


Fig. 2: In-plane MTF for the LR (solid line), the VC (dashed line) and the ZB method (dotted line). The Nyquist frequency is 10 lp/cm.

The VC and ZB methods should have a similar behavior since they both apply a ramp filter. The difference noted here comes from the difference in the current implementations since the ZB method uses a Hamming window in comparison to the Shepp-Logan window used for the VC method. This also explains why the MTF of the VC method extends beyond the Nyquist frequency, thus causing a risk of aliasing. The ZB and LR algorithms cut off at the Nyquist frequency, while LR yields better frequency representations than ZB.

To characterize in-plane resolution, we choose the 5% value (ρ_5) of the in-plane MTF value. The values can be found in Table 1. A Nyquist frequency of 10 lp/cm is expected from the horizontal detector aperture.

To determine axial resolution, we calculate the Slice Sensitivity Profile (SSP) which is derived from the 3D PSF by taking the profile along z at the position x=y=0 (see Fig. 3). Axial resolution is characterized by the value of the Full Width at Half Maximum (FWHM) of the SSP (see Table 1).

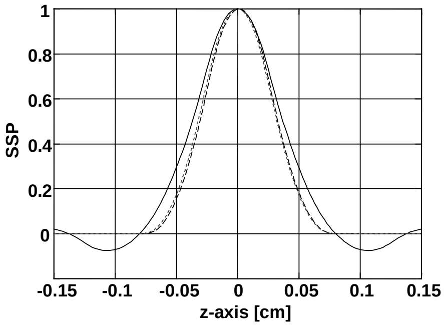


Fig. 3: SSP for the LR (solid line), the VC (dashed line) and the ZB method (dotted line)

The VC and the ZB methods have an identical SSP since they are almost equivalent in the axial direction (no filtering). We obtain a FWHM of about 1.27 collimated slices. The axial resolution of the LR method is less good (about 1.5 collimated slice widths). Moreover, we notice the undershooting already seen in Fig.1.

Note that the PSF's of the algorithms are shift-variant due to the magnification effect of the cone-beam backprojection.

Table 1
In-plane / axial spatial resolutions and noise level
for the three methods

	LR	VC	ZB
ρ_5 [lp/cm]	8.7	12.5	7.9
$\mathrm{FWHM}_{\mathrm{SSP}}$ [cm]	0.074	0.063	0.064
σ_{Noise} [HU]	9.7	14.2	6.7

C. Noise Properties

In order to model real quanta noise, we add attenuation dependent randomized gaussian noise to the original intensity projections before taking the logarithm. To compare the noise performances of the algorithms, we simulate and reconstruct a sphere of diameter 20 cm of constant density. As a measure of noise, the standard deviation of pixel gray values is evaluated in an area of (2 cm × 2 cm) centered in the slice z=0.

The amount of noise is strongly related to spatial resolution (see Table 1). Thus, the VC method with its superior spatial resolution shows more noise than the other methods.

V. PRACTICABILITY

A. Overscan

As shown in Fig. 1.c, the three methods require different overscan ranges to reconstruct a ROI. This overscan depends on the transaxial radius of the ROI (see Fig. 4). For objects of small radius (radius < 0.5 of the scan path radius), the LR method needs a larger overscan than the other two methods. The minimum spiral length required for exact ROI reconstruction is still unclear and is subject to further investigations.

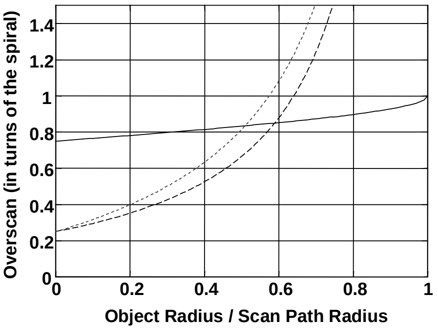


Fig. 4: Overscan expressed in terms of turns of the spiral as a function of the radius of the object: LR (solid line), VC (dashed line) and ZB method (dotted line).

For the Head Phantom (maximal object radius = 12 cm), we find overscan values of 4.6 cm, 5.2 cm and 10 cm along the z-axis for the VC, ZB and LR methods, respectively, which corresponds to a minimal theoretical ROI size of 16.4 cm, 15.2 cm and 5.6 cm. Since the object is not a cylinder but an ellipsoid, the real size of the ROI's might be larger. Note that the ROI that can be exactly reconstructed usually is not cylindrical but has slanted edges with respect to the xy-plane.

B. Reconstruction Time

The reconstruction time of the three non-optimized implementations presented here for the reconstruction of the ROI of the Head phantom are given in Table 2. The times have been obtained on a Pentium 450MHz with 512 MB memory.

The reconstruction times are quite long in comparison to standard 2D CT reconstruction algorithms. The VC and ZB methods are much faster than the LR method because of the simple filtering step. The ZB method requires more effort due to the separate treatment of the projection data on the mask boundaries consisting of 3D re- and backprojection, but on the other hand the backprojection is much faster since the size of the filtered detector is smaller.

Table 2
Reconstruction time in hours for the ROI reconstruction of Fig. 1.c.

LR	VC	ZB
100	45	51

C. Data Handling

The VC and LR methods are quite attractive since they are of real FBP-type. The ZB method in comparison requires many steps for creating the 2 partial images and therefore also a lot of disk memory to save the temporary results. On the other hand, the FBP for VC and LR has to be done on a virtual detector which is much larger than the original detector depending on the reconstructed ROI size, whereas the FBP of ZB can be done on the original detector. However it is possible to perform multi ROI reconstructions as proposed in [1].

VI. CONCLUSION

All three algorithms achieve an image quality good enough to match the requirements of clinical applications. However, the long computation times are prohibitive for any usage beyond research applications.

In terms of image quality, all algorithms investigated here are comparable. It turns out that the three algorithms have a good stability and that the main factor affecting the image quality of the reconstructions is the filter design. In the current implementations, the images of the LR method show less artifacts, while the VC algorithm yields the higher spatial resolution. In terms of practicability, the VC method is superior to the ZB and LR methods. For most clinical applications, it requires fewer projections and less computing time to reconstruct a given ROI. The overscan strongly depends on the size of the objects; while the LR method is not competitive for small objects (e.g. head), the amount of overscan is comparable for larger objects like thorax or abdomen.

VII. ACKNOWLEDGMENT

This work was supported in part by the Bavarian Center of Excellence for Medical Imaging FORBILD.

VIII. REFERENCES

[1] H. Kudo, F. Noo, M. Defrise, "Quasi-exact filtered backprojection algorithm for long-object problem in helical cone-beam tomography", *IEEE Trans. Med. Imag.*, vol. 19, Sept 2000, pp. 902-921

[2] M. Defrise, F. Noo, H. Kudo, "A solution to the long-object problem in helical cone-beam tomography", *Phys. Med. Biol.*, vol. 45, pp. 623-644, March 2000

[3] F. Sauer, S. Samarasekera, K.C. Tam, "Practical cone-beam image reconstruction using local regions-of-interest", U.S. patent 6,009,142, December 28, 1999

[4] K. Sourbelle, G. Lauritsch, K.C. Tam, F. Noo, W.A. Kalender, "Performance evaluation of local ROI algorithms for exact ROI reconstruction in spiral cone-beam computed tomography", *Proc. of IEEE Med. Imag. Conf*, Lyon, 2000

[5] H. Kudo, F. Noo, M. Defrise, "Cone-beam filtered-backprojection algorithm for truncated helical data", *Phys. Med. Biol.*, vol. 43, pp. 2885-2909, 1998

[6] K.C. Tam, B. Ladendorf, F. Sauer, G. Lauritsch, A. Steinmetz, "Backprojection spiral scan region-of-interest cone-beam CT", *Proc. Of SPIE Med. Imag.*, vol 3661, pp 433-441, 1999

[7] K.C. Tam, "Helical and circle scan region of interest computerized tomography", *US. Patent 5,463,666*, October 31, 1995.

[8] S. Schaller, F. Noo, F. Sauer, K.C. Tam, G. Lauritsch, T. Flohr, "Exact Radon rebinning algorithm for the long object problem in helical cone-beam CT", *IEEE Trans. Med. Imag.*, vol. 19, May 2000

[9] K.C. Tam, "Exact local regions-of-interest reconstruction in spiral cone-beam filtered-backprojection CT: theory", *Proc. of SPIE Med. Imag.*, vol. 3979, pp. 506-519, 2000

[10] G. Lauritsch, K.C. Tam, K. Sourbelle, S. Schaller, "Exact local regions-of-interest reconstruction in spiral cone-beam filtered-backprojection CT: numerical implementation and first image results", *Proc. of SPIE Med. Imag.*, vol. 3979, pp. 520-532, 2000

[11] http://www.imp.uni-erlangen.de/forbild

[12] C. Jacobson, "Fourier methods in 3D-reconstruction from cone-beam data", *PhD Thesis*, Linköping University, Sweden, 1996

Filtering point spread function in backprojection cone beam CT and its applications in long object imaging

K.C. Tam[a], G. Lauritsch[b], and K. Sourbelle[c]
[a]Siemens Corporate Research, Inc. Princeton, NJ 08540
[b]Siemens AG, Medical Solutions, Erlangen, Germany
[c]Institute of Medical Physics, University of Erlangen, Erlangen, Germany

Abstract

In backprojection cone beam CT each cone beam image is first filtered, then 3D backprojected into the object space. In this paper the filtering point spread function (PSF) is derived analytically. It is found that the PSF is in the form of 1D Hilbert transforms. The PSF finds applications in a number of aspects in long object imaging, including backprojection implementation of the local ROI algorithm, elimination of the second intersection artifact, reduced pitch spiral scanning for increased S/N, and reduction of spiral overscan in long object imaging.

I. Filtering point spread function

Backprojection cone beam image reconstruction [1,2,8,9] consists of two steps: a 2D step and a 3D step. First, each cone beam image undergoes 2D filtering. Then the filtered image is backprojected into the object space in the 3D step. 2D filtering consists of the following 4 sub-steps: (1) 1D projection of the cone beam image at angle θ; (2) Differentiation of the projections; (3) Backprojection of the projection derivative in the same direction. Sub-steps (1) through (3) are carried out for θ in the angular range [η-π/2, η+π/2] forming a backprojection image, where η is the angular displacement (from the *u* axis) of the direction $\vec{t}$ which is the projection of the scan path tangent on the detector. Finally, (4) take the derivative of the backprojection image in the projected scan path direction $\vec{t}$. The filtered image Y after these 4 sub-steps can be written as:

$$Y = D_t\{\int d\theta[B(\theta)D_{r(\theta)}(\theta)P(\theta)]\}M(X) \qquad (1)$$

where X is the cone beam image, M the data-combination masking operation [5], P(θ) the projection operation in the direction θ, $D_r(\theta)$ the differentiation operation w.r.t. the spatial variable *r* for the projection at angle θ, B(θ) the backprojection operation in the direction θ, and D_t the differentiation (spatial) operation in the projected scan path direction $\vec{t}$. The reason that the projection and backprojection operations in Equation (1) are carried out for θ in the angular range [η-π/2, η+π/2] is to ensure that the product of the derivative d/dt with d/dr has the same sign for all the angles within the range; otherwise the signs of the product at different angles will be inconsistent.

The sub-steps (1) through (3) can be combined into a 1D convolution step as follows:

Theorem 1. The combined operations of sub-steps (1) through (3) in the angular range [χ-π/2, χ+π/2] is equivalent to the 1D Hilbert transform in the direction of the unit vector $\vec{\chi}$ which makes an angle χ with the detector *u* axis:

$$H_{\vec{\chi}} = \int_{\chi-\pi/2}^{\chi+\pi/2} d\theta\, B(\theta) D_{\vec{r}(\theta)} P(\theta)\,, \qquad (2)$$

where $H_{\vec{\chi}}$ is the 1-D Hilbert transform in the direction of the unit vector $\vec{\chi}$. $H_{\vec{\chi}}$ can be expressed as a 2-D function in Fourier space

$$H_{\vec{\chi}}(\vec{k}) = i \cdot \mathrm{sign}(\vec{\chi} \cdot \vec{k})\ . \qquad (3)$$

For the cases where the projection/backprojection angular range is less than π, we have the following results:

Theorem 2. The combined operations of sub-steps (1) through (3) in the angular range [θ_1, θ_2], where $\theta_1 \in$ [η-π/2, η+π/2] and $\theta_2 \in$ [η-π/2, η+π/2], is equivalent to the sum of two 1D Hilbert transforms:

$$\int_{\theta_1}^{\theta_2} d\theta\, B(\theta) D_{\vec{r}(\theta)} P(\theta) = 1/2\left(H_{\overrightarrow{(\theta_2-\pi/2)}} + H_{\overrightarrow{(\theta_1+\pi/2)}}\right) \qquad (4)$$

II Backprojection implementation of local ROI algorithm

An exact backprojection driven spiral scan cone beam CT algorithm for ROI reconstruction using the local ROI technique was reported in [3,4]. For views in the interior of the spiral where the entire cone beam images undergo the filtering operation in Equation (1), the 2D filtering operation can be simplified to the efficient 1D ramp filtering operation in the direction $\vec{t}$. For the views near the two ends of the spiral, the projection operation P(θ) is applied to only part of the masked cone beam image M(X). This is illustrated in Figure 1 which shows the projection operation for the cone beam images in the view angular range [π, 2π] measured from the spiral top. The point C_0 on the *u* axis is determined by the angular displacement of the current source position from the spiral top. During the projection operation, the line integrals on line segments which cross the *u* axis to the left of C_0 are computed between the *u* axis and the bottom mask boundary, and the line integrals on line segments which cross the *u* axis to the right of C_0 are computed between the top and the bottom mask boundary.

Applying Equations (2) and (4) yields the result that the $\int d\theta[B(\theta)D_r(\theta)P(\theta)]$ portion of the 2D filtering can be formulated as a superposition of spatially variant 1-D Hilbert transforms $H_{\vec{t}}$, $H_{\vec{u}}$ and $H_{\vec{\rho}(u,v)}$ in the directions $\vec{t}$ of the projection of the tangent of the spiral path, $\vec{u}$ of the horizontal detector axis and $\vec{\rho}(u,v)$ of the unit vector pointing from the point C_0 to the detector pixel (u,v), respectively [10]. These 1-D Hilbert transforms are illustrated in Figure 2.

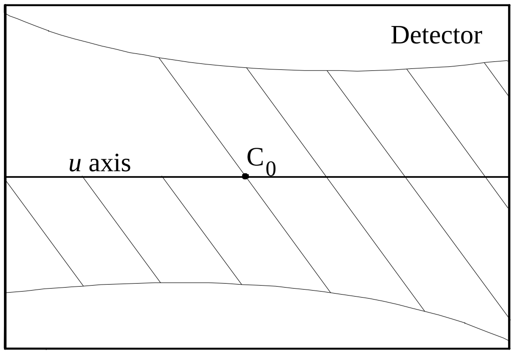


Figure 1. Limits for integration line segments for source positions in the angular range [π,2π] from the spiral top

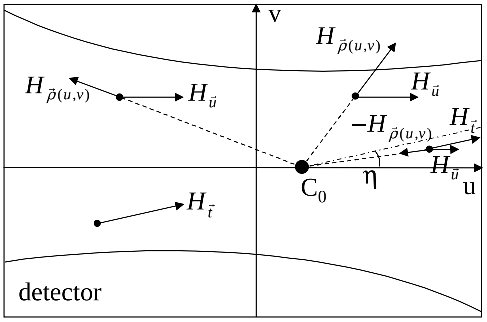


Figure 2: 1-D Hilbert transforms contributing to the filtering of a cone-beam projection around the upper border of the ROI as illustrated in Figure 1

III. Elimination of Second Intersection Artifacts

A key part of the exact spiral cone beam CT reconstruction algorithm is the masking operation M to restrict the projection data to the appropriate angular range required for data combination [5]. The mask consists of a top curve and a bottom curve formed by projecting on the detector the spiral turn above and the turn below from the current source position. Such masking procedure corresponds for the most part to the angular range bound by the prior and the subsequent source positions for data combination. Portions of some line integrals intersecting the mask, however, do not conform to the proper data combination angular range. Consider the top mask boundary and the line L illustrated in Figure 3, where the spiral path which projects onto the mask boundary scans from right to left. Line L intersects the mask boundary at 2 points M_1 and M_2, in other words the integration plane defined by the line L and the current source position intersects the scan path at M_1, M_2, and the current source position. It can be easily seen that M_1 is the next source position after the current one, and M_2 is the next source position after M_1. Thus the portion of the line segment that corresponds to the data combination angular range, i.e. the x-ray data in the angular range bounded by the previous source position below and the next source position above, is the segment to the right of M_1. It is this segment alone that should be included in the projection operation P(θ), i.e. line integration.

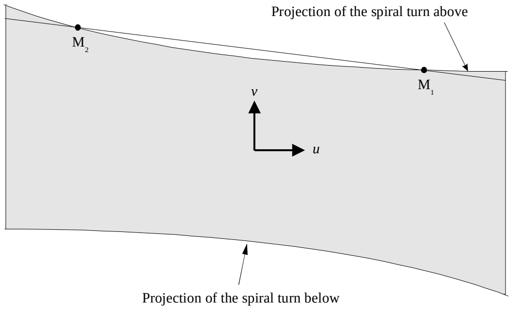
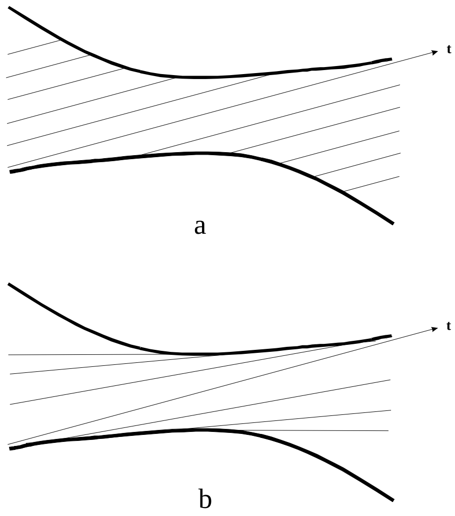


Figure 3. A line of integration intersecting the top mask two times.

In the literature it has been shown [8,9] that the filtering operation of Equation (1) can be simplified using techniques such as 1D ramp filtering, Fourier space convolution, or linogram. In arriving at these results it is assumed that the entire line segment is included in the projection operation. That is to say not only the segment to the right of M_1 but also the segment to the left of M_2 is included in the projection operation. Such unneeded contribution to projection occurs whenever the line of integration intersects the mask boundary more than once. We refer to such cases as second intersection artifacts.

Errors arising from such line integration occur on the line segment after the second intersection point M_2. A method to correct for these artifacts in backprojection cone beam CT image reconstruction was reported in [6]. It was shown that the projection lines that intersect the top mask boundary more than once are those that lie within the angular range $\mathscr{A}_1 = [\tau(-W/2)-\pi/2, \tau(\infty)+\pi/2]$, where $\tau(\infty)$ is the angle the tangent to the top mask boundary at $u = \infty$ makes with the u axis, $\tau(-W/2)$ is the angle the tangent to the top mask boundary at the left edge $u = -W/2$ of the detector makes with the u axis; $\tau(\infty) = \eta = \tan^{-1}(h/2\pi a)$ is also the angle of the projection of the scan path direction on the detector. For the correction of the second intersection artifacts for the upper mask boundary, the combined operation BD_rP is applied to the cone beam image in the limited angular range $\theta \in \mathscr{A}_1$ on the affected portions of the line segments. Applying Equations (2) and (4) yields the result that the $\int d\theta[B(\theta)D_r(\theta)P(\theta)]$ portion of the 2D filtering results in two Hilbert transforms: ($\frac{1}{2} H_{\vec{t}}$) in the directions $\vec{t}$ of the projection of the tangent of the spiral path, and ($\frac{1}{2} H_{\vec{\alpha}}$) in the direction of the unit vector $\vec{\alpha}$ of the line which makes an angle $\alpha \in \mathscr{A}_1$ with the detector u axis, intersects the pixel to be filtered, and tangential to the mask boundary. These two Hilbert transforms are sketched in Figure 4a and 4b respectively.



a

b

Figure 4. 1D Hilbert transforms: a. $H_{\vec{t}}$ b. $H_{\vec{\alpha}}$

IV. Increasing S/N with reduced pitch scanning

Increase the x-ray dosage in spiral cone beam scan to obtain higher S/N can be achieved by reducing the spiral pitch to 1/3, 1/5, 1/7,…, of the original spiral pitch [11]. Take the case of pitch = 1/3. Instead of combining cone beam data using the standard mask [5], the masked formed by projecting on the detector the second spiral turn above and the second turn below from the current source position is used. Data combination performed in this way contains overlapping portions, and the overlap is such that there is a data redundancy of factor 3. With the reduced pitch = 1/3, the modified mask boundary is the cone beam projection from the current source position of the second spiral turn above and the second spiral turn below, rather than the spiral turn above and the spiral turn below. Since there are approximately 3 times the number of source positions in the pitch = 1/3 scan compared to the pitch = 1 scan, the total radiation exposure in the former is thus increased threefold compared to that of the latter, resulting in higher signal-to-noise ratio.

There is a flaw in this method, however, because a small number of integration planes which intersect the spiral path only once also intersect the spiral path with reduced pitch only once rather than 3 times. To correct for this flaw, one solution is to calculate the contribution of these integration planes to the cone beam image filtered with the combined operation BD_rP, multiply this portion by a factor of 2, and add the result to the filtered cone beam image. The procedure is as follows. Construct the two common tangents at angles $\tau(\infty)$ and $\tau(-\infty)$ respectively connecting the top and bottom mask boundary curves diagonally. Then apply the combined operation BD_rP to the cone beam image in the limited angular range $\theta \in \mathscr{A}_2 = [\tau(-\infty)-\pi/2, \tau(\infty)+\pi/2]$, and only to those projections which do not intersect either the top or the bottom mask boundary. Applying Equations (2) and (4) yields the result that the procedure can be simplified as follows: for each angle $\alpha \in \mathscr{A}_2$, 1D Hilbert transform along the 2 lines in the direction of the unit vector $\vec{\alpha}$ which makes an angle α with the detector u axis and are tangential to the top and the bottom mask boundaries respectively. This procedure is illustrated in Figure 5.

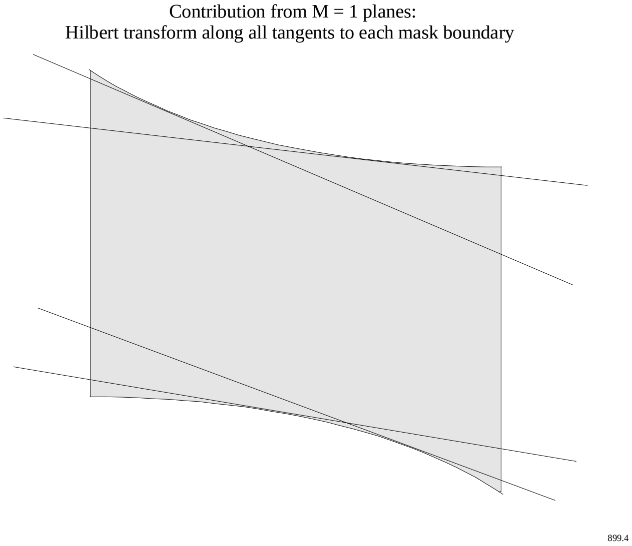


Figure 5. 1D Hilbert transforms illustrated at 2 of the angles $\in \mathscr{A}_2$.

V. Reduction of spiral overscan in long object solution

The results in Equations (2) and (4) are also applied to reduce the amount of overscan in the backprojection local ROI algorithm [3,4]. Briefly, the Hilbert transforms shown in Figure 2 can be used to determine which portion of the ROI is affected by the cone beam data at each source position near the spiral ends; those data that do not affect the ROI will not be needed. The details will be presented in [7].

VI. Acknowledgments

This work was supported in part by Bavarian Center of Excellence for Medical Imaging (FORBILD).

VII . References

1. Defrise, M. and Clack, R., "A Cone-Beam Reconstruction Algorithm Using Shift-Variant Filtering and Cone-Beam Backprojection", IEEE Trans.Med. Imag., MI-13, 186, 1994.
2. H. Kudo and T. Saito, “Derivation and Implementation of a Cone-Beam Reconstruction Algorithm for Nonplanar Orbits”, *IEEE Trans.Med. Imag*., MI-13, pp. 196-211, 1994.
3. K.C. Tam, “Exact local regions-of-interest reconstruction in spiral cone-beam filtered-backprojection CT: theory”, *Proc. of SPIE Medical Imaging Conf.*, vol. 3979, pp. 506-519, 2000
4. G. Lauritsch, K. C. Tam, K. Sourbelle, and S. Schaller, “Exact Local Regions-of-Interest in Spiral Cone-Beam Filtered-Backprojection CT: numerical implementation and first image results,” *Proc. of SPIE Medical Imaging Conf.*, 3979, pp. 520-532, 2000.
5. K.C. Tam, “Helical and circle scan region of interest computerized tomography”, *US. Patent* 5,463,666, Oct 31, 1995.
6. K.C. Tam, G. Lauritsch, and K. Sourbelle, “Eliminating the Second-Intersection Contributions to Spiral Scan Cone Beam CT”, IEEE Nucl. Sci. Symp. and Med. Imag. Conf. Conference Record, October 15-20, 2000, Lyon, France.
7. K.C. Tam, G. Lauritsch, and K. Sourbelle, “Overscan reduction in spiral scan long object problem”, to be submitted to IEEE Nucl. Sci. Symp. and Med. Imag. Conf., November 4-10, 2001, San Diego.
8. H. Kudo, F. Noo, and M. Defrise, "Cone-beam filtered-backprojection algorithm for truncated helical data", *Phys. Med. Biol*., 43, pp. 2885-2909, 1998.
9. K.C. Tam, B. Ladendorf, F. Sauer, G. Lauritsch, and A. Steinmetz, “Backprojection spiral scan region-of-interest cone beam CT”, *Proc. SPIE Medical Imaging 1999: Physics of Medical Imaging*, pp. 433-441, 1999.
10. G. Lauritsch, K. C. Tam, and K. Sourbelle, “Solution to the Long Object Problem by Convolutions with Spatially Variant 1-D Hilbert Transforms in Spiral Cone-Beam Computed Tomography”, IEEE Nucl. Sci. Symp. and Med. Imag. Conf. Conference Record, October 15-20, 2000, Lyon, France.
11. K.C. Tam, “A Method To Increase The Photon Counts For Enhanced Signal-To-Noise In Cone Beam CT”, Siemens Corporate Research invention disclosure IPD 99E9131 US, July 22, 1999.

Segmented Cardiac Volume Reconstruction - A Novel Reconstruction Scheme for Multislice Cardiac Spiral CT

Herbert Bruder, Karl Stierstorfer, Bernd Ohnesorge, Stefan Schaller, Thomas Flohr

Abstract—At the advent of Multislice Computed Tomography (MSCT), cardiac volume imaging is an exciting and promising tool to image the heart's anatomy and dynamics. Volume reconstruction algorithms have been implemented successfully for four slice systems allowing for retrospectively imaging the human heart in spiral scanning mode at any heart phase within the heart cycle. However, this method fails due to cone artifacts for a large number of detector slices. We present a new cardio volume reconstruction technique – called Segmented Cardio Volume Reconstruction (SCVR) – for multislice systems with a large number of detector slices. The method is characterized by computing a multitude of tilted segment images optimally fitted to the spiral path, which in a second step have to be reformatted to axial segment images and added to a complete CT image at the selected phase in the heart cycle.

Excellent image quality using a simulation study will be demonstrated for a 16 slice detector. However, we think this technique will become the method of choice in cardiac CT volume imaging for future multislice detector systems with even larger cone angles. Moreover, the method makes possible an extremely fast generation of a new volume stack for a new freely selectable heart phase by only adding the precalculated axial segment images to a complete CT image dataset. This property is extremely useful for the clinician in particular to adjust the appropiate heart phase for visualizing the coronary arteries.

Index Terms—cardiac volume imaging, spiral scanning, multislice detector

I. INTRODUCTION

Only in the recent past cardiac volume imaging with CT was boosted to high quality imaging using multislice detector technology. Whereas magnetic resonance imaging is highly effective imaging the morphology and functional dynamics of the human heart due to its excellent low contrast detectibility and its capability of diffusion imaging, CT has its domaine in detecting coronary artery disease by visualizing the coronary anatomy. Hence it can be very useful to detect stenotic malformations and in case of negative prediction avoid additional catheter study with classical angiography. Moreover, cardiac CT imaging might have the potential to detect soft plaques in an early stage of development thus providing a powerful tool to diagnose risk for infarction. Early cardiac CT techniques using prospective triggering in sequence mode are now extended by retrospective imaging methods in spiral scanning. The breakthrough of this technology came with multislice detectors in the last two years [1],[2]. Calcium scoring, for example, thus becomes a reliable and efficient tool to predict plaque frequency and increase in the coronary arteries [3].

Cardiac volume imaging currently implemented with four slice detector systems is characterized by continous data sampling and retrospective volume reconstruction of the human heart at any phase in the heart cycle. In principle the temporal resolution is adjustable. In clinical use, however, single and biphasic reconstruction using one or two heart cycles generating one dataset for a single image proved to be sufficient.

The method, however, neglects the cone angle of multislice data and hence is suitable only for a small number of detector slices. In addition, in order to generate a new volume stack for a new selectable heart phase a repeated reconstruction has to be performed for the whole image volume. These disadvantages can be overcome by using the method proposed in this paper. The basic idea is to reconstruct a set of segmented (tilted) volume stacks called booklets consisting of several booklet pages which in a second step, are reformatted to axially segmented volume stacks, and only in a third step these are temporally rebinned to a complete CT dataset for any selected heart phase selected.

II. METHOD

A. Short Review of Segmented Multiple Plane Reconstruction (SMPR)

In this section we will give a short description of the basics of Segmented Multiple Plane Reconstruction (SMPR) which will be presented in detail at this same conference in another presentation. This new reconstruction scheme for

---

H. Bruder, SIEMENS, Medical Engineering Group,
Siemenstr. 1, 91301 Forchheim, Germany,
Phone: +49-9191-188793
Fax: +49-9191-189996
e-mail: herbert.bruder@med.siemens.de

multislice spiral CT is based on the idea that the spiral focus path can be decomposed into small, overlapping segments s_0. One turn of the spiral is divided into N_{seg} subsegments. In each of these segments a stack of image pages $I_{s0,t}(x,y)$ called booklet are defined ($1 \leq t \leq N_{tilt}$; N_{tilt} denotes the number of booklet pages, which should at least equal the number of detector slices). The image pages are tilted in two directions in order to optimally fit the spiral path and to optimally use the multislice detector data. For each of these pages a set of x-rays is identified to be used for convolution and conventional 2D-backprojection. Due to the tilt of the segment pages the foci associated with these x-rays are very close to the corresponding page, thus largely avoiding cone artifacts. We use the parallel geometry in order to employ a 2D- Fourier backprojection algorithm for image reconstruction.

B. Extension of SMPR for Segmented Cardio Volume Imaging (SCVR)

1) Volume reformatting

After 2D-backprojection a set of image pages for each segment booklet is generated which for themselves do not represent CT images, because only a small number of projections according to the length of the predefined segments are utilized for each 2D-backprojection.

In a first step this stack of tilted booklet pages has to be reformatted to axial segment images located at different image positions z_{img}. This is required because the x-rays used for each separate backprojection are adapted to tilted booklet pages. An axial segment page $I_{so,}^{zimg}$ at z-position z_{img} is calculated using the following equation:

$$I_{so}^{zimg}(x,y) = \frac{1}{H(x,y)} \cdot \sum_t h(\left|z_t(x,y) - z_{img}\right|) \cdot I_{s0,t}(x,y) \quad (1)$$

The weighting function h controls the reconstruction slice-width, H represents the sum of all weights and $z_t(x,y)$ denotes the pixel-dependent z-position of the booklet page $I_{s0,t}(x,y)$.

After reformation a stack of segmented axial booklet pages is generated, which for themselves still do not represent complete CT images. Most importantly, each reformattet segment booklet corresponds to a time coordinate of the simultaneously recorded ECG, thus a phase consistent formation of complete CT images is possible as described in the next section.

2) Temporal rebinning

The volume reformation described in the previous section has to be performed for each of the segment booklets to cover the whole image volume of the heart. Thus, for each segment s_0, an image volume stack $V_{z0,s0}$ is generated. In a second step, the axial segment images of the volume stacks have to be temporally rebinned and added to complete CT images.

For this purpose a set of corresponding time coordinates $T_{0,n}$ are extracted in each heart cycle of the patient's ECG, which has been recorded simultaneously to the scanning procedure ($1 \leq n \leq N$; N number of heart cycles contained in the ECG). For each z-position z_{img} we have to identify $N_{seg}/2$ axial segment images which have to complement each other to an angular range of π. Appropriate segments can be retrieved either from one heart cycle in case of single phase rebinning or from a multitude of heart cycles in case of multiphasic rebinning. In principle, using a multitude of heart cycles for temporal rebinning provides adjustable temporal resolution.

However, in this study we restrict ourselves to single and biphasic rebinning algorithms, respectively (Fig 2). For a scanner rotation time of T_{rot}=0.5s which is provided by almost all advanced commercial scanners thus we get a temporal resolution of at least 250 ms, which in case of biphasic rebinning is even superior. We present an algorithm which automatically decides whether single phase or biphasic rebinning is used. The rebinning mode depends on the actual patient's heart rate.

In case of biphasic rebinning the CT dataset is composed of segments acquired in consecutive heart cycles. The number of segments Δs_1, retrieved from the first heart cycle is given by:

$$\Delta s_1 = floor(\frac{N_{seg}}{2} \cdot (\frac{2 \cdot T_H}{T_{rot}} - n_1)) \quad (2)$$

where

$$n_1 = floor(\frac{2 \cdot T_H}{T_{rot}});$$

T_H denotes the length of the first heart cycle.

Hence we conclude that the single phase mode is active, if $\frac{2 \cdot T_H}{T_{rot}} - floor(\frac{2 \cdot T_H}{T_{rot}}) \geq 1$. Then $\Delta s_1 \geq N_{seg}/2$, which means that all of the segments required to complete a CT dataset to an angular range of π are retrieved from only one heart cycle.

The single phase mode is also activated at heart rates below a certain threshold depending on the pitch. The temporal resolution in case of biphasic rebinning is given by

$$\Delta T_{img} = \frac{T_{rot}}{2} \cdot \frac{\max(\Delta s_1, \Delta s_2)}{N_{seg}/2} \quad (3)$$

where Δs_2 denotes the length of the second data sector sampled in the second heart cycle.

If a new heart phase $T_{1,n}$ is selected, the presented algorithm allows for a very fast rebinning of a new volume stack belonging to the desired heart phase by only adding other suitable axial segment images to complete a CT image. Nowadays cardiac volume imaging algorithms repeatedly have to reconstruct a new volume stack for each change of heart phase. A fast change of heart phase, however, is very useful for adjusting the appropriate time instant to image coronary anatomy properly.

3) *Proposal for the table feed*

In order to guarantee gapless volume coverage of the heart volume it is required to adjust the table feed. The z-coverage z_V of each axial segment volume stack depends on the number of detector slices N_r and is given by:

$$z_V = f \cdot (N_r - 1) \cdot S$$

where S is the collimated slice-width and $f \leq 1$ denotes a geometrical factor which restricts the size of the target volume due to the obliqueness of the pages in each booklet. Hence, the precise value of f also depends on the image field of view. As can be seen from fig. 2 in case of the biphasic rebinning mode the spiral feed must not exceed half of the z-coverage z_V in consecutive heart cycles. Due to the possibility of interpolating missing pages at the top or the bottom of a segment booklet this limitation can be relaxed. Putting all together, in case of the biphasic rebinning mode a maximum table feed characterized by the maximum pitch value p_{max} can be calculated (here the pitch is defined as table feed per spiral rotation divided by the collimated slice-width):

$$p_{\max} \leq \frac{T_{rot}}{T_H} \cdot (f \cdot (N_r - 1) + 1) \qquad (4)$$

In case of the single phase rebinning mode p_{max} is given by:

$$p_{\max} \leq \frac{T_{rot}}{T_H} \cdot (f \cdot (N_r - 1) + \frac{1}{2}) \qquad (5)$$

As an example: for a 16 slice detector with gantry rotation time T_{rot}=0.5s and f = 0.85 we get the following pitch values p_{max}:

60 bpm: $p_{max} \leq 6.6$ (single phase rebinning)

90 bpm: $p_{max} \leq 5.5$ (biphasic rebinning)
(bpm=beats per minute).

Using a collimated slice-width of S = 1mm the entire heart can be scanned in about 11s! even if the biphasic rebinning mode is applied. As can be derived from equation (3) the temporal resolution is in the range of 125 – 250 ms for a 0.5s rotation scanner.

III. SIMULATION RESULTS

We present a simulation study based on a semi-anthropomorphic phantom of the human heart and thorax. In order to demonstrate the potential of the Segmented Cardiac Volume Reconstruction we simulated a realistic dynamic behaviour of the coronary arteries. We show axial and Multi-Planar-Reformattings (MPRs) for a 16 slice detector system.
Excellent image quality is achieved both in the single phase and biphasic rebinning mode at different heart rate (Fig 3).

IV. CONCLUSION

A new cardiac volume reconstruction technique, called Segmented Cardiac Volume Reconstruction (SCVR) has been presented. The method is based on an algorithm which takes into account the conical shape of projection data in multislice CT. Thus the new algorithm is well suited for multislice data with a large number of detector slices. In addition, the new algorithm allows a very fast recalculation of an image volume for a new selection of the heart phase within the heart cycle. This is a major advantage when adjusting the appropriate heart phase in order to visualize the coronary arteries. We think that SCVR will become the method of choice for cardiac imaging when larger cone angles in multislice CT, become more common place.

V. REFERENCES

[1] B. Ohnesorge, T. Flohr , Ch. Becker et. al, ‘Cardiac Imaging by Means of Electrocardiographicallly Gated Multisection Spiral CT: Initial Experience’, Radiology 2000, 217:564-571

[2] H. Bruder, S. Schaller, B. Ohnesorge, T. Mertelmeier, ‘High temporal resolution volume heart imaging with multirow computed tomography’, SPIE Med. Imag. Conf., Vol. 3661, 1999, 420-432

[3] C. Becker , T. Kleffel, A. Crispin, A. Knez, Y. Young, U. Schoepf, R. Haberl, M. Reiser (2001). ‘Coronary Artery Calcium Measurement: Agreement of Multirow Detector and Electron Beam CT’ Am J Roentgenol in press.

Fig 1: Parallel projection view of a segment booklet for a 16 slice detector. For better visibility, only four of the booklet pages are shown. In addition the virtual detector located in the center of rotation (D_c) and the physical detector (D_P) are shown. It should be pointed out that a curved physical detector in fan-beam geometry will show

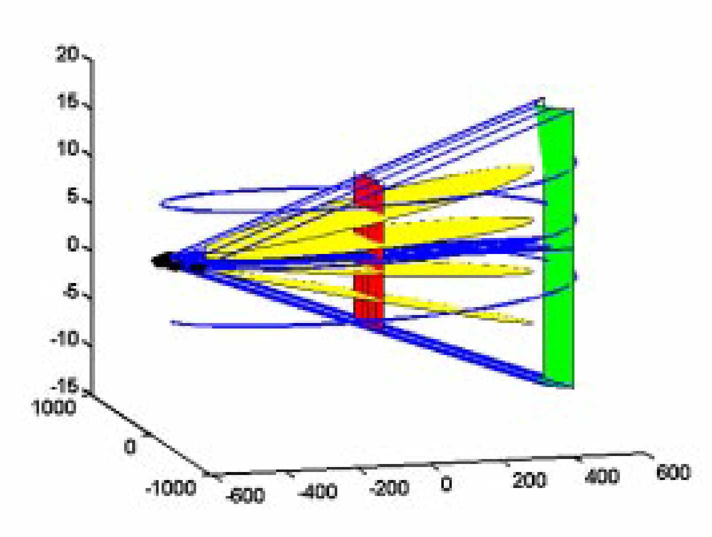


Fig 2 Principle of the Segmented Cardiac Volume Reconstruction (SCVR). The segmented volume stacks are reformatted to axial segmented volume stacks. In a second step a temporal rebinning procedure using temporal information from the patients ECG retrieves phase consistent segmented volume data and adds them to complete CT-images at the desired z-positions z_{img}. The biphasic rebinning mode is shown in the figure below.

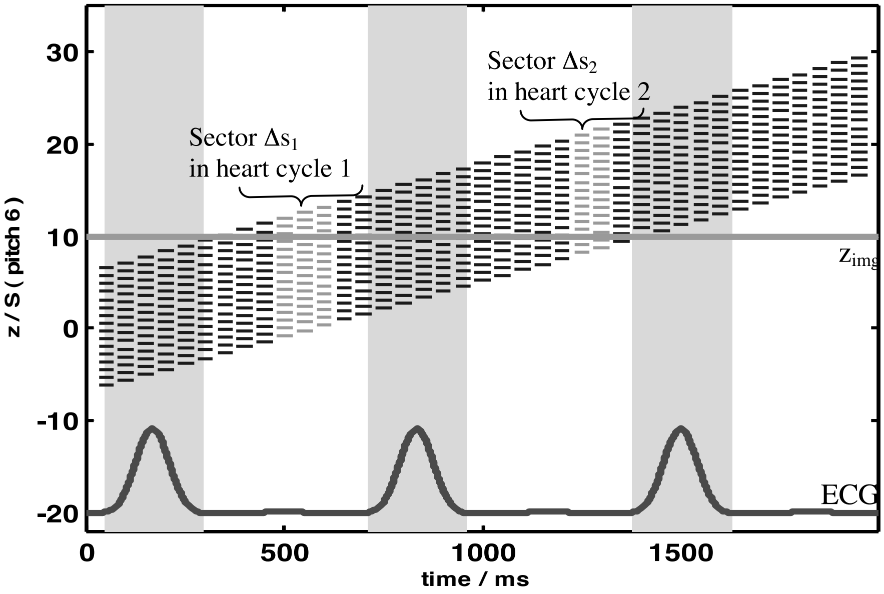


Fig 3: Oblique MPR of anthropomorphic heart phantom. A 16 slice detector with 1 mm collimated slice-width operating at pitch 6 was simulated. At a heart frequency of 90 bpm the biphasic rebinning mode was active. Both calcium and soft plaques are clearly delineated in the model of coronary artery.

Lung Nodule Imaging using Micro CT

N. Niki, Y. Kawata, M. Fujii[a], E. Matsui[b], N. Moriyama[c],Y. Tateno[d]
Univ. of Tokushima, [a]Toshiba IT & Control Systems
[b] Gifu Medical School, [c]National Cancer Center Hospital,
[d] National Institute of Radiological Science,

Abstract-- We are developing a micro-computed tomography for imaging lung nodules. The purpose is to enhance the physician performance in accessing the micro architecture of the nodule for classification between malignant and benign nodules. The basic components of the micro CT consist of a microfocus X-ray source, a specimen manipulator, an image intensifier detector coupled to charge-coupled device camera and an image processing unit. Three-dimensional image reconstruction was performed on a slice by slice technique. A standard fan-beam convolution and back-projection algorithm was used to reconstruct the center plane intersecting the X-ray source. The preprocessing of the 3-D image reconstruction included the correction of the geometrical distortions and the shading artifact introduced by the image intensifier. The main advantage of the system is to obtain a high spatial resolution which ranges between 5 μm and 25 μm. In this work we report on preliminary studies performed with the micro CT for imaging resected tissues of normal and abnormal lung. Experimental results reveal micro architecture of lung tissues, such as alveolar wall, septal wall of pulmonary lobule, and bronchiole. From the results, the micro CT is expected to have interesting potentials for high confidential differential diagnosis.

I. INTRODUCTION

Lung cancer is the leading cause of cancer deaths in the world. Early detection and treatment of lung cancers are crucially important to achieve high survival rate. There is hope in the possibility of early detection of lung cancer with the helical low-dose CT [1]. Computer-aided diagnosis is a promising approach to detect suspicious lesions on the thoracic CT images and alert physicians to these regions [2]. The detailed examinations of the small nodule lesions depend on the malignant potential. The differential diagnosis is ordinarily concluded by histological diagnosis from biopsy. It is often the case that biopsy technique becomes difficult according to nodule size is small. The advance of CT technology, such as a multi-slice CT scanner, provide fully 3-D images of pulmonary nodules with a high spatial resolution which ranges between 300 and 500 μm. There has been a considerable amount of interest in the use of 3-D thoracic CT images to observe small pulmonary nodules for differential diagnosis [3]. A number of investigators have developed feature extraction and classification methods for characterizing pulmonary nodules [4]-[8]. However, these spatial resolutions have limitation for the quantitative diagnosis of micro tissues that form the small pulmonary nodule. Then, it becomes necessary to develop a micro-computed tomography which enables to analyze the micro architecture of lung tissues.

In recent years, there has been an increasing interest for developing the micro CT system. Ruegsegger et al. developed two systems based on multiple fan beam scanner [9]. One of their systems is realized for bone samples and small laboratory animals with a spatial resolution of 20 μm and another system is used for the examination of patients with a spatial resolution of 120 μm. Johson et al. developed a volumetric micro CT with a spatial resolution of 50 μm based on cone-beam scanner for quantification of pulmonary arterial wall [10]. Since the size of lung tissue such as alveolar wall is the order of several μm, a visualization of lung tissues requires the higher spatial resolution. Ritman et al reported that the spatial resolution of 2 μm was achieved with synchrotron X-rays [11]. The spatial instruments require realizing the synchrotron-based system.

We present a micro CT system for analysis the micro architecture of lung tissues. The basic components consisted of a microfocus X-ray source, a specimen manipulator, an image intensifier detector coupled to charge-coupled device camera and an image processing unit. From results of the application to normal and abnormal lung tissues, we will demonstrate the availability of the micro CT for analysis of the lung micro architecture.

II. METHODS

Fig. 1 shows the block diagram of the micro CT. The CT mainly consists of a microfocus X-ray source, a specimen manipulator, an image intensifier detector coupled to CCD camera and an image processing unit. The minimal size of the microfocus X-ray source is 7 μm. The detector size of II can be chosen from three different sizes of 9, 6, 4.5 inch. The CCD array has three different areas corresponding to the II size with 1300 by 1030 pixels or 1300 by 512 pixels. The source-to-specimen distance and source-to-detector distance can be varied to obtain the desired magnification. The micromanipulator precisely positions the specimen table in the X-ray beams and rotates it under computer control through 360 degrees for projection data acquisition. This system can handle objects up to 200mm in diameter and up to 200mm in height. The acquisition time of 1800 views through 360 degrees is approximately one minute. The table position and motions and camera operations are under the control of the host computer with two CPUs (Pentium III Xeon, 500MHz) and 512MB main memory. The host computer receives the projection images and performs all data preprocessing, and image reconstruction is

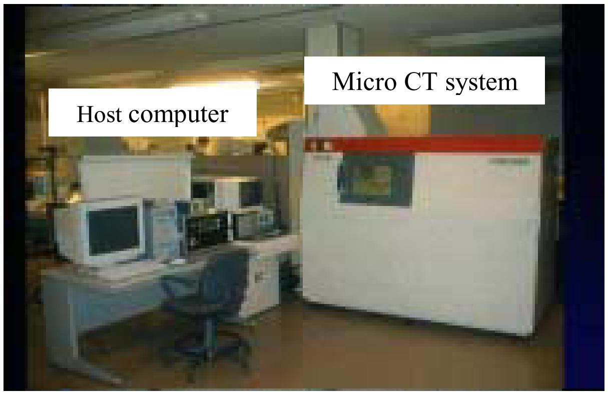


Fig.1. System overview

performed. Since the system geometry is a cone-beam geometry with a single circular orbit of the X-ray source, it is possible to apply cone beam reconstruction techniques [13] based on the Feldkamp algorithm [12]. These techniques provide approximate 3-D reconstructed images because they do not satisfy sufficiently the requirement for exactly 3-D image reconstruction [14]. In this CT, a standard fan-beam convolution and back-projection algorithm is used to precisely reconstruct the center plane intersecting the X-ray source. The preprocessing of the 3-D image reconstruction includes the correction of the geometrical distortions and the shading artifact introduced by II. The spatial resolution on the CT image ranges between 5 μm and 25 μm. The slice thickness is achieved up to the order of several μm.

III. RESULTS

In the preliminary study, we investigate how the spatial resolution affects on quantitative diagnosis of micro lung tissues. We used three kinds of lung tissues, normal lung tissue, peripheral type of lung adenocarcinoma, and pulmonary adenocarcinoma in emphysematous lung. In order to compare between two different spatial resolutions in micro CT images, we reconstructed two volumetric micro CT images with different resolution of 6 μm and 24 μm for each lung tissue specimen. The measurement condition used in this experiment is as follows; the size of the microfocus X-ray : 9 μm, X-ray tube current : 0.1mA, X-ray tube voltage : 60kV, slice thicknesses : 6 μm and 24 μm, and the number of slice : 30slices. Each tissue is reconstructed into 1024 x 1024 x 30 cubic voxels (each voxel has 6 μm or 24 μm on a side).

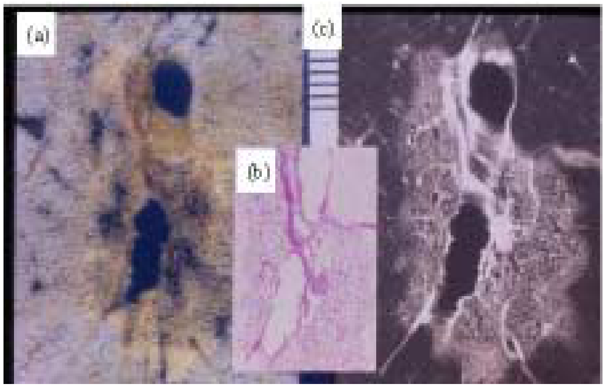
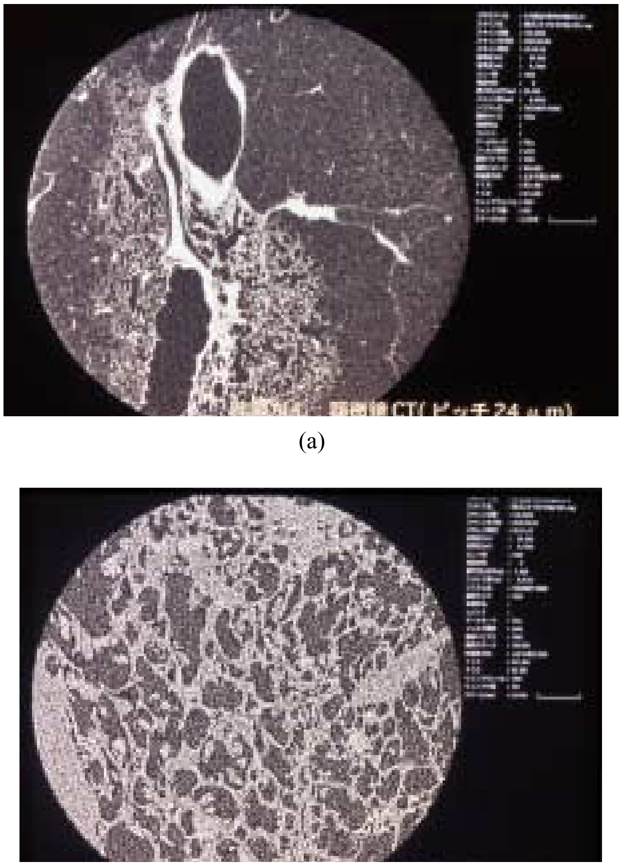


Fig.2. Abnormal tissue (peripheral type of lung adenocarcinoma) (a) Resected lung tissue. (b) Optical microscope image. (c) Soft X-ray image.

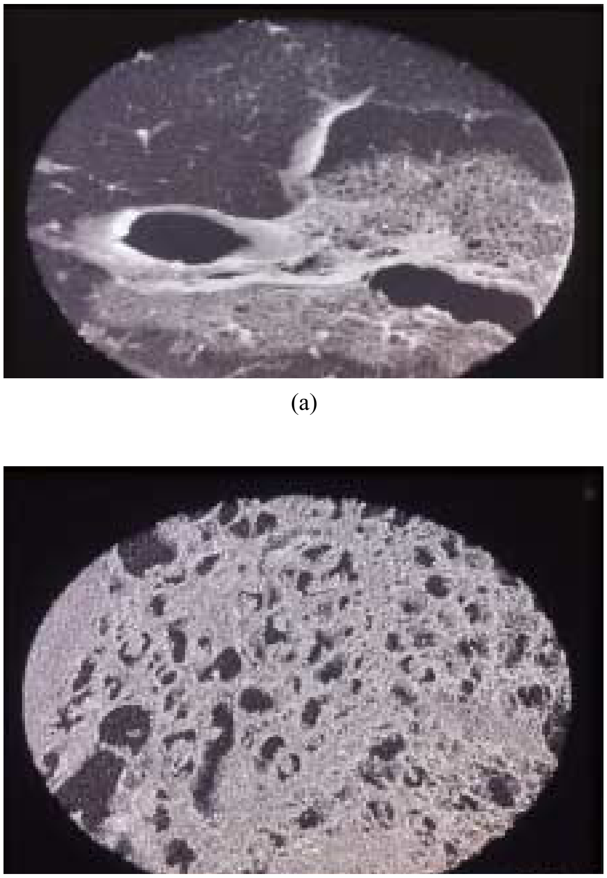


(a)

(b)

Fig.3. Micro CT images of the abnormal lung tissue. (peripheral type of lung adenocarcinoma) (a) Resolution : 24 μm. (b) Resolution : 6 μm.

Fig.2 shows the resected specimen, optical microscopic image, and soft X-ray image of the peripheral type of lung adenocarcinoma. In these figures the width of the alveolar wall where cancer cells are spread ranges from 12 μm to 59 μm . Fig. 3 shows the peripheral type of lung adenocarcinoma with two different spatial resolutions of 24 μm and 6 μm. Using the resolution of 24 μm, the alveolar wall of abnormal area is revealed due to wall thickening, while the alveolar wall of normal area can not be identified. With the resolution of 6 μm, the alveolar walls of both normal and abnormal areas can be easy to identify. Fig. 4 shows the 3-D display of the stacked slices of the peripheral type of lung adenocarcinoma with two spatial resolution of 24 μm and 6 μm. Using the resolution of 24 μm, the 3-D



(a)

(b)

Fig.4. 3-D display of the stacked Micro CT images of the abnormal lung tissue (peripheral type of lung adenocarcinoma) Resolution : 24 μm . (b) Resolution : 6 μm .

structure of the alveolar wall of abnormal area can be easy to identify. While with the resolution of 6 μm , the 3-D structure of the alveolar walls of both normal and abnormal areas are revealed.

Fig. 5 shows the resected specimen, optical microscopic image, and soft X-ray image of the pulmonary adenocarcinoma in emphysematous lung. In these figures the width of the alveolar wall where cancer cells are spread ranges from 12 μm to 59 μm . Fig. 6 shows slice images of the pulmonary adenocarcinoma in emphysematous lung with two different spatial resolution of 24 μm and 6 μm . Using the resolution of 24 μm , the alveolar wall of abnormal area is revealed due to wall thickening, while the alveolar wall of normal area can not be identified. With the resolution of 6 μm , the alveolar walls of both normal and abnormal areas can be easy to identify. Fig. 7 shows 3-D display of the stacked slices of the pulmonary adenocarcinoma in emphysematous lung with two spatial resolutions of 24 μm and 6 μm . Using the resolution of 24 μm , the 3-D structure of the alveolar wall of abnormal area can be easy to identify. While with the resolution of 6 μm , the 3-D structure of the alveolar walls of both normal and abnormal areas are revealed.

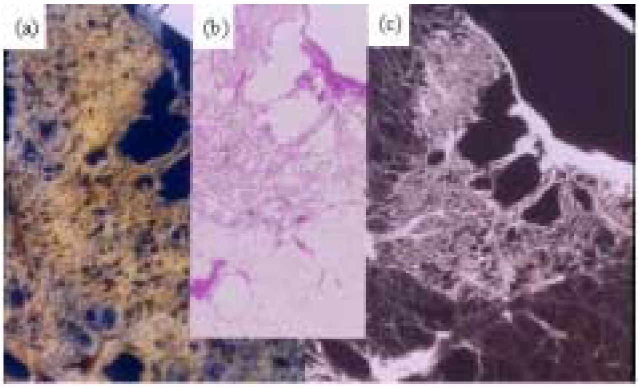
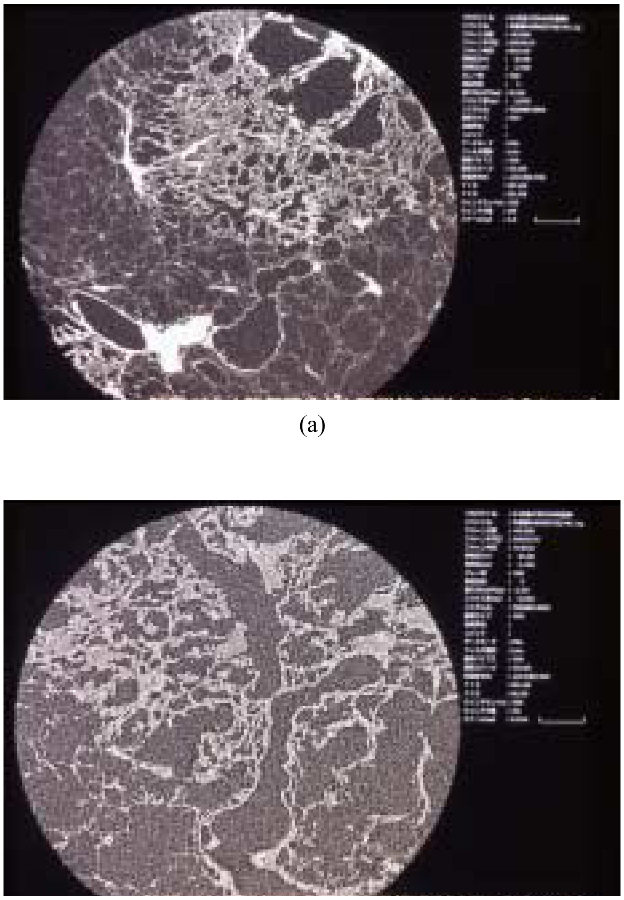


Fig. 5. Abnormal tissue (pulmonary adenocarcinoma in emphysematous lung) (a) Resected lung tissue. (b) Optical microscope image. (c) Soft X-ray image.



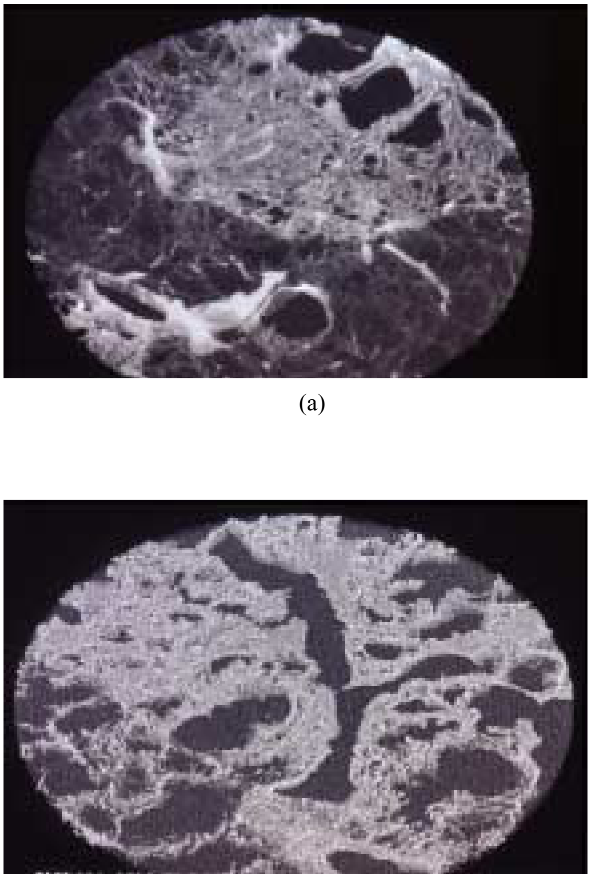
(a)

(b)

Fig.6. Micro CT images of the abnormal lung tissue (pulmonary adenocarcinoma in emphysematous lung) (a) Resolution : 24 μm . (b) Resolution : 6 μm .

IV. CONCLUSION

We have presented a micro CT for imaging pulmonary nodules. The spatial resolution of 5 μm allows physicians to


(a)

(b)

Fig.7. 3-D display of the stacked Micro CT images of the abnormal lung tissue (pulmonary adenocarcinoma in emphysematous lung) Resolution : 24 μm. (b) Resolution : 6 μm.

analyze the micro architecture of lung tissues, such as alveolar wall, septal wall of pulmonary lobule, and bronchiole. This CT allows for the first time to document changes of tissue components between normal and abnormal area. Future work we will have to show if structure indices are sufficient to predict the likelihood of malignancy of the pulmonary nodule. The high spatial resolution micro CT images provide the basic data set and then might help to improve the performance of physician diagnostic decisions.

REFERENCES

[1]M.Kaneko, K.Eguchi, H.Ohmatsu, R.Kakinuma, T.Naruke, K.Suemasu, N. Moriyama, "Peripheral lung cancer: Screening and detection with low-dose spiral CT versus radiography," *Radiology*, 201, pp.798-802, 1996.

[2]K.Kanazawa, Y.Kawata, N. Niki, H. Satoh, H. Ohmatsu, R. Kakinuma, M. Kaneko, N. Moriyama, K. Eguchi, "Computer-aided diagnosis for pulmonary nodules based on helical CT images," *Computerized Medical Imaging and Graphics*, 22, pp.157-167, 1998.

[3]K.Mori, Y.Saitou, K.Tominaga, K.Yokoi, N.Miyazawa, A.Okuyama, M.Sasagawa, "Small nodular legions in the lung periphery: New approach to diagnosis with CT," *Radiology*, 177, pp.843-849, 1990.

[4]S.S.Siegelman, E.A.Zerhouni, F.P.Leo, N.F. Khouri, F.P. Stitik, "CT of the solitary pulmonary nodule," *AJR*, 135, pp.1-13, 1980.

[5]D. Cavouras, P. Prassopoulos and N. Pantelidis, "Image analysis methods for solitary pulmonary nodule characterization by computed tomography," *European Journal of Radiology*, 14, pp.169-172, 1992.

[6]M.F. McNitt-Gray, E.M.Hart, J. Goldin, C.-W, Yao, and D.R. Aberle, " A pattern classification approach to characterizing solitary pulmonary nodules imaged on high resolution computed tomography," *Proc. SPIE*, 2710, pp.1024-1034, 1996.

[7]Y.Kawata, N.Niki, H.Ohmatsu, R.Kakinuma, K.Eguchi, M.Kaneko,N.Moriyama, "Shape analysis of pulmonary nodules based on thin-section CT images," *Proc. SPIE*, 3034, pp.967-974, 1997.

[8]Y.Kawata, N.Niki, H.Ohmatsu, R.Kakinuma, K.Eguchi, M.Kaneko, N.Moriyama, "Quantitative surface characterization of pulmonary nodules based on thin-section CT images," *IEEE Trans. Nuclear Science*, 45, pp.2132-2138, 1998.

[9]P. Ruegsegger, A. Kohlbrenner, D. Ulrich, A. Laib, "First results of a multiple fan-beam approach to analyze bone structure in vivo," Proc. SPIE, 3772, pp.55-62, 1999.

[10]R. H. Johnson, R.C. Molthen, K.L. Karau, C.C. Hanger, S.T. Haworth, C.A. Dawson, and J.H. Linehan, " Analysis of 3D pulmonary micorangiograms," *Computer-Aided Diagnosis in Medical Images*, K. Doi, H. MacMahon, M.L. Giger and K.R. Hoffmann, eds., Elsevier Science, pp.349-376, 1999.

[11]E.L. Ritman, S.M. Jorgensen, P.E. Lund, P.J. Thomas, J.H. Dunsmuir, J.C. Romero, R.T. Turner, and M.E. Bolander, "Synchrotron-based micro-CT of in situ biological basic functional units and their integration," *Proc. SPIE*, 3149, pp.13-24, 1997.

[12]L.A. Feldkamp, L.C. Davis and J.W. Kress, "Practical cone-beam algorithm," *Journal of the Optical Society of America*., 1, pp.612-619, 1984.

[13]Y. Kawata, N. Niki, H. Satoh, and T. Kumazaki, "Three-dimensional blood vessels reconstruction using a high-speed X-ray rotational projection system," *Trans. IEICE*, J76-D-II, pp.2133-2142, 1993.

[14]B.D. Smith, "Image reconstruction from cone-beam projections: Necessary and sufficient conditions and reconstruction methods," *IEEE Trans. Medical Imaging*, MI-4, pp.14-25, 1985.

Theoretical framework for a dynamic cone-beam reconstruction algorithm based on a dynamic particle model

Pierre Grangeat, Anne Koenig, Thomas Rodet, Stéphane Bonnet

Abstract—**Dynamic cone-beam reconstruction algorithms are required to reconstruct 3D image sequences on dynamic 3D CT combining multi-row 2D detectors and ultra-fast rotating gantry. In order to compensate for time evolution and motion artifacts, we propose to use a dynamic particle model to describe the object evolution. One main interest is to process data acquisition on several half-turns in order to reduce the dose delivered per rotation with the same signal to noise ratio. We describe the dynamic particle model and its approximations, the dynamic cone-beam CT acquisition model and the dynamic cone-beam reconstruction algorithm based on a cone-beam to fan-parallel beam rebinning approach.**

Index Terms—**Dynamic tomographic imaging, fully four-dimensional image reconstruction, CT Fluoroscopy, cone-beam, particle model, motion compensation, time evolution.**

I. INTRODUCTION

THE purpose of dynamic computed tomography (CT) imaging is to reconstruct tomographic image sequences of dynamic organs in order to take into account the dynamic nature of a living human body. The description of dynamic organs includes both time evolution and motion. In this publication, we mainly focus on dynamic 3D Computed Tomography combining multi-row 2D detectors and ultra-fast rotating gantry. The main applications are 3D CT Fluoroscopy for interventionnal radiology, to help the radiologist to guide biopsy needles through soft tissues like the lung, radiotherapy planning to better delineate the tumour and healthy tissues during motion, heart diagnostic imaging to study heart kinetic or to reconstruct coronary arteries.

LETI is involved in the European project DynCT (IST - 1999 - 10515) dedicated to both real time and off-line motion compensated reconstruction and visualisation for dynamic computed tomography. Only the real time case is described here. In this presentation, we introduce the theoretical framework of a new dynamic cone-beam reconstruction algorithm based on a dynamic particle model. Standard approach tends to use a short scan acquisition over one half-turn without motion compensation [Taguchi, 2000]. But it implies to increase the dose delivered per rotation to preserve the signal to noise ratio. Using the dynamic particle model to compensate for dynamic evolution, it becomes possible to increase the acquisition time window over several half-turns, up to 4 in our case, in order to reduce the dose delivered per rotation to the patient with the same signal to noise ratio.

In [Hsieh, 1997] the impact of various reconstruction algorithms on 2D CT Fluoroscopy is investigated and inherent limitations of the CT Fluoroscopy in terms of time lag and delay is demonstrated. In [Taguchi, 1998], the authors propose to use a feathering technique to suppress an image artifact which rotates like a radar search line and sometimes hinders accurate observations. In [Ritchie, 1996], the reduction of in plane motion artifacts is achieved using a pixel-specific back projection technique in 2D. In [Schäffter, 1999] a motion compensated projection reconstruction algorithm is proposed for the reduction of blurring artifacts in MRI using motion estimation applied on a first set of low resolution reconstructed images. The new algorithm proposed here uses also this double reconstruction principle. It allows to compensate for both time evolution and motion.

II. THE DYNAMIC PARTICLE MODEL

We consider the image function f as the map of the physical property we want to study, in the present case the X-ray linear attenuation coefficient or the density. This function f is defined in a (O, x, y z) cartesian coordinate system. We represent by f (M, t) = f (x, y, z, t) the value of the function f at the point M of coordinate (x, y, z), at the time instant t. We suppose here that the function f is sufficiently smooth and vanishes outside a ball Ω of radius R_0.

In order to define the particle model, we describe at the reference instant t_0 the object as a continuum of particles associated with each point M of the support Ω. We define the trajectory $\Gamma(M, t)$ associated with this point M as the set of positions this point M will take at each time value t within Ω. By definition, M corresponds to the initial position along the trajectory $\Gamma(M, t)$:

$$\Gamma(M, t_0) = M \qquad (1)$$

In the following, the image sequence will be reconstructed at discrete time samples t_i.

In the more general case where both time evolution and motion have to be taken into account, the function f may vary along the trajectory Γ. Thus, the general function expression associated with the dynamic particle model is given by the

---

The authors are with the Laboratoire d'Electronique et de Technologie de l'Information (LETI), Département Systèmes pour l'Information et la Santé (DSIS), Direction des Recherches technologiques (DRT), Commissariat à l'Energie Atomique (CEA), 38054 Grenoble cedex 9, France.
Pierre Grangeat telephone is : +33 (0)4 76 88 43 73
E-mail : firstname.lastname @cea.fr

.

formula $f(\Gamma(M,t),t)$, where $\Gamma(M,t)\in\Omega$ and $t\in R$. This corresponds to a continuous expression of the discrete particle models used in computer graphics to describe deformable objects [Lombardo, 1996].

One example is a flying ball of constant radius which content might vary along time due to contrast product injection or to matter exchange with the neighbouring flying balls. In the special case of matter conservation and non-compressible material, it would be equivalent to assume the content remains constant along time.

The time evolution compensation will be based on the first order approximation of the expression $f(\Gamma(M,t),t)$, in the neighborhood of the discrete time sample t_i:

$$f(\Gamma(M,t),t)\approx f(\Gamma(M,t_i),t_i)+[<\nabla f, \frac{\partial \Gamma}{\partial t}(M,t_i)>+\frac{\partial f}{\partial t}(\Gamma(M,t_i),t_i](t-t_i) \qquad (2)$$

where ∇f is the gradient of f and < , > the scalar product.

When the object fulfills the mass conservation principle and is irreducible, the term in bracket [] is null and we get:

$$f(\Gamma(M,t),t)=f(\Gamma(M,t_i),t_i) \qquad (3)$$

However, in the reality, some important tissues like the lung are reducible and since the organs may move outside the field of view, the mass conservation principle is not always fulfilled. Thus we need to consider the general case.

In order to later simplify the computation, we introduce a cartoon like step-by-step motion law :

$$\Gamma(M,t) = \Gamma(M,t_i) \text{ for } t_i\leq t<t_{i+1} \qquad (4)$$

Then, since $\frac{\partial \Gamma}{\partial t}$ is null, the first order approximation (2) of the dynamic model becomes:

$$f(\Gamma(M,t),t)=f(\Gamma(M,t_i),t_i)+\frac{\partial f}{\partial t}(\Gamma(M,t_i),t_i)(t-t_i) \qquad (5)$$

In this case, this formula (6) can be approximated by a linear prediction law with respect to t between two time samples, for instance t_i and $t_i - T$, where T is the rotation period of the continuously rotating scanner:

$$f(\Gamma(M,t),t)\approx f(\Gamma(M,t_i),t_i)+\left[\frac{f(\Gamma(M,t_i),t_i)-f(\Gamma(M,t_i),t_i-T)}{T}\right]\cdot(t-t_i) \qquad (6)$$

III. THE DYNAMIC CONE-BEAM CT ACQUISITION MODEL

1. Dynamic cone-beam projection

We consider the cone-beam geometry associated with a curved multi-row detector centred on the X-ray source F as described on the figure 1.

The dynamic cone-beam geometry is parametrized by the angle β between (F,O) and y axis, the angle γ between the detector column and the (F,O) axis, the row height q_d with respect to the trajectory plane, and the acquisition time t (see fig. 1). It is defined as follows :

$$X_c f(\beta,\gamma,q_d,t)=\int_{M\in D(\beta,\gamma,q_d)} f(M,t)dM \qquad (7)$$

where $D(\beta,\gamma,q_d)$ is the straight line between the source point F and the detector cell A_d. In order to model a continuously rotating acquisition process, we state the following relation :

$$\beta(t) = \beta_0 + \omega.t \qquad (8)$$

where ω defines the angular rotation speed. For a CT scanner with a 0.5 s rotation period, $\omega = 4\pi$ rad.s^{-1}.

In the following, we assume the angle γ belongs to $[-\gamma_m,+\gamma_m]$ where γ_m is the half fan angle. In our case we choose $\gamma_m=\frac{\pi}{6}$rad.

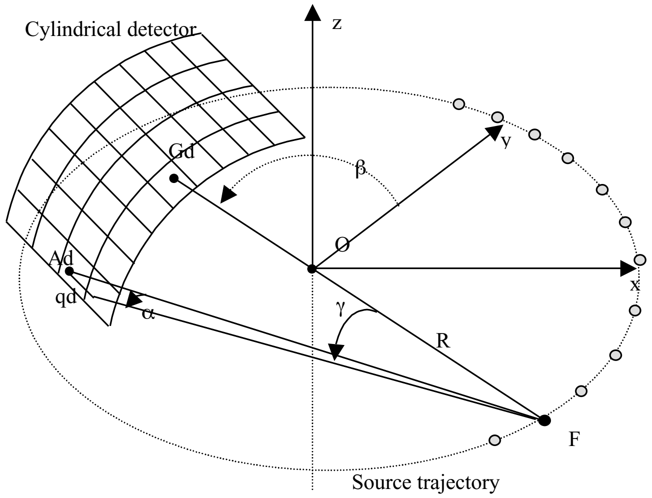


Figure 1 : cone-beam geometry.

2. Dynamic fan-parallel projection

The reconstruction will be done via the line rebinning approach in fan-parallel projection as suggested in [Grass, 2000].

The fan-parallel geometry is defined by a virtual planar detector placed on the rotation axis with coordinates p parallel to the trajectory and q parallel to the rotation axis. We define φ as the angle between the virtual detector and the (0,x) axis.

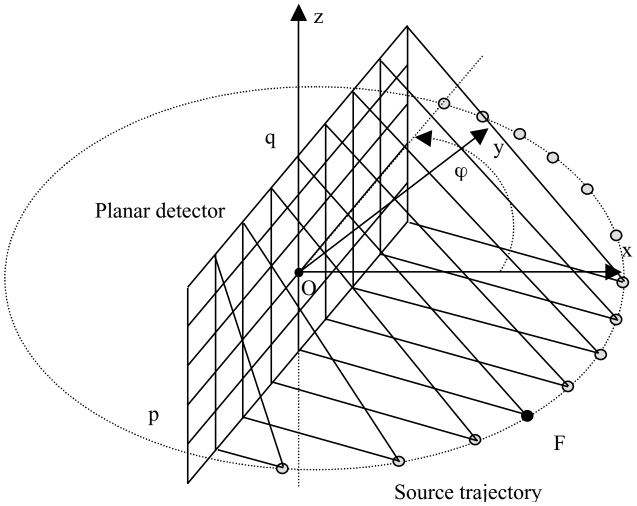


Figure 2 : fan-parallel-beam geometry.

The dynamic fan-parallel projection is defined as follows :

$$X_{fp} f(\varphi,p,q,t)=\int_{M\in D(\varphi,p,q)} f(M,t)dM \qquad (9)$$

In the ideal case, the full fan-parallel projection data set is available for each t. Then, the short scan dynamic fan-parallel reconstruction formula is:

$$f(\Gamma(M,t),t) = \int_0^{\pi} HDYf\,(\varphi, A\,[\Gamma\,(M, t)], t)\; d\varphi \qquad (10)$$

where Yf is the weighted fan-parallel projection:

$$Yf(\varphi,A,t) = X_{fp},f(\varphi,A,t).w(A) \qquad (11)$$

and $w(A) = \dfrac{\sqrt{R^2 - p^2}}{\sqrt{R^2 - p^2 + q^2}}$

A [Γ (M, t)] is the fan-parallel projection of the point Γ(M,t) onto the virtual detector, with detector coordinates p,q.

HDYf is the weighted projection Yf convoluted along the transverse row by the ramp filter HD(p).

IV. THE DYNAMIC CONE-BEAM RECONSTRUCTION ALGORITHM

The reconstruction algorithm described here is dedicated to 3D CT fluoroscopy, assuming a real-time reconstruction processing at a frame rate of 12 frames per second for a gantry rotation period T of 0.5 s.

1. The sliding window principle

Let us consider the discrete fan-parallel angles :

$$\varphi_i = i\ \Delta\varphi \qquad (12)$$

where Δφ is the angular step between two reconstructed frames. In the following, we assume $\Delta\varphi = 2.\,\gamma_m$. It means we divide the full rotation in constant angular positions separated from the full fan angle $2.\,\gamma_m$ equal to $\frac{\pi}{3}$. For a 0.5 s gantry rotation period, this corresponds to a frame rate of 12 images per second.

The associated discrete reconstruction time t_i associated with the last projection of the angular range $[\varphi_i, \varphi_i+\Delta\varphi]$ is :

$$t_i = \frac{\left(\varphi_i + \Delta\varphi\right) + \gamma_m - \beta_0}{\omega} \qquad (13)$$

where $\dfrac{\gamma_m}{\omega}$ corresponds to the rebinning latency delay to begin the fan-parallel reconstruction.

We define the overscan φ angular range to compute $f(\Gamma(M,t_i),t_i)$ as the φ sliding window $[\varphi_i- n.\pi, \varphi_i+\Delta\varphi]$ where n represents the number of half-turns on which we want to smooth the data to improve the signal to noise ratio or to reduce the dose.

The associated β angular range defined by the following rebinning equation (14) is the β sliding window $[\varphi_i-(n.\pi+\frac{\Delta\beta}{2}),\varphi_i+\frac{3}{2}\,\Delta\beta]$, where $\Delta\beta=\Delta\varphi=2.\gamma_m$.

The sliding window principle is that a new frame is computed for each new t_i value. This new frame corresponds to a shift of the β sliding window from Δβ, and the associated shift of the φ sliding window from Δφ.

2. The cone-beam to fan-parall-beam rebinning

We get the following rebinning equation :

$$X_{fp}f(\varphi, p, q, t) = X_c f(\beta, \gamma, q_d, t) \qquad (14)$$

with :

$$\begin{cases} \beta = \varphi - \gamma \\ \gamma = \arcsin\left(-\dfrac{p}{R}\right) \\ q_d = q\left(\dfrac{FGd}{\sqrt{R^2 - p^2}}\right) \end{cases} \qquad (15)$$

However, we need to compute $X_{fp}f(\varphi, p, q\ t_i)$ at t_i, the reconstruction time associated with the time range $[t_{i-1}, t_i]$.

When the time difference (t_i-t) is small, we can use a nearest neighbour interpolation:

$$X_{fp}f(\varphi,p,q,t_i) = X_{fp}f(\varphi,p,q,t) \qquad (16)$$

Otherwise, if the sliding window is larger than 2π, using the linear interpolation model associated with equation (6) we get the following extrapolation formula :

$$X_{fp}f(\varphi,p,q,t_i)=\frac{t_i - \left(t - T\right)}{T}\,X_c f(\beta,\gamma,q_d,t) + \frac{t_i - t}{T}\,X_c f(\beta,\gamma,q_d,t-T) \qquad (17)$$

where T is the gantry rotation period.

3. Block reconstruction

We split the φ angular range into elementary projection blocks of size $\Delta\varphi = \frac{\pi}{3}$. We denote $BHDYf(\Gamma(M,t),t,\varphi_i)$ the partial block backprojection over the projection angular $[\varphi_i, \varphi_i + \Delta\varphi]$:

$$BHDYf(\Gamma(M, t),t,\varphi_i) = \int_{\varphi = \varphi_i}^{\varphi_i + \Delta\varphi} HDYf\left(\varphi, A\left[\Gamma\left(M, t\right)\right] t\right) d\,\varphi \qquad (18)$$

For seak of simplicity, we assume here we want to reconstruct the function at the instant $t = t_0$. Thus $\Gamma(M,t) = M$. The following result can then be generalized for each t_i time by shifting the sliding window.

From the reconstruction formula (10), we get:

$$f(M,t_0) = \int_0^{\pi} HDY\,f\left(\varphi, A\,[M], t_0\right) d\varphi \qquad (19)$$

Since $\Delta\varphi = \frac{\pi}{3}$, we can decompose this integral into 3 terms associated with 3 partial block backprojections sectors $[0, \frac{\pi}{3}]$, $[\frac{\pi}{3}, \frac{2\pi}{3}]$, $[\frac{2\pi}{3},\pi]$:

$$f(M,t_0) = \sum_{i=0}^{2} BHDY\,f\,(M,t_0, \varphi_i) \qquad (20)$$

The dynamic evolution compensation will take place in the estimation of each partial block backprojection as described in the next section.

4. Dynamic evolution compensation

Given an angular sector i, let us take a set of N_b partial block projection acquired in the past for the same angular range modulo π :

$$\varphi_{ij} = \varphi_i - j.\pi \qquad j = 0,\ldots, N_b - 1$$

We get for each M point a set of values along the $\Gamma(M,t)$ trajectory for each associated block instant $BHDYf(\Gamma(M,t_{ij}),t_{ij}\ \varphi_{ij})$, where t_{ij} is the time associated with φ_{ij} according to equation (13). Using the first order approximation (2) of the dynamic particle model, we get :

$$f(\Gamma(M,t_{ij}),t_{ij}) \approx f(M,t_0)+[<\nabla f, \frac{\partial \Gamma}{\partial t}(M,t_0)>+ \frac{\partial f}{\partial t}(M,t_0)](t_{ij}-t_0) \tag{21}$$

Under the piecewise constant motion hypothesis (4), as no motion occurs during the block angular range, the same relation holds for the partial backprojection:

$$BHDYf(\Gamma(M,t_{ij}),t_{ij},\varphi_{ij}) \approx BHDYf(M,t_0,\varphi_{i0}) + a(M,\varphi_{i0}).(t_{ij}-t_0) \tag{22}$$

Thus, the terms $BHDYf(M,t_0,\varphi_{i0})$ and $a(M,\varphi_{i0})$ can be computed by linear regression on the discrete sample set:

$$\left\{BHDYf\left(\Gamma\left(M, t_{ij}\right), t_{ij}, \varphi_{ij}\right)\right\} \ j\in\left\{0,\ldots, Nb\text{-}1\right\}.$$

In order to apply such a linear regression, we need to have at least two samples belonging to the sliding window. Otherwise using the same relation, we can only compensate for motion, and not for time evolution. The motion compensation equation is given by:

$$BHDYf(\Gamma(M,t_{ij}),t_{ij},\varphi_{ij}) = BHDYf(M,t_0,\varphi_{i0}) \tag{23}$$

This is equivalent to choose a zero order regression model. In every case, it is important to note that using an overscan range within the sliding window, the dynamic evolution compensation by regression introduces a principle equivalent to the feathering technique used by different authors in CT fluoroscopy [Taguchi, 1998].

5. *Motion estimation*

In the previous sections, we have assumed we can measure the motion field between M and Γ (M, t_{ij}). However, as this motion field is unknown, we need to estimate it. We present here only the basic idea. Detailed explanation will be given in the final version.

The first order approximation (2) can be interpreted as the sum of a spatial shift term and a time evolution term. In first approximation, we will assume here that the shift can be approximated by a translation motion at constant speed:

$$\Gamma(M,t) = M + D(M).\frac{2\left(t - t_0\right)}{T} \tag{24}$$

where D(M) is the displacement of the point M after a π rotation of the gantry and T is the gantry rotation period.

The reconstruction process without motion compensation will produce a blurred image. However, the reconstruction after a π rotation will produce the same blurred image, shifted from the displacement vector D(M). Thus, using a correlation principle, it will be possible to estimate this displacement vector. Such a correlation approach should not be too much disturbed by time evolution.

In order to reduce the blurring spread, the image sequence needed to evaluate the motion should be reconstructed with the smallest temporal resolution corresponding to a half-turn rotation.

In fact, we need to get the full motion field over the region to reconstruct. The actual technique we use is a block matching approach as for motion estimation in MPEG coding [Sikora, 1997].

One other important issue is the ability to detect when the particle trajectory goes outside the region of interest. In the approach described here, we can detect it since the correlation fails. The prediction should be applied only to those points which are inside the region of interest. To manage this issue, we estimate for each motion vector a confidence factor associated with the correlation factor, and we take it into account within the estimation of the prediction rule coefficients given by the equation (22).

V. CONCLUSION

In this paper, we present the outline of the theoretical framework for time evolution and motion compensation in dynamic cone-beam reconstruction algorithm using a dynamic particle model. The principle described here for a line rebinning algorithm can also be applied to direct cone-beam reconstruction algorithm such as Feldkamp algorithm.

However, according to [Grass, 2000], this line rebinning approach can be generalized to reconstruct a larger region of interest by correctly handling the line shadow zone region. The approach described here can also be extended to such a case. Further investigation are needed to extend this dynamic approach to indirect plane rebinning algorithm via the first derivative of the Radon transform using Grangeat formula.

Preliminary results will be presented at the conference.

ACKNOWLEDGMENT

This work is supported by the European Commission under grant IST - 1999 - 10515, associated with the DynCT project.

The authors would like to thank the DynCT partners and specially Roland Proksa and Michael Grass from Philips Research Laboratory in Hamburg (PRL) for their collaboration.

REFERENCES

[Grass, 2000] Grass M., Köhler Th., Proksa R., "3D cone-beam CT reconstruction for circular trajectories", *Phys. Med. Biol.*, 45:329-347, 2000.

[Hsieh, 1997] Hsieh J.,"Analysis of the temporal response of computed tomography fluoroscopy", *Med. Phys.*, 24(5) :665-675, 1997.

[Lombardo, 1996] Lombardo J.-C., "Modélisation d'objets déformables avec un système de particules orientées", *Ph. D Thesis*, Université Joseph Fourier, 1996.

[Ritchie, 1996] Ritchie C. J., Crawford CR et al, "Correction of computed tomography motion artifacts using pixel-specific back-projection", *IEEE Trans. Med. Ima.,* 15(3):333-342, 1996.

[Schäffter, 1999] Schäffter T., Rasche V., Carlsen IG, "Motion compensated projection reconstruction", *Mag. Res. in Med*, 41:954-963, 1999.

[Sikora, 1997] Sikhora T., "MPEG digital video-coding standards", IEEE Signal Proc. Mag., 14(5):82-100, 1997.

[Taguchi, 1998] Taguchi K., Otawara M. S., "Improvement of CT Fluoroscopy image quality by feathering the data to cover the lag between acquisition times", *Radiology*, November 98:435, 1998.

[Taguchi, 2000] Taguchi K., Anno H., "High temporal resolution for multislice helical computed tomography", Med. Phys., 27(5):861-872, 2000.

Image to Volume Weighting Generalized ASSR for Arbitrary Pitch 3D and Phase–Correlated 4D Spiral Cone–Beam CT Reconstruction

Marc Kachelrieß*, Theo Fuchs, Robert Lapp, Dirk–Alexander Sennst, Stefan Schaller, Willi Kalender

Abstract—**The next generation of medical CT scanners will measure up to 16 slices or more simultaneously which will require dedicated cone–beam reconstruction algorithms. The basic requirements for medical CT are high image quality and fast reconstruction (reconstruction of a complete volume within a few minutes is desired). Due to the small cone–angle (typically a few degrees only) approximate cone–beam reconstruction will be the method of choice. A very promising candidate is the advanced single–slice rebinning (ASSR) which rebins the cone–beam data to parallel beam data on tilted reconstruction planes (R–planes) and utilizes 2D reconstruction algorithms to obtain tilted images [Med. Phys. 27(4), 754–772 (2000)].**

In its original form ASSR allows to reconstruct 3D data for a fixed pitch obtained with a non–tilted gantry. However, medical demands are manifold: the pitch must be freely selectable, gantry tilt scans are required and, last but not least, cardiac applications require phase–correlated 4D reconstructions. We have therefore generalized the ASSR algorithm by adding the following three attributes: a) the table increment per rotation is now correctly taken into account as the vector $\boldsymbol{d}$, b) the restriction on the optimality of the R–planes is loosened to allow for more than one R–plane per reconstruction position and c) the final volume is generated using adjustable weigths for the tilted images. These weights are used to balance between image quality and dose usage and to select a desired cardiac phase in the final volume. This (tilted) image to (cartesian) volume (I2V) weighting approach can be performed in real–time. To evaluate the new method we have simulated cone–beam rawdata of a thorax and a cardiac motion phantom.

The generalized ASSR approach in combination with I2V shows very good results even for low pitch ($p < 1.5$) scans. Since slightly more artifacts appear for the low pitch case with full dose usage (equal weights for all planes) it is necessary to provide real–time access to the weights. The cardiac reconstructions are of high image quality but slightly lower temporal resolution as compared to the gold standard 180° MCI.

Keywords—**Computed tomography (CT), Cone–beam Spiral CT (CBCT), 3D reconstruction, 4D reconstruction**

I. INTRODUCTION

FUTURE medical CT scanners will scan significantly more than four slices simultaneously. Major CT manufacturers have announced scanners of up to 16 slices for the end of 2001. Neglecting the cone–angle of the scanner as it is done in today's 4–slice reconstruction algorithms will then yield unacceptable image artifacts [1]. Therefore, there is a need for fast and efficient cone–beam reconstruction algorithms. Although a large number of more or less efficient algorithms have been developed in the last decade [2], none of them meets all requirements of medical CT: a) correct handling of the gantry tilt, b) arbitrary spiral pitch while using the full detector area and c) the ability of performing a 4D cardiac reconstruction. Whereas the gantry tilt problem can be solved in principle for all the existing algorithms by simply reformulating the coordinate transformations solving the other restrictions is not straightforward.

We therefore generalize the ASSR (advanced single–slice rebinning) algorithm [3] to fulfill these requirements. The ASSR algorithm, as an advancement of Noo's single–slice rebinning algorithm [4], fits tilted reconstruction planes to the spiral trajectory to perform a rebinning from the 3D cone–beam data to 2D parallel–beam data on these R–planes. The reconstruction then uses a standard 2D method (e.g. filtered backprojection) and the set of reconstructed tilted images is interpolated in the z–direction to obtain the final volume [3] (a similar method which has never been evaluated was proposed in [5]). Since ASSR has turned out to be very promising [6], [7] two generalizations thereof have been proposed: one for the case of arbitrary gantry tilt [8] and one to allow for arbitrary pitch [9]. The results are encouraging and led to the development of an even further generalization to combine these approaches. The new approach presented here uses the rebinning equations of reference [8] together with the idea to losen the restriction of the R–planes to allow for more than one R–plane per reconstruction position [9]. A novel idea of our approach is to weight each available image before it is interpolated into the final volume. The weights can be chosen to trade off between high image quality and high noise (i.e. assigning smaller weights to R–planes which are less optimal) or lower image quality and lower noise (i.e. assigning the same weight for each plane regardless of its optimality).

As a spin–off, the generalized ASSR combined with I2V allows to perform 4D reconstructions for low pitch scans with periodically moving objects. The weights of images corresponding to cardiac phases which should not appear in the volume are simply set to zero. Details of cardiac CT scanning and the restrictions on the maximum pitch as a function of the patient's heart rate can be found in references [10], [11], [12].

In this paper, we will outline the generalized ASSR algorithm and the weight selection and give some descriptive examples.

Institute of Medical Physics, University of Erlangen–Nürnberg, Krankenhausstr. 12, 91054 Erlangen. Corresponding author: Marc Kachelrieß, E–mail: marc.kachelriess@imp.uni–erlangen.de

II. SIMULATIONS

To evaluate our new approach we have simulated spiral cone–beam data corresponding to the in–plane geometry of a typical medical CT Scanner (1160 projections per rotation, 672 detector channels per detector row, and a fan angle $\Phi = 52^\circ$) using a dedicated x–ray simulation tool (ImpactSim, VAMP GmbH, Möhrendorf, Germany). Two phantoms have been simulated: the thorax phantom described in the phantom data base http://www.imp.uni-erlangen.de/forbild and the cardiac motion phantom described in [11].

For the thorax scan in standard mode we have chosen a collimation of 16×1 mm and we have performed simulations for 4 mm, 8 mm, 16 mm and 24 mm table increment per rotation.

Since shorter rotation times and thinner slices are expected for the future we have simulated the cardiac motion phantom for a wide range of heart rates f_H with 0.375 s rotation time (160 rpm), 12×0.5 mm collimation and a table increment of $d = 2$ mm. This allows to cover the heart (typically 12 cm to 15 cm axial length) within less than 30 s (single breath–hold). Evaluating the thorax phantom in cardiac mode yields too huge data sets due to the highly overlapping and fine sampling in z. Thus we scaled the thorax scan by a factor of 2 in the z–direction, i.e. 12×1 mm and $d = 4$ mm (even so, the rawdata file size is 2.5 Gigs).

III. RECONSTRUCTION

The reconstructions shall be centered about the views $n\Delta\alpha_\mathrm{R}$ where $\Delta\alpha_\mathrm{R}$ is the so–called reconstruction increment and should be chosen small enough to ensure full detector usage and resolution [3]. The integer n counts the reconstruction positions. For each reconstruction position α_R an optimal R–plane which minimizes the mean square deviation Δ_mean of the plane to the spiral trajectory within the interval $\alpha_\mathrm{R} - \pi/2$ to $\alpha_\mathrm{R} + \pi/2$ can be computed [8]. Using these planes only, reconstruction for low pitch would be possible, but parts of the detector would remain unused and thus would have to be collimated out. Since we want to be able to use the full detector, our new approach allows to reconstruct more than one tilted image from a given reconstruction position. Assuming M images per reconstruction position the set of all R–planes is given (in normal representation) as

$$\mathrm{R}_{nm} : \boldsymbol{n}_{nm} \cdot \boldsymbol{r} - a_{nm} = 0$$

with $m = 1, \ldots, M$. The a_{nm} are chosen equidistant in m as $a_{nm} = a_n + m\Delta a_M$. The increments $\Delta\alpha_\mathrm{R}$ and Δa_M together with the value of M are chosen to ensure full detector usage. The normal vectors $\boldsymbol{n}_{nm}$ are chosen for a given a_{nm} to minimize the mean square deviation

$$\Delta_{nm}^2 = \frac{1}{\pi} \int\limits_{\alpha_\mathrm{R} - \frac{1}{2}\pi}^{\alpha_\mathrm{R} + \frac{1}{2}\pi} d\alpha \left(\boldsymbol{n}_{nm} \cdot \boldsymbol{s}(\alpha) - a_{nm}\right)^2$$

with $\boldsymbol{s}(\alpha)$ being the spiral source trajectory and $\alpha_\mathrm{R} = n\Delta\alpha_\mathrm{R}$. The minimization procedure is described in [8].

IV. IMAGE WEIGHTS

We make use of two possibilities of weighting the individual images prior to volume interpolation. For a standard reconstruction, the image weights are chosen as

$$w_{nm} = \left(\frac{1}{\Delta_{nm}}\right)^q \quad \text{with} \quad q \geq 0 \; .$$

The quality parameter q is used to balance between best dose usage ($q = 0$, making full use of non–optimal R–planes) and best image quality ($q = \infty$, only using the optimal R–planes and neglecting non–optimal ones).

For cardiac 4D reconstruction we additionally use the cardiac phase $c(\alpha) \in [0, 1)$, which is a function of the view angle α and describes the cardiac motion relative to R-R, to weight the images. The user desires to reconstruct the images at the reconstruction phase c_R. The mean square deviation of the cardiac phases contributing to reconstruction position $\alpha_\mathrm{R} = n\Delta\alpha_\mathrm{R}$ from the target phase c_R is defined as

$$\Delta_n^2 = \frac{1}{\pi} \int\limits_{\alpha_\mathrm{R} - \frac{1}{2}\pi}^{\alpha_\mathrm{R} + \frac{1}{2}\pi} d\alpha \left(c(\alpha) - c_\mathrm{R}\right)^2 .$$

Here, we implicitly assume a proper handling of the modulo property of the cardiac phase. We define the cardiac weight as

$$w_n = \left(\frac{1}{\Delta_n}\right)^{q'} \quad \text{with} \quad q' \geq 0$$

and again we have a quality parameter available to adjust the image quality.

V. IMAGE TO VOLUME WEIGHTING

The final step is to perform an interpolation from a set $\{f_{nm}(x, y, z)\}$ of tilted images (each with weight w_{nm}) to a cartesian volume $f(x, y, z)$. In general, this can be achieved by convolving the reconstructed planes with a three dimensional interpolation kernel $k(x, y, z)$ followed by proper normalization:

$$f(x, y, z) = \frac{\sum\limits_{nm} w_{nm} f_{nm}(x, y, z) * k(x, y, z)}{\sum\limits_{nm} w_{nm} 1_{nm}(x, y, z) * k(x, y, z)} \; . \tag{1}$$

The indicator function 1_{nm} is 1 if $(x, y, z) \in \mathrm{R}_{nm}$ and 0 elsewhere. Since ASSR does not require interpolations between image pixels in the x and y direction, the interpolation kernel reduces to a function of z only: $k(z)$. Its shape is currently chosen triangular and care is taken that the z–kernel is wide enough to avoid gaps in the final volume: the denominator of (1) must be positive $\forall\, x, y, z$. Using local kernels $k(x, y, z, x_0, y_0, z_0)$ and spatially varying weights $w_{nm}(x, y, z)$ may be of advantage but is not discussed here because this is beyond this short paper's scope. More sophisticated I2V methods such as modifying the kernel and the weights as a function of the cardiac information are under further investigation.

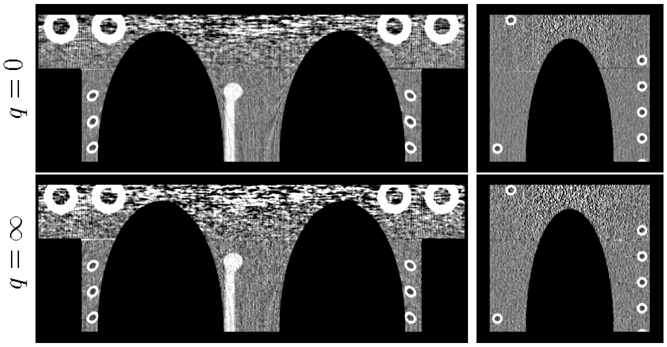


Fig. 1. ASSR, 16 x 1 mm collimation, 8 mm table increment. Coronal and sagittal MPRs of the thorax phantom. I2V using all planes ($q = 0$, top) vs. I2V using optimal planes only ($q = \infty$, bottom). Noise is increased by a factor of 1.7 for the latter case. (0/100)

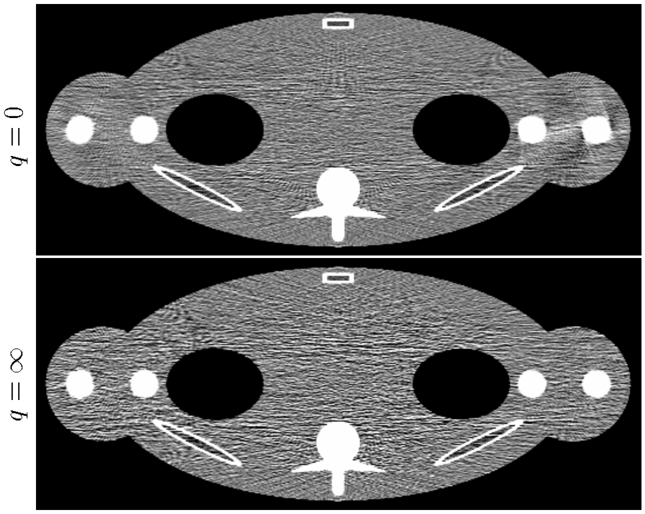


Fig. 2. ASSR, 16 x 1 mm collimation, $d = 8$ mm. The slices show a slight decrease of noise and an increase of artifacts when using all planes with equal weights ($q = 0$, top) in comparison to using only the optimal planes ($q = \infty$, bottom). (0/100)

VI. RESULTS

In general, using non–optimal planes due to overlapping data acquisition shows no significant disadvantage as compared to the maximum pitch ($p = 1.5$) approach (original ASSR) except for the longer reconstruction time. This is demonstrated in figure 1 showing differences only in image noise which is increased by a factor of 1.7 for the reconstruction using only the optimal planes (this reconstruction is equivalent to a $p = 1.5$ scan with fewer slices). The increase in noise can be understood as follows: to achieve a table increment of 8 mm, a collimation of 5 × 1 mm would suffice (pitch 1.5). Thus, only 5/16–th of the detector are used by $q = \infty$ which is roughly $1/1.7^2$.

To demonstrate differences apart from the image noise between the ASSR approach only using the optimal planes ($q = \infty$) and the one making full use of the patient dose ($q = 0$) it is necessary to look at transaxial planes (as we have seen that no significant differences can be observed in MPR displays for the geometry we simulated). Therefore, figure 2 shows a slice in the shoulder region where attention should be paid to the four spheres representing the humerus. Especially one of the spheres is surrounded by artifacts for $q = 0$ whereas they are imaged artifact–free for the optimal $q = \infty$ case. A similar behavior can be observed for the ribs (no images shown). The streaks emerging from these spheres may lead to misdiagnosis: the real patient anatomy contains more complex objects close to the lung. An example may be the heart which is often filled with contrast agent. Making use of I2V's real–time capabilities and evaluating the volume as a function of q helps to resolve such ambiguities.

The reconstructions of our virtual heart phantom (phantom and motion function are defined in [11]) in figure 3 ($f_{\mathrm{H}} = 70\ \mathrm{min}^{-1}$ and $130\ \mathrm{min}^{-1}$) demonstrate the capability of I2V ASSR to resolve motion. We have chosen $q' = \infty$ for these reconstructions, except for the non–cardiac image which has been produced with $q' = 0$ to force equal weights for all planes. The images show optimal quality for low heart rates and less optimal quality for the high heart rate case. This is not surprising, since ASSR is a partial scan algorithm that makes use of 180° data always and its temporal resolution is $t_{\mathrm{rot}}/2$. In that respect, it is similar to the partial scan cardiac algorithm 180°MCD [11].

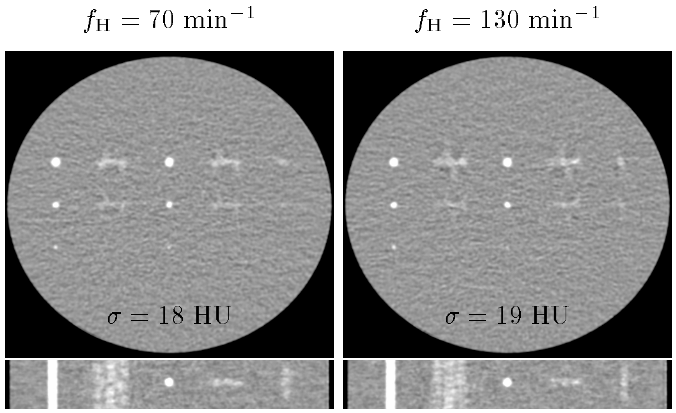
Reconstructions of the same data and the same slices using the multi-slice gold standard 180°MCI (Multi-slice Cardio Interpolation, images not shown here) are slightly superior due to the better temporal resolution of the multi–phase approach 180°MCI. However, 180°MCI is not suited for cone–beam scanners since 180°MCI does not take into account the cone–beam nature of the x–rays. To demonstrate this, we have reconstructed the thorax phantom in the cardio mode ($q' = \infty$) using ASSR and 180°MCI. The results (figure 4) show clearly the advantages of ASSR over 180°MCI. Especially in the humerus severe geometric distortions appear in 180°MCI. Some of the spheres appear to be egg–shaped. Artifacts are also apparent in the ASSR reconstruction. These are due to using only a few reconstruction positions ($q' = \infty$). In general, for ASSR no geometric distortions appear and the artifacts are less severe than for 180°MCI.

VII. DISCUSSION

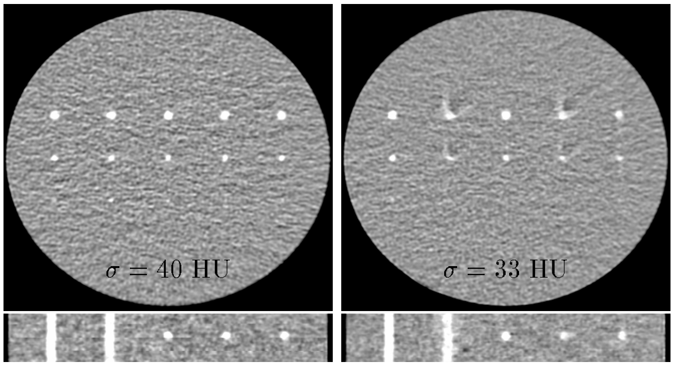
Our results indicate that there is no significant disadvantage of performing overlapping data acquisition and using non–optimal reconstruction planes instead of doing a high–pitch scan with optimal planes only. For transaxial planes, however, it seems to be of interest to be able to change the weight strategy on–line.

Of course, if the tube current is not the restricting factor of the scan, overlapping data acquisition should be avoided and the pitch should be set to 1.5, which is the optimal pitch for ASSR. Low pitch should be used only if one intends to accumulate dose to further decrease the image noise.

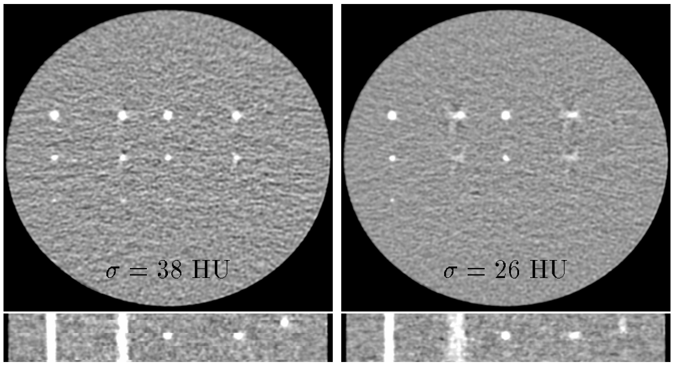
For the new cardiac approach, promising results can be obtained for the cardiac motion phantom within a wide range of heart rates.



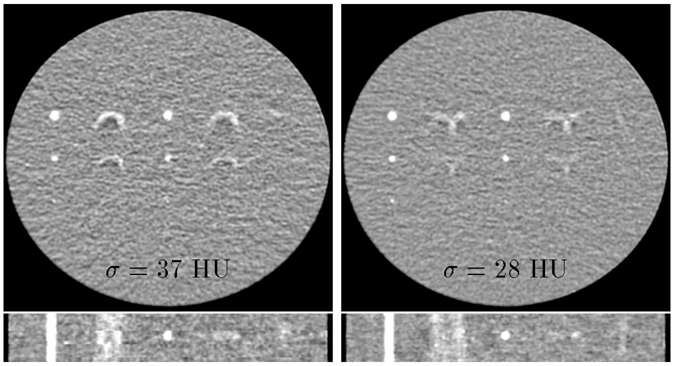
Non–cardiac (zero exponent)



Slow motion phase ($c_{\mathrm{R}} = 0\%$)



Medium motion phase ($c_{\mathrm{R}} = 30\%$)



High motion phase ($c_{\mathrm{R}} = 50\%$)

Fig. 3. ASSR for various weighting strategies and low (left column, $70\ \mathrm{min}^{-1}$) and high (right column, $130\ \mathrm{min}^{-1}$) heart rates. The MPRs extend over 17 mm in the z–direction and show the 3 mm calcifications. Image noise σ is given for a circular ROI centered in the lower heart region. (0/500)

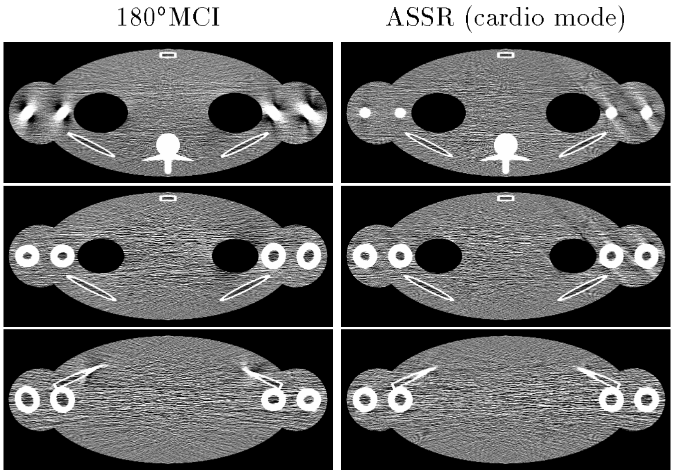


Fig. 4. 180°MCI and ASSR for $f_{\mathrm{H}} = 70\ \mathrm{min}^{-1}$. (0/100)

The conclusions that can be drawn from a simulation study are restricted, however. The algorithms, the weighting strategies and the cardiac reconstruction will undergo further evaluation using real patient data and the expertise of radiologists.

ACKNOWLEDGEMENTS

This work was supported by grant AZ 262/98 of "Bayerische Forschungsstiftung, D–80333 München, Germany".

REFERENCES

[1] W. A. Kalender, *Computed Tomography.* Wiley & Sons, 2000.

[2] H. Turbell, *Cone-Beam Reconstruction using Filtered Backprojection.* PhD Thesis, Linköping, 2001.

[3] M. Kachelrieß, S. Schaller, and W. A. Kalender, "Advanced single–slice rebinning in cone–beam spiral CT," *Med. Phys.*, vol. 27, pp. 754–772, Apr. 2000.

[4] F. Noo, M. Defrise, and R. Clackdoyle, "Single–slice rebinning method for helical cone–beam CT," *Phys. Med. Biol*, vol. 44, pp. 561–570, 1999.

[5] G. Larson, C. Ruth, and C. Crawford, "Nutating slice CT image reconstruction apparatus and method," 1998. United States Patent 5,802,134.

[6] H. Bruder, M. Kachelrieß, S. Schaller, and T. Mertelmeier, "Performance of approximate cone–beam reconstruction in multislice computed tomography," *SPIE Medical Imaging Conference Proc.*, vol. 3979, pp. 541–555, 2000.

[7] H. Bruder, M. Kachelrieß, S. Schaller, K. Stierstorfer, and T. Flohr, "Single–slice rebinning reconstruction in spiral cone–beam computed tomography," *IEEE Transactions on Medical Imaging*, vol. 19, pp. 873–887, Sept. 2000.

[8] M. Kachelrieß, T. Fuchs, S. Schaller, and W. A. Kalender, "Advanced single–slice rebinning for tilted spiral cone–beam CT," *Med. Phys.*, vol. 28, pp. 1033–1041, June 2001.

[9] S. Schaller, K. Stierstorfer, H. Bruder, M. Kachelrieß, and T. Flohr, "Novel approximate approach for high-quality image reconstruction in helical cone beam CT at arbitrary pitch," *SPIE Medical Imaging Conference Proc.*, 2001.

[10] M. Kachelrieß and W. A. Kalender, "Electrocardiogram–correlated image reconstruction from subsecond spiral computed tomography scans of the heart," *Med. Phys.*, vol. 25, pp. 2417–2431, Dec. 1998.

[11] M. Kachelrieß, S. Ulzheimer, and W. A. Kalender, "ECG–correlated image reconstruction from subsecond multi–slice spiral CT scans of the heart," *Med. Phys.*, vol. 27, pp. 1881–1902, Aug. 2000.

[12] M. Kachelrieß, S. Ulzheimer, and W. A. Kalender, "ECG–correlated imaging of the heart with subsecond multislice CT," *IEEE Transactions on Medical Imaging*, vol. 19, pp. 888–901, Sept. 2000.

HeinzelCluster: accelerated reconstruction for FORE and OSEM3D

St. Vollmar[1], C. Michel[2], J.T. Treffert[3], D. Newport[3],
C. Knöss[1], K. Wienhard[1] and W.-D. Heiss[1]
[1]Max-Planck-Institut für neurologische Forschung, Gleueler Str. 50,
50931 Köln, Email: vollmar@pet.mpin-koeln.mpg.de
[2]PET Laboratory, Catholic University of Louvain, Louvain-la-Neuve, Belgium
[3]CTI Inc., Knoxville, TN, USA

I. Introduction

Iterative reconstruction techniques for reconstruction of Positron Emission Tomography (PET) data are usually too time consuming on most single processor machines that are affordable. This is especially true for the HRRT (High Resolution Research Tomograph) which demands sinogram dimension of unsurpassed size (presently one 3D data set consists of 2209 sinograms with 256 radial elements and 288 views), [1].

One strategy to drastically improve reconstruction time is the use of Fourier Rebinning (FORE), [4]: the 3D scan is transformed into the format of a 2D scan with 207 sinograms (in case of the HRRT) preserving the information of the 3D data. Thus the reconstruction problem is reduced to reconstructing independent 2D slices and offers a very convenient approach to cluster computing.

In a previous work we used this approach to scale down the reconstruction time with implementations utilizing RPC or Syngo[1] communication facilities on a Windows NT network of commodity PCs and with the first version of our dedicated reconstruction cluster (seven four-processor-systems, Intel PIII @ 700 MHz, 1 GB RAM, switched fast ethernet), [6].

These previous results have encouraged us to upgrade our dedicated cluster to Myrinet networking equipment [8] as we could identify fast ethernet bandwidth as the limiting factor (less so for special purpose network topologies with multiple fast ethernet cards per node). We think that we now have a good basis to tackle a more complex reconstruction method for cluster adaption: OSEM3D in the implementation of C. Michel [9]. This reconstruction method has produced the best results for the HRRT data so far and is more suitable for an adequate treatment of the sinogram gaps that result from the detector geometry of the HRRT.

We are also in the process of developing a complementary suite of tools to integrate cluster reconstruction for HRRT and ECAT7 data into our clinical routine.

II. Material and Methods

A. Previous work based on RPC and Syngo

The *BeeHive* package was mainly developed to utilize idle Windows NT user work stations for distributed computing of FORE-preprocessed sinograms, [5]. It consists of three components: (a) the “busy bees” (slaves) which are installed (automatically) on all NT machines, (b) the “BeeKeeper” component that is responsible for adding new bees to the beehive, and (c) the “QueenBee” (master) which distributes work among the “bees” and collects results. The idea is roughly this: have a single scheduling thread on the master maintain a list of jobs and a list of workers (a single thread to keep the algorithm simple and to avoid network bottlenecks/deadlocks). Distribute the jobs in a round-robin way, use redundancy where possible and allow for runtime reconfigurations of the slaves, have one single binary (90 KB) that is easy to maintain (network update, single click de/installation). As the original reconstruction problem by virtue of FORE-preprocessing is rendered “embarrassingly parallel”, we did not see the need to rely on the usual message passing libraries (MPI, PVM, s. section *B*).

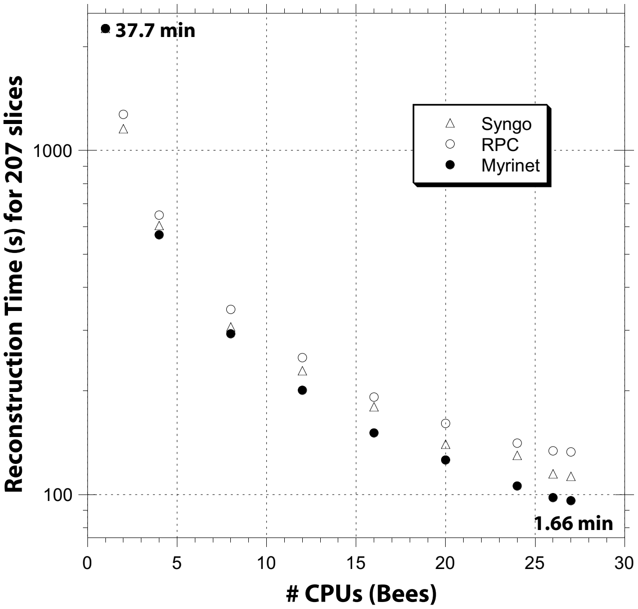


Figure 1: Influence of cluster size on reconstruction time for 207 slices (approx. 70 ML-EM iterations each), lin-log-plot, on our dedicated reconstruction cluster. Reconstruction time drops from 38 min in this example to well under 2 min for a full (FORE preprocessed) HRRT data set.

As described above, by using FORE the 3D-data sets are transformed (“rebinned”) to independent 2D-data sets that can be reconstructed by any fast 2D reconstruction method, e.g. OSEM. Currently, we are using Cologne-HOSP, [3], that is based on Schmidlin’s HOSP algorithm, [2], and has been port-

---

[1] Syngo by Siemens AG is “a common language for software applications in the field of medical technology“ (www.syngo.com).

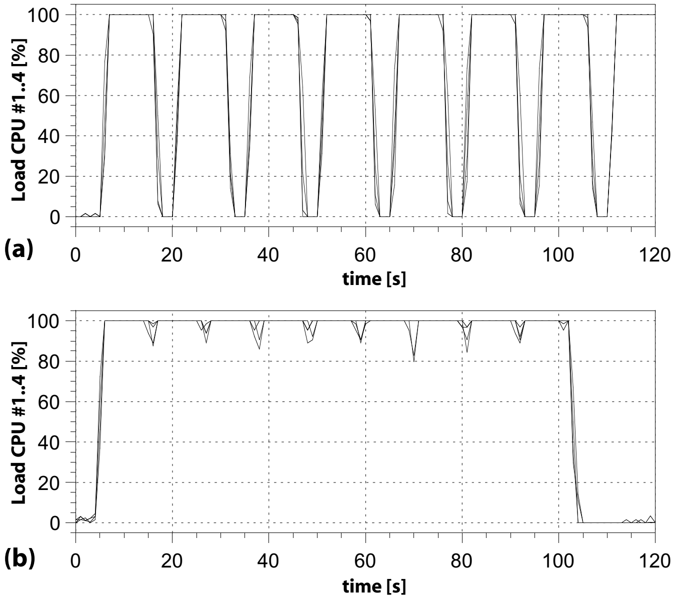


Figure 2: (a) CPU usage of one SGI 1450 node (four processors) of the *HeinzelCluster* during Syngo-based reconstruction of FORE-pre-processed data. This reconstruction involves 26 CPUs on seven nodes. (b) Same reconstruction job with Myrinet-GM-based implementation of *BeeHive*: observe the difference in idle time between productive intervals.

ed to *Solaris Sparc*, *IRIX*, *Linux x86*, *PowerMac* and *Windows NT*.

One disadvantage of the original RPC version is the polling approach of the scheduler which can create a lot of network overhead: each *Bee* is queried periodically for status information which can be a serious handicap when the cluster is inhomogeneous (too much time spent on slow machines). Syngo offers a powerful set of tools based on the ACE framework, [10], for implementing a pushing mechanism instead: the scheduler is notified about the status of a reconstruction bee automatically (proxy-return objects). Using Syngo also facilitates integration into the HRRT software environment (database access, reconstruction queues, visualization tools) as this is also based on the Syngo framework. Machines that only run a Syngo hosted reconstruction engine (backend) do not need a full Syngo distribution, a very small and easily maintainable subset suffices that does not create any overhead if not running.

We found that the simple RPC-based approach works fine for small homogeneous clusters (8 machines, speedup approx. a factor of 7) and that the more refined Syngo approach still gives good performance for homogeneous clusters of more than twice that size, s. Figure 1 and Table 1.

B. HeinzelCluster, BeeHive over Myrinet GM

The *HeinzelCluster* consists of seven SGI 1450 nodes, [15]: each has four Intel PIII Xeon processors @ 700 MHz, 1 MB L2-Cache, 1 GB RAM; all networked with full duplex fast ethernet on a Bay Stack 28115 switch; Myrinet M2M-PCI 64 A-2 cards and a Myrinet M2M SW16 switch and currently running Window NT. We decided against using Gigabit Ethernet as an alternative to Myrinet because this technology is limited to about 35 MB/s sustained data transfer by the IP stack handling, [11]. Myrinet is not much more expensive and gives much better performance: with our own Myrinet GM-based software (s.b.) we can transfer blocks of 52 MB size with a sustained rate of 130 MB/s and still expect more with the new generation "Myrinet 2000" hardware.

However, superior performance requires using special implementations of the message passing standard MPIch, [8], or adapting existing applications to GM, [12], a Myrinet proprietary message passing driver that is available for all major platforms.

We favor GM over standard message passing interfaces such as MPI and PVM (parallel virtual machine) because (apart from one scatter-gather function) we think that GM offers all the MPI functionality we need (s. *C*)—and on a standard NT system it just requires installing a single driver (like another regular ethernet card). As is pointed out in [13], MPI and PVM implementations handle network problems and hardware malfunctions usually quite ungracefully: GM is more robust and can resend lost packets automatically.

# CPUs (Bees)	Speedup Syngo	Speedup Myrinet
8	7.36	7.72
12	9.86	11.25
16	12.55	14.98
20	16.16	17.81
24	17.40	21.34
26	19.67	23.32

Table 1: Comparison of speedup *BeeHive* Syngo vs *BeeHive* Myrinet.

We have developed a little API that just requires linking one library and using one initialization sequence in order to send GM messages to any node of the cluster. The result is a Myrinet-GM based version of *BeeHive* that is fast, slim and conceptually very simple: on the *QueenBee* (master) and on each *Bee* (slave), i.e. on each of the seven cluster nodes, one GM event loop is running in its own thread and continuously polling for messages from other nodes (latency is known to be extremely good). Surprisingly, this creates very little overhead and allows for straight-forward event handling by implementing hooks that start appropriate worker threads, e.g. we send a message string that starts with a predefined identifyer (arbitrary integer): "4711"—receive sinogram of known dimensions (rest of message) and start reconstruction thread.

As can be seen from Fig. 2, with Myrinet-based *BeeHive* we visibly reduce idle time between "number crunching" periods.

C. OSEM3D acceleration

Porting the original Alpha version of OSEM3D (mostly ANSI C) to Windows NT was a straight-forward task. Our strategy to reduce reconstruction time (1 iteration for current HRRT data sets using the direct port on one Intel Pentium III@700 MHz takes about 11 h) consists of three steps: (a) parallelization for one node with four processors (shared memory), (b) higher level parallelization for the seven nodes to involve a reasonable number of CPUs (probably all) in the OSEM3D reconstruction, (c) platform specific optimizations.

We chose to work on (a) first, because a profile run clearly shows that more than 80% of the total wall time is spent inside the routines for forward and backprojection (which is the case for many iterative algorithms).

Thus we adapted the forward and back projection routines for multi-threaded execution with a wrapper routine (outer loop) leaving most of the original (unoptimized) engine untouched. This was an easy choice (and something Fortran compilers do by default on high perfomance platforms, [13]), e.g. the back projection algorithm is voxel driven so every thread takes over the voxels of 207/threads planes.

execution time [min]	**PC PIII 600 Mhz**	**SGI 1450 1 Thread**	**SGI 1450 4 Threads**	**Compaq Alpha EV67**
Normfac	259	212	73	76
1 Iteration	569	453	156	132
Total	828	665	229	208

Table 2: Multi-threaded implementation vs direct port. “Normfac” refers to a file of precomputed factors, s.b.

In case of the SGI 1450 running 4 threads, approx. 75% of the 229 min total time have been spent in either the forward or back projection routine, i.e. have been executed in parallel (which is acceptable for modifying only two routines and leaving almost everything else “as is”). The speedup of about 3 correlates nicely to the number of processors (4) and the time spent in parallel mode.

Apart from details in the thread synchronization (POSIX vs NT/2000), this approach should work for all multi-processor systems.

Approach (b) tackles parallelization with a much coarser granularity and is similar for the calculation of the “Normfac” data (precomputed factors needed only once for all iterations) and the main iteration loop. We propose to run the innermost loop (over the segments) of the pseudocode listing in Figure 3 in parallel. It follows that the theoretical speedup can not exceed the number of segments, which is 15 for the current configuration of the HRRT. After completion of the `isegment`-loop, the `cimage` volumes of all nodes must be collected and added (52 MB per cluster node) and for parallel execution of the next segment all nodes have to be synchronized by updating their local image buffer (another 52 MB per cluster node).

Finally, we believe that with (c) there is a significant potential for platform specific tuning we should investigate. However, we want to perform these optimizations last in order to stay as close to the original code as long as possible. Here the most relevant optimizations will be directly linked to the challenge of using the Pentium III’s advanced (and often neglected) features more efficiently. First tests with the new Intel 5.0 compiler plugin for the Microsoft Visual 6 C++ compiler have been encouraging. In specific, we have successfully vectorized portions of code using compiler intrinsics (rather than Assembler code). With the Streaming SIMD extensions (SSE) of the PIII that can give a speedup of up to 4 on single precision float operations. Furthermore, the PIII processor is rather dependent on proper alignment of operands for many of its more efficient operations (the Alpha processor is much more lenient, admittedly). All these changes will also be beneficial to potential Linux applications (on the Intel platform) of the code.

```
for ( iter ) {
  for ( isubset ) {
    zero ( cimage )
    for ( isegment ) {
        forward3D ( image, &estimate )
        corrections ( trues, norm, atten,
                        &estimate )
        backproj3D ( estimate, &cimage )
    }
    image = image * cimage
  }
}
```

Figure 3: Pseudocode for OSEM3D (non-parallel) main loop. `image` and `cimage` are image volumes, `trues`, `norm` and `atten` hold the appropriate projection data for segment `isegment`.

D. *BeeQ, Vinci and clinical routine*

In previous versions of *BeeHive* we used the *BeeKeeper* component for automatic network enabled update of the *Bee* service. With *BeeQ* we have extended the original concept: *BeeQ* is running as a service (similar to a Unix daemon) on every node of the *HeinzelCluster* and can run any program locally that does not require a graphical user interface. Communications have been realized with named pipes (but could easily be changed to TCP/IP sockets). *BeeQ* redirects *stdin* and *stdout* to log files and/or reports output directly (much in the fashion of telnet). It maintains a dynamic list of all processes it controls and allows to suspend and kill them. It has a built-in ftp client; processes information which jobs should run concurrently and which not and in what sequence; it can communicate with other instances of *BeeQ* on the cluster for load balancing. We intend to use *BeeQ* for monitoring a “heart beat” function and if necessary restart essential services.

In addition to the obvious need for a reconstruction cluster with HRRT data sets, we intend to use the *HeinzelCluster* also for reconstruction of especially demanding data sets from our ECAT HR and ECAT EXACT scanners (e.g. multi-frame, multi-bed studies). Data transfer from the Sun workstations running CTI’s ECAT 7 software has been realized via ftp and gives a performace of 9.5 MB/s. For routine handling, it is necessary to incorporate the cluster reconstruction option into the console’s user interface. This can be realized with a CAPP module that sends commands to the cluster via rsh to a redirector module that forwards *stdin* to *BeeQ*.

We currently can communicate with instances of *BeeQ* through a terminal-like tool that runs on the Windows platform from anywhere in the network (utilizing the network protocols TCP/IP or NetBUI), a subset of its functionality will be available as an IDL module (call external).

Reconstructions of HRRT data usually produce image volumes exceeding 50 MB: there is a new challenge to visualize these image volumes and the even more demanding sinograms

(> 300 MB, short integer format) efficiently. With *Vinci*, we have developed an application for Windows NT/2000 (C++ with MFC) that is entirely true color based and takes advantage of some features every modern PC graphic board offers. We have incorporated the reslicing engine that drives the MPITool, [16], which has been the institute's workhorse for visualization so far. *Vinci* also supports the ECAT7 data format and can be integrated into the Syngo frontend on the HRRT console; it only requires moderate hardware resources (memory requirements depend on how many image volumes you want to look at simultaneously) and was designed to run well on midrange laptop systems. In addition, we use parts of its true color engine for online visualization of reconstruction results.

III. RESULTS

With *BeeHive* over Myrinet and HRRT data we are quite close to the theoretical maximum speedup of the cluster while having the benefits of a simple scheduling approach. However, there is still some work ahead to speedup the preprocessing steps: we have started work on running FORE on more than one CPU. This is even more important for the more computationally demanding FORE-J, [17], and FORE-X, [18], flavours which we hope to evaluate for HRRT data in a collaboration with M. Defrise and X. Liu of the Department of Nuclear Medicine, Free University Brussels, Belgium.

With approach (a) of our strategy to accelerate OSEM3D one cluster node already performs in the range of an Alpha EV67. It remains to be seen how much more acceleration approaches (b) and (c) will provide.

IV. CONCLUSION

HeinzelCluster with *BeeHive* over Myrinet is a suitable platform for several reconstruction strategies, especially for HRRT data. Future work with this platform might include list-mode reconstruction with correction of motion artefacts.

V. ACKNOWLEDGMENTS

S.V. thanks M. Schmand, M. Lenox, J. Bao, P. Luk, M. Casey and R. Nutt for fruitful discussions and a most stimulating environment during S.V.'s stay at CTI.

VI. REFERENCES

[1] K. Wienhard., M. Schmand, M.E. Casey, K. Baker, J. Bao, L. Eriksson, W.F. Jones, C. Knoess, M. Lenox, M. Lercher, P. Luk, C. Michel, J.H. Reed, N. Richerzhagen, J. Treffert, S. Vollmar, J. Young, W.D. Heiss, R. Nutt, "The ECAT HRRT: Performance and First Clinical Application of the New High Resolution Research Tomograph.," IEEE MIC, Conference Record, 2000.

[2] P. Schmidlin, M.E. Bellemann, G. Brix, "Iterative reconstruction of PET images using a high-overrelaxation single-projection algorithm," *Phys. Med. Biol.* 42: 569-582, 1997.

[3] S. Vollmar, W. Eschner, K. Wienhard, U. Pietrzyk, "Iterative reconstruction of emission tomography data with a-priori-information," *Conference proceedings of the 1998 IEEE nuclear science symposium and medical imaging conference*, Toronto 1998.

[4] M. Defrise, P.E. Kinahan, D.W. Townsend, M. Sibomana, D.F. Newport, "Exact and Approximate Rebinning Algorithms for 3-D PET Data," IEEE Transactions on Medical Imaging, Vol. 16, No. 2: 145-158, 1997.

[5] S. Vollmar, K. Wienhard, "BeeHive: Cluster-Computing im NT-Netzwerk", *Conference Proceedings of the 16th DV meeting of the Max-Planck-Society*, Göttingen, November 1999.

[6] S. Vollmar, M. Lercher, C. Knöss, C. Michel, K. Wienhard, W.D. Heiss, "BeeHive: Cluster Reconstruction of 3-D PET Data in a Windows NT network using FORE", IEEE 2000, Lyon.

[7] Intel Corporation, Intel Architecture Optimization Reference Manual, Order Number: 730795-001.

[8] Myrinet networking hardware, http://www.myri.com.

[9] C. Michel, M. Schmand, X. Liu, M. Sibomana, S. Vollmar, C. Knöss, M. Lercher, C. Watson, D. Newport, M. Casey, M. Defrise, K. Wienhard, W.D. Heiss, "Reconstruction strategies for the HRRT", IEEE 2000, Lyon.

[10] D. Schmidt, "The ADAPTIVE Communication Environment (ACE)", http://www.cs.wustl.edu/~schmidt/ACE.html.

[11] c't, 16/99, p. 152, http://www.heise.de/ct

[12] Myricom, Inc., "The GM Message Passing System", http://www.myri.com/scs/GM/doc/gm.pdf

[13] K. Dowd, C. Severance, "High Performance Computing", O'Reilly, 2nd edition 1998.

[14] LAM/MPI Org, University of Notre-Dame, "Top 10 reasons to prefer MPI over PVM", http://www.lam-mpi.org/mpi/mpi_top10.php.

[15] SGI 1450 overview, http://www.sgi.com/servers/1450/

[16] U. Pietrzyk, A. Thiel, K. Herholz, W.D. Heiss, "A hybrid image registration method employing interactive and automated techniques", Neuroimage, 7:789, 1998.

[17] M. Defrise, X. Liu, "A fast rebinning algorithm for 3D PET using John's equation", Inverse Problems 15 1047-1065, 1999.

[18] X. Liu et al, "Exact Rebinning Methods for 3D PET", IEEE Trans Med Imag MI-18 657-664, 1999.

Evaluation of an exact Fourier rebinning algorithm for a large aperture PET scanner

Xuan Liu, Christian Michel, Stefan Vollmar, Klaus Wienhard, Mike Casey, Michel Defrise

INTRODUCTION

The reconstruction of 3D PET data is a computationally intensive task, especially for dynamic or whole-body studies involving several scans of the same patient. Depending on the available computational hardware, on the type of studies and on the required patient throughput, reconstruction time may still limits the clinical applications of 3D PET. This motivates continuing research for faster reconstruction algorithms for 3D PET data. The fastest methods to date are based on *rebinning methods* [1-7] which factor the 3D field-of-view into a stack of parallel *transaxial slices* and estimate, starting from the measured data, the 2D Radon transform (the sinogram) of the tracer distribution in each of these slices. Rebinning thereby reduces the redundant 3D data to a stack of independent 2D sinograms which can be reconstructed using either 2D filtered-backprojection (FBP) or 2D iterative algorithms such as OSEM.

A practical rebinning method is the *Fourier rebinning algorithm* [FORE 4,5,6,8]. This algorithm is based on the application of the frequency-distance relation to the 2D Fourier transform of each oblique sinogram. The frequency-distance relation is accurate only for large values of the frequency. This is why the accuracy of FORE decreases with increasing values of the angle θ between the LORs and the transaxial slices. Nevertheless, this algorithm is used routinely by several groups and was shown to be sufficiently accurate for most clinical studies with the current multi-ring PET scanners which have axial apertures not exceeding about $\theta_{max} \approx 15$ degrees. In addition to its approximate character, a second limitation of FORE is that it modifies the statistical properties of the data: contrary to the measured 3D data, the rebinned sinograms are not distributed as independent Poisson variables and hence should not in principle be reconstructed using standard statistical algorithms such as OSEM. In fact, OSEM is a very robust algorithm which has been applied with good results to 2D data rebinned using FORE [8,9,10].

Owing to these two limitations -the approximate character of FORE and its complex effect on data statistics- it is likely that fully 3D iterative algorithms [see e.g. 7] will eventually replace the rebinning algorithms, at least if the available computation power increases more rapidly than the number of LORs acquired by the scanners. Todate, however, fast reconstruction using rebinning algorithms remains attractive for scanners such as the ECAT HHRT [14,15] which acquire data sets which are still too large for a routine application of 3D iterative algorithms. But these scanners are precisely those for which the axial aperture is largest (it exceeds 20 degrees), and for which the approximation in FORE may degrade axial resolution. This was recently demonstrated with phantom studies on the ECAT HRRT scanner [15,16].

These observations suggest to investigate the potential benefit of applying an exact rebinning algorithm, FORE-J [12,13], to the HRRT scanner. This algorithm is exact but can be applied directly to axially truncated data, avoiding the reprojection step required by the reprojection method [3DRP, 17] or by the other exact rebinning algorithm described in the literature, FOREX [18]. The algorithm FORE-J is based on the property that the 3D X-ray transform of a function -the quantity sampled by a 3D PET acquisition- must be solution to the 2nd order partial differential equation (PDE) first studied by F. John. From a practical point of view, FORE-J is easy to implement since it has the same structure as FORE. The only modification is a small correction added to each oblique sinogram prior to rebinning. The calculation of this correction is fast numerically but involves a second derivative of the data with respect to the axial variable z, a quantity which is sensitive to noise.

Until now, FORE-J has only been applied to simulated data which did not model the axial undersampling used in practice or the effect of *gaps* in the detectors.

The aim of the present paper is to evaluate the practical usefulness of the algorithm using data from the ECAT HRRT scanner. In particular we

- compare FORE-J with FORE in terms of spatial resolution and clinical image quality,
- investigate the influence of the detector gaps on the use of rebinning algorithms,
- investigate stable yet accurate methods to estimate the second derivative in the FORE-J correction term.

IMPLEMENTATION OF FORE-J

The HRRT scanner has an axial field-of-view of 25.2 cm and comprises 104 "rings" of detectors. Each "ring" has an octogonal geometry. The list mode data from the HRRT are reorganized into sinograms with 256 radial elements and 288

X.L and M.D. is with the Department of Nuclear Medicine, AZ-VUB, Free University Brussels, Belgium (mdefrise@minf.vub.ac.be).

C.M. and M.C. are with CTI Inc, Knoxville, TN, USA

S.V. and K.W. are with the Max Plank Institute for Neurological Research, Cologne, Germany

angular samples. The radial sinogram sampling interval is 1.22 mm. Because of the junctions between the eight sides of the octogonal detector, the sinograms cannot be measured in eight diagonal bands called "gaps". In the angular variable, each gap covers about 6 samples out of the 288.

To reduce the amount of data to store and process, *axial angular undersampling* (compression) is normally used in the HRRT scanner. In this paper we use the current sinogram set which consists of 2209 sinograms (Span 9, ring difference 67, 325.7 Mbyte) and recovers 88% of the LOR information.

Before applying Fourier based rebinning the gaps in each sinograms are filled with a linear interpolation in the angular direction, using the standard routine in the ECAT software.

The basic equation of the FORE-J algorithm is [12,13]

$$P_{reb}(\omega,k,z_0)=\frac{1}{\delta_{\max}}\int_0^{\delta_{\max}} d\delta\;\Big(P(\omega,k,z_0+\frac{k\delta}{\omega},\delta)$$

$$+\frac{\delta\,(\delta_{\max}-\delta)}{\omega}\,\frac{\partial^3 P(\omega,k,z_0+\frac{k\delta}{\omega},\delta)}{\partial\omega\partial z^2}\Big)\quad \omega\neq 0$$

In this equation, $P_{reb}(\omega,k,z_0)$ is the 2D Fourier transform of the rebinned sinogram for the transaxial slice z_0 and $P(\omega,k,z,\delta)$ is the 2D Fourier transform of a measured oblique sinogram, where δ is the tangent of the angle θ between the oblique LORs and the transaxial plane, and z is the axial coordinate of the mid-point of the LORs. The variables ω and k denote respectively the radial and azimuthal frequencies. The first term in the RHS corresponds to the FORE algorithm, and the second term (second line) is the additive correction in the FORE-J method. The frequencies $\omega=0$ (in practice smaller than some small threshold) are rebinned using a single-slice rebinning approach, using only small values of θ.

For the result presented in this abstract, the second derivative of the sinograms with respect to the axial variable z (the slice index) was estimated using a three point mask (1,-2,1). Alternative methods will be tested.

After rebinning, each sinogram was reconstructed using 2D filtered-backprojection with a ramp filter (rectangular window cut-off at Nyquist's frequency). For the brain scan shown below, the resulting image was smoothed with a 3D gaussian filter with a FWHM of 1.6 pixel.

RESULTS

Simulated line source

A line source located in the central transaxial slice (FWHM 2.4 mm) was simulated as a digital 3D image, and 3D data were generated by ray tracing (Joseph's reprojection method with linear interpolation) through this digital image. The axial and transaxial FWHM of the reconstructed line source was estimated as a function of the position along the line.

Figure 1 shows the results obtained with FORE-J, with and without simulated gaps. In the absence of gaps, the transaxial and axial resolutions are almost independent of the radial position, as could be expected since the algorithm is exact and the simulation did not model the position dependence of the detector response. When the gaps are present and are filled by interpolation before rebinning, the axial resolution is degraded by about 0.5 mm. Note that the effect of the gaps on the resolution depends on the radial position.

The comparison with the approximate algorithm FORE (figure 2) illustrates the potential improvement of image quality allowed by the use of an exact rebinning algorithm, at least for a large aperture scanner such as the HRRT.

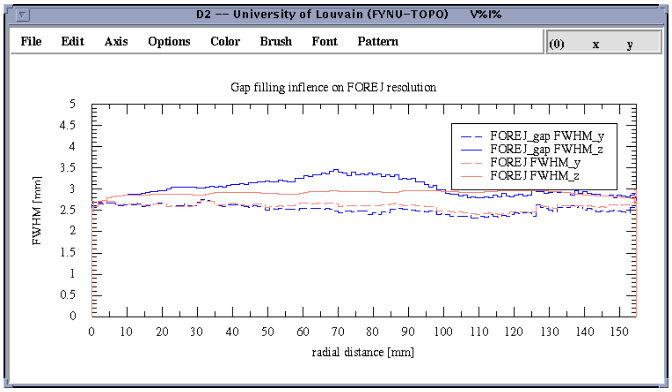


Figure 1. FWHM of the reconstructed simulated line source as function of the radial distance from the axis of the scanner. The axial (full curve) and transaxial (dashed curve) resolutions are shown for the case where the detector gaps are not simulated (red) and for a simulation including the gaps (blue).

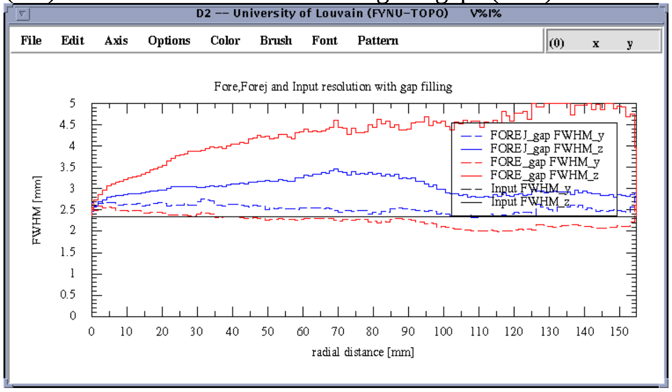


Figure 2. FWHM of the reconstructed simulated line source as function of the radial distance from the axis of the scanner. The axial (full curve) and transaxial (dashed curve) are shown for the FORE (red) and FORE-J algorithms (blue, same data as in figure 1) . The simulation includes the detector gaps. The full horizontal line indicates the FWHM of the simulated line.

Brain scan

An FDG brain scan was reconstructed using both FORE and FORE-J, with a voxel size of 1.22 mm. The total reconstruction time for the two algorithms was respectively 500 s and 717 s. The reconstructions are very similar but FORE-J introduces faint artefacts which appear as horizontal lines in the coronal and sagittal slices, and as rings in the transaxial sections (figures 3). These artefacts are not visible with FORE (not shown) and are tentatively attributed to noise propagation in the calculation of the second derivative. These artefacts are hardly visible on the smoothed images in figure 4 and 5, which would be used in practice. A close inspection of these images, and a comparison with a 3D OSEM reconstruction (8 subsets, 2 iterations, about 4 hours per iteration) reveals several details in the cortical structures which are better reproduced by the FORE-J reconstruction than by FORE. One example shown in figure 6 is the separation between the cortex and the cerebellum (tentorium cerebelli).

CONCLUSION

Even though the spatial resolution is also limited by the presence of gaps and by the axial undersampling, the exact rebinning algorithm FORE-J significantly improves the axial resolution for the ECAT HRRT scanner. The implementation of FORE-J only requires a minor modification of FORE, which only slightly increases the total reconstruction time. FORE-J appears to better reveal some fine details in the brain scan. However, the preliminary implementation of FORE-J introduces low level artefacts. Alternative methods to discretize the second derivative term in FORE-J will be investigated, and additional tests with phantom ("point" source) and clinical data will be presented at the conference.

REFERENCES

[1] Daube-Witherspoon M E and Muehllehner G, 1987, Treatment of axial data in three-dimensional PET, J Nucl Med 28, 1717-1724

[2] Erlandsson K, Esser P D, Strand S-E, van Heertum R L, 1994, 3D reconstruction for a multi-ring PET scanner by single-slice rebinning and axial deconvolution, Phys Med Biol 39, 619-629.

[3] Lewitt R M, Muehllehner G and Karp J S, 1994, Three-dimensional reconstruction for PET by multi-slice rebinning and axial image filtering, Phys Med Biol 39, 321-340.

[4] Defrise M, 1995, A factorization method for the 3D X-ray transform, Inverse Problems 11, 983-994.

[5] Tanaka E and Amo Y, 1998, A Fourier rebinning algorithm incorporating spectral transfer efficiency for 3D PET, Phys Med Biol 43, 739-746.

[6] Matej S, Karp J, Lewitt R, Becher A, 1998 Performance of the Fourier rebinning algorithm for PET with large acceptance angles, Phys Med Biol **43** 787-795

[7] Qi J, Leahy R, Hsu C, Fraquar T, Cherry S R, 1997, Fully 3D Bayesian Reconstruction for the ECAT EXACT HR+, Proc IEEE Med Imag Symposium

[8] Liow J-S, Anderson J R, Strother S C 1999, Comparing Reconstruction Algorithms using a Multivariate Analysis, records of the 1999 IEEE Medical Imaging Symposium, Seattle, WA. paper M06-001

[9] Comtat C, Kinahan P E, Defrise M, Michel C and Townsend D W, 1998, Fast reconstruction of 3D PET with accurate statistical modelling, IEEE Trans Nucl Sc NS-45, 1083-1089

[10] Liu X, Comtat C, Michel C, Kinahan P, Defrise M, Townsend D 2001 Comparison of 3D Reconstruction with 3D OSEM and with FORE+OSEM for PET, to be published in IEEE Trans Med Imag

[11] Defrise M, Kinahan P E, Townsend D W, Michel C, Sibomana M and Newport D F, 1997, Exact and approximate rebinning algorithms for 3D PET data, IEEE Trans Med Imag MI-16, 145-158.

[12] Defrise M, Liu X 1999 A fast rebinning algorithm for 3D PET using John's equation, Inverse Problems **15** 1047-1065

[13] Defrise M, Liu X 1999 Fast and exact Fourier rebinning using John's equation, records of the 1999 IEEE Medical Imaging Symposium, Seattle, WA. paper M03-001.

[14] Wienhard K, et al, 1999, Performance Evaluation of a new LSO High Resolution Research Tomograph - HRRT , records of the 1999 IEEE Medical Imaging Symposium, Seattle, WA. paper M04-002.

[15] Wienhard K et al 2000, The ECAT HRRT: Performance and First Clinical Application of a New High Resolution Research Tomograph, records of the 2000 IEEE Medical Imaging Symposium, Lyon, France

[16] Michel C et al 2000, Reconstruction Strategies for the HRRT, records of the 2000 IEEE Medical Imaging Symposium, Lyon, France

[17] Kinahan P E and Rogers J G, 1990, Analytic three-dimensional image reconstruction using all detected events, IEEE Trans Nucl Sci NS-36, 964-968.

[18] Liu X et al 1999, Exact Rebinning Methods for 3D PET, IEEE Trans Med Imag **MI-18** 657-664

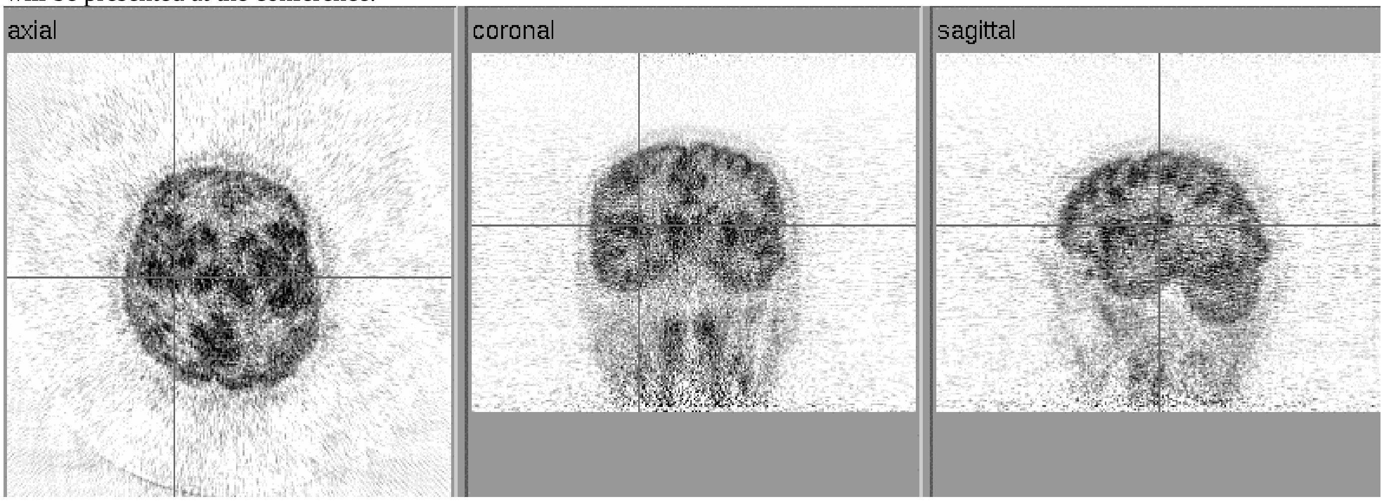


Figure 3. Reconstruction of the brain scan with the exact rebinning algorithm FORE-J, followed by 2D filtered-bacprojection with a rectangular window. Note the ringing artefacts in the transaxial slice.

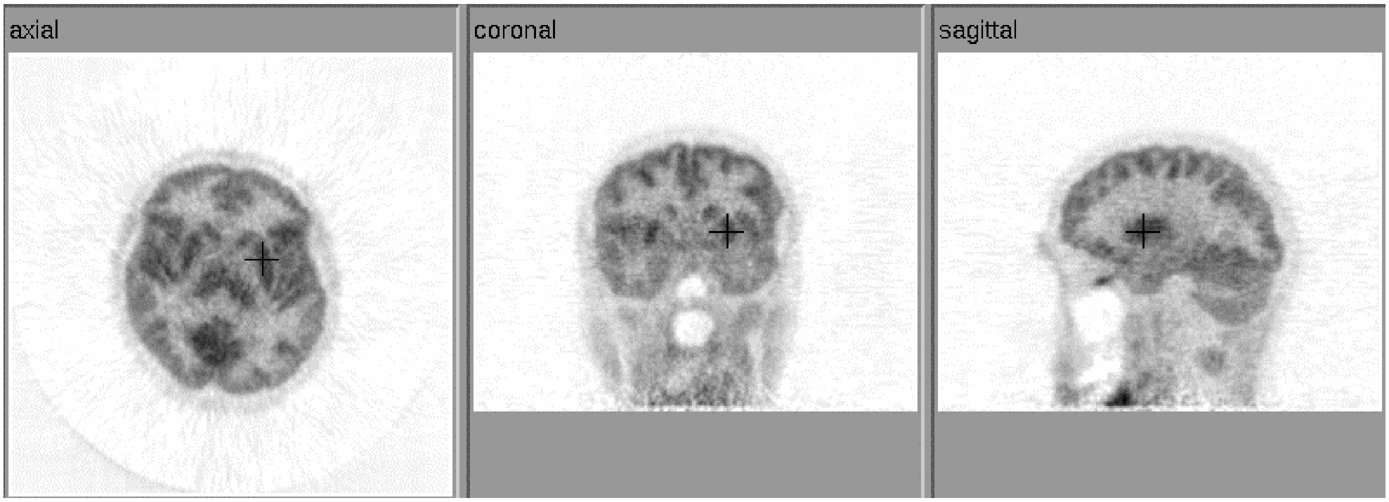


Figure 4. Reconstruction of the brain scan with the exact rebinning algorithm FORE-J, followed by 2D filtered-bacprojection and 3D Gaussian post smoothing (FWHM 1.6 pixel).

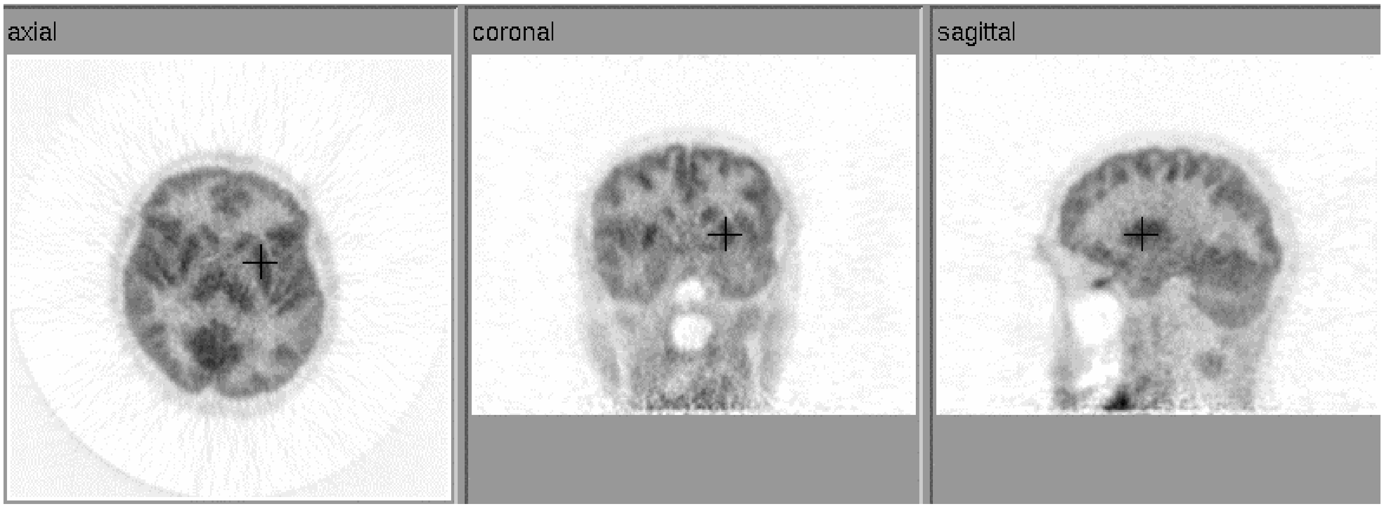


Figure 5. Reconstruction of the brain scan with the approximate rebinning algorithm FORE, followed by 2D filtered-bacprojection and 3D Gaussian post smoothing (FWHM 1.6 pixel).

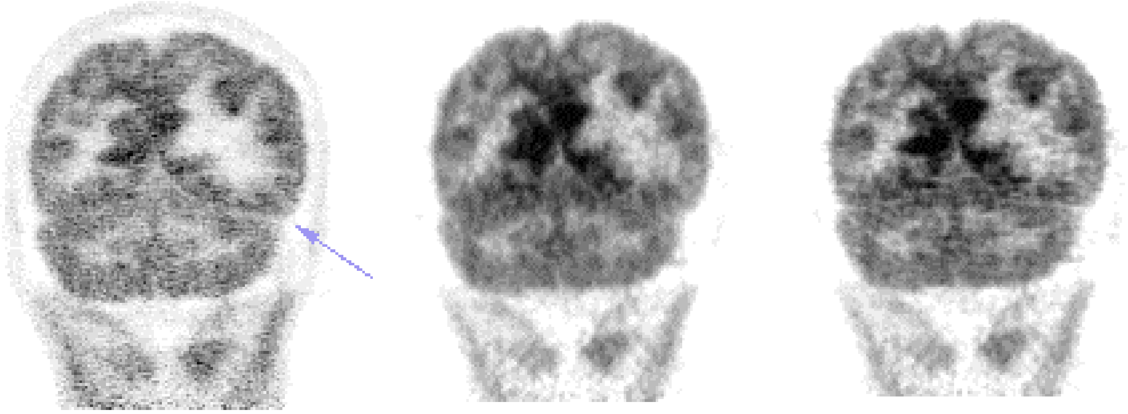


Figure 6. Coronal slice of the brain scan. Left: 3D OSEM with scatter correction, 8 subsets, 2 iterations. Middle: FORE+FBP as in figure 5, Right: FORE-J+FBP as in figure 4 (both without scatter correction). The gray scale is the same for FORE and FORE-J, but not for 3D OSEM. Note the different noise structure and better contrast in the scatter corrected 3D-OSEM image, and the finer detail (arrow) seen in FORE-J than with FORE.

Performance of 3D RAMLA with Smooth Basis Functions on Fully 3D PET Data

Samuel Matej, Margaret E. Daube-Witherspoon, and Joel S. Karp

Abstract— **3D reconstructions from fully 3D PET data can yield high-quality images but often at a high computational cost. To obtain practical data processing and reconstruction times, simplified and less precise approaches are used in the routine clinical use. We studied the feasibility of using the 3D row action maximum likelihood algorithm (3D RAMLA) with 3D spherically-symmetric basis functions (blobs) located on an efficient spatial (body centered cubic) grid for clinical PET data. The BCC grid provides more uniform distribution of the basis functions that represent the reconstructed object and decrease the computational time. Another development used in our study is a fast Fourier based forward projector that provides very fast calculation of the attenuation coefficients in fully 3D data space. These two developments move fully 3D reconstruction using appropriate data processing approaches toward clinically practical times. We are studying the practical effects of the use of these more precise approaches to fully 3D reconstruction on clinical data.**

Keywords— **Fully 3D reconstruction, PET, RAMLA, 2.5D reconstruction, smooth basis functions, attenuation correction.**

I. Introduction

MODERN Positron Emission Tomography (PET) scanners are characterized by a large axial Field Of View (FOV) which enables acquisition of data from a large range of oblique angles. Efficient 3D algorithms are needed to process the fully 3D data provided by those scanners in clinically reasonable times. In our previous studies [1–3] on 3D image reconstruction for PET we used various iterative algorithms operating on a series expansion representation of the volume, where the spatial distribution to be reconstructed was represented by the superposition of 3D basis functions [4]. These basis functions, which we call "blobs," are spherically-symmetric with bell-shaped radial profiles. Using a blob basis function inside the reconstruction process preserves the consistent component (true signal) of the data, thus preserving the resolution of the measured data, while at the same time suppressing the stochastic part (noise) of the data, as experimentally confirmed in [1]. Although filtering also suppresses noise, it can not preserve the spatial resolution of the data at the same time. In all of our studies, the 3D iterative reconstructions using blobs provided substantial and consistent improvement over the methods using classical basis functions - voxels. Until recently, the computational demands of the 3D iterative algorithms were found to be too high for routine clinical use.

To decrease reconstruction times, clinical reconstruction approaches usually involve rebinning of the measured fully 3D data into non-oblique sinogram data within slices, followed by slice by slice reconstruction. The most popular, and clinically used, rebinning technique is Fourier rebinning (FORE) [5]. Rebinning substantially decreases the volume of the data at the expense of reduced quality caused by using 2D reconstruction techniques instead of the 3D methods. In our recent studies [6, 7] we evaluated a 2.5D reconstruction approach (used after FORE) that considerably improves image quality over a pure 2D approach but with computational demands of the same order as those of 2D techniques. The 2.5D reconstruction considerably decreases the computational burden because it uses rebinned data while, at the same time, it takes into consideration volume nature of the measured data (by using 3D blob basis functions) and thus keeps some advantages of the 3D techniques. More precisely, in the 2.5D approach the reconstructions of the individual slices are coupled, and iteration calculations for each projection line are influenced by, and contribute to, several image slices. In our previous studies, images reconstructed by 2.5D algorithms were found to be superior in terms of a number of figures of merit related to resolution and noise to those produced by 2D algorithms, in which the individual slices are handled separately [6, 7]. However, both 2D and 2.5D approaches are affected by rebinning approximation errors as the axial acceptance angle increases [8, 9]. For axial acceptance angles around (or exceeding) ±15°, as occurs with some of the recent commercial PET scanners, FORE approximation errors already introduce noticeable image artifacts and nonuniform deterioration of axial resolution. There are other more precise rebinning approaches [5], they are, however, computationally more demanding. In addition to the approximation errors, approaches using rebinned data are limited because of the lack of proper reconstruction models taking into account rebinning (and data acquisition) effects.

Image representation using smooth basis functions provides an alternative approach that reduces the computational demands of 3D iterative approaches. This particular image representation allows utilization of a more advantageous spatial grid, compared to the classical voxel (simple cubic) grid. It also leads to a more uniform 3D distribution of the basis functions (grid points) throughout the 3D space, using Body Centered Cubic (BCC) grid based on the effective spatial sampling [10]. This, in turn, allows one to decrease the grid density (number of grid points) without compromising the quality of the image representation. This result has been demonstrated in our previous

This work was supported by The Whitaker Foundation and DOE Grant FG02-88ER60642.

Authors are with the Department of Radiology, University of Pennsylvania, 423 Guardian Drive, 4th floor Blockley Hall, Philadelphia, PA 19104, USA (e-mail: matej@mipg.upenn.edu).

studies using simulated data [10] and is reevaluated in this current study using measured PET data.

Proper treatment and utilization of attenuation information plays an important role in whole body PET imaging [11]. For the attenuation information to be treated properly, the attenuation coefficients are needed on the same set of LORs as those of the acquired emission data. In the fully 3D data case, the measured emission data has to be corrected for attenuation before being processed by rebinning or 3D reconstruction. For 3D iterative techniques, the attenuation coefficients can be utilized directly within the reconstruction model. Most typically, the attenuation coefficients are calculated by the forward projection of (preprocessed) transmission images. However, forward projection into 4D parameter (fully 3D data) space is a time consuming operation which might take considerably more time than the reconstruction itself, as in the case of FORE+2D/2.5D reconstruction approaches. For this reason, simplified approaches are often used in routine clinical use, such as rebinning of raw emission data without attenuation correction and applying (2D) attenuation correction afterwards. This substantially speeds-up the attenuation correction calculations, since only non-oblique attenuation factors are needed, but introduces additional errors into the reconstruction process. As part of the development of Direct Fourier Reconstruction with Fourier Reprojection (3D_FRP) [12], we have implemented fast Fourier based forward projector (FoProj) allowing for very fast calculation of fully 3D attenuation data and making proper treatment of the attenuation information within the fully 3D reconstruction process more practical.

In this work, we compare performance of the 3D and 2.5D iterative Row Action Maximum Likelihood Algorithm (RAMLA) [3,13] using blobs on the efficient spatial (BCC) grid. We also evaluate the impact of using attenuation correction information in the reconstruction algorithm. 3D PET data obtained from a clinical C-PET (ADAC UGM) whole body scanner are used in our study.

II. Methods

A. Reconstruction Algorithms

The row action maximum likelihood algorithm (RAMLA) [13] was developed as a faster alternative to the maximum likelihood expectation maximization (ML-EM) algorithm for maximizing the Poisson likelihood in PET. In RAMLA, the reconstructed image is updated for each projection line (row of the system matrix) in a controlled way using a relaxation parameter. The $(k+1)$'th update step, where $k \geq 0$, produces an image represented by a set of basis function coefficients $\{ c_j^{(k+1)} \}_{j=1}^{J}$ using the formula

$$c_j^{(k+1)} = c_j^{(k)} + \lambda_k c_j^{(k)} \left(\frac{g_{i_k}}{\langle a_{i_k}, c^{(k)} \rangle} - 1 \right) a_{i_k,j} \ , \qquad (1)$$

where $i_k = [k(mod I) + 1]$, g_{i_k} and $\langle a_{i_k}, c^{(k)} \rangle$ represent the measured data and forward projection for the i_k'th line, respectively, and λ_k is the relaxation parameter (including normalization factor) fulfilling the condition $\lambda_k a_{i_k,j} \leq 1$ for every i_k, j [13]. The data are accessed using a special ordering scheme [14] to ensure that the sequential projection lines are as orthogonal as possible, thus considerably speeding-up the rate of convergence. RAMLA achieves stable performance and relative independence on the starting point via an appropriately chosen relaxation parameter, controlling the amount of updates/corrections in each iteration step.

In the studies reported in this paper we used a 3D implementation of RAMLA using blobs - modified Kaiser-Bessel basis functions of second order [4] located on the standard cubic (voxel) grid and on the efficient BCC grid [10] of various grid step sizes. To study the improvement of the true 3D reconstruction over the (pseudo 3D) reconstruction from rebinned data we employed 2.5D RAMLA [6], which is currently our best reconstruction technique for rebinned PET data.

We applied 3D RAMLA and 2.5D RAMLA (after FORE) to phantom and patient datasets. To determine optimal parameters for each reconstruction algorithm for a given activity distribution, the data were reconstructed and analyzed for a range of parameters and blob sizes (FWHM) by varying the blob radius a (2-3, relative to voxel size of 4mm), blob shape parameter α (chosen as described by Matej and Lewitt [1, 10]), grid spacing g for the BCC grid (1.8-2.26, relative to voxel size), and relaxation parameter λ (0.001-0.1). In most cases, performing more than one iteration of RAMLA did not significantly improve image quality. The representative reconstruction times for a set of comparable reconstruction parameters ($a = 2.5$, $\alpha = 8.63$, $g = 2.00$, 1 iteration, data: 128 radial bins, 64 slices, 96 views, 7 tils, image: 144x144x64 with 4x4x4mm^3 voxels) obtained on single processor Sparc Ultra 10 (440MHz) were: 3D RAMLA (regular grid) 35:20 (min:sec), 3D RAMLA (BCC grid) 6:45, 2.5D RAMLA 1:56 (including 1min for FORE).

B. Attenuation Effects

We studied various ways of utilization of attenuation information: reconstructions from data precorrected for attenuation in the fully 3D data space (4DAC), date corrected in the rebinned data space (2DAC), and reconstructions using attenuation information within the system model (sysAC). The attenuation coefficients were obtained by forward projection of the transmission images. Transmission images were reconstructed from a ^{137}Cs single-photon source transmission data, rescaled to 511-keV, and segmented [15]. The transmission data were rebinned using the single-slice rebinning algorithm (SSRB) [16] and corrected for emission contamination [17].

Forward projection was calculated using fast Fourier based projector (FoProj) [12]. The computation time for complete 4D attenuation coefficient/correction data (image:144x144x64, data:128x64x96x7) was 0.8 min, as compared to 3.7 min for the forward projection based on the Siddon algorithm [18]. Although, both times are clinically reasonable for this image size, the ratio between them increases approximately proportionally to the image size increase (based on the computation complexity of $O(N^3 logN)$ for FoProj versus $O(N^4)$ for Siddon). Additional substantial speed-up of the FoProj algorithm can be obtained by using (off-the-shelf) FFT processor boards.

C. Torso Phantom Studies

The IEC phantom distribution recently adopted by the NEMA Coincidence Imaging Task Force for the measurement of image quality [19] was used to mimic patient imaging of the torso. The distribution consists of a torso phantom, containing hot and cold spheres in a warm background. The hot spheres have diameters of 1.0, 1.3, 1.7, and 2.2 cm; the cold spheres have diameters of 2.8 and 3.7 cm. A 5 cm diameter lung-like insert is also placed in the center of the phantom to provide a nonuniform attenuation distribution. The background was filled with ^{18}F at an activity concentration typical of what is seen in patient FDG studies (250MBq in 70kg, or 3.6kBq/cc). The hot spheres were filled with an activity concentration of 30 kBq/cc, for a "tumor":background activity ratio of 8:1. Scan durations of 3, 6, and 12 min were selected to generate a wide range of count densities. The ADAC UGM C-PET scanner with a maximum axial acceptance angle of ±15°, sorted into 7 tilt angles, was used.

To quantitate the performance of the algorithms, regions of interest (ROIs) with diameters equal to the physical inner diameter of each sphere were drawn on the slice through the centers of the spheres. Twelve ROIs of the same sizes as those for the spheres were drawn throughout the background in the central slice, as well as in slices ±8 mm and ±20 mm away. The coefficient of variation of the means in these 60 background ROIs was determined for each sphere size as a measure of the background variability. The hot sphere contrast recovery coefficient (CRC_{hot}) was calculated as

$$CRC_{hot} = (C_{hot}/C_{bkg} - 1)/(a_{hot}/a_{bkg} - 1) \qquad (2)$$

where C_{hot} and C_{bkg} are the average of the counts measured in the hot sphere ROI and the average of the counts in all 60 background ROIs, respectively, and a_{hot}/a_{bkg} is the ratio of the activities in the hot sphere and background (8 in this study). The cold sphere contrast recovery coefficient (CRC_{cold}) was calculated as

$$CRC_{cold} = (1 - C_{cold}/C_{bkg}) \qquad (3)$$

where C_{cold} is the average of the counts measured in the cold sphere ROI.

D. Whole-body Patient Studies

Several clinical whole-body patient studies were acquired after injection of 250MBq/70kg of [^{18}F]-FDG on the ADAC UGM C-PET scanner. The five-frame studies covered 70 cm axially. The acquisition followed an interleaved emission (6 min) / transmission (1.5 min) protocol.

III. Results and Discussion

Figure 1 shows for the IEC phantom data a plot of contrast vs. background variability for the 6-min scan for 3D RAMLA_BCC using various blob, grid and λ parameters. The plots for each feature size form approximately a single curve. The points representing results for 3D RAMLA using a regular cubic grid (solid symbols) and optimal parameters (based on our previous studies) lie on the same curves, showing that the two methods provide similar contrast versus noise tradeoff for an appropriate choice of parameters. At the same time, RAMLA_BCC provide about 5.2-times shorter reconstruction time. The comparable quality of the two methods was also confirmed by the visual observation of the reconstructed phantom and patient images.

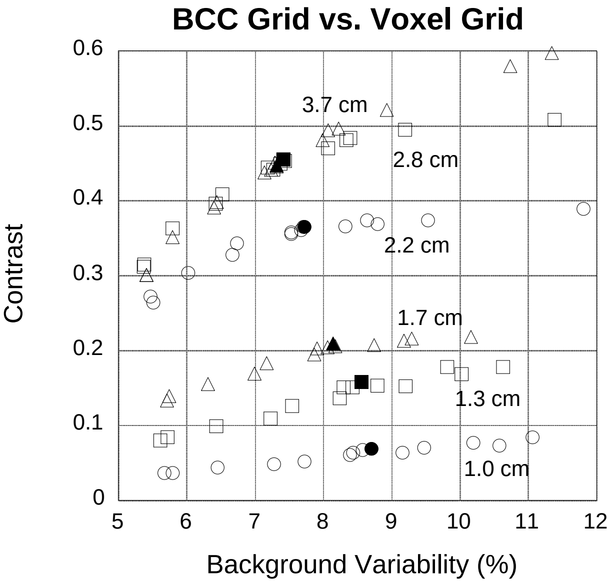


Fig. 1. Plots for the 6 spheres (4 smallest are hot, 2 largest are cold) for 6-min scan. The open symbols are for various blob and grid parameters (a:α:g = 2:3.6:2.26, 3:9.5:2.26, 2.5:8.63:2.00, 3:13.06:1.8) and various λ for 3D RAMLA using BCC grid. The closed symbols are for one representative reconstruction by slower version of RAMLA (a:α = 3.0:12.95) using standard cubic grid. We can obtain comparable contrast/variability performance with RAMLA_BCC for an appropriate choice of blob/grid parameters as with slower version of 3D RAMLA, but in about 1/5 the reconstruction time.

Figure 2 shows example of the study using IEC phantom comparing 2.5D reconstruction with 2D attenuation precorrection done after FORE to the 3D RAMLA (BCC) with system attenuation correction within the model. The first approach represents a simplification of the theory, but as can be seen in the left image, it still provides reasonable images and is often employed in the routine clinical use because of its low computational demands. The improvement due to proper use of the attenuation information in the fully 3D reconstruction algorithm is demonstrated by the increase of the background uniformity and the increase of the contrast of the features.

Figure 3 shows similar comparison for the whole-body patient study. On the left is coronal image of 2.5D RAMLA with pre-correction for attenuation after FORE using 2D attenuation correction factors. On the right is the corresponding slice for 3D RAMLA with system-modeling of attenuation effects using 4D attenuation correction factors generated using FoProj. It is evident that the fully 3D

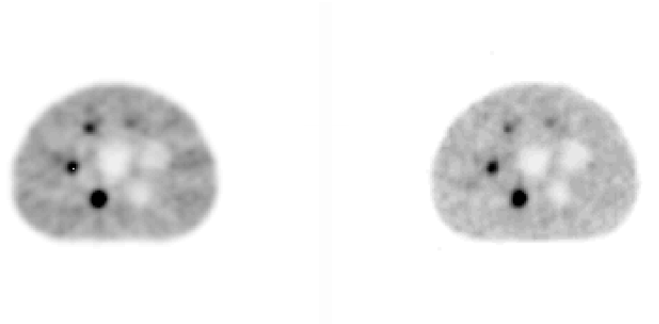


Fig. 2. 6-min IEC phantom study reconstructed using 2.5D RAMLA with 2DAC (left side) and 3D RAMLA with sysAC (right side). It can be seen that the background is more uniform with 3D RAMLA and system-modeling of attenuation effects.

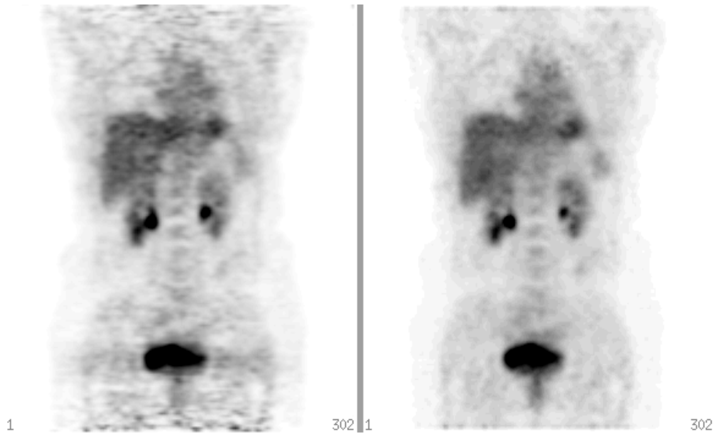


Fig. 3. Patient whole-body study (coronal images) reconstructed using 2.5D RAMLA with 2DAC (left side) and 3D RAMLA with sysAC (right side). It can be clearly seen that 3D RAMLA provides more uniform and less noisy reconstruction. The bladder streak artifacts seen throughout the slices of the 2.5D/2DAC reconstruction are completely absent in the 3D/sysAC reconstruction.

reconstruction together with the system modeling of the attenuation provides a definite improvement of the image quality. This is demonstrated by a combination of increased background uniformity, decreased noise and suppression of the bladder artifacts.

Examples shown in Figures 2 and 3 represent two ways to utilize attenuation information. The first - 2DAC after FORE - is oversimplified but is a very fast approach, while the second - system modeling with 4DAC - is more theoretically sound but is significantly slower. Note, that due to the developments discussed in this paper, even the second approach is now becoming feasible for the clinical use. There are several intermediate possibilities between these two extremes, such as using 4DAC+FORE+2D iterative reconstruction (OS-EM) which is in use at other clinical sites [11]. Our observations suggest that both the system modeling of attenuation and the use of a fully 3D reconstruction algorithm (rather than rebinning) lead to improvements in image quality. We are currently investigating the influence of individual steps on the reconstruction quality of the PET images in a more rigorous way employing measures of contrast and noise, as represented by the study in Figure 1. The results will be used to guide the data processing and image reconstruction protocols for the whole body patient studies.

REFERENCES

[1] S. Matej and R. M. Lewitt, "Practical considerations for 3D image reconstruction using spherically-symmetric volume elements," *IEEE Trans. Med. Imaging*, vol. 15, no. 1, pp. 68–78, 1996.

[2] S. Matej, G. T. Herman, T. K. Narayan, S. S. Furuie, R. M. Lewitt, and P. E. Kinahan, "Evaluation of task-oriented performance of several fully 3D PET reconstruction algorithms," *Phys. Med. Biol.*, vol. 39, no. 3, pp. 355–367, 1994.

[3] S. Matej and J. A. Browne, "Performance of a fast maximum likelihood algorithm for fully 3D PET reconstruction," in *Series Computational Imaging and Vision: Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, P. Grangeat and J.-L. Amans, Eds., pp. 297–316. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.

[4] R. M. Lewitt, "Multidimensional digital image representations using generalized Kaiser-Bessel window functions," *J. Opt. Soc. Am. A*, vol. 7, no. 10, pp. 1834–1846, 1990.

[5] M. Defrise, P. E. Kinahan, D. W. Townsend, C. Michel, M. Sibomana, and D. F. Newport, "Exact and approximate rebinning algorithms for 3D PET data," *IEEE Trans. Med. Imaging*, vol. 16, no. 2, pp. 145–158, 1997.

[6] T. Obi, S. Matej, R. M. Lewitt, and G. T. Herman, "2.5D simultaneous multislice reconstruction by series expansion methods from Fourier-rebinned PET data," *IEEE Trans. Med. Imaging*, vol. 19, no. 5, pp. 474–484, 2000.

[7] M. E. Daube-Witherspoon, S. Matej, J. S. Karp, and R. M. Lewitt, "Application of the row action maximum likelihood algorithm with spherical basis functions to clinical PET imaging," *IEEE Trans. Nucl. Sci.*, vol. 48, no. 1, pp. 24–30, 2001.

[8] S. Matej, J. S. Karp, R. M. Lewitt, and A. J. Becher, "Performance of the Fourier rebinning algorithm for PET with large acceptance angles," *Phys. Med. Biol.*, vol. 43, no. 4, pp. 787–795, 1998.

[9] J. S. Karp, A. J. Becher, S. Matej, and P. E. Kinahan, "Data processing and image reconstruction methods for the HEAD PENN-PET scanner," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 3, pp. 1144–1151, 1998.

[10] S. Matej and R. M. Lewitt, "Efficient 3D grids for image reconstruction using spherically-symmetric volume elements," *IEEE Trans. Nucl. Sci.*, vol. 42, no. 4, pp. 1361–1370, 1995.

[11] C. Comtat, P. E. Kinahan, M. Defrise, C. Michel, and D. W. Townsend, "Fast reconstruction of 3D PET data with accurate statistical modeling," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 4, Part 2, pp. 1083–1089, 1998.

[12] S. Matej, "3D-FRP: Direct Fourier reconstruction with Fourier reprojection for fully 3D PET," in *Proceedings of the 2000 IEEE Nuclear Science Symposium and Medical Imaging Conference. CDROM, IEEE572.* Lyon, France, October 15-20, 2000, To be published.

[13] J. A. Browne and A. R. De Pierro, "A row-action alternative to the EM algorithm for maximizing likelihoods in emission tomography," *IEEE Trans. Med. Imaging*, vol. 15, no. 5, pp. 687–699, 1996.

[14] G. T. Herman and L. B. Meyer, "Algebraic reconstruction techniques can be made computationally efficient," *IEEE Trans. Med. Imaging*, vol. 12, no. 3, pp. 600–609, 1993.

[15] R. J. Smith, J. S. Karp, G. Muehllehner, E. Gualtieri, and F. Bénard, "Singles transmission scans performed post-injection for quantitative whole body PET imaging," *IEEE Trans. Nucl. Sci.*, vol. 44, no. 3, pp. 1329–1335, 1997.

[16] M. E. Daube-Witherspoon and G. Muehllehner, "Treatment of axial data in three-dimensional PET," *J. Nucl. Med.*, vol. 28, pp. 1717–1724, 1987.

[17] J. S. Karp, G. Muehllehner, H. Qu, and X. H. Yan, "Singles transmission in volume-imaging PET with a ^{137}Cs source," *Phys. Med. Biol.*, vol. 40, pp. 929–944, 1995.

[18] R. L. Siddon, "Fast calculation of the exact radiological path for a three-dimensional CT array," *Med. Phys.*, vol. 12, no. 2, pp. 252–255, 1985.

[19] J. S. Karp, R. J. Smith, G. Muehllehner, M. E. Daube-Witherspoon, and H. Hines, "Image quality measurement for evaluation of PET scanner performance (abstract)," *J. Nucl. Med.*, vol. 39, no. 5, pp. 133P–134P, 1998.

Movement Artifacts in Helical CT Cone-Beam Reconstruction

Claas Bontus[1], Roland Proksa[1], Jan Timmer[2], Thomas Köhler[1], Michael Grass[1]

Abstract— **We designed different dynamical CT phantoms and simulated a CT Cone-Beam scanner. The obtained projection data were used as input for three different reconstruction algorithms. Since the images show movement artifacts the comparison of different reconstruction algorithms for helical scanning yields information about the sensitivity of these algorithms to movement during the scanning procedure.**

Keywords— **CT reconstruction, Cone-Beam algorithms, Helical CT scanning, Movement artifacts.**

I. INTRODUCTION

THE next generation of medical CT scanners, viz. Cone-Beam scanners, necessitates sophisticated reconstruction algorithms. Those algorithms which are considered for practical purposes have to be evaluated for numerical stability, artifacts, reconstruction times and the sensitivity to all kinds of system imperfections. In a realistic environment movement is always involved during the scanning procedure. The patient and his/her organs never stay completely motionless. The evaluation of the different reconstruction algorithms with respect to motion artifacts is therefore another important aspect. We have designed different dynamical phantoms and tested the resulting level of artifacts for different helical reconstruction algorithms. In particular we restricted ourselves to the evaluation of approximative algorithms, which are of the back-projection type. For the comparison with a well known acquisition scheme we used single detector row circular reconstruction (filtered back-projection) as a gold standard.

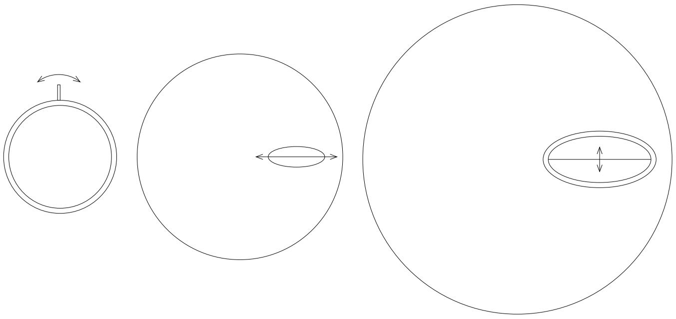
II. SCANNER GEOMETRY

For the simulations we used a CT scanner with 40 detector lines and 1024 elements per line. We set the detector height equal to 35.34 mm and its width to 904.7 mm, which corresponds to a fan-angle of 56.96 degrees. The distance of the focal spot to the rotation axis was chosen to be 515 mm and the distance of the focal spot to the detector center was equal to 910 mm. The scanner's rotation time was set to 0.5 s (2 Hz).

For helical scanning we performed simulations with two different acquisition schemes, viz. 1-PI and 3-PI mode. In the 1-PI mode the pitch was set equal to 32.64 mm per rotation for the head-phantom (see below) and to 26.04 mm for the abdomen phantoms. In the 3-PI mode it was set equal to 11.97 mm per rotation for the head-phantom and to 10.34 mm for the abdomen phantoms.

III. PHANTOM DESIGN

In the following subsections we describe the phantoms in detail. They are composed of mathematical objects with different densities. Some of these objects oscillate during the scanning procedure. For our analysis we have chosen relatively high frequencies and amplitudes compared to a realistic case. By this way we realized a scenario that covers the worst possible case. The following images indicate the location of the mathematical objects within the phantoms. The arrows illustrate the movement.



A. Head phantom

This phantom is intended to approximate movement of the septum, which can result in severe artifacts in clinical scanning. The phantom is composed of a cylinder with a radius of 80 mm consisting of water. The cylinder is surrounded by a shell with a thickness of 10 mm consisting of bone. The symmetry axis of the cylinder coincides with the scanner's rotation axis. Moreover, a box with a thickness of 2 mm and a length of 22.5 mm is connected to the outer shell. This box also consists of bone and is supposed to approximate the septum. For the simulation of the movement we let this box oscillate with an amplitude of five degrees and a frequency of 1.51 Hz.

B. Abdomen phantom, colon movement

During inspections of the abdomen colon movement can also result in artifacts. We therefore designed a phantom, which is composed of a cylinder consisting of water. This cylinder has a radius of 170 mm and its symmetry axis coincides with the rotation axis. For the simulation of the colon we placed an ellipsoid consisting of air inside the cylinder. It has a horizontal radius of 50 mm, a vertical radius of 20 mm and was placed such that its center is located on the horizontal axis 90 mm from the center of rotation. For the movement simulation we let this ellipsoid

---

[1] Philips Research Laboratories, Division Technical Systems, Röntgenstraße 24-26, 22335 Hamburg, Germany

[2] Philips Medical Systems, Veenpluis 4-6, 5680 DA Best, The Netherlands

oscillate horizontally with an amplitude of 10 mm and a frequency of 2.23 Hz.

C. Abdomen phantom, bladder movement

In a third simulation we studied changes of the water level in the bladder. For this we approximated the abdomen by a cylinder with radius 240 mm consisting of water. The bladder was approximated by an ellipsoid with horizontal radius 90 mm, vertical radius 30 mm and longitudinal radius 30 mm. It was placed on the horizontal axis with its center located 130 mm from the center of rotation. The bladder is surrounded by an outer shell consisting of muscle with a thickness of 10 mm. Its interior consists of water in the lower half and of air in the upper one. For the simulation of the movement we made the water level oscillate vertically with an amplitude of 5 mm and a frequency of 1.75 Hz.

IV. Reconstruction algorithms

For our analysis we compared the results of three different helical reconstruction algorithms with classical circular reconstruction. In the following we give a short summary of the reconstruction algorithms used. Each algorithm is an approximative algorithm and of the filtered back-projection (FBP) type.

A. The PI-Method

The PI-Method [1] is based on the so-called PI-sufficiency condition, which requires that each object point gets illuminated over an angular range of 180 degrees. The pitch mentioned above is chosen such that the acquisition fulfills this requirement.

For the reconstruction the first step, which has to be performed, is a parallel rebinning row by row. This yields data on a virtual planar detector containing the rotation axis. In the next steps the rebinned data are weighted, filtered row wise and back-projected.

B. The Three-PI-Method

The n-PI-Methods [2] are a generalization of the PI-Method. They provide the possibility to choose a smaller pitch and to use the resulting redundant data. In particular the algorithm necessitates that each voxel gets illuminated over an angular range of $n \times 180$ degrees. For the study presented here we have restricted ourselves to the case $n = 3$, i.e. the 3-PI-Method.

Rebinning is again the first step of the reconstruction algorithm. It yields data on a virtual rectangular detector containing the rotation axis. Weighting and ramp filtering of the data have to be performed before the back-projection step results in the reconstructed images.

C. Advanced single slice rebinning

Advanced single slice rebinning (ASSR) [3] is a reconstruction algorithm for which tilted planes are reconstructed using classical two-dimensional filtered back-projection. The system pitch can be chosen equal to the value of the PI-Method.

In particular the algorithm requires the projection data to be rebinned onto tilted virtual planes. These data have to be weighted with factors taking the tilt angle and a length correction into account. Finally 2D filtered back-projection yields image data on slices which are not orthogonal to the rotation axis. The longitudinal distance between these slices is chosen smaller than the obtainable resolution. Using an interpolation step a certain number of these planes are averaged in order to improve the image quality.

V. Results and discussion

Movement during the scanning procedure results in artifacts in the reconstructed images. A closer look onto the details can yield valuable information about the origin of the artifacts. In the following we summarize our observations.

Figs. 1–4 contain reconstruction results of the head phantom. The two images within each figure show the best and worst obtained results for each reconstruction method under consideration. Fig. 1 shows the artifacts which result from the same phantom in classical circular scanning in combination with a single-line scanner and filtered back-projection.

Figs. 5–8 contain reconstruction results of the first abdomen phantom. We restrict ourselves to images showing the worst obtained artifacts for each reconstruction method.

Comparing the images we draw the following conclusions:

1. The 3-PI-Method yields the best images. Circular filtered back-projection proves to perform better than ASSR and 1-PI. The latter two yield images of comparable quality. In order to put these observations on a more stable ground we computed the root-mean-squared (RMS) of the hounsfield values in the central parts, i.e. the water cylinder, in Figs. 1–4. The corresponding numbers for those slices which contain the worst artifacts are 3.82, 6.84, 10.3 and 12.8 for 3-PI, circular, ASSR and 1-PI, resp.
2. For each considered reconstruction algorithm we can define a relation between the first projection which contributes to one particular reconstruction plane and the position of the moving object. As it turns out the severity of the artifacts strongly depends on this position. This holds true for each algorithm considered.
3. Movement artifacts can have different shapes. In the best case they are hardly visible at all. In the worst case they can result in the non-visibility of small details or in curved streaks in the reconstructed images.
4. As mentioned above the phantoms are exaggerated concerning the movements' frequencies and amplitudes. As an additional result we observed that smaller frequencies and/or amplitudes tend to decrease the movement artifacts only slightly.

In order to understand why the artifacts are differently distinct we have to remember the algorithms' details. For ASSR and the 1-PI-Method projection data taken over an angular range of 180 degrees are back-projected. For circular filtered back-projection the angular range corresponds to 360 degrees and for the 3-PI-Method it corresponds to 540 degrees. We therefore conclude than a larger angular range provides the possibility than inconsistencies due to movement get averaged out.

In summary we observe that the shape of the artifacts varies significantly with the used reconstruction algorithms and the strength of the artifacts depends also on the method used. Before the decision if a particular algorithm is suitable for clinical scanning can be taken, the comparison with well-known protocols is necessary. For this study we used circular scanning as gold standard. The considered helical algorithms result in artifacts which are comparable to those of this gold standard. While the 3-PI-Method seems to be even less sensitive to movement, ASSR and the 1-PI-Method yield only slightly worse images than circular filtered back-projection.

REFERENCES

[1] P.E. Danielsson, P. Edholm, J. Eriksson, M. Magnusson-Seger, "Towards exact 3D-reconstruction for helical cone-beam scanning of long objects", *Proceedings of the 3D'97 conference*, Nemacolin, Pennsylvania, USA 1997, pp. 141–144.

[2] R. Proksa, Th. Köhler, M. Grass, J. Timmer, *IEEE Transactions on medical imaging*, Vol. 19, No. 9, pp. 848–863, 2000.

[3] M. Kachelrieß, S. Schaller, W.A. Kalender, *Medical Physics*, Vol. 27, No. 4, pp. 754–772, 2000.

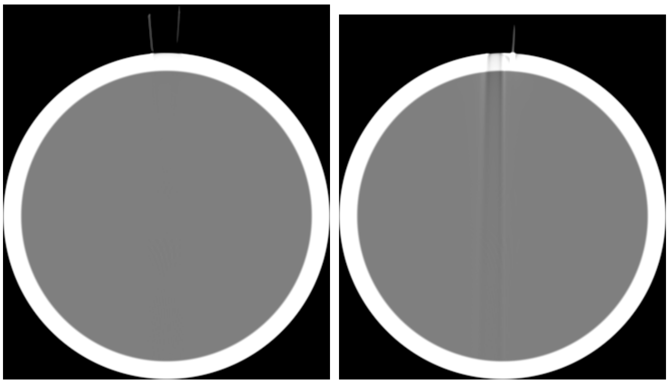


Fig. 1. The head phantom scanned in single-slice circular mode. Level: 0 HU, Window: 1000 HU.

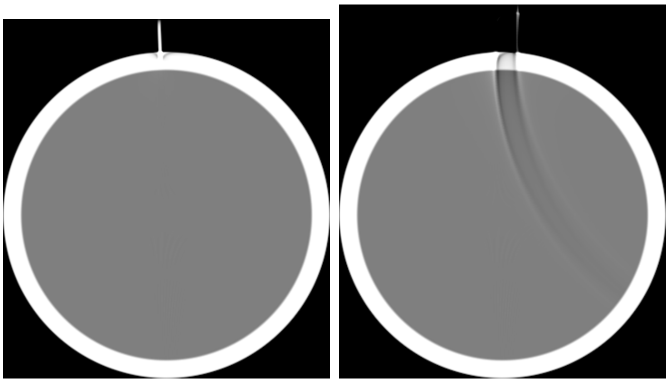


Fig. 2. Head phantom: 1-PI-Method.

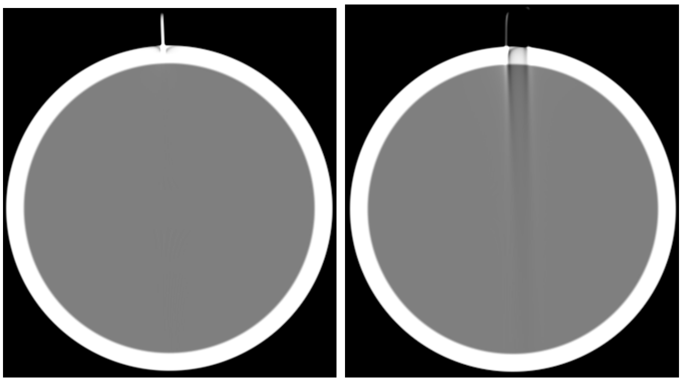


Fig. 3. Same as Fig. 2 but for the ASSR method.

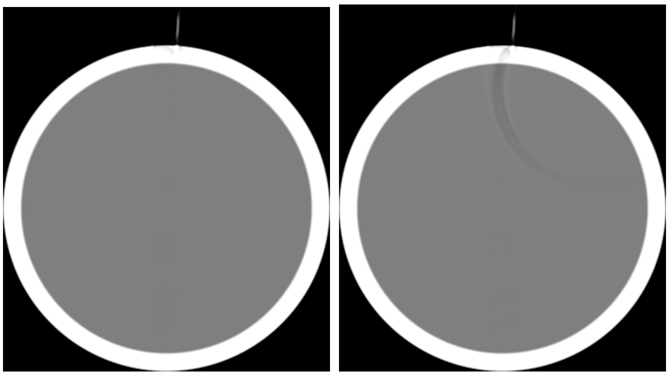


Fig. 4. Same as Fig. 2 but for the 3-PI-Method.

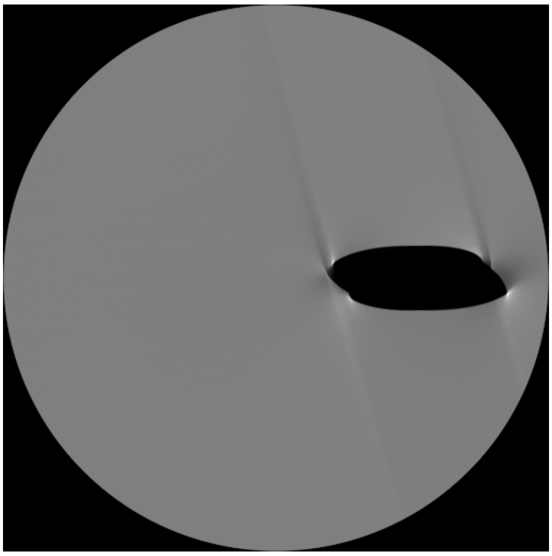


Fig. 5. Simulation of colon movement. Circular scanning, Level: 0 HU, Window: 1000 HU.

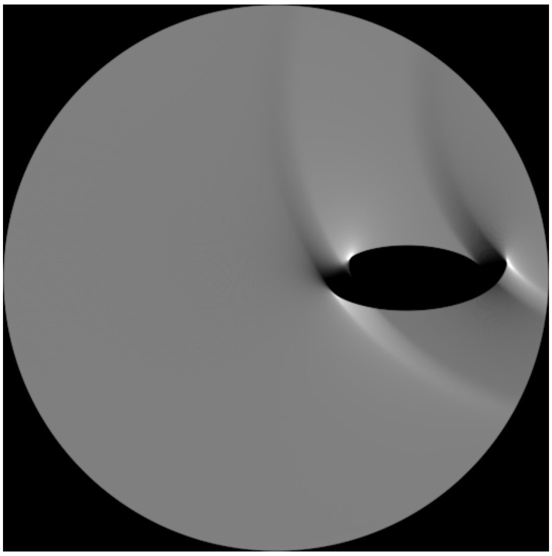


Fig. 7. Same as Fig. 5 but for the ASSR Method.

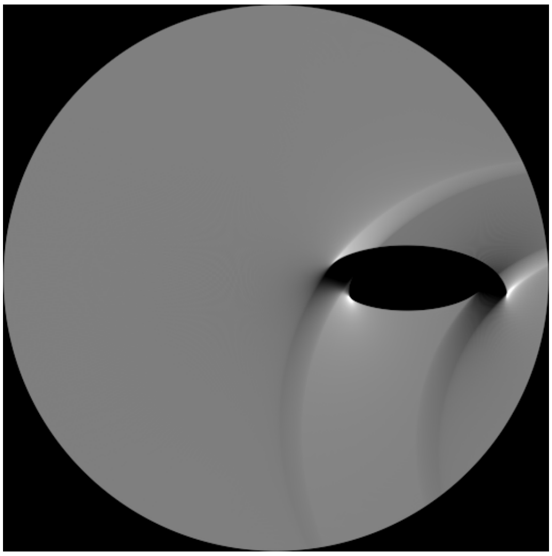


Fig. 6. Same as Fig. 5 but for the 1-PI-Method.

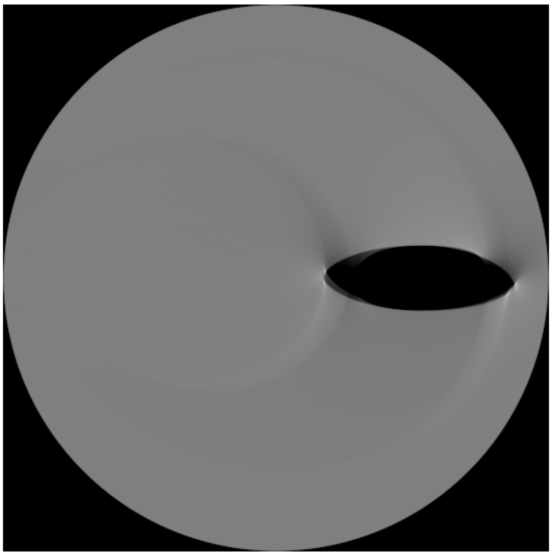


Fig. 8. Same as Fig. 5 but for the 3-PI-Method.

Longitudinal sampling and aliasing properties in multi-slice helical computed tomography

Patrick J. La Rivière and Xiaochuan Pan

Abstract— **In this work, we investigate longitudinal sampling and aliasing effects in multi-slice helical CT. We demonstrate that longitudinal aliasing can be a significant, complicated, and potentially detrimental effect in multi-slice helical CT reconstructions. Multi-slice helical CT scans are generally undersampled longitudinally for all pitches of clinical interest, and the resulting aliasing effects are spatially variant. As in the single-slice case, aliasing is shown to be negligible at the isocenter for circularly symmetric objects due to a fortuitous aliasing cancellation phenomenon. However, away from the isocenter, aliasing effects can be significant, spatially variant, and highly pitch-dependent. This implies that measures more sophisticated than isocenter slice sensitivity profiles are needed to characterize longitudinal properties of multi-slice helical CT systems adequately. Such measures are particularly important in assessing the question of whether there are preferred pitches in helical CT. Previous analyses have generally focused only on isocenter sampling patterns, and our more global analysis leads to somewhat different conclusions than have been reached before, suggesting that pitches 3, 4, 5, and 6 are favorable, and that half-integer pitches are somewhat suboptimal.**

I. INTRODUCTION

With the advent of helical systems and particularly with the advent of the multi-slice helical systems, computed tomography (CT) has become a truly volumetric imaging modality. Image acquisition now involves complicated three-dimensional sampling patterns, and volumetric visualization and analysis techniques have become essential tools for viewing and analyzing the huge amount of reconstructed data produced by each scan.

This move toward volumetric CT imaging brings with it a need for a more complete understanding of the modality's three-dimensional image quality properties. In-plane sampling and resolution properties have been extensively studied in the context of step-and-shoot CT, and most of those properties carry over directly to the helical case. In the longitudinal direction, however, several novel effects of the single-slice helical scan have been identified and studied by Yen *et al* [1, 2]. They showed both analytically and experimentally that single-slice helical CT scans are generally undersampled longitudinally by a factor of at least 2, and that the resulting aliasing effects are highly spatially variant across the field of view due to the peculiarities of the longitudinal sampling engendered by the helical scan. These spatially variant aliasing effects can alter object appearance and contrast across the field view, can make it difficult to set display windows properly, and can generally degrade image quality. Not surprisingly, in single-slice helical CT, the severity of aliasing effects increases monotonically with the helical pitch.

In this work, we investigate longitudinal sampling and aliasing effects in multi-slice helical CT. While the helical scan is expected to have a similar influence on aliasing as in the single-slice case, the longitudinal interlacing of the multiple detector rows produces very complicated sampling patterns whose effect on aliasing requires careful study. In particular, we expect the severity of aliasing effects (and the transmission of principal frequencies) to vary in complicated, non-monotonic ways as a function of helical pitch. Indeed, we hope to use this study to shed further light on the question of whether there are certain pitches—"preferred pitches"—in multi-slice helical CT that lead to inherently more favorable sampling patterns than do others.

This question has been addressed in many ways and often with a different answer. Hu [3] has argued that pitches 3 and 6 are preferred because the bands of projection-ray dependent complementary samples (discussed below) are centered between the direct samples and thus produce average sampling intervals equivalent to those in single-slice helical CT operating at pitches 1 and 2, respectively. Wang and Vannier [4] performed a "sensitivity analysis" of central detector channel sampling patterns in multi-slice helical CT and reached a nearly opposite conclusion, arguing that pitch 6 is distinctly suboptimal relative to other nearby pitches and that a pitch slightly less than 3 is to be preferred to pitch 3 itself. Both of these sampling analyses are limited in their scope, however, in that they only truly apply to the longitudinal sampling for the central detector channel ($\gamma = 0$), and thus really only predict performance at the system isocenter. Given the extreme variation in longitudinal sampling patterns between central and peripheral detector channels and the spatially variant nature of aliasing effects, we felt it important to develop a more global analysis accounting for the entirety of the longitudinal sampling that arises at each pitch.

II. METHODS

A. Theoretical analysis of sampling and aliasing

We begin with a theoretical analysis of longitudinal sampling and aliasing effects in reconstructed multi-slice helical CT volumes. To facilitate comparison with the analysis of single-slice helical CT sampling and aliasing, we adopt and, where necessary, extend the notation of Yen *et al.* [1].

In N-slice helical CT under the multiple parallel fanbeam approximation, we can regard the measured data as samples of a 3D fanbeam sinogram $p_\beta(\gamma, z)$, where β is the projection angle, γ is the angle between the projection ray in question and the central ray of the fanbeam, and z is

---

The authors are with the Department of Radiology, The University of Chicago, Chi cago IL 60637 USA.

the longitudinal position along the object being imaged. In practice, the imaging is performed with detectors of finite longitudinal extent; thus the function being sampled is more appropriately modeled as the convolution of the ideal sinogram and the detector's longitudinal response function, which we denote $a(z)$ (in this work we ignore blurring and sampling effects in the γ coordinate in order to focus on longitudinal effects). Thus we denote by

$$p'_{\beta}(\gamma, z) = p_{\beta}(\gamma, z) * a(z)$$

the 3D function whose samples we acquire. Often $a(z)$ is modeled as an ideal rectangular function of width D, where D is the longitudinal detector collimation at the isocenter.

At each of M projection angles $\beta_i = (2\pi i/M)$, $i = 0, \ldots, M-1$, and at each value of γ, the data acquired by the mth detector row in an R-revolution scan at helical pitch P_h comprises samples in z that can be written as

$$s^{(m)}_{\beta_i}(\gamma, z) = p'_{\beta_i}(\gamma, z) \sum_{k=0}^{R-1} \delta\left[z - \left(i\frac{d_{s1}}{M} + kd_{s1} + mD\right)\right], \tag{1}$$

where $d_{s1} = P_h D$. The delta function contains three terms of interest. The first term, $i\frac{d_{s1}}{M}$, reflects the overall longitudinal offset of the sequence at each projection angle; this arises due to the helical nature of the scan. The second term, kd_{s1}, is the actual sampling step, as k is the summation index and d_{s1} is the longitudinal interval between samples at (β_i, γ) and $(\beta_i + 2\pi, \gamma)$. The third term, mD, reflects the overall longitudinal offset of the sequence for the mth detector.

Given these sequences of longitudinal samples at each β_i and γ for each detector row, it is common practice to interpolate a complete sinogram $p^{(\mathrm{int})}_{\beta_i}(\gamma, z)$ at a fixed value of z so that reconstruction can proceed by use of conventional fanbeam reconstruction algorithms. In addition to making use of the *direct* samples of Eq. 1 for interpolation, it is generally advantageous to exploit the redundancy of fanbeam data acquired over 2π to augment these samples with the *complementary* samples $s^{(m)}_{\beta_i+\pi+2\gamma}(-\gamma, z)$ that correspond geometrically to the same projection ray as $s^{(m)}_{\beta_i}(\gamma, z)$, only shifted longitudinally by $\left(\frac{\pi+2\gamma}{2\pi}\right) d_{s1}$. The most straightforward interpolation approach, 180MLI, makes use of linear interpolation among these direct and complementary samples, and it can be expressed as

$$\begin{aligned} p^{(\mathrm{int})}_{\beta_i}(\gamma, z) &= \sum_{m=0}^{N-1} \Big[s^{(m)}_{\beta_i}(\gamma, z) *_z h^{(m)}_d(\gamma, z) \\ &+ s^{(m)}_{\beta_i+\pi+2\gamma}(-\gamma, z) *_z h^{(m)}_c(\gamma, z) \Big] \end{aligned} \tag{2}$$

where $*_z$ denotes a 1D convolution in the variable z, and $h^{(m)}_d(\gamma, z)$ and $h^{(m)}_c(\gamma, z)$ are the linear interpolation kernels applied to the direct and complementary samples, respectively, of the mth detector row. In the z variable, these kernels have the form of an asymmetric triangle function. The half-widths of these asymmetric triangles can be computed empirically from the sampling patterns that arise when the direct and complementary samples for all rows are interlaced appropriately at each γ.

Given a stationary fanbeam sinogram $p^{(\mathrm{int})}_{\beta_i}(\gamma, z)$, the penultimate step of the reconstruction process is to filter it in the γ direction,

$$q_{\beta_i}(\gamma, z) = \left[D_0 \cos(\gamma) p^{(\mathrm{int})}_{\beta_i}(\gamma, z)\right] *_{\gamma} g(\gamma), \tag{3}$$

where D_0 is the fanbeam focal length, and $g(\gamma)$ is the fanbeam reconstruction filter. Reconstruction of an image $f(r, \phi, z)$, expressed here in polar coordinates, is then achieved by use of

$$f(r, \phi, z) \approx \Delta\beta \sum_{i=0}^{M-1} \frac{1}{L^2(r, \phi, \beta_i)} q_{\beta_i}(\gamma', z), \tag{4}$$

where $L(r, \phi, \beta_i)$ and γ' are known functions of r, ϕ, and β_i defined in [1, 5].

Equation 4 can be viewed as an expression for a continuous 3D volume $f(r, \phi, z)$ reconstructed from the measured, sampled data $s^{(m)}_{\beta_i}(\gamma, z)$. We are interested in the effect that the longitudinal sampling during the measurement process has on the longitudinal properties of the reconstructed volume. Thus we regard $f(r, \phi, z)$, for fixed (r, ϕ) as a continuous 1D function of z, which we denote $f_{(r,\phi)}(z)$, and will examine the spectrum of the resulting profiles. In practice, of course, we do not reconstruct volumes that are continuous in the z coordinate, but rather that comprise a finite set of slices spaced by a reconstruction interval d_{s2}. However, because we are interested in the effect of the acquisition sampling, we can safely disregard this resampling step.

Computing the Fourier transform with respect to z of Eq. 2, as in [1], yields

$$\begin{aligned} \mathcal{F}_z\left\{f_{(r,\phi)}(z)\right\} = \Delta\beta D_0 R \sum_{i=0}^{M-1} \Bigg\{ \frac{1}{L^2(r, \phi, \beta_i)} \Big[\cos(\gamma) \\ \times \sum_{k=-\infty}^{\infty} P'_{\beta_i}\left(\gamma, f_z - \frac{k}{d_{s1}}\right) e^{-j2\pi i \frac{k}{M}} H'_k(\gamma, f_z)\Big] *_{\gamma} g(\gamma) \Bigg\} \Bigg|_{\gamma'} \end{aligned} \tag{5}$$

where

$$\begin{aligned} H'_k(\gamma, f_z) &= \sum_{m=0}^{N-1} e^{-j2\pi k \frac{m}{P_h}} \Big[H^{(m)}_d(\gamma, f_z) \\ &+ H^{(m)}_c(-\gamma, f_z) e^{-jk(\pi+2\gamma)} \Big], \end{aligned} \tag{6}$$

and where $H^{(m)}_d(\gamma, f_z)$, $H^{(m)}_c(\gamma, f_z)$, and $P'_{\beta_i}(\gamma, f_z)$ are the Fourier transforms with respect to z of $h^{(m)}_d(\gamma, z)$ and $h^{(m)}_c(\gamma, z)$, and $p'_{\beta_i}(\gamma, z)$, respectively.

Equation 5 allows us to draw some conclusions about potential aliasing effects. It can be argued that the helical CT data is at least approximately bandlimited longitudinally by the first zero of the rectangular detector response function, which occurs at $f_z = 1/D$ [1]. Satisfying the Nyquist sampling condition would thus require a sampling interval $d_{s1} = D/2$. From Eq. 5, it can be seen that the fundamental sampling interval, which determines the spacing of the spectral replications $P'_{\beta_i}\left(\gamma, f_z - \frac{k}{d_{s1}}\right)$ is $d_{s1} = P_h D$, and thus the Nyquist condition is not satisfied for reasonable values of P_h. In general, then, we might expect substantial aliasing effects characteristic of sampling at interval $P_h D$. However, the phase factors in Eqs. 5 and 6 can lead to partial or complete cancellation of aliasing in some circumstances.

The first such case arises when the helical pitch P_h is an integer less than or equal to the number of detector rows N. In this case, the $H^{(m)}(\gamma, f_z)$ are all of the form $w_m H(\gamma, f_z)$, where the w_m are normalization weights chosen so that the w_m for detector rows that follow the same trajectories sum to 1. Given this, it is possible to show that the only remaining m dependence in Eq. 6 is contained in a term of the form:

$$\sum_{m=0}^{P_h-1} e^{-j2\pi k \frac{m}{P_h}} = \begin{cases} P_h, & \text{for } k = 0, \pm P_h, \pm 2P_h, \ldots \\ 0, & \text{otherwise} \end{cases}$$

In this case, it is possible to replace all the k in Eq. 5 by $P_h k$, and thus the effective sampling interval becomes $d_{s1}/P_h = D$. This interval still does not satisfy the Nyquist condition, but the nature and severity of the aliasing effects will be substantially altered relative to the general, non-integer pitch case.

The second situation in which we might expect outright aliasing cancellation is at the system isocenter when imaging circularly symmetric objects, a phenomenon demonstrated by Yen *et al.* for single-slice helical CT. In this situation, it is possible to manipulate Eq. 5 to isolate a factor of the form

$$\sum_{i=0}^{M-1} e^{-j2\pi i \frac{k}{M}} = \begin{cases} M, & \text{for } k = 0, \pm M, \pm 2M, \ldots \\ 0, & \text{otherwise,} \end{cases} \tag{7}$$

and thus the effective isocenter sampling interval becomes d_{s1}/M, a reduction by the number of projection angles per revolution!

B. Numerical evaluation of analytic results

B.1 Spatial distribution of aliasing effects

In order to study the spatial distribution of aliasing effects, we evaluated Eq. 5 for a circularly symmetric, cylindrical object with radius R_c and with sinusoidal longitudinal variation of frequency f_0. That is, we assumed

$$p_\beta(\gamma, z) = p_\gamma(\gamma)\cos(2\pi f_0 z),$$

where

$$p_\gamma(\gamma) = \begin{cases} \sqrt{R_c^2 - D_0^2 \sin^2(\gamma)}, & \gamma \leq \sin^{-1}(R_c/D_0) \\ 0, & \gamma > \sin^{-1}(R_c/D_0) \end{cases} .$$

Thus

$$P_\beta(\gamma, f_z) = p_\gamma(\gamma)\left\{\frac{1}{2}\left[\delta(f_z - f_0) + \delta(f_z + f_0)\right]\right\}.$$

This object is similar to the cylindrical, square-wave phantom employed by Yen *et al.* [1], and it allows us to isolate and quantify aliasing effects easily. An ideal reconstruction would have energy only at the frequencies $f_z = \pm f_0$. However, due to aliasing effects we actually expect energy at $f_z = \pm f_0 + \frac{k}{d_{s1}}$, for $k \in \mathbb{Z}$. For one such frequency of interest f_z we evaluate the factor in square brackets in Eq. 5 by computing the sum over the appropriate values of k (in most cases, only one k contributes to each f_z of interest). We evaluate $H^{(m)}(\gamma, f_z)$ by computing the asymmetric triangles $h^{(m)}(\gamma, z)$ on a discrete (γ, z) grid, heavily zero-padding in z, and then taking a discrete Fourier transform to obtain estimates of $H^{(m)}(\gamma, f_z)$ on a discrete (γ, f_z) grid. The quantity in square brackets can be expressed as a 2D, complex-valued array having dimensions of a sinogram, and thus $\mathcal{F}_z\left\{f_{(r,\phi)}(z)\right\}$ can be computed on a Cartesian grid simply by applying a fanbeam filtered backprojection routine to the real and imaginary components of this sinogram-like quantity. The magnitude of the resulting complex-valued image gives the magnitude of the longitudinal spectrum at frequency f_z at each transverse point in the reconstructed volume.

We evaluated such images for a 4-row system at a number of pitches, using a cylindrical phantom of radius 230 mm and fundamental frequency $f_0 = 1/3.175$ mm^{-1}. We modeled the detector response as rectangular, with longitudinal collimation 2.5 mm at the isocenter. We used 128 projection angles and 128 equiangular projection rays, spaced by 4 mm at the isocenter, and a focal length of 540 mm. Typical results are given in Sec. III-A.

B.2 Contrast to aliased noise as a function of pitch

In addition to examining the spatial distribution of aliasing effects, we also sought to quantify the severity of aliasing effects as a function of helical pitch, in an effort to shed some light on the issue of whether there are preferred pitches in multi-slice helical CT. To do so, we adopted the point of view of Park *et al.* [6], in which aliasing is regarded as structured, signal-dependent noise that can interfere with the detection of a signal. We then computed a *contrast-to-aliased-noise ratio* (CN_aR) for reconstructions of the object discussed in Sec. II-B, using the geometry discussed there, for pitches ranging from 1.0 to 8.0 in increments of 0.1. Yen *et al.* also used a CN_aR figure of merit in their consideration of single-slice helical CT [2]. However, in that case, the main frequency produces only a single low-frequency aliased peak of interest, and so they define CN_aR for a single longitudinal profile simply in terms of the amplitudes of the reconstructed main and aliased spectral peaks. In the multi-slice case, there will, in general, be numerous aliased peaks at low frequencies. We are also interested in a more global measure of aliasing content in an entire reconstructed volume. Thus we define

$$CN_aR(P_h) = 20 * \log_{10} \frac{\sqrt{\sum_i \sum_j |F_{P_h}(x_i, y_j, f_0)|^2}}{\sqrt{\sum_{f_a} \sum_i \sum_j |F_{P_h}(x_i, y_j, f_a)|^2}},$$

where $F_{P_h}(x_i, y_j f_z)$ is the spectral magnitude image for frequency f_z on a Cartesian grid. The i and j sums run over pixels less of distance less than R_c from the image center and the f_a sum runs over aliased frequencies between 0 and $2/D$. Note that CN_aR is expressed in decibels.

III. Results

A. Spatial variance of aliasing effects

Figure 1 depicts the magnitude $|F_{P_h}(x_i, y_j, f_z)|$ for $i = 64, \ldots 127$, $j = 64$, and three values of f_z: $f_z = f_0$, $f_z = f_{a1} = \left(-f_0 + \frac{1}{D}\right)$, and $f_z = f_{a2} = \left(f_0 + \frac{1}{D}\right)$. These are the only three values of $f_z \in \left[0, \frac{2}{D}\right)$ for which the spectral magnitude is non-negligible. This fact alone confirms the finding in Sec. II-A that the effective sampling interval is D and not DP_h for integer pitches less than or equal N. Otherwise, we would have expected to find non-negligible spectral magnitude at, for instance, $f_z = \left(-f_0 + \frac{1}{3D}\right)$. The figure itself confirms the spatially variant nature of the aliasing effects, which increase in magnitude from zero at the isocenter to a maximum near the periphery of the phantom.

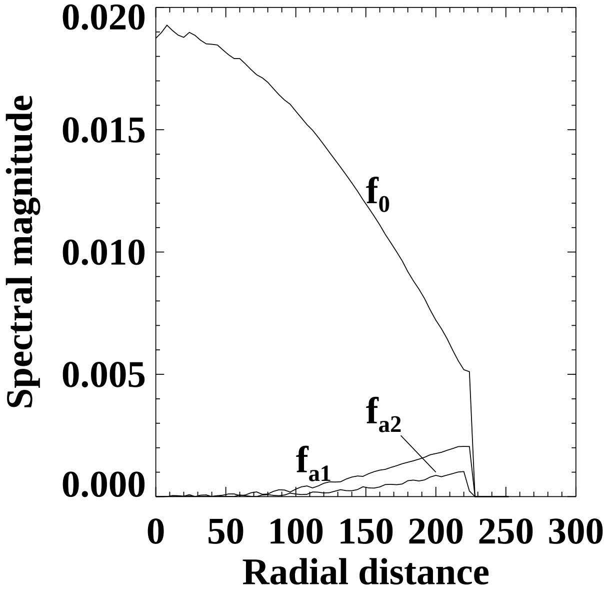


Fig. 1. Spatial dependence of longitudinal spectral magnitude for three frequencies of interest along a line from the isocenter to the edge of the phantom.

B. Contrast to aliased noise as a function of pitch

Figure 2 plots the calculated CN_aR versus helical pitch; higher values of CN_aR are better. As expected, the curve is far from monotonic. Local maxima are evident around pitches 3, 4, 5, and 6. Some complicated variation is evident between pitches 1 and 2, with pitch 2 being a clear minimum in the curve. Other minima occur at half-integer pitches.

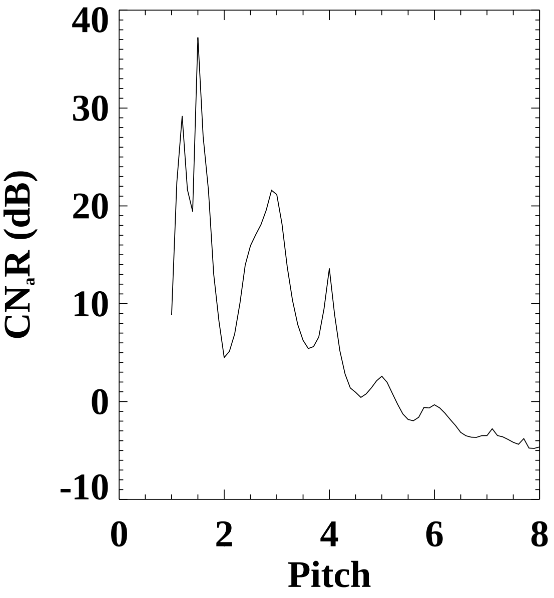


Fig. 2. CN_aR, in decibels, versus helical pitch for a four-slice scanner with longitudinal collimation width 2.5 mm at the isocenter imaging an object with sinusoidal longitudinal variation of frequency $f_0 = 1/3.175$ mm^{-1}. Higher values of CN_aR are better.

An intuitive explanation of these results can be obtained by considering the longitudinal sampling patterns that arise at various pitches and at various values of γ. For example, at pitch 3, the direct samples of the four detector rows interlace to form a uniform sampling pattern with interval D. At $\gamma = 0$, the complementary data reside midway between these direct samples, and thus provides effective sampling interval of $D/2$, which allows relatively narrow interpolation kernels to be applied. At larger values of γ, the complementary data approaches but do not, in general, cross over the direct samples, and thus do continue to provide some measure of stable aliasing suppression.

Perhaps the most surprising aspect of Fig. 2 is the peak at pitch 4, which is generally considered a poor choice because of its unfavorable isocenter SSP performance. As at pitch 3, the direct samples of the four detector rows interlace to form a uniform sampling pattern with interval D. However, now at $\gamma = 0$, the complementary data are coincident with these direct samples, and relatively broad interpolation kernels must be applied; this explains the poor isocenter SSP performance. However, for larger values of γ, which contribute most to areas of the reconstructed volume where aliasing is likely to be problematic, the complementary data reside nearly midway between the direct samples and thus provide a measure of aliasing suppression where it is needed most.

IV. DISCUSSION AND CONCLUSIONS

We have demonstrated that longitudinal aliasing can be a significant, complicated, and potentially detrimental effect in multi-slice helical CT reconstructions. Multi-slice helical CT scans are generally undersampled for all pitches of clinical interest, and the resulting aliasing effects are spatially variant.

As in the single-slice case, aliasing is negligible at the isocenter for circularly symmetric objects due to a fortuitous aliasing cancellation phenomenon. The effective oversampling at the isocenter of such objects explains why it is even possible to obtain reasonable looking slice sensitivity profiles (SSP) and longitudinal modulation transfer functions of an ostensibly undersampled system: these measurements are always performed at the isocenter by use of a circularly symmetric phantom. However, this phenomenon in no way licenses the use of these measures to characterize longitudinal resolution properties of multi-slice helical CT systems. Away from the isocenter, aliasing effects can be significant, spatially variant, and highly pitch dependent. More sophisticated measures of longitudinal properties are needed to characterize multi-slice helical CT systems adequately.

Such measures are particularly important in assessing the question of whether there are preferred pitches in helical CT. Previous analyses have generally focused only on isocenter sampling patterns, and thus predict isocenter SSP performance, but neglect aliasing effects away from the isocenter. Our more global analysis suggests that the much maligned pitch 4, despite its poor isocenter SSP performance, actually yields a very favorable global CN_aR because of favorably uniform longitudinal sampling for its outer detector channels. The analysis did confirm the advantages of pitches 3 and 6, and also suggested that pitch 5 would be favorable. Half-integer pitches appeared to be poor performers by this measure.

The analysis performed above implicitly blends sampling and interpolation effects. In this work, we examined the use of straightforward linear interpolation. In practice, straightforward linear interpolation is rarely used because frequent "changeovers" in the pairs of detector rows contributing to a given slice tend to produce artifacts in reconstructed images. In general, either attention is restricted to pitches where linear or quasi-linear approaches can be applied safely [3] or broader, adaptive z-filtering interpolation approaches are employed [7,8]. We do not expect that the use of a different interpolation approach would fundamentally alter the conclusions of this study. The presence of longitudinal aliasing effects and their spatial distribution are effectively inherent properties of the helical scan geometry. Altering the interpolation approach might alter the transmission of principal, unaliased frequencies somewhat, but would not be expected to affect aliasing effects profoundly. Nonetheless, we do intend to extend our analysis to these alternative approaches, as their longitudinal properties have only been characterized through use of SSP measurements, whose limitations should be evident from the preceding discussions.

One effect that was not considered in great detail in this summary is the small cone angle that arises in multi-slice helical CT. This cone angle has generally been ignored in deriving reconstruction algorithms and sampling analyses by making the multiple parallel fanbeam assumption made here, *i.e.* by assuming that the multi-row data comprises multiple, parallel transverse projections of the object, not multiple, differently oblique ones. This approximation is regarded as reasonably sound for systems of 4 or fewer rows, although we have shown elsewhere that the small cone beam angle does influence longitudinal sampling and aliasing properties by introducing inconsistencies among the data measured by the different detector rows [9]. We felt justified in ignoring the effect in the present summary because we felt that we could contribute a novel, thorough analysis of sampling in multi-slice helical CT under the approximation that has been the foundation of all previous analyses. The development of a more complete analysis that does account for the cone angle effect is the subject of ongoing work.

REFERENCES

[1] S. Y. Yen, C. H. Yan, G. D. Rubin, and S. Napel, "Longitudinal sampling and aliasing in spiral CT," *IEEE Trans. Med. Imag.*, vol. 18, pp. 43–58, 1999.

[2] S. Y. Yen, G. D. Rubin, and S. Napel, "Spatially varying longitudinal aliasing and resolution in spiral computed tomography," *Med. Phys.*, vol. 26, pp. 2617–2625, 1999.

[3] H. Hu, "Multi-slice helical CT," *Med. Phys.*, vol. 26, pp. 5–18, 1999.

[4] G. Wang and M. W. Vannier, "The effect of pitch in multislice spiral/helical CT," *Med. Phys.*, vol. 26, pp. 2648–2653, 1999.

[5] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, Academic Press, New York, 1982.

[6] S. K. Park and R. Hazra, "Image restoration versus aliased noise enhancement," in *Proc. SPIE*, 1994, vol. 2239, pp. 52–62.

[7] K. Taguchi and H. Aradate, "Algorithm for image reconstruction in multi-slice helical CT," *Med. Phys.*, vol. 25, pp. 550–560, 1998.

[8] S. Schaller, T. Flohr, K. Klingenbeck, J. Krause, T. Fuchs, and W. Kalendar, "Spiral interpolation algorithms for multi-slice spiral CT part I: Theory," *IEEE Trans. Med. Imag.*, vol. 19, pp. 822–834, 2000.

[9] P. J. La Rivière and X. Pan, "Longitudinal sampling and aliasing in multi-slice helical CT," in *Proc. 2000 IEEE Nucl. Sci. Symp. and Med. Imag. Conf.*, 2000, To appear.

Aspects of 3-D Imaging by Classical Tomography for Dual Detector PEM

M. F. Smith[1], S. Majewski[1], A. G. Weisenberger[1], D. Kieper[1], J. D. Kalen[2], P. P. Fatouros[2]

[1]Thomas Jefferson National Accelerator Facility, Newport News, VA 23606
[2]Department of Radiology, Virginia Commonwealth University Health System, Richmond, VA 23298

ABSTRACT

Images from dual detector positron emission mammography (PEM) systems are commonly reconstructed by backprojection methods of classical tomography. Characteristics of 3–D PEM images were investigated using analytic models and computer simulations, in particular depth resolution and the quality of images in the third dimension normal to the detectors. These modeling tools provide insight into the depth blurring observed in 3-D images from experimental line source and breast phantom studies acquired with detectors built using pixellated arrays of LGSO crystals. Approaches to improved 3-D breast imaging are discussed.

I. INTRODUCTION

Positron emission mammography (PEM) with F-18 fluorodeoxyglucose (FDG) has the capability to image regions of increased metabolic activity in the breast, which may prove useful in the detection, staging and treatment of breast cancer. Several groups, including our own, have built or are developing dedicated PEM imaging systems. System designs include parallel planar detectors [1-3], curved plate detectors [4] and small ring configurations [5, 6]. Clinical breast imaging with FDG is an area of active research [7-9].

With dual detector PEM the breast is imaged between two static parallel planar detectors. There is incomplete angular sampling in planes perpendicular to the detectors. Image reconstruction is commonly performed in planes parallel to the detectors using the backprojection method of classical tomography [10]. In a simulation study 2-D iterative reconstruction in perpendicular planes was implemented after single slice rebinning [5]. For both methods there is blurring between the detectors due to incomplete angular sampling.

The goal of this work is to provide a better understanding of some aspects of classical tomography reconstruction performance for 3-D PEM image formation. In particular, depth resolution and the quality of images normal to the detectors will be investigated. Analytic and computer modeling will be used and results from physical experiments will be presented.

II. ANALYTIC MODEL

In this section a model of the angle-dependent point source sensitivity will be used to estimate the intensity of a point source in different reconstructed image planes. These values will be used to estimate resolution normal to the detectors.

A. Point Source Sensitivity

Consider a simple PEM detector model where the two planar detector heads are parallel and separated by a distance D (Figure 1). Assume that each detector has an efficiency s for detection of an annihilation photon and that photons are detected at the front surface of the detectors.

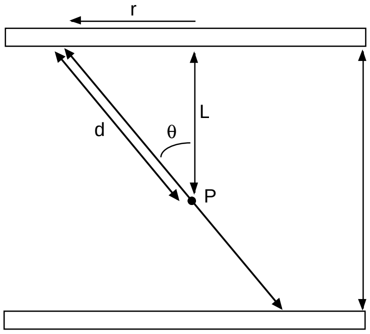


Figure 1. Diagram of the geometry of the point source between the two PEM detectors.

Now suppose that a point source of activity A (positron emissions/sec) is located at position P that is a distance L from one of the detector heads. The photon flux density from annihilation photons (counts/sec/unit area) on the surface of that detector is

$$f(r) = (A/(2\pi d^2))\cos\theta = (A/2\pi)(L/(L^2+r^2)^{3/2}) \quad (1)$$

where r is the distance from the orthogonal projection of point P onto the detector surface and θ is the angle from normal incidence. Equation (1) can be integrated for the photon flux within an acceptance angle $\theta_{\max}$ from normal incidence,

$$F(\theta_{\max}) = A(1-\cos\theta_{\max}) \quad (2)$$

The total flux will be the same on the opposing detector.

The geometric efficiency for detection of annihilation events is $1-\cos\theta_{\max}$ and the total sensitivity is $s^2(1-\cos\theta_{\max})$. Use of a larger acceptance angle in image reconstruction will increase sensitivity, which should decrease statistical noise in the resulting images and enable better detection of smaller and lower contrast tumors. Increasing the acceptance angle will degrade image uniformity, however [3].

B. Depth Resolution for an Ideal Detector and Circular Pixels

Reconstructed images are formed in planes parallel to the detectors by backprojecting counts along the lines of response connecting the detection locations of the annihilation photons. The resulting images are the sum of a focused image of activity in the desired plane and blurred images of activity in other planes. The point spread function normal to the detectors can be derived for the case of an ideal detector with infinitely small pixels and backprojection into circular image pixels.

In the image plane containing the point source, all events within the axial acceptance angle will contribute to the pixel

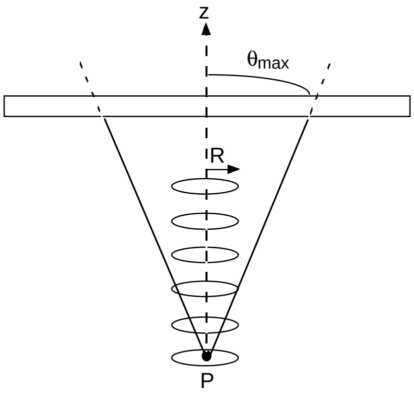


Figure 2. Diagram of circular pixels of radius R in different image planes.

containing the source point (Figure 2). Now consider a circular pixel of radius R that is located a distance z from the point source along a line perpendicular to the detector face. The maximum angle from the normal of annihilation photons originating at point P that pass through this circle is $\theta_z = \arctan(R/z)$. If the maximum acceptance angle for image reconstruction is $\theta_{\max}$, then the photon flux from point P contributing to reconstructed counts in this pixel is

$$F(\theta_z) = A\,(1 - \cos(\min\{\theta_z, \theta_{\max}\})) \qquad (3)$$

The normalized point spread function (PSF) perpendicular to the detector follows from equations (2) and (3) as

$$\begin{aligned} PSF(z) &= (1 - \cos(\min\{\theta_z, \theta_{\max}\}))/(1 - \cos\theta_{\max}) \\ &= \min\{(1 - 1/[1 + (R/z)^2]^{1/2})/(1 - \cos\theta_{\max}), 1\} \end{aligned} \qquad (4)$$

A measure of the imaging resolution perpendicular to the detector face is given by the full-width at half-maximum (FWHM) of the PSF. From equation (4) this is

$$FWHM(\theta_{\max}) = R(1 + \cos\theta_{\max})/[1 - (1/4)(1 + \cos\theta_{\max})^2]^{1/2} \qquad (5)$$

For the discrete case where the image of a point source is reconstructed into square pixels that are the same size as the detector pixels, the image counts are usually spread among a few pixels in the focal plane. The maximum pixel value is a factor $\beta < 1$ times the total image counts and is dependent on the location of the point source with respect to the discretization grid. For this more general case an expression for the PSF normal to the detectors has not yet been derived, however we have found empirically that the FWHM of small point sources often can be modeled by

$$\begin{aligned} FWHM(\theta_{\max}) = 2R[1 - (\beta/2)(1 - \cos\theta_{\max})]/ \\ \{1 - [1 - (\beta/2)(1 - \cos\theta_{\max})]^2\}^{1/2} \end{aligned} \qquad (6)$$

which reduces to equation (5) when β=1.

As an example, consider the case of circular pixels 3.39 mm in diameter (area 9 mm^2). The dependence of the PSFs on maximum acceptance angle is shown in Figure 3. The resolution normal to the detectors improves when the acceptance angle increases because the point sources are increasingly blurred in non-focal planes (Figure 4).

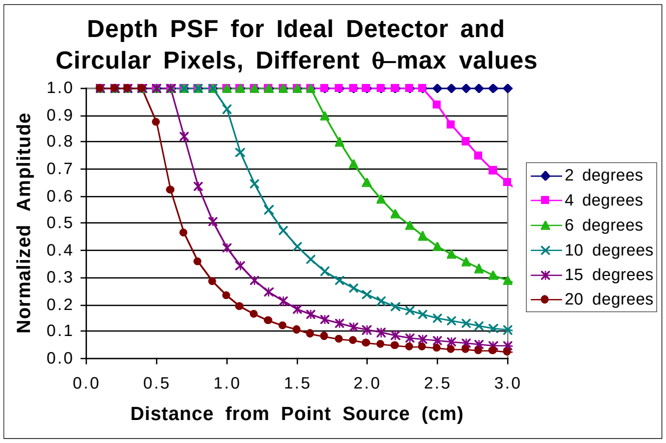


Figure 3. Point spread functions normal to the detector for different maximum acceptance angles ($\theta_{\max}$) of coincidence events (equation (4)). An ideal detector is modeled and the circular image pixels are 3.4 mm in diameter.

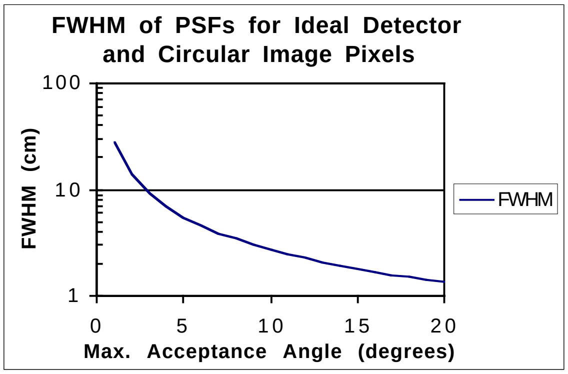


Figure 4. Full-width at half-maximum of the PSFs normal to the detector as a function of the maximum acceptance angle of coincidence events contributing to the backprojection images (equation (5)). An ideal detector is modeled and the circular image pixels are 3.4 mm in diameter.

III. NUMERICAL SIMULATIONS AND PHYSICAL EXPERIMENTS

A. Point Source

Point source acquisitions were simulated for a PEM system with pixellated detectors. The computer program for these simulations employed ray-tracing and numerical integration methods. Photons were assumed to be detected at the surface of the detectors. The detector heads were modeled as 29 x 29 arrays of 3.3 mm pixels, the same geometry as PEM detectors we have built [3]. The detector separation was 18 cm.

Image reconstruction was by classical tomography with a pixel size of 3.3 mm and image plane spacing of 3.3 mm. Images were reconstructed with acceptance angles of 5, 10, 15 and 20 degrees. Depth resolution improves with increasing acceptance angle, though there is considerable blurring in the images (Figure 5) and in vertical profiles through the point sources (Figure 6). The FWHM of the profiles are 10.6, 5.2, 3.6 and 2.7 cm for 5, 10, 15 and 20 degree acceptance angles, respectively. The widths predicted by equation (6) with β = 0.29 are 11.2, 5.6, 3.7 and 2.8 cm, respectively.

B. Line Source

Coincidence data were acquired with a PEM system built with a pixellated detector array of LGSO crystals (Hitachi, Inc.). Each detector head was 10 cm x 10 cm with a 29 x 29 crystal array; the individual crystals were 3 mm x 3 mm x 10 mm and the crystal pitch was 3.3 mm. The crystals were coupled to an array of 4 x 4 Hamamatsu R7600-00-C8 position sensitive photomultiplier tubes. The rest of the detector design has been described elsewhere [3].

Three line sources were filled with F-18 and placed midway between the detector heads, which were separated by 18 cm. Images were reconstructed in the central 21 slices with a spacing of 3.3 mm using 10 and 20 degree acceptance angles. Blurring normal to the detectors is more severe for the smaller acceptance angle (Figure 7). Depth resolution differs for the line sources because of the more limited angular range of lines of response near the edge of the detector.

C. Breast Phantom with Tumors

A 6 cm thick box phantom simulating a compressed breast was filled with 33 nCi/cc F-18 and imaged for 10 min with the previously described PEM system. The distance between the detector heads was 7.5 cm. Simulated tumors 12 mm, 10 mm, 8 mm and 4 mm diameter were filled with activity in a 10:1 tumor:background activity concentration ratio. Images were reconstructed for 21 slices with a 3.3 mm spacing using a coincidence acceptance angle of 20 degrees. The three larger tumors are visible in their focal plane, but there is appreciable blurring normal to the detectors (Figure 8).

IV. DISCUSSION

The analyses and simulations of this paper were simple ones and they represent an initial effort toward quantifying depth-dependent blurring for PEM. This is an important issue since classical tomography is widely used for dual detector PEM image reconstruction. Depth-dependent blurring as well as sensitivity, image uniformity and image noise are all affected by the maximum acceptance angle chosen for reconstruction.

The analytic expression for the PSF normal to the detector could be improved by generalizing it to model discretization of the detectors and of the backprojected images. The computer simulations of the PEM system could be made more realistic by better modeling photon interactions in the source region and detector.

Image reconstruction by classical tomography has advantages and disadvantages for dual detector PEM. One advantage is high sensitivity since all of the coincidence data within a given acceptance angle can be used in image reconstruction, reducing statistical noise. It is fast enough that image reconstruction can be implemented in real-time as part of data acquisition [2].

The major disadvantage of classical tomography is that activity from neighboring planes is blurred into the image plane, not removed as for computed tomography. It would be interesting to investigate whether low contrast lesions can be detected or localized better in classical tomography images or in limited angle computed tomography images reconstructed from the same dataset, particularly for compressed breasts. A broader question is whether there are better methods for extracting 3-D information from coincidence data acquired with dual detector PEM.

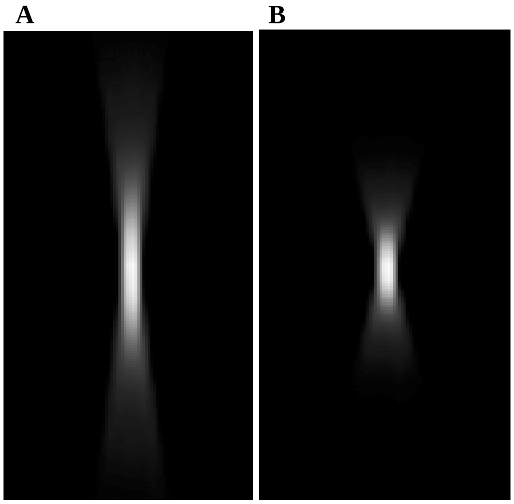


Figure 5. Images normal to the detector face of reconstructions of a point source for (a) 10 degree and (b) 20 degree acceptance angles. The detectors would be positioned at the top and bottom of each image.

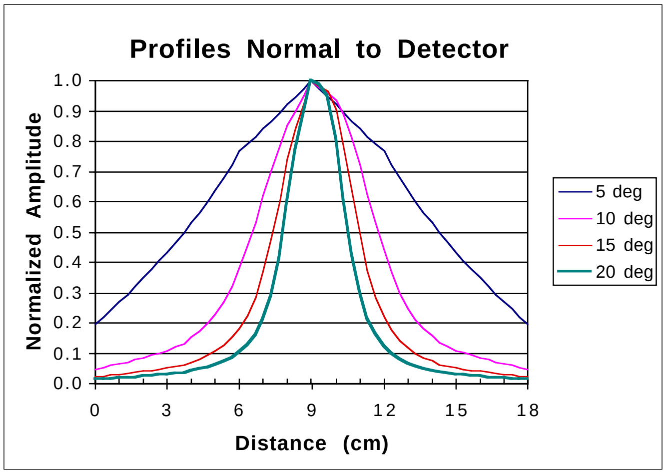
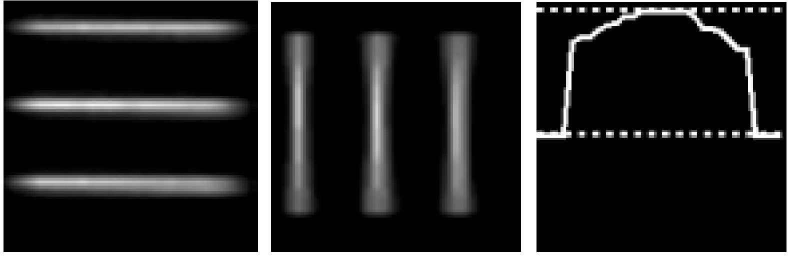
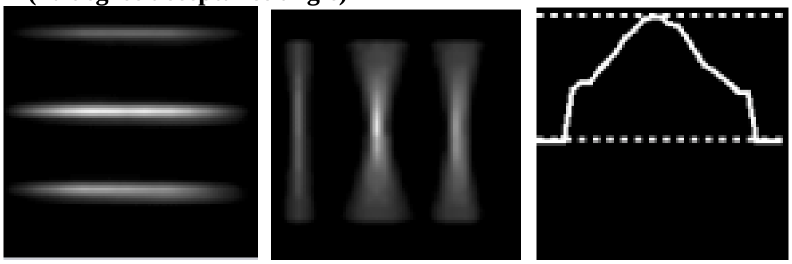


Figure 6. Profiles normal to the detectors for a point source simulation and image reconstruction with different acceptance angles. The images corresponding to the 10 and 20 degree curves are shown in Figure 5.

A more general challenge for PEM is the optimization of detector design (e.g. planar detectors, curved detectors, ring detectors), detector motion (static, few or many detector positions), image reconstruction method (classical tomography, computed tomography, iterative reconstruction), breast positioning (uncompressed vs. compressed) and other factors for the desired detection or quantitation task.



A (10 degree acceptance angle)



B (20 degree acceptance angle)

Figure 7. Images of line sources from an experimental acquisition for (a) 10 degree and (b) 20 degree acceptance angles. (left) Focal plane image parallel to the detectors, (center) image normal to the detectors and (right) vertical profile through the center line source of the image normal to the detectors. Only the center 21 slices spaced at 3.3 mm were reconstructed.

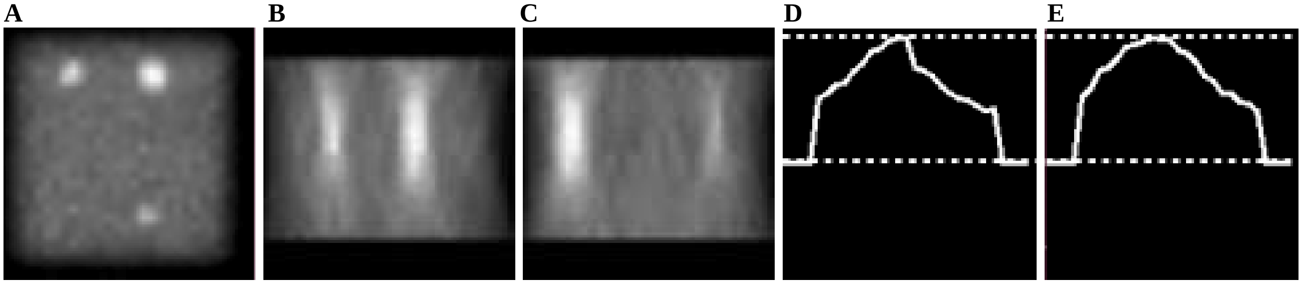


Figure 8. Images of a compressed breast phantom with simulated tumors from an experimental acquisition. Image reconstruction was with a 20 degree acceptance angle. (a) Focal plane image parallel to the detectors, (b) image normal to the detectors through the upper two tumors of (a), (c) image normal to the detectors through the tumors on the right side of (a), (d) vertical profile through the left tumor of (b), (e) vertical profile through the right tumor of (b).

V. CONCLUSION

Three-dimensional imaging for a dual detector positron emission mammography system has been investigated for image reconstruction by classical tomography. Analytic models and computer simulations of a PEM system provide insight into the blurring normal to the detectors observed in reconstructed PEM images from experimental line source and breast phantom acquisitions. Improved methods of extracting 3-D information from PEM coincidence data would be beneficial.

VI. REFERENCES

[1]C. J. Thompson, K. Murthy, Y. Picard, I. N. Weinberg, and R. Mako, "Positron emission mammography (PEM): a promising technique for detecting breast cancer," *IEEE Trans. Nucl. Sci.*, vol. 42, pp. 1012-1017, 1995.

[2]I. N. Weinberg, S. Majewski, A. G. Weisenberger, A. Markowitz, L. Aloj, L. Majewski, D. Danforth, J. Mulshine, K. Cowan, J. Zujewski, C. Chow, E. Jones, V. Chang, W. Berg, and J. Frank, "Preliminary results for positron emission mammography: real-time functional breast imaging in a conventional mammographic gantry," *Eur. J. Nucl. Med.*, vol. 23, pp. 804-806, 1996.

[3]R. R. Raylman, S. Majewski, R. Wojcik, A. G. Weisenberger, B. Kross, V. Popov, and H. A. Bishop, "The potential role of positron emission mammography for detection of breast cancer. A phantom study," *Med. Phys.*, vol. 27, pp. 1943-1954, 2000.

[4]R. Freifelder, J. S. Karp, S. Surti, and J. A. Wear, "A dedicated PET scanner for breast imaging using two curve-plate NaI(Tl) detectors," *J. Nucl. Med.*, vol. 39, p. 171P, 1998.

[5]R. Freifelder and J. S. Karp, "Dedicated PET scanners for breast imaging," *Phys. Med. Biol.*, vol. 42, pp. 2463-2480, 1997.

[6]H. Baghaei, W.-H. Wong, J. Uribe, H. Li, N. Zhang, and Y. Wang, "Breast cancer imaging studies with a variable field of view PET camera," *IEEE Trans. Nucl. Sci.*, vol. 47, pp. 1080-1084, 2000.

[7]R. L. Wahl, M. A. Helvie, A. E. Chang, and I. Andersson, "Detection of breast cancer in women after augmentation mammoplasty using fluorine-18-fluorodeoxyglucose-PET," *J. Nucl. Med.*, vol. 35, pp. 872-875, 1994.

[8]K. Yutani, M. Tatsumi, E. Shiba, H. Kusuoka, and T. Nishimura, "Comparison of dual-head coincidence gamma camera FDG imaging with FDG PET in detection of breast cancer and axillary lymph node metastasis," *J. Nucl. Med.*,vol. 40,pp. 1003-1008, 1999.

[9]K. Murthy, M. Aznar, C. J. Thompson, A. Loutfi, R. Lisbona, and J. H. Gagnon, "Results of preliminary clinical trials of the positron emission mammography system PEM-I: a dedicated breast imaging system producing glucose metabolic images using FDG," *J. Nucl. Med.*, vol. 41, pp. 1851-1858, 2000.

[10]H. H. Barrett and W. Swindell, *Radiological Imaging: The Theory of Image Formation, Detection and Processing*. New York: Academic, 1981.

Diffusion Tensor MR Imaging of Principal Directions: A Tensor Tomography Approach

V. Y. Panin, G. L. Zeng, M. Defrise, and G. T. Gullberg
University of Utah, Department of Radiology
729 Arapeen Dr., Salt Lake City, UT 84108-1218, USA.
AZ-VUB University Hospital, Free University, Division of Nuclear Medicine
B-1090 Brussels,Belgium

Abstract

This paper investigates a novel approach of reconstructing the principal directions of a diffusion tensor field directly from magnetic resonance imaging (MRI) data using a tensor tomography data acquisition approach. Because tensor eigenvalues are assumed to be known, the reconstruction of principal directions requires less measurements than the reconstruction of the full tensor field. The tensor tomography data acquisition method (rotated diffusion gradients) leads to a unique reconstruction of principal directions while the conventional MRI acquisition technique (stationary diffusion gradients) leads to an ambiguous reconstruction of principal directions if the same number of measurements are used. A computer generated phantom was used to simulate the diffusion tensor field in the mid-ventricular region of the myocardium. The diffusion model in this study depends upon the fiber structure of the myocardium. An iterative algorithm was used in the reconstruction. Computer simulations verify that the proposed method provides accurate reconstructions of the principal directions of a diffusion tensor field.

I. INTRODUCTION

Magnetic resonance imaging (MRI) has been shown to be effective for imaging diffusion tensor fields through a process known as diffusion-weighted imaging (DWI). In some imaging applications such as cardiac imaging, the primary objective is to use DWI to determine the principal directions of the diffusion tensor field where a complete understanding of the tensor field itself is of secondary interest. It has been established for isolated perfused myocardium that water diffusion anisotropy measured by MRI faithfully parallels histologic anisotropy. In a cardiac study the knowledge of the principal directions of the tensor field provides myocardial fiber organization [1]. Myocardial fiber architecture is a key determinant of the electrical and mechanical properties of the myocardium. On the other hand, the eigenvalues of the diffusion tensor in cardiac tissue may be assumed to be known. These values are similar to diffusivities reported in other human tissues, which are less than half the diffusivity of water at 37° [1].

This paper is concentrated on two tasks. The **first** is presentation of a novel approach of reconstruction of principal directions directly from diffusion-weighted MRI data, assuming known eigenvalues. Because *a priori* information is incorporated, the reconstruction of principal directions requires less measurements than in the case of the reconstruction of a full tensor field. This can be an important asset, because some DWI measurements strongly suffer from systematic errors such as eddy current artifacts [2]. Reducing the number of measurements may be beneficial when it is desired to use only artifact-free measurements. The **second** goal of this paper is to show that a MR tensor tomography diffusion-weighted imaging (TTDWI) approach [4,5] is more effective for reconstructing principal directions than the standard MR DWI techniques.

II. METHODS

A. DWI imaging: Tensor tomography versus the conventional MRI technique.

One of the approaches used in DWI is projection reconstruction (PR) imaging [3]. In PR imaging radial lines are acquired in Fourier space instead of rectilinear lines as in 2D Fourier transform imaging methods, for example, echo-planar imaging (EPI). The PR signal during readout can be expressed as

$$s_{\vec{\varpi}}(t) = \int \rho(\vec{x}) e^{i\gamma \vec{x} \cdot \vec{G}_r t} e^{-\gamma^2 G^2 (\vec{\varpi}^T D(\vec{x}) \vec{\varpi}) \Lambda^2 (\Delta - \Lambda/3)} d\vec{x} \ , \quad (1)$$

where $\vec{G}_r$ is the readout gradient, ρ is the spin density, γ is the gyromagnetic ratio, G is the amplitude of the diffusion weighting gradient, $\vec{\varpi}$ is the direction of the applied diffusion weighting gradient, Λ is the length of one lobe of the diffusion pulse, D is the diffusion tensor, and Δ is the separation between each start of the two gradient pulses.

Taking the Fourier transform of $s_{\vec{\varpi}}(t)$, (1) can be rewritten as

$$p_{\vec{\varpi}\vec{\varpi}}(\theta, t) = \int \rho(\vec{x}) e^{-g^2 (\vec{\varpi}^T D(\vec{x}) \vec{\varpi})} \delta(\vec{x} \cdot \vec{\theta} - t) d\vec{x} \ , \quad (2)$$

where g^2 is a constant and $\vec{\theta}$ is the direction of the readout gradient [4]. According to (2), the PR signal can be presented as a "projection" $p_{\vec{\varpi}\vec{\varpi}}(\theta, t)$ [which is proportional to the Fourier transform of $s_{\vec{\varpi}}(t)$ in (1)] of the two functions ρ and D.

The transition from (1) to (2) is made by applying the well-known Fourier section theorem. Our goal is to reconstruct a 3D diffusion tensor field using fully 3D reconstruction [5]. However, in the following analysis we will restrict our attention to slice-by-slice data acquisition, where the readout direction is defined by the projection angle θ. The necessary information is provided by data sets with different choices of $\vec{\varpi}$. This vector has a direction in 3D as shown in Figure 1.

The goal behind DWI is the reconstruction of D, assuming that ρ is a known function. In practice, ρ is reconstructed using (2) when g is set to zero. In the conventional MR DWI technique measurements are made with stationary diffusion gradients, *i. e.* $\vec{\varpi}$ is a constant vector. The same function is either projected at every angle θ, as in the case of PR, or acquired line by line in Fourier space, as in the case of EPI. The standard reconstruction technique for PR DWI is the filtered backprojection (FBP) method, for EPI it is the Fourier inversion formula. Exponential terms of D are reconstructed and D is obtained by taking a logarithm. Therefore, conventional DWI provides reconstruction of $\vec{\varpi}^T D \vec{\varpi}$ for some set of fixed $\vec{\varpi}$. For 3D symmetric tensor imaging it is necessary to use six different $\vec{\varpi}$ to obtain six different data sets. By appropriate choices of $\vec{\varpi}$, conventional MR DWI reconstructs the diffusion tensor components D_{xx}, D_{yy}, etc. from six data sets.

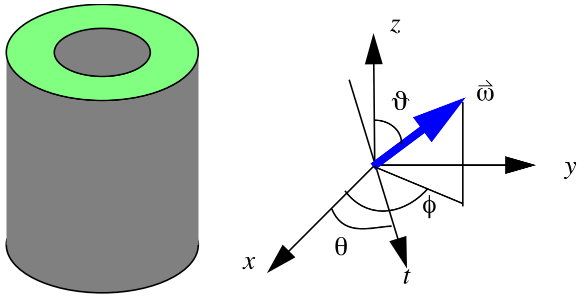


Figure 1. The projection geometry

The tensor tomography DWI (TTDWI) approach is similar to the PR technique but uses the rotating diffusion gradients when $\vec{\varpi}$ is a function of θ. In order to reconstruct a 3D tensor field, the six different $\vec{\varpi}$ should also be used and all data sets are used simultaneously during the reconstruction process. Because a different function is projected depending on the projection angle, a special reconstruction technique is required [4, 5,6].

B. Parameterization and reconstruction of principal directions when eigenvalues are known

1) 2D case

The main idea can easily be demonstrated in 2D case. The 2D tensor for each pixel can be represented in terms of known eigenvalues and eigenvectors:

$$D_{ij} = \lambda_1 X_i^1 X_j^1 + \lambda_2 X_i^2 X_j^2 ,\ \lambda_1 > \lambda_2 ,\ i,j = 1,2\ . \qquad (3)$$

The unknown eigenvectors can be parameterized by the angle Φ, which ensures orthogonality and normalization:

$$\vec{X}^1 = [\cos\Phi,\sin\Phi]^T ,\ \vec{X}^2 = [-\sin\Phi,\cos\Phi]^T ,\ 0 \le \Phi < \pi\,. \qquad (4)$$

Note that Φ and Φ+π are equivalent, because the principal eigenvectors are defined up to sign. Because the function $\Phi(\vec{x})$ is the unknown variable, one can expect that only a single set of measurements of diffusion-weighted gradients is necessary to reconstruct principal directions.

Conventional DWI provides reconstruction of $\vec{\varpi}^T D\vec{\varpi}$, where $\vec{\varpi} = [\cos\phi,\sin\phi]^T$ at some fixed ϕ. The expression $\vec{\varpi}^T D\vec{\varpi}$ can be rewritten:

$$\vec{\varpi}^T D\vec{\varpi} = \lambda_1\cos(\Phi-\phi)^2 + \lambda_2\sin(\Phi-\phi)^2\,. \qquad (5)$$

Equation (5) can be solved with respect to Φ:

$$\begin{aligned}
\Phi_1 &= \phi + \operatorname{atan}\left(\sqrt{\frac{\lambda_1-\vec{\varpi}^T D\vec{\varpi}}{\vec{\varpi}^T D\vec{\varpi}-\lambda_2}}\right) + \pi k_1,\ \ \vec{\varpi}^T D\vec{\varpi}-\lambda_2 \ne 0\\
\Phi_2 &= \phi + \operatorname{atan}\left(-\sqrt{\frac{\lambda_1-\vec{\varpi}^T D\vec{\varpi}}{\vec{\varpi}^T D\vec{\varpi}-\lambda_2}}\right) + \pi k_2,\ \ \vec{\varpi}^T D\vec{\varpi}-\lambda_2 \ne 0T \qquad (6)\\
\Phi_{1,\,2} &= \phi + \frac{\pi}{2} + \pi k_3,\ \ \vec{\varpi}^T D\vec{\varpi}-\lambda_2 = 0,
\end{aligned}$$

where integers k_1, k_2 and k_3 are chosen to ensure that $0 \le \Phi_1,\Phi_2 < \pi$. A single measurement of diffusion-weighted gradients provides two different Φ in general. Therefore, the principal directions are not uniquely defined.

The source of this ambiguity can be seen from other points of view. To define all eigenvalues and eigenvectors it is necessary to know all components of a symmetrical tensor and standard MRI DWI provides reconstruction of them. Suppose we know *a priori* λ_1, λ_2 and reconstruct, for example, D_{xx} (from measurement when ϕ=0). Can we then define D_{yy} and D_{xy}? It is known that two invariants exist for 2D second order tensors:

$$\begin{aligned}
D_{xx} + D_{yy} &= \lambda_1 + \lambda_2\\
D_{xx}^2 + D_{yy}^2 + 2D_{xy}^2 &= \lambda_1^2 + \lambda_2^2
\end{aligned}\ . \qquad (7)$$

Given λ_1, λ_2, and D_{xx} it is clear that the sign of D_{xy} is not defined.

The source of ambiguity can also be presented graphically. Because $(\vec{\varpi}^T\vec{X}^1)^2 + (\vec{\varpi}^T\vec{X}^2)^2 = 1$, the value of $\vec{\varpi}^T D\vec{\varpi} = (\lambda_1-\lambda_2)(\vec{\varpi}^T\vec{X}^1)^2 + \lambda_2$ is the same for an equivalent $\left|\vec{\varpi}^T\vec{X}^1\right|$. According to Figure 2(a) there are two $\vec{X}^1$ that provide the same $\left|\vec{\varpi}^T\vec{X}^1\right|$ in the general case.

We present here an empirical conjecture that the TTDWI method removes this ambiguity. Our results show that the TTDWI method derives a unique estimate of the direction of the principal component in fewer measurements than are required in conventional DWI. It can be shown that only one root in (6) satisfies (5) simultaneously for all ϕ. In the TTDWI method we do not reconstruct $\vec{\varpi}^T D\vec{\varpi}$, because ϕ is not fixed, but instead we reconstruct D. We can only fit Φ to projection measurements at the angles θ corresponding to one $\vec{\varpi}(\phi(\theta))$. However, we should not have duality of Φ, because only one Φ satisfies (5) for all ϕ. In the TTDWI method only projection measurements over θ corresponding to one $\vec{\varpi}(\phi(\theta))$ is necessary to define Φ for any given pixel whereas from our above arguments measurements for multiple $\vec{\varpi}(\phi)$ are needed for the conventional DWI method. Later we present a simulation that seems to suggest that our conjecture is correct for the case we present in the next section.

2) 3D case, primary anisotropy $\lambda_1 > \lambda_2 = \lambda_3$

This case has a simple graphical interpretation. Only the first principal direction is to be reconstructed. The other two eigenvectors that correspond to the same eigenvalues can be defined arbitrarily in the plane perpendicular to the first principal direction. The tensor components are independent of this particular choice. We can define three eigenvectors for a given voxel using two angles Θ and Φ (similarly ϑ and ϕ are illustrated on Figure 1).

$$\begin{aligned}
\vec{X}^1 &= [\sin\Theta\cos\Phi,\sin\Theta\sin\Phi,\cos\Theta]^T\\
\vec{X}^2 &= [-\sin\Phi,\cos\Phi,0]^T \qquad (8)\\
\vec{X}^3 &= [-\cos\Theta\cos\Phi,-\cos\Theta\sin\Phi,\sin\Theta]^T
\end{aligned}$$

where $0 \le \Phi < 2\pi$ and $0 \le \Theta \le \pi/2$ is only considered. Then $\vec{\varpi}$ is defined as $\vec{\varpi} = [\sin\vartheta\cos\phi,\sin\vartheta\sin\phi,\cos\vartheta]^T$ and

$$\begin{aligned}
\vec{\varpi}^T D\vec{\varpi} = {} & \lambda_1[\sin\vartheta\sin\Theta\cos(\phi-\Phi) + \cos\vartheta\cos\Theta]^2 +\\
& \lambda_2[\sin\vartheta^2\sin(\phi-\Phi)^2 + (\cos\vartheta\sin\Theta - \sin\vartheta\cos\Theta\cos(\phi-\Phi))^2]
\end{aligned} \qquad (9)$$

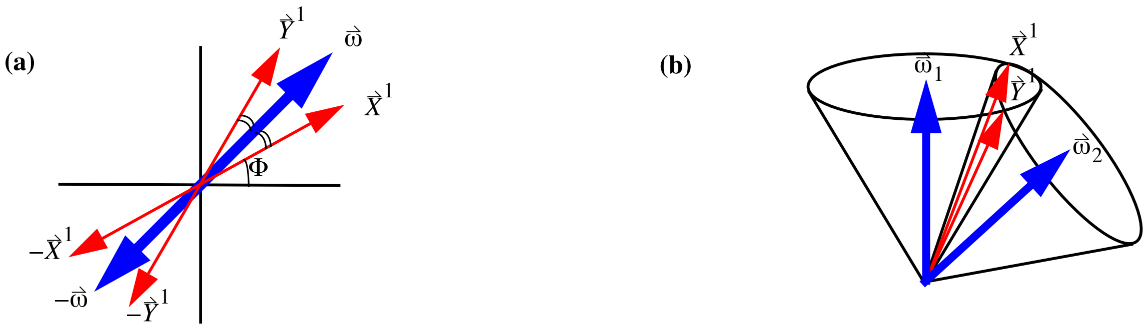


Figure 2. The source of ambiguity in reconstruction of principal directions from data acquired by standard MR DWI when $\vec{\omega}$ is fixed. (a) 2D case. Given $d = |\vec{\omega}^T\vec{X}^1|$, $\vec{X}^1$ and $\vec{Y}^1$ provide the same d. The principal vectors located in the upper half can be chosen. (b) 3D primary anisotropy case. Two values $d_1 = |\vec{\omega}_1^T\vec{X}^1|$ and $d_2 = |\vec{\omega}_2^T\vec{X}^1|$ define two cones. The intersection of these cones in general provides the different principal vector $\vec{Y}^1$, that satisfies d_1 and d_2. The only upper hemisphere is shown.

Because the two functions $\Phi(\vec{x})$ and $\Theta(\vec{x})$ are unknown, one can expect that two measurements with two different $\vec{\omega}$ are necessary when reconstructing the principal directions. This is true in the case of TTDWI, because each $\vec{\omega}$ is a function of θ. However in the case of conventional MR DWI, where $\vec{\omega}$ are fixed, it is necessary to take three measurements to uniquely define the principal directions. Taking into account that $(\vec{\omega}^T\vec{X}^1)^2 + (\vec{\omega}^T\vec{X}^2)^2 + (\vec{\omega}^T\vec{X}^3)^2 = 1$ and $\lambda_2 = \lambda_3$, the equation $\vec{\omega}^T D\vec{\omega} = (\lambda_1 - \lambda_2)(\vec{\omega}^T\vec{X}^1)^2 + \lambda_2$ is satisfied for equal vales of $|\vec{\omega}^T\vec{X}^1|$. As it can be seen in Figure 2(b), each $\vec{\omega}$ and $\vec{X}^1$ define a cone. Intersection of these two cones defines $\vec{X}^1$ which provides the same $\vec{\omega}_1^T D\vec{\omega}_1$ and $\vec{\omega}_2^T D\vec{\omega}_2$. In general, there are two such $\vec{X}^1$ in the upper hemisphere.

3) 3D case, secondary anisotropy $\lambda_1 > \lambda_2 > \lambda_3$

The existence of secondary myocardial anisotropy in the cross-fiber direction was previously established [1]. In this case it is necessary to estimate all three principal directions. They can be parameterized by three Euler's angles: $\alpha(\vec{x})$, $\beta(\vec{x})$, and $\gamma(\vec{x})$ for each voxel. We omit the corresponding notation here. The three different measurements (three different nonstationary $\vec{\omega}$) may be required to reconstruct these three functions in TTDWI. In the case of a stationary $\vec{\omega}$, that is conventional MR DWI, three measurements are not enough to obtain unique estimates of the principal directions. This can be understood in terms of 3D tensor invariants:

$$\begin{aligned}
&Tr(D) = D_{xx} + D_{yy} + D_{zz} = \lambda_1 + \lambda_2 + \lambda_3\\
&D_{xx}D_{yy} + D_{xx}D_{zz} + D_{yy}D_{zz} - (D_{xy}^2 + D_{xz}^2 + D_{yz}^2) =\\
&\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3\\
&det(D) = D_{xx}D_{yy}D_{zz} + 2D_{xy}D_{xz}D_{yz} -\\
&(D_{zz}D_{xy}^2 + D_{yy}D_{xz}^2 + D_{xx}D_{yz}^2) = \lambda_1\lambda_2\lambda_3\,. \qquad (10)
\end{aligned}$$

Providing, for example, D_{xx}, D_{yy} and D_{xz}, D_{yz} and D_{xy} are defined up to sign change. Four measurements may be enough. (Measurement components must be chosen carefully. It is a bad choice to measure D_{xx}, D_{yy} and D_{zz} among these four measurements. The tensor estimation still will be ambiguous.)

C. Reconstruction algorithm

For the purposes of this abstract we restricted our attention to the more simple case of primary anisotropy. Only the first principal direction was estimated. Since the MR DWI model in (2) is nonlinear, the reconstruction with nonstationary $\vec{\omega}$ requires use of iterative methods [6,7]. In order to estimate principal directions, the least squares differences between modeled and measured projections can be minimized:

$$L(\Phi(\vec{x}),\Theta(\vec{x})) = \sum_{\theta,\,t}\sum_{\vec{\omega}}\left\| p_{\vec{\omega}\vec{\omega}}^{model} - p_{\vec{\omega}\vec{\omega}}^{measured}\right\|^2 \qquad (11)$$

where

$$p_{\vec{\omega}\vec{\omega}}^{model}(\theta, t) = \int\rho(\vec{x})e^{-g^2(\vec{\omega}^T D(\vec{x})\vec{\omega})}\delta(\vec{x}\cdot\vec{\theta} - t)d\vec{x} \qquad (12)$$

and

$$\vec{\omega}^T D(\vec{x})\vec{\omega} = \lambda_1[\sin\vartheta\sin\Theta\cos(\phi - \Phi) + \cos\vartheta\cos\Theta]^2 +$$
$$\lambda_2[\sin\vartheta^2\sin(\phi - \Phi)^2 + (\cos\vartheta\sin\Theta - \sin\vartheta\cos\Theta\cos(\phi - \Phi))^2] \qquad (13)$$

In order to minimize (11) with respect to the angle functions, a gradient-type algorithm was applied. Note that the minimization problem is complicated due to the fact that the objective function is periodic with respect to the angle functions; therefore, *L* has an infinite number of minima. We implemented the gradient descent (GD) algorithm to minimize *L*. At each iteration this algorithm updates each angle function for a given voxel by its corresponding derivative of *L*. This algorithm relies on an arbitrarily chosen relaxation parameter ε, that defines the step size in the gradient direction. This parameter should be small enough to not over shoot the downhill direction. The choice of a very small value of ε leads to slow convergence, however, it allows to stay near one of the many minima.

III. RESULTS

A computer generated phantom was used to simulate the diffusion that might be expected in a cardiac study. The phantom is comprised of a circular cylindrical tube. The phantom simulates the mid-ventricular wall of the left ventricle. The spin density ρ is assumed to be uniform inside the phantom and zero outside. The fiber structure of the myocardium is helical. The principal vectors of the diffusion tensor are referenced to a helical fiber structure with material coordinates (X_R, X_F, X_C), which are orthogonal. The fiber axis X_F is located on the plane of the wall normal to the radial axis R. The fiber angle in the circumferential direction has a variation which is continuous and linear. The angle changes from 60° to -60°, varying from the endocardial to the epicardial wall in a radial direction. The axis X_C is the cross-fiber in-plane axis and the axis X_R coincides with R. The phantom was chosen to be independent of the z coordinate so that one slice of the reconstruction was enough to represent the entire phantom.

The phantom represented a 32x32 slice image of a cylinder with an inner radius $R_1 = 7$ and an outer radius R_2=14. This grid size was chosen in order to achieve a comprehensive visualization of the vector field of the principal vector. The eigenvalues of the myocardial diffusion tensor were $\lambda_1 = 1.6$, $\lambda_2 = \lambda_3 = 0.7$. The spin density ρ was equal to 1 inside the cylinder and zero outside the cylinder. The parameter g^2 was 0.7, so $g^2\lambda_1 > 1$.

The projection and backprojection operation of the GD algorithm were implemented, using a ray-driven operator. Thirty two projections of the slice were generated over $\theta = [0, \pi)$. The sampling bin width was equal to the reconstructed pixel width. Two projection data sets were used with two fully 3D rotated diffusion gradient directions

$$\vec{\omega}_1 = \left\{ \phi = \theta;\ \vartheta = \begin{matrix} \theta, & \theta \le \pi/2 \\ \pi/2 - \theta, & \theta \ge \pi/2 \end{matrix} \right\} \tag{14}$$

$$\vec{\omega}_2 = \left\{ \phi = \theta + \pi/2;\ \vartheta = \begin{matrix} \theta, & \theta \le \pi/2 \\ \pi/2 - \theta, & \theta \ge \pi/2 \end{matrix} \right\}. \tag{15}$$

The initial condition for the iterative algorithm was uniform: $\Phi = \pi$ and $\Theta = \pi/4$ for every voxel.

Figure 3 shows the behavior of the LS norm in (11) as a function of the iteration number. Because the data is noise-free, the algorithm converged to *L=0*. We chose 4000 iterations for the final reconstruction of the noise-free data. This reconstruction is nearly identical to the original phantom. (The angular difference between phantom and reconstruction principal directions was $1.6^{\circ} \pm 1.4^{\circ}$.) A smaller iteration number, however, can be used. Figure 4 presents the reconstructed vector field of the first principal vector showing the fiber structure of the heart for one transaxial slice. The first principal direction is well reconstructed in the case of noise-free data from only two diffusion-weighted projection data sets.

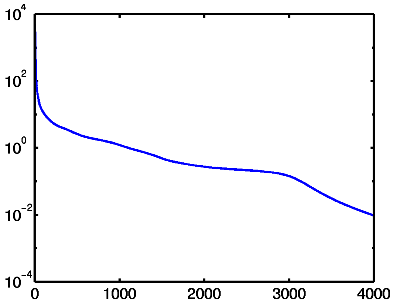


Figure 3. LS norm as a function of the iteration number.

IV. DISCUSSION AND CONCLUSION

We have developed a novel approach to reconstructing the principal directions of the diffusion tensor field. We have also demonstrated that MR TTDWI data acquisition is more efficient than standard MR DWI data acquisition. Further work is needed to prove this result mathematically. Our approach requires use of an iterative algorithm. Further work is required to increase the convergence rate. The choice of an optimal direction of the diffusion gradient direction during acquisition needs further investigation. Acquisition and reconstruction of real data is currently underway.

V. REFERENCES

[1] T. G. Reese, R. M. Weisskoff, R. N. Smith, B. R. Rosen, R. E. Dinsmore, and V. J. Wedeen, "Imaging myocardial fiber architecture *in vivo* with magnetic resonance," *Magn. Reson. Med.*, vol. 34. pp.786-791, 1995.

[2] M. H. Seifert, P. M. Jakob, V. Jellus, A. Haase, and C. Hillenbrand, "High-resolution diffusion imaging using a radial turbo-spin-echo sequence: implementation, eddy current compensation, and self-navigation," *J. Magn. Reson.*, vol. 144, pp. 243-254, 2000.

[3] A. F. Gmitro and A. L. Alexander "Use of a projection reconstruction method to decrease motion sensitivity in diffusion-weighted MRI," *Magn. Reson. Med.*, vol. 9. pp. 835-838, 1993.

[4] G. T. Gullberg, D. N. Ghosh Roy, G. L. Zeng, A. L. Alexander, and D. L. Parker, "Tensor tomography," *IEEE Nucl. Sci.*, vol. 46, pp. 991-1000, 1999.

[5] G. T. Gullberg, M. Defrise, V. Y. Panin and G. L. Zeng, "Efficient cardiac diffusion tensor MRI by three-dimensional reconstruction of solenoidal tensor fields," *Magn. Reson. Imag.*, in press, 2001.

[6] V. Y. Panin, G. L. Zeng and G. T. Gullberg, "An iterative approach to tensor tomography," *Proceeding of the 2000 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Oct. 15-20, 2000, Lyon, France, (in press).

[7] V. Y. Panin, G. L. Zeng, G. T. Gullberg, A. L. Alexander, and D. L. Parker, "An iterative regularized algorithm for tensor tomography in MRI," In *Proceedings of the International Society for Magnetic Resonance in Medicine, 9th Scientific Meeting and Exhibition*, April 21-27, 2001, Glasgow, Scotland, pp. 765.

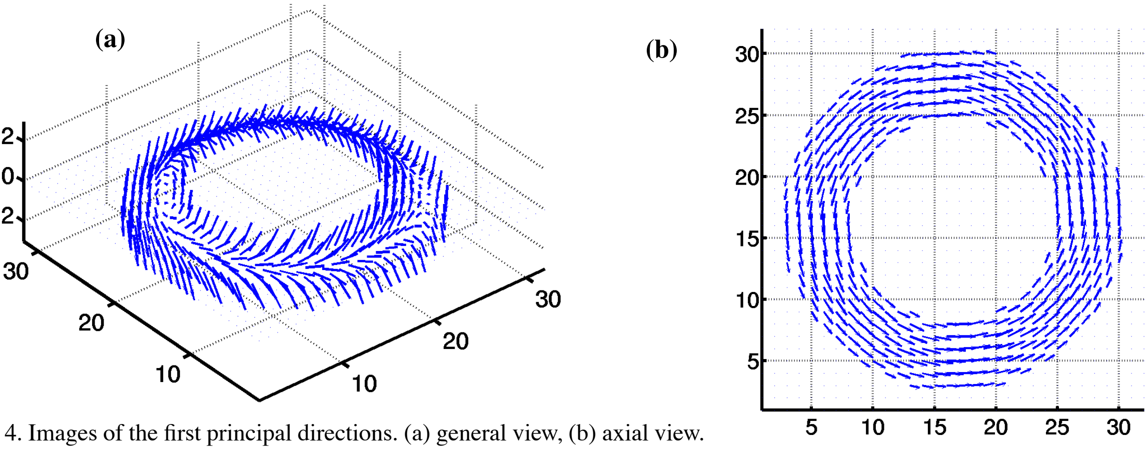


Figure 4. Images of the first principal directions. (a) general view, (b) axial view.

Complete Source Trajectories for C-Arm Systems and a Method for Coping with Truncated Cone-Beam Projections

Hermann Schomberg
Philips Research Hamburg, Röntgenstraße 24–26, 22335 Hamburg, Germany

1 Introduction

A C-arm system, such as the one shown in Figure 1, may be used to acquire X-ray cone-beam projections of a patient's three-dimensional (3D) region of interest (ROI) while the X-ray source moves along some trajectory around the ROI. Then, a 3D image of the X-ray attenuation coefficient within the ROI may be reconstructed from the acquired cone-beam projections.

In fact, C-arm systems are being used in this way, see e.g. [1, 2]. Typically, the source moves along a circular arc spanning an angle of 180°–200°, and the image is reconstructed by a variant of the algorithm of Feldkamp, Davis, and Kress [3]. As the cone-beam does not cover the whole patient, the projections are truncated. This is usually handled by extending the projections in a simple manner parallel to the plane containing the source trajectory. In practice, the true trajectory differs slightly from its ideal, but this deviation can be measured [5, 6] and taken into account during the reconstruction. The reconstructed image is subjected to a surface or volume rendering process designed to extract and visualize only the high contrast structures of the object under examination.

The reconstructed image itself is usually cluttered by severe artifacts. Nevertheless, the high contrast structures, such as bones or blood vessels filled with intraarterially injected contrast agents, are well recovered. This is because such structures stand out well against the background and are also reconstructed at the correct geometric locations.

Accurate images of medium contrast structures, such as blood vessels filled with intravenously injected contrast agents, or low contrast structures, such as soft tissue organs, are not obtained in this way. If this is to be improved upon, the following conditions will have to be met: First, the data acquired by the C-arm system must provide (after some preprocessing) accurate, though sampled, cone-beam projections of the object function (the X-ray attenuation coefficient). Second, the sampling density along the trajectory and on the detector surface must be sufficiently high. Third, the source trajectory must be complete in the sense that every plane that intersects the ROI contains a source point. Fourth, the cone beam projections must not be truncated. Under these conditions, any standard exact cone-beam reconstruction algorithm will produce an accurate estimate of the object function within the ROI. It has been tacitly assumed that the object function varies only spatially. Imaging moving parts of the human body, such as the beating heart, is not considered here.

Measuring accurate cone-beam projections of the object function presents a number of technical challenges, but no fundamental obstacles, and is also not considered here. Making the sampling density sufficiently high is not a fundamental problem either. The problem of designing source trajectories that are both complete and realizable by appropriate C-arm systems is discussed and solved in Section 4. The remaining requirement for non-truncated cone beam projections cannot be satisfied by any realistic C-arm system when the object to be imaged is part of a human body. As in the fan-beam case [4] one can show that truncated cone-beam projections leave the object undetermined. To cope with this fundamental limitation we suggest in Section 3 to extend the truncated projections in a relatively simple fashion so as to fake non-truncated projections of an imaginary object that is somewhat bigger than the ROI, and to apply a suitable cone-beam reconstruction algorithm to the extended projections. It will be argued in Section 3 that an image thus obtained can be expected to differ, inside the ROI, from the true image merely by some unknown, but smooth, nearly constant, and fairly weak ghost image. The reconstructed image would therefore still allow the detection of fine anatomical details.

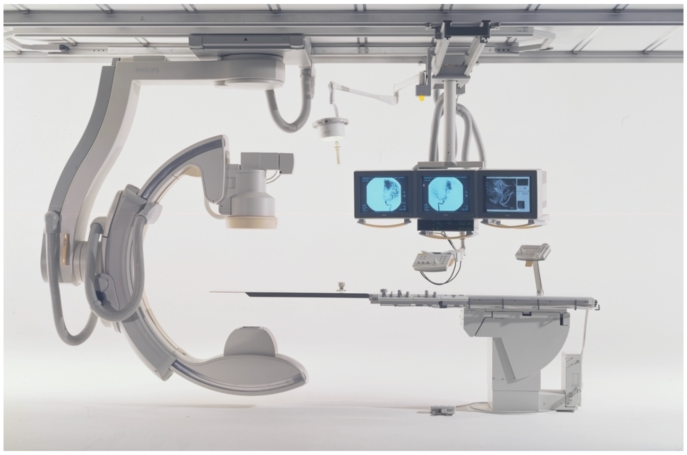


Figure 1. Integris Allura C-arm system (Philips Medical Systems, Best, The Netherlands).

2 C-Arm Systems

A typical C-arm system, such as the one shown in Figure 1, is equipped with a point like X-ray source and a planar X-ray detector, usually an image intensifier, which are mounted to the ends of a C-arm. The C-arm is held by another arm, which we refer to as the C-arm suspension. The C-arm suspension is attached to an L-arm, which is mounted to the ceiling (or floor). The L-arm can be rotated about a vertical axis. The C-arm suspension can be rotated about a horizontal axis attached to the L-arm. By rotating L-arm and C-arm suspension about their axes, the orientation of the plane containing the C-arm may be changed. The C-arm itself may be rotated within this plane about an axis perpendicular to this plane. All three rotation axes meet in a single point, the isocenter. Also, the straight line from the X-ray source to the center of the detector surface passes through the isocenter. The rotations themselves are effected by servo motors. When one or more of the the arms rotate, the X-ray source moves along a corresponding trajectory. The trajectory is confined to the surface of an isocentric sphere and further constrained by the electromechanical design of the joints connecting the arms.

3 The Reconstruction Problem

Points in space will be referenced with respect to a right-handed Cartesian coordinate system, the laboratory system, which we attach to the isocenter of the C-arm system such that the z-axis points upwards and the y-axis along the patient table. The object to be imaged is represented by a function $f : \mathbf{R}^3 \to \mathbf{R}$. The source trajectory is represented by a smooth mapping $\boldsymbol{a} : \Lambda \to \mathbf{R}^3$, where $\Lambda = [\lambda_-, \lambda_+]$ is a bounded interval. When the source is at position $\boldsymbol{a}(\lambda)$, the sensitive area of the detector defines a plane $D(\lambda) \subset \mathbf{R}^3$. The sensitive area of the detector itself is represented by the disk $D_0(\lambda) \subset D(\lambda)$. The set of unit vectors $\boldsymbol{\theta} \in S^2$ such that the ray $\{\boldsymbol{a}(\lambda) + s\boldsymbol{\theta} \mid s \geq 0\}$ hits $D_0(\lambda)$ is denoted by $S_0(\lambda)$. The source point $\boldsymbol{a}(\lambda)$ and the unit vectors in $S_0(\lambda)$ define the cone $C_0(\lambda) = \{\boldsymbol{a}(\lambda) + s\boldsymbol{\theta} \mid s \geq 0,\ \boldsymbol{\theta} \in S_0(\lambda)\}$. The biggest centered ball contained in all cones $C_0(\lambda)$, $\lambda \in \Lambda$, is denoted by B_0, its radius by r_0. This ball is also the biggest centered ball within which we can hope to obtain an accurate reconstruction. In practice, r_0 lies between 100 and 150 mm. We define the function $g : \Lambda \times S^2 \to \mathbf{R}$ by

$$g(\lambda, \boldsymbol{\theta}) = \int_0^\infty f(\boldsymbol{a}(\lambda) + s\boldsymbol{\theta})\, ds. \tag{1}$$

The function $g(\lambda, \cdot)$ represents the (non-truncated) cone-beam projection of f with vertex $\boldsymbol{a}(\lambda)$. The data acquisition process provides (a sampled and noise contaminated version of) the function $\tilde{g}(\lambda, u, v) = g(\lambda, \tilde{\boldsymbol{\theta}}(\lambda, u, v))$, where (u, v) are local coordinates in the detector plane $D(\lambda)$ and $\tilde{\boldsymbol{\theta}}(\lambda, u, v) \in S_0(\lambda)$ is the unit vector pointing from $\boldsymbol{a}(\lambda)$ to the point (u, v) on $D_0(\lambda)$. Thus, $g(\lambda, \boldsymbol{\theta})$ is available only for $\lambda \in \Lambda$, $\boldsymbol{\theta} \in S_0(\lambda)$. Using these data, we wish to reconstruct f in the ball B_0. When the object is a part of a human being, the acquired cone-beam projections are inevitably truncated, i.e., the support of f extends beyond B_0, and there exist $\lambda \in \Lambda$ and $\theta \in S^2 \setminus S_0(\lambda)$ such that $g(\lambda, \boldsymbol{\theta}) \neq 0$.

For a stable and accurate reconstruction of f in B_0 it is necessary that the source trajectory $\boldsymbol{a}$ satisfy the following completeness condition (see [7] and the references cited therein): *Every plane that intersects the ball* B_0 *contains a source point* $\boldsymbol{a}(\lambda)$. A source trajectory satisfying this condition will be called complete with respect to B_0. The normals of the planes that contain a source point $\boldsymbol{a}(\lambda)$ and intersect the ball B_0 form an umbrella-shaped surface

$$U(\lambda, r_0) = \{r\boldsymbol{\theta} \mid |r| \leq r_0,\ \boldsymbol{\theta} \in S^2,\ \boldsymbol{a}(\lambda) \cdot \boldsymbol{\theta} = r\} \tag{2}$$

within B_0. Using the correspondence between planes and their normals, the completeness condition may be rephrased as follows: *The collection of the surfaces* $U(\lambda, r_0)$, $\lambda \in \Lambda$, *fills the ball* B_0 *completely.*

In general, it is hard to prove by argument that a given trajectory is complete with respect to B_0. However, a computer may be used to draw a large number of surfaces $U(\lambda_0, r_0), \ldots, U(\lambda_N, r_0)$ with $\lambda_n = n(\lambda_+ - \lambda_-)/N$, $n = 0, 1, \ldots, N$, and it may be checked visually whether these surfaces will densely fill B_0 as $N \to \infty$.

In the remainder of this section, we assume that the source trajectory is complete with respect to B_0. Then, if the cone-beam projections were not truncated, f could be reconstructed inside B_0 using an appropriate cone-beam reconstruction algorithm, such as the filtered backprojection algorithm described in [8]. This algorithm is designed to act on $\tilde{g}(\lambda, u, v)$ rather than on $g(\lambda, \boldsymbol{\theta})$.

The development presented in [8] also allows one to derive the explicit reconstruction formula

$$f_{\mathrm{rec}}(\boldsymbol{x}) = \frac{1}{8\pi^2} \int_\Lambda \int_{S^2} K(\boldsymbol{x}, \lambda, \boldsymbol{\theta}) g(\lambda, \boldsymbol{\theta})\, d\boldsymbol{\theta}\, d\lambda \tag{3}$$

with the kernel

$$K(\boldsymbol{x}, \lambda, \boldsymbol{\theta}) = \int_{S^2} \frac{|\boldsymbol{a}'(\lambda) \cdot \boldsymbol{\beta}|}{\|\boldsymbol{x} - \boldsymbol{a}(\lambda)\|^2} M(\lambda, \boldsymbol{\beta}) \delta'_\epsilon(\boldsymbol{\beta} \cdot \boldsymbol{\phi}(\lambda, \boldsymbol{x})) \delta'_\epsilon(\boldsymbol{\beta} \cdot \boldsymbol{\theta})\, d\boldsymbol{\beta}. \tag{4}$$

Here, $\boldsymbol{a}'$ is the derivative of $\boldsymbol{a}$; δ'_ϵ is a regular and smooth approximation to the derivative of the δ-function; M is a smooth weighting function accounting for the fact that a plane can contain several source points [7, 8]; and

$$\boldsymbol{\phi}(\lambda, \boldsymbol{x}) = \frac{\boldsymbol{x} - \boldsymbol{a}(\lambda)}{\|\boldsymbol{x} - \boldsymbol{a}(\lambda)\|}. \tag{5}$$

When the conditions are right, one expects f_{rec} to tend to f inside B_0 as δ'_ϵ tends to δ'. A discretized version of (3) could provide the basis for a (computationally inefficient) reconstruction algorithm.

More interestingly, the kernel (4) highlights an important property of the reconstruction problem: Since δ'_ϵ is concentrated near zero, $\delta'_\epsilon(\boldsymbol{\beta} \cdot \boldsymbol{\phi}(\lambda, \boldsymbol{x}))$ is appreciably nonzero only when $\boldsymbol{\beta}$ belongs to a neighborhood of the circle $S^2 \cap \boldsymbol{\phi}(\lambda, \boldsymbol{x})^\perp$, and $\delta'_\epsilon(\boldsymbol{\beta} \cdot \boldsymbol{\theta})$ is appreciably nonzero only when $\boldsymbol{\beta}$ belongs to a neighborhood of the circle $S^2 \cap \boldsymbol{\theta}^\perp$. The mentioned two circles are identical if $\boldsymbol{\theta}$ and $\boldsymbol{\phi}(\lambda, \boldsymbol{x})$ are parallel; otherwise they have only two points in common. As a result, $K(\boldsymbol{x}, \lambda, \boldsymbol{\theta})$ is appreciably nonzero at most when $\boldsymbol{\theta}$ and $\boldsymbol{\phi}(\lambda, \boldsymbol{x})$ are nearly parallel, i.e., when the ray $\{\boldsymbol{a}(\lambda) + s\boldsymbol{\theta} \mid s \geq 0\}$ comes close to $\boldsymbol{x}$. In addition, $K(\boldsymbol{x}, \lambda, \boldsymbol{\theta})$ depends smoothly and weakly on λ and $\boldsymbol{\theta}$ whenever $\boldsymbol{a}$ depends smoothly on λ and the ray $\{\boldsymbol{a}(\lambda) + s\boldsymbol{\theta} \mid s \geq 0\}$ does not come close to $\boldsymbol{x}$. Although the reconstruction process is not strictly local, it is still semi-local.

The reconstruction problem of 2D CT has similar properties [9]. It is also possible to extend the ideas presented in [9] to the 3D case: From (3) we find that

$$f_{\mathrm{rec}}(\boldsymbol{x}) = f_0(\boldsymbol{x}) + f_1(\boldsymbol{x}) \tag{6}$$

with

$$f_0(\boldsymbol{x}) = \frac{1}{8\pi^2} \int_\Lambda \int_{S_0(\lambda)} K(\boldsymbol{x}, \lambda, \boldsymbol{\theta}) g(\lambda, \boldsymbol{\theta})\, d\boldsymbol{\theta}\, d\lambda, \tag{7}$$

$$f_1(\boldsymbol{x}) = \frac{1}{8\pi^2} \int_\Lambda \int_{S^2 \setminus S_0(\lambda)} K(\boldsymbol{x}, \lambda, \boldsymbol{\theta}) g(\lambda, \boldsymbol{\theta})\, d\boldsymbol{\theta}\, d\lambda. \tag{8}$$

The function f_0 is well determined by the available data, but f_1 is completely undetermined. Owing to the semi-local nature of the reconstruction process, f_0 will be fairly smooth well outside B_0 and decay quickly there. Conversely, f_1 will be fairly smooth well within B_0. Also, a crude guess of the truncated portions of the cone-beam projections should suffice to compute a fair approximation to f_1 inside B_0. Already a crude guess of the

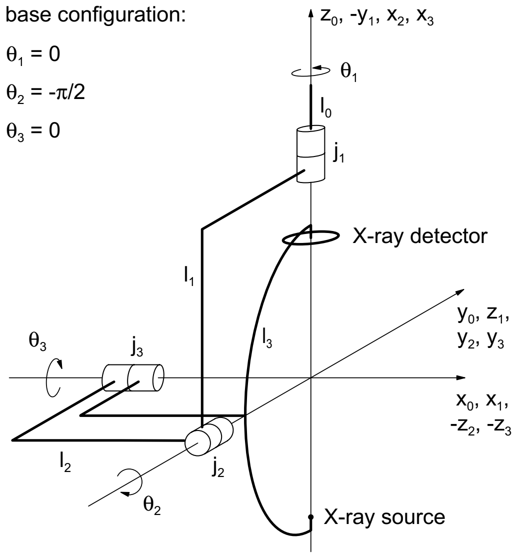


Figure 2. A C-arm system drawn as a series of links and joints.

integrals along the lines that pass through a centered ball B_1 a little larger than B_0 should suffice for this purpose.

These observations suggest the following method for coping with truncated cone-beam projections: 1. Choose a centered ball B_1 somewhat bigger than B_0. For each $\lambda \in \Lambda$, let $D_1(\lambda)$ be the perspective projection of B_1 from $\boldsymbol{a}(\lambda)$ onto $D(\lambda)$. 2. Extend each truncated cone-beam projection $\tilde{g}(\lambda, \cdot, \cdot)$ radially from $D_0(\lambda)$ to $D_1(\lambda)$ using, for example, the 1D extension method proposed in [9] in radial direction. 3. Reconstruct f in B_0 by applying the filtered backprojection algorithm described in [8], or some other suitable cone-beam reconstruction algorithm, to the extended projections.

An image reconstructed in this way can be expected to differ inside B_0 from the true image merely by some unknown, but smooth, nearly constant, and fairly weak ghost image.

4 Complete Source Trajectories

To specify and design source trajectories, we adopt a method commonly used in robotics [10]. The first step is to model the C-arm system by a series of rigid links connected by revolute joints. Figure 2 illustrates the idea. There are four links, denoted by l_0, l_1, l_2, l_3, and three joints, denoted by j_1, j_2, j_3. Link l_0 is fixed to the laboratory. Link l_1 represents the L-arm, link l_2 the C-arm suspension, and l_3 the C-arm. Joint j_k connects links l_{k-1} and l_k, $k = 1, 2, 3$. Each joint defines an axis of rotation, and these axes intersect in the isocenter.

Next, a right-handed Cartesian (x_k, y_k, z_k)-coordinate system is attached to link l_k, $k = 0, 1, 2, 3$. There is some freedom in the choice of the origins and orientations of these coordinate systems; we make the choices indicated in Figure 2. In the jargon of robotics, the coordinate system attached to link l_3 is called the tool frame. The coordinate system attached to link l_0 is called the base frame and coincides with the laboratory system introduced in Section 3. The angle between the x_{k-1}-axis and the x_k-axis, measured about the z_{k-1} axis, is denoted by θ_k, $k = 1, 2, 3$. Each triple $(\theta_1, \theta_2, \theta_3)$ defines a configuration of the C-arm system. Figure 2 illustrates the configuration associated with $(\theta_1, \theta_3, \theta_3) = (0, -\pi/2, 0)$.

A point in space may be specified by its coordinates in either of the four coordinate systems. It is explained in [10] how to transform the coordinates of the point from the tool frame to its coordinates in the base frame: If the point is represented by $\boldsymbol{x}_3$ in the tool frame, then it is represented by

$$\boldsymbol{x}_0 = \boldsymbol{R}(\theta_1, \theta_2, \theta_3)\boldsymbol{x}_3 \tag{9}$$

in the base frame, where $\boldsymbol{R}(\theta_1, \theta_2, \theta_3)$ is a rotation matrix. This matrix can be derived from the information provided in Figure 2 and is given by (see also example 3.3.3 in [10])

$$\boldsymbol{R}(\theta_1, \theta_2, \theta_3) = \begin{pmatrix} c_1c_2c_3 - s_1s_3 & -c_1c_2s_3 - s_1c_3 & c_1s_2 \\ s_1c_2c_3 + c_1s_3 & -s_1c_2s_3 + c_1c_3 & s_1s_2 \\ -s_2c_3 & s_2s_3 & c_2 \end{pmatrix}$$

with $c_k = \cos\theta_k$ and $s_k = \sin\theta_k$, $k = 1, 2, 3$.

In the tool frame, the source is always at position $\boldsymbol{x}_{\text{src}} = (-r_{\text{src}}, 0, 0)^T$, where r_{src} is the distance between source and isocenter. When the angles θ_1, θ_2, θ_3 are chosen as functions of the parameter $\lambda \in \Lambda$, then $\boldsymbol{x}_{\text{src}}$ moves along the trajectory

$$\boldsymbol{a}(\lambda) = \boldsymbol{R}(\theta_1(\lambda), \theta_2(\lambda), \theta_3(\lambda))\boldsymbol{x}_{\text{src}}, \quad \lambda \in \Lambda \tag{10}$$

in the base frame. Thus a trajectory may be specified by three angular functions $\theta_k(\lambda)$, $k = 1, 2, 3$.

To design trajectories that are complete with respect to B_0, we may simply guess appropriate angular functions. Whether a candidate trajectory is complete with respect to B_0 may be checked as described in Section 3. The mechanical and electromechanical constraints of the C-arm system will have to be obeyed. Angle θ_1 is preferably chosen constant, but for a complete trajectory the other two angles must vary. Accordingly, the C-arm system must allow a simultaneous rotation of the C-arm and the C-arm suspension while cone-beam projections are being taken.

To realize a trajectory, the parameter λ is replaced by a smooth, monotonically increasing function of time, say $\tau : [t_-, t_+] \to \Lambda$, and the resulting functions $\theta_k(\tau(t))$ are used to drive the corresponding servo motors of the C-arm system. For physical reasons the first and second derivatives of the entailing trajectory $\boldsymbol{a}(\tau(t))$ must vanish as $t \downarrow t_-$ and $t \uparrow t_+$. This may be achieved by a proper choice of the function τ. Alternatively, the original trajectory may be augmented at both ends with extra segments for acceleration and deceleration.

The left panel of Figure 3 illustrates a favorable source trajectory found in this way. In this and the following examples, the distance between source and isocenter is 810 mm, and the radius of B_0 is about 125 mm. The right panel of Figure 3 shows 28 umbrella-shaped surfaces within B_0, as described in Section 3. As can be seen, this trajectory is complete. Figure 4 illustrates four additional complete trajectories. The angular functions for all five example trajectories are given in Table 1. The definitions involve some magic constants which depend on the geometrical parameters of the underlying C-arm system. As indicated in the table, these example trajectories have various merits and may impose additional technical requirements on the C-arm system.

As with a circular arc, the true trajectory will deviate a little from the prescribed trajectory, but these deviations can be measured and taken into account during the reconstruction. With high-end C-arm systems the deviations are also reproducible.

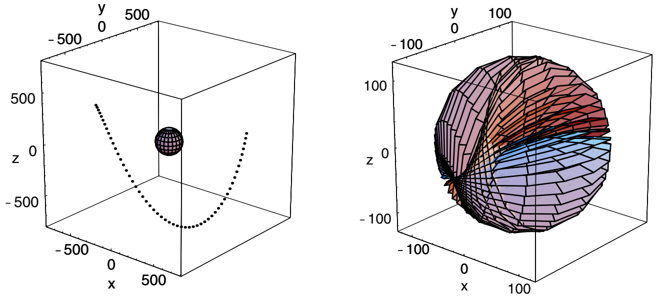


Figure 3. A complete source trajectory. The small ball in the left panel represents B_0. The right panel illustrates the filling of B_0 with umbrella-shaped surfaces, as described in Section 3.

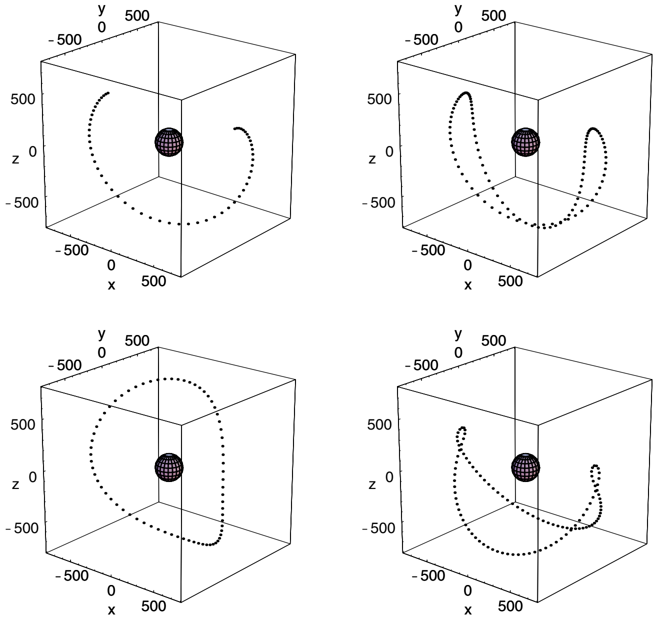


Figure 4. Four additional complete source trajectories.

Illustration	Definition ($0 \leq \lambda \leq 1$)	Comments
Figure 3	$\theta_1(\lambda) = -\frac{\pi}{180}36$ $\theta_2(\lambda) = \frac{\pi}{180}(22.5 - 225\lambda)$ $\theta_3(\lambda) = \frac{\pi}{180}(45 - 90\lambda)$	L-arm at −36° θ_1, θ_2 linear
Figure 4 top left	$\theta_1(\lambda) = 0$ $\theta_2(\lambda) = \frac{\pi}{180}(22.5 - 225\sin^2(\pi\lambda/2))$ $\theta_3(\lambda) = \frac{\pi}{180}18\sin(2\pi\lambda)$	L-arm at 0°
Figure 4 top right	$\theta_1(\lambda) = 0$ $\theta_2(\lambda) = \frac{\pi}{180}(22.5 - 225\sin^2(\pi\lambda))$ $\theta_3(\lambda) = \frac{\pi}{180}12\sin(4\pi\lambda)$	L-arm at 0° trajectory closed
Figure 4 bottom left	$\theta_1(\lambda) = 0$ $\theta_2(\lambda) = \frac{\pi}{180}(-90 - 360\lambda)$ $\theta_3(\lambda) = \frac{\pi}{180}15\cos(4\pi\lambda)$	L-arm at 0° trajectory closed requires 360° rotation for j_2
Figure 4 bottom right	$\theta_1(\lambda) = -\frac{\pi}{180}90$ $\theta_2(\lambda) = \frac{\pi}{180}a_2(\lambda)$ $\theta_3(\lambda) = \frac{\pi}{180}a_3(\lambda)$	L-arm at −90° trajectory closed a_2, a_3 shown in Figure 5 requires 210° rotation for j_3

Table 1. The definitions of the five example trajectories.

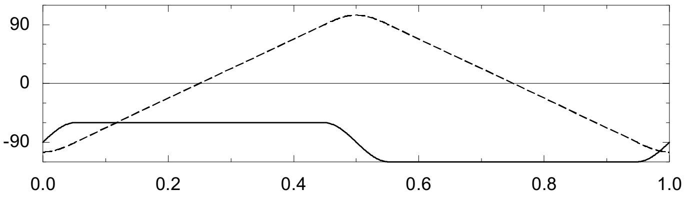


Figure 5. The functions a_2 (solid) and a_3 (dashed) referenced in the last row of Table 1.

References

[1] M. Grass, R. Koppe, E. Klotz, R. Proksa, M. H. Kuhn, H. Aerts, and J. Op de Beek, "Three-dimensional reconstruction of high contrast objects using C-arm image intensifier projection data," *Computerized Medical Imaging and Graphics*, vol. 23, pp. 311–321, 1999.

[2] K. Wiesent, K. Barth, N. Navab, T. Brunner, O. Schuetz, and W. Seissler, "Enhanced 3D-reconstruction algorithms for C-arm systems suitable for interventional procedures," *IEEE Transactions on Medical Imaging*, vol. 19, no. 5, pp. 391–403, May 2000.

[3] L. A. Feldkamp, L. C. Davis, and W. J. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Amer. A*, vol. 1, no. 6, pp. 612–619, June 1984.

[4] C. Hamaker, K. T. Smith, D. C. Solmon, and S. L. Wagner, "The divergent beam X-ray transform," *Rocky Mountain Journal of Mathematics*, vol. 10, no. 1, pp. 253–283, 1980.

[5] R. Koppe, E. Klotz, J. Op de Beek, and H. Aerts, "3D vessel reconstruction based on rotational angiography," in *CAR '95: Computer Assisted Radiology and Surgery*, H. U. Lemke, K. Inamura, C. C. Jaffe, and M. W. Vannier, Eds. 1995, pp. 101–107, Springer.

[6] R. Fahrig and D. W. Holdsworth, "Three-dimensional computed tomographic reconstruction using a C-arm mounted XRII: Image-based correctin of gantry motion nonidealities," *Medical Physics*, vol. 27, no. 1, pp. 30–38, January 2000.

[7] R. Clack and M. Defrise, "Cone-beam reconstruction by the use of Radon transform intermediate functions," *J. Opt. Soc. Amer. A*, vol. 11, no. 2, pp. 580–585, February 1994.

[8] M. Defrise and R. Clack, "A cone-beam reconstruction algorithm using shift-variant filtering and cone-beam backprojection," *IEEE Trans. Med. Imag.*, vol. 13, no. 1, pp. 186–195, March 1994.

[9] R. M. Lewitt, "Processing of incomplete measurement data in computed tomography," *Medical Physics.*, vol. 6, no. 5, pp. 412–417, 1979.

[10] M. W. Spong and M. Vidyasagar, *Robot Dynamics and Control*, John Wiley & Sons, 1989.

Computation of Unmeasured 3^{rd} Generation VCT Views from Measured Views - Preliminary Results

S. K. Patch

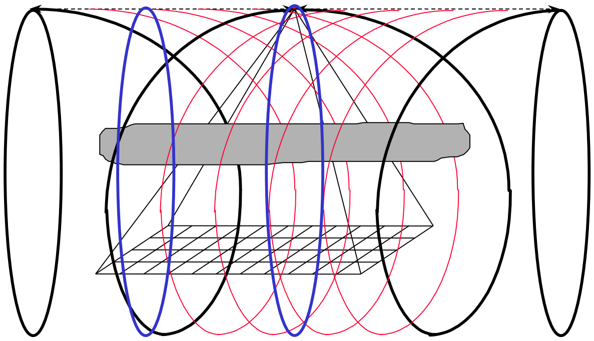


Figure 1. Measured source trajectory shown in black. Axial scans immediately precede and follow a helix, providing mathematically complete data for unbounded objects. Views corresponding to focal spot positions on a neighboring helix, shown in red, are computed at each iteration of our numerical PDE solver. Data corresponding to axial scans shown in blue at $z = 0$ and $z = -0.063$m were saved to reconstruct in those planes by standard 2D filtered backprojection.

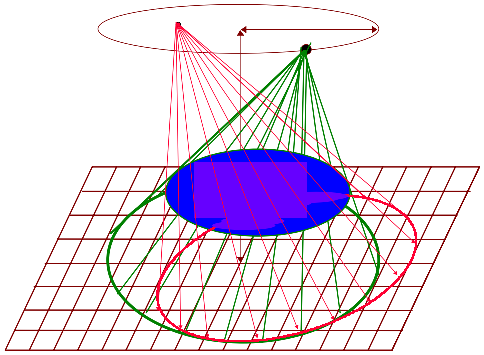


Fig. 2. Source trajectory is a horizontal circle above the patient. Unmeasured cone-beam projections are computed for focal spot positions within the circle

Abstract— **We generalize work done in [1] by computing unmeasured cone beam projections from measured projections. We do this by solving a characteristic boundary value problem for an ultrahyperbolic differential equation [2]. One potential use for this technique is reduction of cone-angle artifacts suffered by approximate volumetric reconstruction techniques, including Feldkamp. By working in the Fourier domain, we convert the 2^{nd} order PDE into a family of 1^{st} order ODE's. A simple 1^{st} order integration is used to solve the ODEs.**

Keywords— **ultrahyperbolic partial differential equation, cone-beam CT**

I. Introduction

We compute unmeasured volumetric computed tomography (VCT) views from measured views by enforcing range conditions [2] requiring that VCT data satisfy the ultrahyperbolic partial differential equations:

$$\left(\frac{\partial^2}{\partial\eta_i\,\partial\xi_j} - \frac{\partial^2}{\partial\eta_j\,\partial\xi_i}\right) u(\xi;\eta) = 0 \qquad \text{for } i,j = 1,2,3 \tag{1}$$

By solving a characteristic boundary value problem for these equations, unmeasured views corresponding to unmeasured axial scans are computed. See Figure 1.

Fritz John's range conditions in equation 1 were published in '38 and the idea of computing unmeasured views has been considered for tomosynthesis type systems. In [1] Edholm & Danielsson showed that *cone beam* projections measured by a circular source trajectory lying in a single plane could be used to compute *parallel beam* projections of a *different* object. Their derivation uses fundamental geometric arguments, as does the derivation in [3] showing the necessity of John's equation 1 for xray-transforms of smooth, compactly supported functions. The same measured cone beam projections are used to compute *unmeasured cone beam* projections of the *same* object. See Figure 2. The projections are computed in [3] much as we do in this paper, by solving John's equation 1 in the Fourier domain. The tomosynthesis geometry allowed us to solve analytically, whereas the helical system we consider here requires a change of variables which creates a nasty forcing term in the right-hand side of the transformed equation. We have not found an analytical solution, so the results presented here were computed using crude first-order integration.

John first introduced the ultrahyperbolic equation in a different form, equivalent to 1 by a linear change of variables. The original equation's variables did not correspond to the xray transform used in computed tomography, but did permit a mean-value theorem [4] and even analytic solutions for special geometries [5],[6]. A mathematically exact inversion formula exploiting John's equation was presented for "bounded" objects in [7].

Normalized VCT data measures line integrals of a three dimensional imaging object's linear attenuation coefficient (LAC). The 1-dimensional source trajectory along which we measure cone beam projections with a 2-dimensional

Applied Science Lab, GE Medical Systems, PO Box 414, Milwaukee, WI 53201 email: sarah.patch@med.ge.com

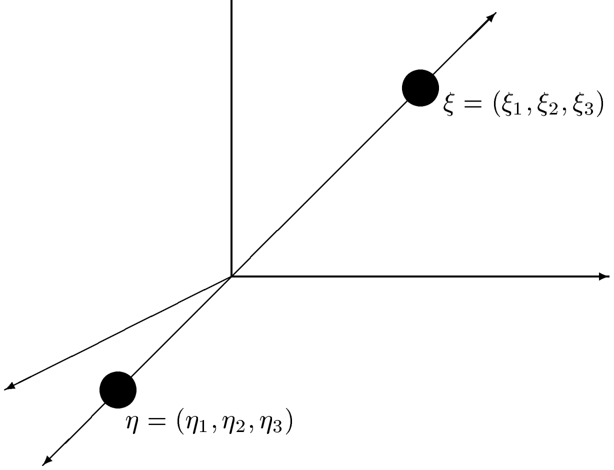


Fig. 3. $X\rho(\eta;\xi)$ integrates ρ along the line passing through η and ξ.

xray detector with pixels parameterized by (α_1, α_2), providing $1 + 2 = 3$ dimensions worth of boundary value (BV) data on a characteristic surface. For a constant-pitch helix of pitch $p \geq 0$ where the FS moves a distance $\frac{p}{2\pi}$ during a single rotation, we measure on the $1 + 2 = 3$ dimensional surface defined by

$$z \quad = \quad p\theta \qquad \text{where } \alpha_1, \alpha_2 \text{ are free}$$

We change variables to parameterize 3^{rd}-Gen data and rewrite the consistency conditions in the new coordinate system. Notice that 1 is written in terms of six variables. Figure 3. Our measured data represents line integrals of a function defined in R^3 and therefore should be a function of three independent variables. However, in a 3^{rd}-Gen system, the radii of rotation are fixed so we only measure u for varying θ, z, α_1, and α_2, where (θ, z) parametrize focal spot (FS) positions and (α_1, α_2) correspond to pixel location on a flat panel xray detector. See Figure 1. In order for our dimension count to be correct there can be only one *independent* constraint upon u. Modulo first order identities on u, all three of the conditions in 1 boil down to the single constraint:

$$\frac{\partial u_{\alpha_2}}{\partial \theta} - \rho\frac{\partial u_{\alpha_1}}{\partial z} = \frac{-1}{(\rho+d)}\left(2\alpha_1 u_{\alpha_2} + \alpha_1\alpha_2 u_{\alpha_2,\alpha_2} + \left[(\rho+d)^2 + \alpha_1^2\right] u_{\alpha_2,\alpha_1}\right) \tag{2}$$

subject to the boundary conditions *which we measure*

$$u(\theta, p\theta; \alpha_1, \alpha_2) \quad = \quad f(\theta; \alpha_1, \alpha_2) \tag{3}$$

where ρ and d represent source-to-iso and iso-to-detector distances, respectively. Notice that the left hand side of 2 is first order with respect to z and θ. Although standard numerical solvers for partial differential equations (PDEs) can be used to solve 2, it is also possible to transform this single PDE into a coupled system of ordinary differential equations (ODEs), which we solve with a first-order integration. Numerical results are presented in Section II.

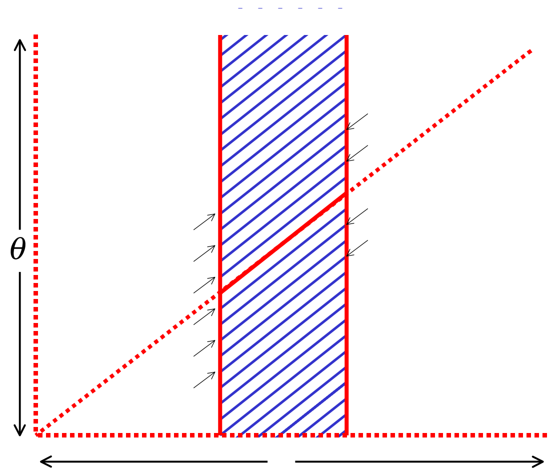


Fig. 4. Source trajectory for results presented here is comprised of circles connected by a short helical segment, shown in bold red. Each iteration of the PDE solver computed views along one of the blue helical segments. Ideally, the helix is longer, but computational costs prohibited us from using a full rotation helix to generate these preliminary results.

II. Numerical Results

Preliminary results are presented here for a source trajectory failing the Kirillov-Tuy completeness condition. We consider here only data measured on a short segment of a constant pitch helical segment source trajectory. Results are computed both with and without noise in the measured data. The incompleteness of our measured data implies that a wedge of Fourier components comprising less than 2% of the Fourier components of our computed projections are inaccessible. We therefore fill in these missing Fourier data both with and without noise. We hope to extend our numerical solver to handle complete source trajectories consisting of two axial scan trajectories connected by a constant pitch helical scan. (Remember, an axial scan corresponds to pitch $p = 0$!)

A. System Configuration

Noise-free test data of a simple phantom object was simulated assuming a 512×512 array with pixel pitch of $1.5e - 3\ m$, source trajectory with helical pitch 150 measured at isocenter, and 984 views per 2π gantry rotation. The FS moves on a radius of $0.541m$, whereas the detector is slightly closer to the z-axis, rotating at a distance of $0.408m$. With this system geometry and helical pitch, the FS moves $0.128m$ in a single rotation. Note that we are working nowhere near the theoretical maximum pitch for exact reconstruction. With this (perhaps suboptimal) discretization, $dz <<< d\theta$ and at each iteration of our PDE solver, only a few neighboring projections contribute to the next step. See Figure 4. In the tests presented here we solved only for projections on a thin ring around the center of our phantom to reduce computation time.

Our phantom object is unbounded, consisting of one large ellipsoid full of water and containing several homo-

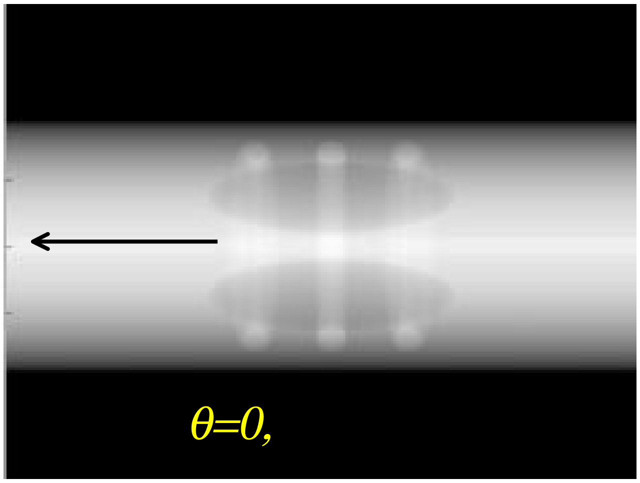


Fig. 5. Projection measured with FS at $(\theta, z) = (0, 0)$. The spine parallel to the z-axis is barely visible.

geneous density inclusions: one thin rod parallel to the patient axis, two smaller ellipsoidal air pockets which are surrounded by three higher density rings. The rod simulates a spine, the air pockets simulate lungs, and the rings simulate ribs. Object centers, dimensions, and densities listed below:

tissue type	center (cm)	eccentricities/radii (cm)	density = LAC
water body	$(0, 0, 0)$	$(20, 10, 1e6)$	190
spine	$(-4, 0, 0)$	$(1, 1, 1e6)$	304
lung_1	$(2, -4, 0)$	$(3, 3, 8)$	0
lung_2	$(2, 4, 0)$	$(3, 3, 8)$	0
rib_1	$(0, 0, -5)$	$1/8.5$	304
rib_2	$(0, 0, 0)$	$1/8.5$	304
rib_3	$(0, 0, 5)$	$1/8.5$	304

Table 1. Different tissue types, their sizes, positions, and attenuation coefficients are listed above.

The linear attenuation coefficient (LAC), $u \in C_0$ is bounded and for each (θ, z) pair, has compact support in α_1, and α_2. *However,* the LAC is "unbounded" with respect to the detector, since the water body and spine are longer than the detector. We should note that this is a relatively benign with little frequency content in the z-direction. A projection taken with focal spot in the central plane is shown in Figure 5.

B. *Noise-Free Results*

Boundary value data as described in 3 from source positions on the helical source trajectory were first simulated without noise and used to solve equation 2. The central columns from each projection with a full of spot position in the middle blue circle shown in Figure 1 were compiled to create a 2D axial sinogram. The reconstruction of that central plane is compared to the mathematically exact image in Figure 6. Notice the slight shading across the image and broadening of the rib. These errors are due to inaccuraies in our first-order numerical solver, since our data is noise-free. Vertical and horizontal profiles across this same reconstructed slice are shown in Figure 7.

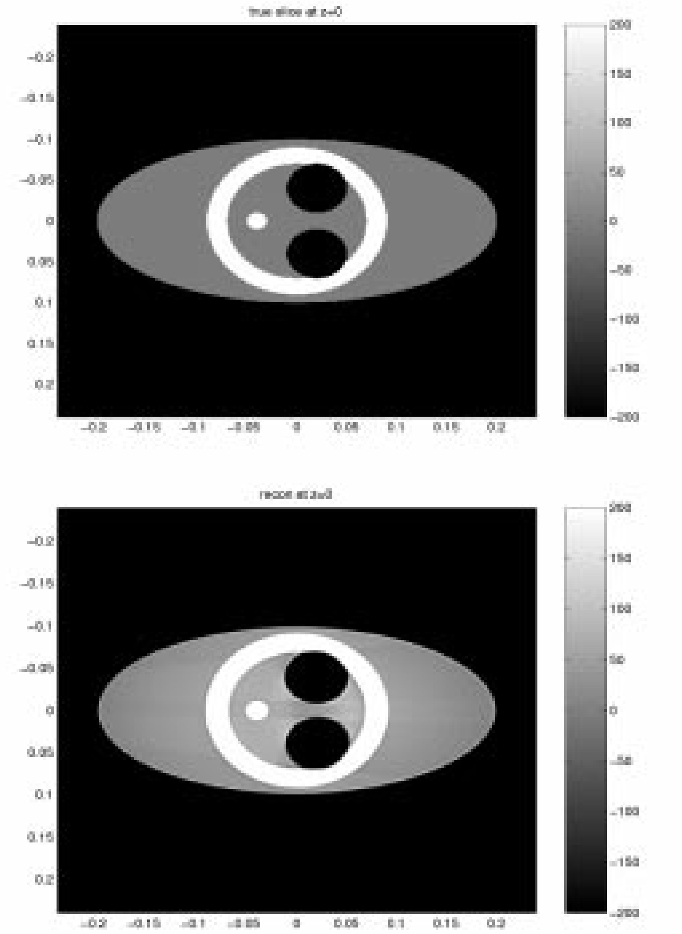


Fig. 6. Reconstruction via standard 2D filtered-backprojection of true and computed sinograms. No noise was added to the projections from which the second sinogram was computed.

C. *Noisy Results*

The same code was run to test robustness to noise where both additive and multiplicative noise were added to all "measured" projections as follows:

$$u_{noisy}(\theta, \alpha_1, \alpha_2) = u_{true}(\theta, \alpha_1, \alpha_2)(1 + 0.005X) + 0.05Y \tag{4}$$

where $X, Y \in N(0, 1)$. Projections in the axial sinogram computed from noisy BV data are compared to noisy sinogram projections with the same noise levels in Figure 8. Whether the differences are repeatable for different realizations of the experiment remains to be determined.

III. Conclusion

For "bounded" objects, 3^{rd}-Gen helical VCT data can be reconstructed exactly once the 3^{rd}-Gen version of John's equation has been solved. To reconstruct a volume using

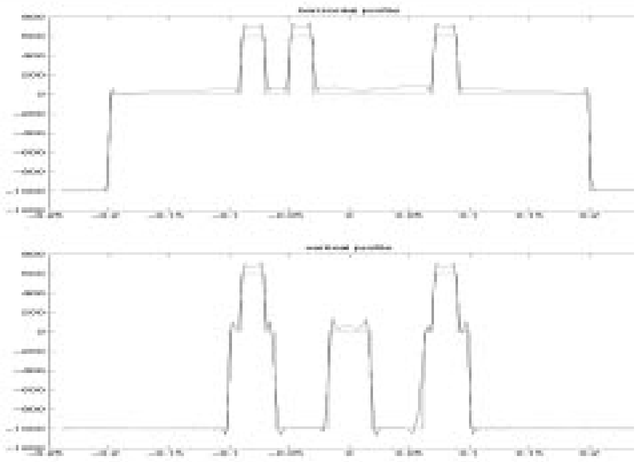


Fig. 7. Notice that Gibbs ringing is the largest source of error in the vertical profile (bottom) but that the shading artifact can be seen in the horizontal profile (top). In both cases, locations of edges are accurately recovered.

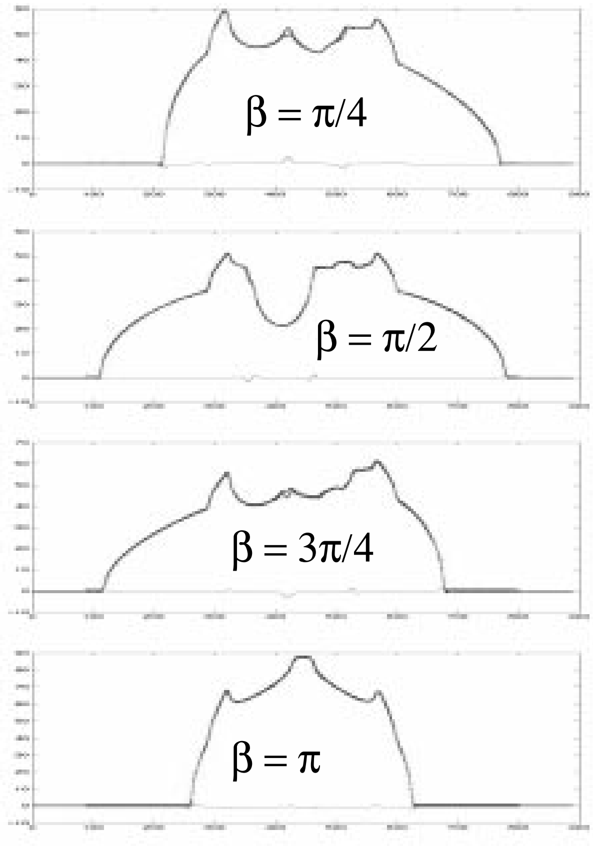


Fig. 8. Profiles of projections computed from noisy boundary value data as well as noisy projections measured directly are both plotted in heavy dots. The difference between these noisy profiles is plotted below in a thin black line.

2D filtered backprojection, we solve for all projections corresponding to focal spot positions on the cylinder, saving at each iteration only data on the central columns to generate 2D axial sinograms. Because our numerical scheme was low order, we were forced to solve for all projections anyway. A higher-order solver would allow us to take larger step sizes, and more importantly, improve our solution accuracy. However, a high-order scheme is elusive [10].

Our next task in this effort is to incorporate boundary value data from 2 circles + helix source trajectory, eliminating the need to "fill in" inaccessible Fourier components. This is straightforward numerical work. Our next task, developing an exact method for "unbounded objects" will require more effort. Views of unbounded objects are not exactly recovered using this technique. Whoever as the size of the flat panel detector rate increases errors due to the unbounded object problem decrease. Although the errors are likely to be small, a mathematically exact solution analogous to that in [8],[9] is required to complete our analysis.

REFERENCES

[1] Edholm, P.R., Danielsson, P.E., "A Theorem on Divergent Projections," *Three Dimensional Image Reconstruction in Radiation and Nuclaer Medicine*, P. Grangeat & J.L. Amans, eds., pp. 35-45, (1996).

[2] John, F., "The Ultrahyperbolic equation with 4 independent variables", *Duke Math. Journal*, pp. 300-322, (1938).

[3] Patch, S. K., "Almost-Everywhere Extrapolation using 2D Transforms from Cone Beam Data", U.S. Patent #6,173,030, (2001).

[4] Asgiersson, L., ""Ueber eine Mittelwertseigenschaft von Loesungen homogener linearer partieller Differentialgleichungen zweiter Ordnung mit konstanten Koeffizienten", Mathematische Analen, 13, pp. 321-346, (1938).

[5] Owens, G., "An Explicit formula for the solution of the Ultrahyperbolic Equation in four Variables", Duke Math Journal, 9, pp. 272-282, (1942).

[6] Owens, G., "A Boundary-Value Problem for Analytic Solutions of an Ultrahyperbolic Equation" Duke Math Journal, 20, pp. 29-38, (1953).

[7] D. Finch, "Cone Beam Reconstruction with Sources on a Curve," *SIAM J. Appl. Math*, **45** #4, pp. 665-673, (1985).

[8] Tam, K., US Patent #5,390,111, "Method and System for Processing Cone Beam Data for Reconstructing Free of Boundary-Induced Artifacts a Three Dimensional Computerized Tomography Image," 1993.

[9] Kudo H., Noo, F., Defrise, M., "Cone-beam filtered-backprojection algorithm for truncated helical data," Phys. Med. Biol., 43, pp. 2885-2909, (1998).

[10] J. Hu, C. Ingrassia, S. Lowitsch, J. Park, A. Pineda, D. Reynolds, N. Valdivia, *Final Report of the IMA Summer Workshop of Mathematical Modeling in Industry*, July, 2000, posted at www.ima.umn.edu.

[11] Greenleaf, A., Uhlmann, G., "Non-Local Inversion Formulas for the X-ray Transform," *Duke Mathematical Journal,* **58**, no. 1, pp. 205-240, (1989).

[12] Kirillov, A.A., "On a Problem of I.M. Gelfand", Sov. Doklady, pp. 268-269.

[13] Tuy, H.K., "An inversion formula for cone-beam reconstruction," SIAM J. Applied Math, 43, pp. 546-552, (1983).

[14] Grangeat, P., "Mathematical Framework of cone beam 3D reconstruction via the first derivative of the Radon transform," Mathematical Methods in Tomography - Lecture Notes in Mathematics, pp. 66-97, (1991).

[15] Helgason, S., *The Radon Transform,* Birkhauser, 1980.

[16] Natterer, F., *The Mathematics of Computerized Tomography,* Wiley, 1986.

A Generalized Model for 3D Compton Scatter in Single Photon Emission Computed Tomography Using Slice-by-Slice Blurring and Scatter Projection Rebinning

Chuanyong Bai[1], Gengsheng L. Zeng[2], Grant T. Gullberg[2]
[1]ADAC Laboratories, 540 Alder Drive, Milpitas, CA95035
[2]MIRL/CAMT, 729 Arapeen Drive, SLC, UT84108

Abstract

The modeling of 3D Compton scatter in SPECT is complicated by the use of a variety of collimator geometry. To date, effective and efficient modeling of 3D Compton scatter in parallel-beam SPECT has been made possible by Frey and Tsui [1], Zeng et al [2], and Bai et al [3]. For fan-beam and cone-beam SPECT, the model in [3] can effectively model first order 3D Compton scatter. But for varying-focal-length fan-beam and cone-beam SPECT, no effective scatter models have been developed so far.

We propose in this work a generalized model for 3D first order Compton scatter in SPECT for different collimator geometry. The slice-by-slice blurring technique in [3] is used to generate a complete set of first order Compton scatter projections; the scatter projections are then properly rebinned according to the specific collimator geometry to obtain the scatter estimation. When generating the complete set of scatter projections, a set of virtual planar detectors with parallel-beam collimation are used to orbit around the object for scatter detection. Detailed discussion of this generalized model is provided in the METHODS section.

I. INTRODUCTION

In SPECT data acquisition, Compton scattered events are detected inevitably along with the primary (non-scattered) events due to the finite energy resolution and finite spatial resolution of the detectors. A photon usually loses part of its energy and changes its direction after being Compton scattered. Improper handling of the scattered photons can thus lead to the incorrect positioning of the source photons, and as a consequence, the reconstructed images can be degraded both quantitatively and qualitatively.

The primary difficulty for Compton scatter compensation in SPECT imaging is to obtain accurate scatter estimation. There are two major categories of techniques to obtain scatter estimation. One is the direct scatter estimation from data acquired using scatter energy window(s). Techniques in this category are easy to implement but not accurate. The other one is the modeling of the scatter from the attenuation map and the emission map during image reconstruction. Techniques in this category can be accurate but are difficult to implement, especially when complicate collimator geometry is used.

There are many techniques developed for the modeling of Compton scatter in SPECT. 1) Using Monte Carlo simulation for each specific scan [4-6]. 2) Using physical measurement to obtain the scatter response of a specific scan setup, the measured response is used as a characteristic of the given SPECT system [7]. 3) Using ideal integration based on the scatter cross-section described by the Klein-Nishina formula [8-9]. 4) Using ray driven projector/backprojector to model first order Compton scatter [10-12]. 5) Using a slab derived scatter model [13-18]. 6) Using an effective scatter source estimation model [1,19]. And 7) using a slice-by-slice blurring technique to model 3D first order Compton scatter [2-3, 21]. While almost all these techniques are effective to model scatter for parallel-beam SPECT, only [3] and [20] have shown to be effective and efficient for fixed-focal-length fan-beam and [3] for fixed-focal-length cone-beam SPECT.

In this work, we develop a generalized model for 3D first order Compton scatter in SPECT, using the slice-by-slice blurring technique [3] and scatter projection rebinning technique. The application of this model in SPECT with slant-hole collimator, annular collimator, and fan-beam and cone-beam collimators with both fixed-focal-length and varying-focal-length are discussed.

II. METHODS

A. The slice-by-slice blurring technique for first order Compton scatter

There are three basic principles for the slice-by-slice blurring model for first order Compton scatter described in [3]. (1) For each projection angle, photons emitted from a point source and scattered at a scattering point can only be detected at a certain position at the detector surface due to the assumption of perfect collimation. This means that when the position of point source is given, there is a one-to-one relationship between the positions of the scattering point and the detecting point, and thus the scattering angle for the detected scattered photons is unique for each scattering point. Figure 1 shows an example in fan-beam geometry. (2) For a given point source, on a scattering slice which is parallel to the detector surface, the photon energy and scattering angle dependent factor of the scatter probability (given by the Klein-Nishina formula) has a 2D distribution. This 2D distribution can be approximated using a 2D Gaussian function, considering the energy detection probability, the energy of the incident photon, and the detection energy window. (3) The 2D Gaussian function of the scattering slice which is closer to the detector surface (and thus further from the point source) can be obtained by a slice-by-slice blurring from the 2D Gaussian function of a scattering slice which is further away from the detector surface.

The generation of scatter estimation using this technique includes three steps. Step 1: the generation of the intermediate scatter source image (ISSI) from the attenuation map using slice-by-slice blurring. The value of each element of ISSI is the production of the source intensity, the effective attenuation from the point source to the scattering voxel, and the scattering angle dependent factor of the scatter probability at the scattering voxel. The solid angle subtended by the element with the origin at the point source is also considered. Step 2: the generation of the effective scatter source image (ESSI) from ISSI, which is a voxel-by-voxel multiplication operation of the ISSI and the linear attenuation coefficient map. And step 3: the projection of the ESSI using a slab-by-slab blurring model to generate the final scatter estimation [21].

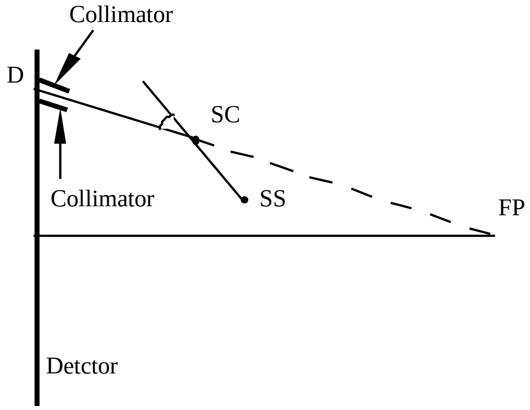


Figure 1. Photons that are originated from source SS and scattered at scattering point SC can only be detected at point D on the detector surface under the assumption of perfect collimation.

The slice-by-slice blurring model includes the non-uniform attenuation from the source to scattering point in Step 1 and from the scattering point to the detecting point in Step 3. The geometric point response effect of the SPECT system can also be included in Step 3. In Step 2, the effective non-uniform scatter probability of a scattering voxel is considered. Using this model, one can accurately estimate 3D first order Compton scatter along a give direction for any scattering media.

B. A generalized model for 3D Compton scatter

When photons shine on a scattering voxel, they can be scattered to a 4• solid angle, theoretically. We can divide this 4• solid angle to a series of sub-solid angles. The scattered photons in each sub-solid angle can then be detected by using a planar detector with parallel beam collimation. In this way, we can obtain a complete detection of the photons that are scattered by the scattering voxel. Figure 2 illustrates the basic idea of generating a complete set of scatter projections.

The generalized model includes the following two steps. **Step 1.** Using the slice-by-slice blurring model to generate a set of scatter projections, assuming that we use a set of virtual planar detectors with parallel collimation to detect the scattered photons. Each parallel scatter projection contains the scatter estimation for the corresponding sub-solid angle, and thus the set of parallel scatter projections form a complete scatter estimation. **Step 2.** For a given SPECT geometry, using an appropriate rebinning process to obtain the scatter estimation for all the detector bins. In this step, the set of generated scatter projections in Step 1 is used for rebinning.

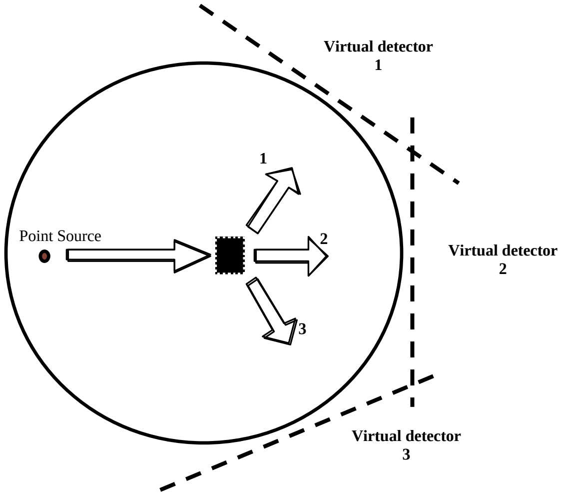


Figure 2. A complete scatter estimation can be obtained through slice-by-slice blurring model, using one virtual planar detector with parallel collimation at each sub-solid angle.

C. Implementation of the model

1. Slant hole collimation, planar detectors

For planar detectors with slant hole collimation [22], the collimation in each segment is essentially parallel. Figure 3 shows the geometry of a bilateral slant-hole collimator. The collimation directions of different segments are different. For each segment, one only needs to generate a series of parallel projections, which have different distances in the direction of the collimation, using virtual planar detectors (dashed lines in Figure 3) with parallel collimation. For a detector bin *i* on the slant detector, the rebinning procedure is to first determine its interception position (***k***, i) with one of the parallel scatter projections ***k*** , and then estimate its scatter using the scatter values of the bins closest to (***k***, i) of projection ***k***. If we use $bin\_size \cdot \cos(\beta)$ as the bin size in the slant direction, and *bin_size* in the parallel direction, no interpolation is needed for the rebinning. Here *bin_size* is the slant-hole collimator bin size and β is the slant angle. For this special collimation geometry, one can incorporate the asymetric geometric point response of the detector in Step 3 of the slice-by-slice blurring model during the generation of complete scatter estimation.

2. Parallel collimation, annular detectors

Figure 4 shows the geometry (transaxial cross section) of a SPECT system with annular detector and parallel collimation. In the axial direction of the detector, the collimation is also parallel. A complete set of scatter projection can be obtained by using a set of virtual parallel detectors evenly distibuted along the annular ring. For a detector bin corresponding to angle C in Figure 3, an accurate way to estimate its scatter is to generate a virtual parallel scatter projection (with same bin size as the annular detector) for this angle and pick up the scatter value directly. If only the parallel projection in the neighboring angles A and B are generated for the sake of decreasing computational burden, an interpolation step should be taken for the rebinning. Explicitly, we first find the bins (A, i) and (B, j) on the planar parallel detectors at angles A and B that correspond to the detector bin in direction C. If their values are S(A, i) and S(B, j), then the scatter at the detector bin in angle C can be expressed as:

$$\left[S(A,\ i)\cdot(B-C)+S(B,\ j)\cdot(C-A)\right]/(B-A)\,. \qquad (1)$$

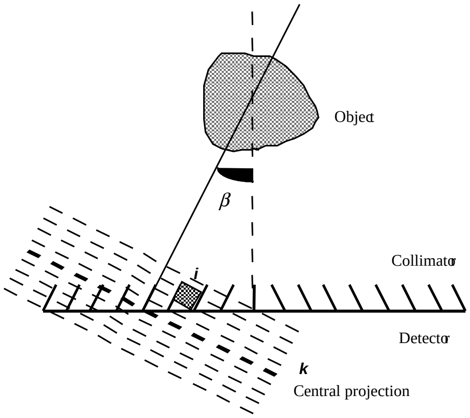


Figure 3. Bilateral slant-hole parallel collimation.

Note that here we simply consider the angular interpolations between neighboring projections. To be more accurate, a small compensation of solid angle difference should be made due to the longer distances from (A, i), and (B, j) to the center of rotation than the detector bin at angle C. One can project the bin to parallel projections A and B and then sum up the values of the projected bins. Thus equation (1) becomes:

$$\left[\sum_{i} S(A,i)\cdot(B-C)+\sum_{j} S(B,j)\cdot(C-A)\right]/(B-A) \qquad (2)$$

3. Fixed-focal-length fan-beam and cone-beam collimation, plane detectors

Figure 5 illustrates the geometry of a SPECT system with fixed-focal-length fan-beam collimation in the fan direction and parallel collimation in parallel direction. For each angle from the focal point to a detection bin in the fan direction, we generate a scatter projection using a virtual planar detector (dashed lines in Figure 5) with parallel collimation. The rebinning step assigns to the detector bin the scatter value of the central bin of the virtual parallel scatter projection. If one uses less number of angles than the number of bins to generate the scatter responses, the scattered value at each detector bin can be obtained with a simple angular interpolation as is the same as described in equation (1). However, when compensating for the changed solid angle, equation (2) should be slightly modified to reflect the fact that the projected bin(s) on one of the neighboring virtual projection are closer to the center of rotation, and thus the subtended angles are increased.

When performing SPECT scans using fixed-focal-length fan beam SPECT systems, sometimes a non-circular orbit is used [23] to avoid truncation or to make the line of focus remain fixed to the center of the organ of interest, etc. For these scans, the focal-point orbit function should be used to determine the relative position of the center of object matrix to the focal point and the detector surface at different projection angle for proper scatter projection rebinning and solid angle compensation.

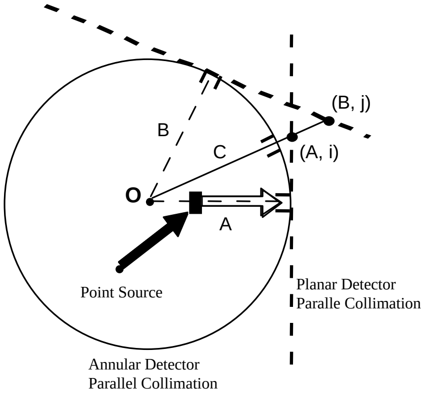


Figure 4. An annular detector with parallel collimation.

For planar detector with fixed-focal-length collimation, the generation of the complete set of scatter projection is three-dimensional. The scatter projection rebinning is also three-dimensional.

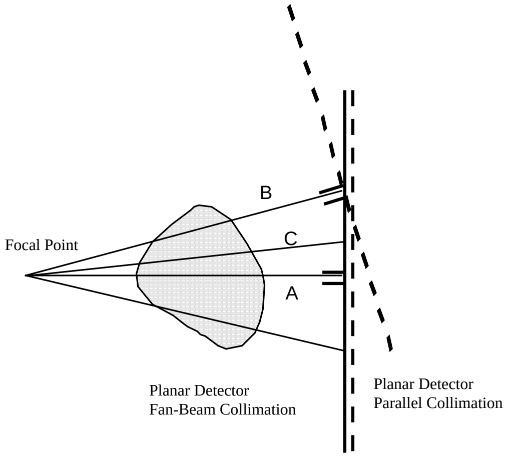


Figure 5. Fixed-focal-length fan-beam geometry.

4. Varying-focal-length fan beam and cone beam collimation, planar detectors

For varying-focal-length fan beam collimation and planar detectors, one can use the same scheme as for fixed-focal-length fan beam collimation and planar detectors to obtain the scatter estimation. The difference is that when rebinning, one should use the focal-length function [24, 25] to determine which parallel projections should be used for rebinning. The direction of these scatter projections are determined by the detector bins and the position of the focal points corresponding to these bins.

For varying-focal-length cone-beam collimation, the generation of complete set of scatter projections should be three-dimensional, and the interpolation described in equation (2) should also be three-dimensional.

It should be noted that for SPECT imaging using varying focal-length fan-beam or cone-beam collimation and planar detector, the focal point closest to the detector surface should be out of the object of interest to avoid multiple images. However, this restriction is not needed for scatter estimation using this generalized scatter model.

III. DISCUSSION

The scatter model proposed in this work is based on the assumption that a complete first order 3D Compton scatter projection can be obtained by using a set of virtual planar detectors with parallel collimation. This assumption is accurate when enough number of parallel scatter projections are generated. Simulation studies for fixed-focal-length fan beam SPECT will be presented. Simulation studies for varying-focal-length fan-beam and cone-beam collimation will be performed when a proper simulator is available.

IV. REFERENCES

[1] Frey EC and Tsui BMW 1996 A new method for modeling the spatially-variant, object-dependent scatter response function in SPECT Record 1996 *IEEE Nucl. Sci. Symp. Med. Imag. Conf.* 1082-1086

[2] Zeng GL, Bai C, and Gullberg GT 1999 A projector/ backprojector with slice-to-slice blurring for efficient 3D scatter modeling *IEEE Trans. Med. Img.* **18** 722-273

[3] Bai C, Zeng GL, and Gullberg GT 2000 A slice-by-slice blurring model and kernel evaluation using the Klein-Nishina formula for 3D scatter compensation in parallel and converging beam SPECT *Phys. Med. Biol.* **45** 1275-1307

[4] Floyd CE, Jaszczak RJ, and Coleman RE 1989 Inverse Monte Carlo: A unified reconstruction algorithm *IEEE Trans. Nucl. Sci.* **36** 779-785

[5] Frey EC and Tsui BMW 1990 Parameterization of the scatter response function in SPECT imaging using Monte Carlo simulation *IEEE Trans. Nucl. Sci.* **37** 1308-1315

[6] Bowsher JE and Floyd CE, Jr. 1991 Treatment of Compton scatter in maximum-likelihood, expectation-maximization reconstruction of SPECT images *J. Nucl. Med.* **32** 1285-1291

[7] Beekman FJ, Frey EC, Kamphuis C, Tsui BMW, and Viergever MA 1994 A new phantom for fast determination of scatter response of a Gamma camera *IEEE Trans. Nucl. Sci.* **41** 1481-1488

[8] Cao ZJ, Frey EC, and Tsui BMW 1994 A scatter model for parallel and converging beam SPECT based on the Klein-Nishina formula *IEEE Trans. Nucl. Sci*. **41** 1594-1600

[9] Wells RG, Celler A, and Harrop R 1998 Analytical calculation of photon distributions in SPECT projections *IEEE Trans. Nucl. Sci.* **45** 3202-3214

[10] Welch A, Gullberg GT, Christian PE, and Datz FL 1995 A transmission-map-based scatter correction technique for SPECT in inhomogeneous media *Med. Phys.* **22** 1627-1635

[11] Welch A and Gullberg GT 1997 Implementation of a model-based non-uniform Scatter correction scheme for SPECT *IEEE Trans. Med. Imag.* **16** 717-726

[12] Laurette I, Welch A, Christian PE, and Gullberg GT A three-dimensional transmission-map-based scatter correction technique for SPECT in inhomogeneous media (submitted to) *Phys. Med. Biol.*

[13] Frey EC, Ju ZW, and Tsui BMW 1993 A fast projector-backprojector pair modeling the asymmetric, spatially varying scatter response function for scatter compensation in SPECT imaging *IEEE Trans. Nucl. Sci.* **40** 1192-1197

[14] Frey EC and Tsui BMW 1993 A practical method for incorporating scatter in a projector-backprojector for accurate scatter compensation in SPECT *IEEE Trans. Nucl. Sci*. **40** 1107-1116

[15] Beekman FJ, Eijkman EGJ, Viergever MA, Borm GF, and Slijpen ETP 1993 Object Shape dependent PSF model for SPECT imaging *IEEE Trans. Nucl. Sci.* **40** 31-39

[16] Beekman FJ and Viergever MA 1995 Fast SPECT simulation including object Shape dependent scatter *IEEE Trans. Med. Imag.* **14** 271-282

[17] Beekman FJ, Kamphuis C, and Viergever MA 1996 Improved SPECT quantitation Using fully three-dimensional iterative spatially variant scatter response compensation *IEEE Trans. Med. Img*. **15** 491-499

[18] Beekman FJ, Kamphuis C, and Frey EC 1997 Scatter compensation methods in 3D iterative SPECT reconstruction: A simulation study *Phys. Med. Biol.* **42** 1619-1632

[19] Kadrmas DJ, Frey EC, Karimi SS, and Tsui BMW 1998 Fast implementation of reconstruction-based scatter compensation in fully 3D SPECT image reconstruction *Phys. Med. Biol*. **43** 857-873

[20] Bai C, Zeng GL, and Gullberg GT 1998b A fan-beam slice-by-slice blurring model for scatter, geometric point response, and attenuation corrections in SPECT using the iterative OS-EM algorithm (abstract) *J. Nucl. Med*. **39** 120P

[21] Bai C, Zeng GL, Gullberg GT, DiFilippo F, and Miller S 1998a Slab-by-slab Blurring model for geometric point response correction and attenuation correction using iterative reconstruction algorithm *IEEE Trans. Nucl. Sci.* **45** 2168-2173

[22] Clack R, Christian PE, Defrise M, and Welch AE 1996 Image reconstruction for a novel SPECT system with rotating slant-hole collimators Conf. Rec. 1995 *IEEE Med. Imag. Conf*. 1948-1952

[23] Gullberg GT and Zeng GL 1995 Backprojection filtering for variable orbit fan-beam tomography *IEEE Trans. Nucl. Sci*. **42** 1257-1266

[24] Zeng GL and Gullberg GT 1994 A backprojection filtering algorithm for a spatially varying focal length collimator *IEEE Trans. Med. Imag*. **13** 549-556

[25] Zeng GL, Gullberg GT 1998 Iterative and analytical reconstruction algorithms for varying-focal-length cone-beam projections *Phys. Med. Biol*. **43** 811-821

Computational Determination of Orlov Volumes

S.D. Metzler, J.E. Bowsher and R.J. Jaszczak

Abstract— **Objective: Orlov has derived the sufficient conditions for adequate two-dimensional projection sampling of a three-dimensional density function in order to reconstruct that density function. This condition may be represented as a curve of vantage angles on a unit sphere of directions. Orlov's condition states that the density function can be unambiguously reconstructed if the curve of vantage angles intersects all great circles on the sphere. The set of points for which Orlov's condition is met may be termed the Orlov volume. Although this volume can be intuitively determined for simple sampling orbits, there is no known algorithm in the literature for determining the volume for the generalized case. Further, the Orlov volume principle may be applied to converging and diverging collimators, in addition to parallel-hole collimators. Methods: We consider a voxelized representation of the volume inside an orbit of a given collimator type. We then construct a digitized version of the vantage points of the voxel for a given camera orbit. We then determine if any great circles can exist on the Orlov sphere without intersecting the vantage curve. Results: We have implemented this algorithm in C++ using Object-Oriented programming techniques. The algorithm considers generic collimator types, of which we have currently implemented slant-hole, parallel-hole and pinhole collimators. Other collimator types can be added without modification to the algorithm. Multiple orbits can be simultaneously considered. Multiple collimator types can also be simultaneously considered. The output is the voxelized volume that meets Orlov's condition. Conclusions: The algorithm has successfully determined the Orlov volume in cases that are easily verified intuitively. It has been used to study the more complex scenarios of pinhole collimators following spiral orbits and simultaneous acquisition of parallel-hole and slant-hole collimators. This technique may be useful for studying sufficient orbits and understanding sampling artifacts.**

Keywords— **Orlov, Reconstruction, Sufficient Sampling**

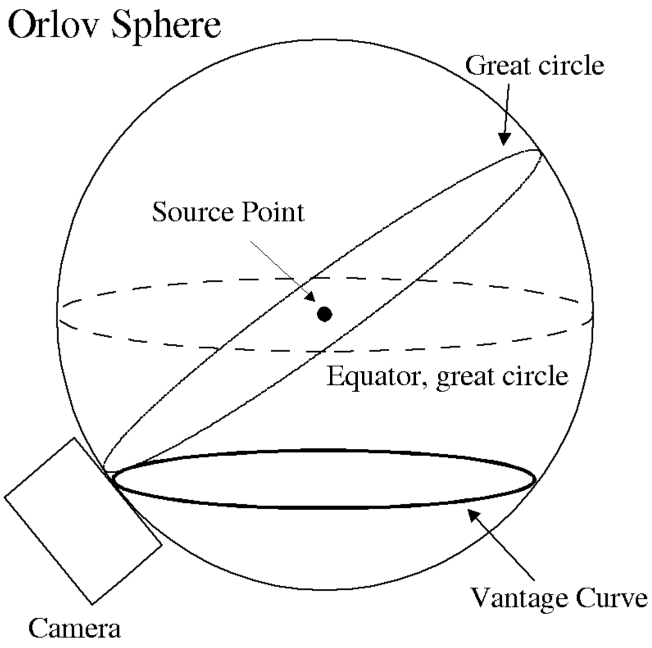


Fig. 1. Orlov Sphere and Vantage Curve for Slant-hole Collimator. The figure depicts the vantage curve for a slant-hole collimator as it observes a source point at the center of the sphere. The camera follows a circular orbit that coincides with the vantage curve. It is possible to draw great circles on this sphere that do not intersect the vantage curve. This vantage curve is the same as the vantage curve of a tilted parallel-hole collimator following the same orbit.

I. INTRODUCTION

In his work with electron microscopy, Orlov derived the sufficient-sampling condition for three-dimensional reconstruction from projection data [1]. He stated his condition geometrically: the curve of vantage angles on a unit sphere of directions must "have points in common with any arc of a great circle [1]." If this condition is met, the sampled density function can be unambiguously reconstructed.

Two Orlov spheres are shown in Figures 1 and 2. Figure 1 shows the vantage curve for a slant-hole or tilted parallel-hole collimator following an orbit that coincides with the vantage curve. It is possible to draw great circles on the Orlov sphere that do not intersect the vantage curve. The projection data derived from this orbit would be insufficient to unambiguously reconstruct the source. In Figure 2, the vantage curve for a parallel-hole collimator is shown. The collimator is not tilted. The vantage curve intersects all great circles on the sphere.

Tuy and Smith realized that Orlov's condition is met by a limited set of points that will be referred to in this paper as the Orlov volume [2,3]. It has been observed that reconstruction artifacts occur in regions outside the Orlov volume [4].

The Orlov volume can be determined intuitively in the cases of simple collimator types, such as parallel-hole, following simple circular orbits. The symmetry of the sampling makes the Orlov volume cylindrical in this case. However, there is no known algorithm in the literature for determining the algorithm for the generalized case of non-parallel-hole collimators and non-circular orbits.

A method has been developed to computationally determine the Orlov volume for any set of collimators following any set of orbits. Herein this method will be described.

II. ALGORITHM

A voxelized representation of the volume to evaluate is created. Then a set of collimator models is constructed to simulate the positions, orientations and spatial extents of the collimators. For example, a single parallel-hole collimator following an orbit that includes m projection views would be represented by m collimator models. Two collimators with m and n projection views, respectively, would

This work was supported by DOE Grant DE-FG-02-96ER62150, NIH Grants CA76006 and CA33541 and the Whitaker Foundation Grant RG-99-0305.

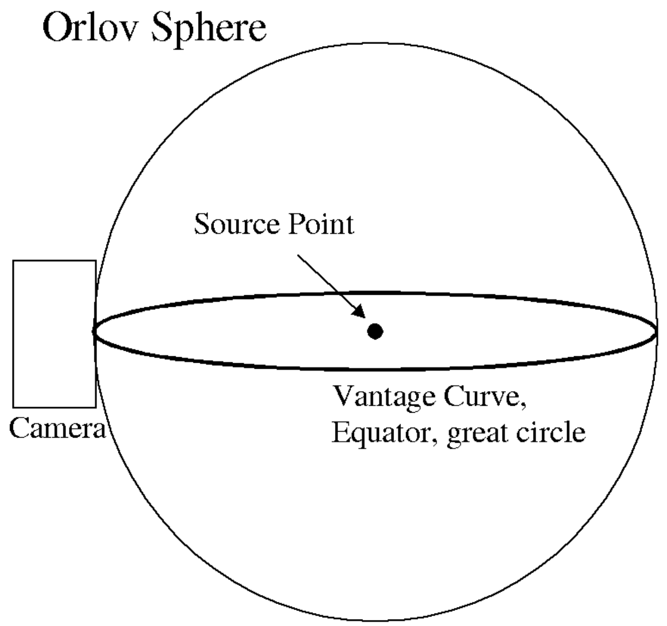


Fig. 2. Orlov Sphere and Vantage Curve for Parallel-hole Collimator. The figure depicts the vantage curve for the parallel-hole collimator as it observes a source point at the center of the sphere. The camera follows a circular orbit that coincides with the the equator of the sphere. The vantage curve coincides with the equator. All great circles on this sphere intersect the vantage curve.

be represented by $m+n$ collimator models. For each voxel, a digitized version of the vantage points of the voxel are determined from the set of collimator models. The digitized vantage curve is then evaluated to determine if any great circles can exist on the Orlov sphere without intersecting the vantage curve.

This algorithm has been implemented in C++ using Object-Oriented programming techniques. When started, the program allocates and initializes a boolean matrix representing the voxelized volume to consider. It then reads one or several orbit files and constructs a set of collimator representations. All detector representations obey an abstract interface that determines if a given voxel is within the field-of-view of the collimator and the vantage angles for that voxel.

The use of an abstract interface makes it possible to model multiple collimator types simultaneously and extend the program to consider new collimator types without any change to the algorithm. The new collimator simply needs to implement this interface. Multiple collimators can be used because each projection view in the orbit file or files results in the construction of a new collimator representation. All the representations are stored and considered together when evaluating whether a voxel meets Orlov's criteria.

Each voxel is fully evaluated by considering the vantage curve generated by all the projection views before the next voxel is considered. This reduces the required memory considerably. For each projection view, it is determined if the voxel in question is seen by the detector at that view. If it is seen, the vantage angles are recorded to make a vantage curve.

After all the projection views have been considered for a voxel, the vantage curve is evaluated to determine if any great circles can exist on the sphere without intersecting the vantage curve. The parameterization of a generic great circle (Fig. 3) can be found by considering that the points with the minimum and maximum z values are

$$\begin{aligned} \vec{r}_{\max} &= (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \\ \vec{r}_{\min} &= (-\sin\theta\cos\phi, -\sin\theta\sin\phi, -\cos\theta). \end{aligned} \tag{1}$$

A normal to the circle can be parameterized as

$$\begin{aligned} \hat{N} &= \left(\sin\left(\theta+\frac{\pi}{2}\right)\cos\phi, \sin\left(\theta+\frac{\pi}{2}\right)\sin\phi, \cos\left(\theta+\frac{\pi}{2}\right)\right) \\ &= (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta). \end{aligned} \tag{2}$$

The normal can be used to determine the basis vector orthogonal to $\vec{r}_{\max}$, $\vec{r}_{\min}$, and $\hat{N}$.

$$\begin{aligned} \hat{b} &= \vec{r}_{\max} \times \hat{N} \\ &= \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \end{vmatrix} \\ &= -\hat{x}\sin\phi + \hat{y}\cos\phi \end{aligned} \tag{3}$$

Any point on the great circle can be parameterized as

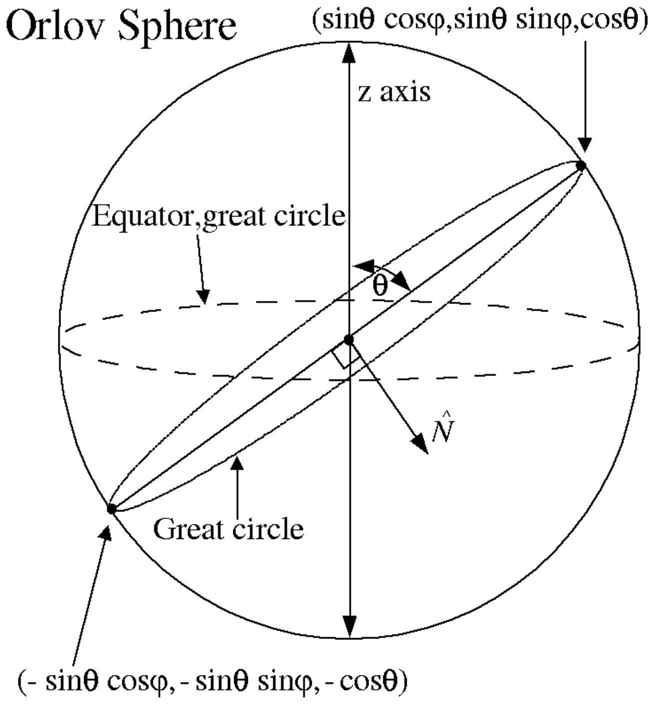


Fig. 3. Generic Great Circle on an Orlov Sphere. The point $(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ has the maximum value of z on this curve. The point $(-\sin\theta\cos\phi, -\sin\theta\sin\phi, -\cos\theta)$ has the minimal value of z on this curve. The normal to the plane of this great circle is $\hat{N}$.

$$\overrightarrow{r} = \alpha \overrightarrow{r}_{\max} + \beta \hat{b}. \quad (4)$$

Since $|\overrightarrow{r}| = 1$, $|\overrightarrow{r}_{\max}| = 1$, $|\hat{b}| = 1$, and $\overrightarrow{r}_{\max}$ and $\hat{b}$ are orthogonal,

$$\alpha^2 + \beta^2 = 1. \quad (5)$$

Letting $\alpha = \sin\gamma$ and $\beta = \cos\gamma$,

$$\begin{aligned}\overrightarrow{r} = (&\sin\gamma\sin\theta\cos\phi - \cos\gamma\sin\phi, \\ &\sin\gamma\sin\theta\sin\phi + \cos\gamma\cos\phi, \sin\gamma\cos\theta),\end{aligned} \quad (6)$$

where θ and ϕ are the coordinates of the point of maximum z on the curve and γ parameterizes the curve. Equation 6 is used to evaluate if any great circles can exist on the sphere without intersecting the vantage curve.

III. RESULTS

The program has been tested using parallel-hole collimators, slant-hole collimators, pinhole collimators and combinations of the above. Example volumes are described below.

A. Parallel-hole Collimators with Circular Orbits

Parallel-hole collimators with circular orbits have been evaluated and give cylindrical Orlov volumes (Fig. 4), as expected. The radius of the cylindrical Orlov volume in this case is larger than the collimator's radius of rotation (ROR). Clinically, there generally would not be activity outside the ROR since the gamma camera passes through that region, but, as the program indicates, that region is sufficiently sampled. This is because the Orlov volume of a parallel-hole collimator following a circular orbit depends only on the detector dimensions and not the orbit dimensions. That volume is cylindrical with extent given by $\pi w^2 d/4$, where w is the width of the detector in the plane of the orbit and d is the depth of the detector normal to the plane of the orbit.

B. Pinhole Collimators

The algorithm can also be used to determine sufficiently sampled volumes for converging-beam and diverging-beam collimators. The pinhole collimator has been tested with a circular orbit, giving a circular slice as the Orlov volume. This was expected, because all voxels that are off-axis have insufficient data. The pinhole-collimator has also been tested with a spiral orbit (Fig. 5) and gives a nearly cylindrical volume. The defects in the cylinder are at the ends, as expected because of the strong dependence on the initial orientation of the camera.

C. Combined Parallel-Hole and Tilted Parallel-Hole

Combined parallel-hole and tilted parallel-hole have been used to understand Orlov volumes in conjunction with breast-imaging research using vertical-axis-of-rotation (VAOR) orbits [4]. A digital phantom with a single pendulous breast was used to study the utility of parallel-hole acquisition around the torso using a standard horizontal-axis-of-rotation (HAOR) orbit versus tilted parallel-hole acquisition around the breast using a VAOR orbit. The VAOR orbit allows the camera to be positioned nearer the breast to improve spatial resolution, but the HAOR orbit gives better Orlov sampling. Fig. 6 shows the Orlov volume for a tilted parallel-hole acquisition with two additional arcs. Notice that the VAOR orbit does not yield a volume that includes the entire breast. Fig. 7 shows the Orlov volume for the tilted parallel-hole acquisition combined with a parallel-hole orbit extending around the posterior side of the patient. The VAOR orbit gives a much larger Orlov volume.

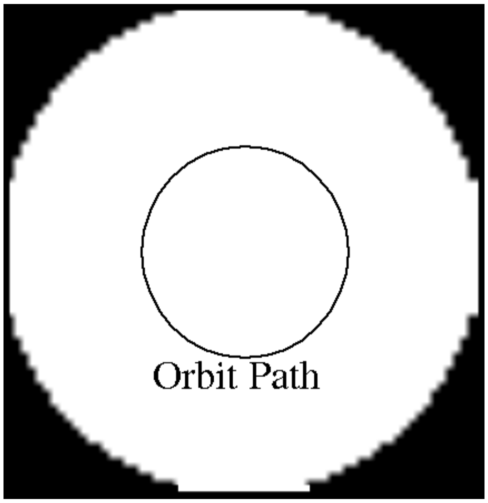


Fig. 4. Orlov Volume for Parallel-hole Collimator with a Circular Orbit. The voxelized volume meeting the Orlov condition for a circular orbit of an untilted parallel-hole collimator is shown in white. The collimator's orbit is shown by the superimposed circle.

IV. DISCUSSION

Orlov's condition assumes continuous angular sampling and infinitesimal sampling bins. Both of these conditions

Fig. 5. Orlov Volume for Helical Pinhole-Collimator Orbit. A slice of the voxelized volume meeting the Orlov condition for a helical orbit of a pinhole collimator is shown in white. The volume is nearly cylindrical, except for defects at the ends (not shown) where the valid volume depends on the initial angle of acquisition.

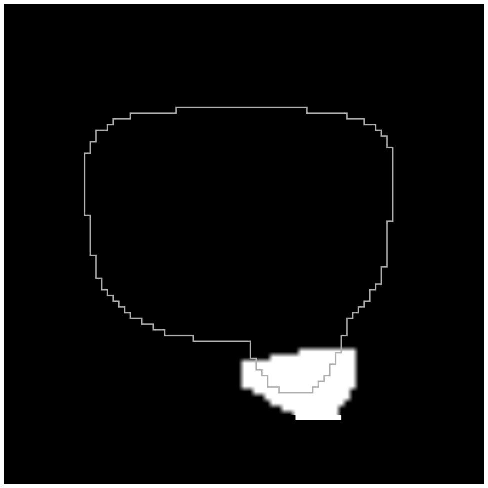


Fig. 6. Orlov Volume for Vertical-Axis-of-Rotation Tilted Parallel-Hole Collimator Orbit. A parallel-hole collimator was used to sample regions in and near the breast. The tilted acquisition was combined with additional arcs near the sternum and lateral side of the breast. A slice of the volume meeting the Orlov condition is shown as white. The outline of the digital phantom is also shown.

are invalid in realistic acquisitions. By Nyquist's theorem, there are limitations in reconstruction resolution due to discrete sampling. The discrete nature of the sampling merges the concepts of Nyquist frequency and the Orlov volume. This work represents an important future development.

The current algorithm has successfully determined the Orlov volume in cases that are easily verified intuitively. It has been used to study the more complex scenarios of pinhole collimators following spiral orbits and simultaneous acquisition of parallel-hole and slant-hole collimators. Further, it has improved understanding of artifacts found in VAOR acquisitions of the breast. This technique may be useful for studying sufficient orbits and understanding sampling artifacts for complex projection acquisitions.

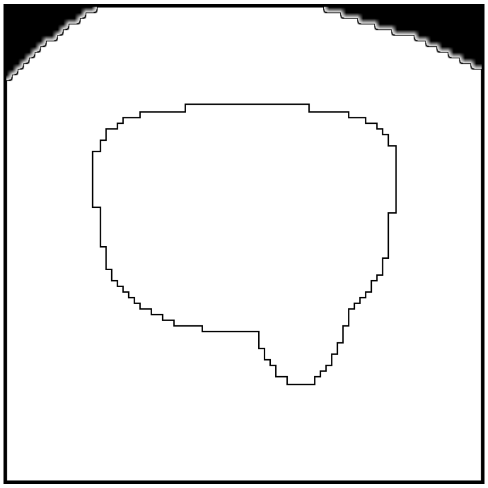


Fig. 7. Orlov Volume for Combined HAOR-VAOR Orbits. The same orbit described for Fig. 6 was supplemented with a standard HAOR orbit around the patient. The additional orbit greatly increases the size of the sufficiently sampled volume. The outline of the digital phantom is also shown.

REFERENCES

[1] S.S. Orlov, "Theory of three dimensional reconstruction. I. Conditions for a complete set of projections," *Sov. Phys. Crystallogr.*, vol. 20, no. 3., pp. 312–314, 1975.

[2] H.K. Tuy, "An Inversion Formula for Cone-Beam Reconstruction," *SIAM J. Appl. Math.*, vol. 43, no. 3, pp. 546–552, 1983.

[3] B.D. Smith, "Image Reconstruction from Cone-Beam Projections: Necessary and Sufficient Conditions and Reconstruction Methods," *IEEE Trans. Med. Imag.*, vol. MI-4, no. 1, 1985.

[4] S.D. Metzler, J.E. Bowsher, M.P. Tornai, B.C. Pieper and R.J. Jaszczak, "SPECT Breast Imaging Combining Horizontal and Vertical Axes of Rotation," *IEEE Trans. Nucl. Sci.*, submitted 2001.

Inversion of the 3-D Exponential X-ray Transform for a Semi Equatorial Band

Frédéric Noo[1], Rolf Clackdoyle[1], Jean-Marc Wagner[2] *

[1]University of Utah, USA [2]University of Liège, Belgium

1 Introduction

In this work, the reconstruction of a 3-D image from attenuated parallel-beam projections is investigated for fully 3-D data acquisition geometries.

In a partially simplified model, the measurements provided by a SPECT scanner are considered to be attenuated parallel-beam projections of the activity distribution f. They are mathematically described by the formula

$$p(\underline{\theta}, \underline{s}) = \int_{-\infty}^{+\infty} dt\, f(\underline{s}+t\underline{\theta})\, \exp\Big(-\int_{t}^{+\infty} dl\, \mu(\underline{s}+l\underline{\theta})\Big) \quad (1)$$

where $\underline{\theta}$ is the direction of projection defined by the orientation of the camera and the collimator holes and μ is the attenuation function. Vector $\underline{s}$ is orthogonal to $\underline{\theta}$ and is used to specify detector locations for the $p(\underline{\theta}, \cdot)$ projection.

A simplication of the relation (1) between the data and the image f occurs when the activity is contained in some convex region where μ is constant. We will assume that this condition holds. In this case, the data $p(\underline{\theta}, \underline{s})$ can be modified into

$$g(\underline{\theta}, \underline{s}) = \int_{-\infty}^{+\infty} dt\, f(\underline{s} + t\,\underline{\theta})\, e^{\mu_0 t}, \qquad \underline{s} \cdot \underline{\theta} = 0 \quad (2)$$

where μ_0 is the value of μ in the activity region. In the literature, g is referred to as the exponential X-ray (parallel-beam) projection of f in the direction $\underline{\theta}$. The relation between p and g can be written in the form

$$g(\underline{\theta}, \underline{s}) = p(\underline{\theta}, \underline{s})\, m_\mu(\underline{\theta}, \underline{s}) \quad (3)$$

where $m_\mu(\underline{\theta}, \underline{s})$ is calculated from the attenuation map. See [1] for details. Even if the attenuation map is not known, p can be converted into g in a reasonably accurate way using the consistency conditions for the exponential X-ray transform [2].

The set of directions $\underline{\theta}$ for which p is measured defines the data acquisition geometry. We use Ω to denote this set. By definition, Ω is a subset of the unit sphere. The most common set Ω encountered in SPECT imaging is the great circle (360° scan) or half great circle (180° scan) of directions orthogonal to the patient bed. However, fully 3-D geometries are also possible, such as the RSH-SPECT scanner [3].

Image reconstruction from exponential X-ray projections on a great circle has been widely studied over the last twenty years and is now well-understood, especially thanks to the significant work of Tretiak and Metz [4] and Pan and Metz [5, 6]. In fully 3-D geometries, the situation is very different. To our knowledge, only three works concerning exact fully 3-D reconstruction from exponential X-ray projections have been published [7, 8, 9]. The most general of these assumes that the set Ω is a union of great circles.

Currently, the class of data sets for which a closed-form inversion formula of the imaging equation (2) exists is unknown. It is not even known what conditions a data set must satisfy to be complete. The 3-D reconstruction theory for X-ray projections ($\mu = 0$) [10, 11] is not easily modified to handle exponential X-ray projections.

We have derived a closed-form inversion formula for the 3-D exponential X-ray transform which is valid for any data set Ω made up of half great circles. A basic example of such a set is the half equatorial band illustrated in figure 1a; our presentation concentrates on this example.

Our results generalize all previously published results on the exponential X-ray transform. They constitute one step further towards a full understanding of this trans-

---

*This work was partially supported by the National Institutes of Health, grant number R01 HL55610. The work of F. Noo was partially supported by Marconi Medical Systems.

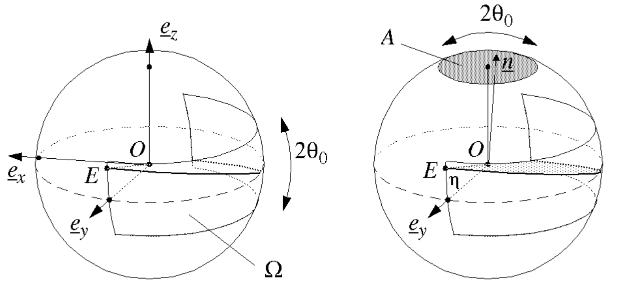


Figure 1: (a-left) Illustration of the half equatorial band including a half great circle. (b-right) Set A of vectors $\underline{n}$ corresponding to all the half great circles in Ω.

form[1] which is an important mathematical tool in SPECT imaging and in Intensity Modulated Radiation Therapy (see [12] for more details on this latter application).

The basic idea is to combine a recent result on inversion of 180° scans [13] with the TTR concept of Ra *et al.* [14].

2 A half great circle

Let $\mathcal{C}(\underline{n}, E)$ be the right-oriented half great circle which starts at E and is orthogonal to the unit vector $\underline{n}$ on the unit sphere. See figure 1. Let $\underline{b} = \underline{OE}$ where O is the center of the sphere and let $\underline{a} = \underline{n} \times \underline{b}$. From the results in [9] and [13], it can be shown that f satisfies the following integral equation

$$f(\underline{x}) = f_0(\underline{x}, \underline{n}, E) + w(\underline{x}, \underline{n}, E) * f(\underline{x}) \tag{4}$$

where $f_0(\underline{x}, \underline{n}, E)$ and $w(\underline{x}, \underline{n}, E)$ are defined as follows. The function $f_0(\underline{x}, \underline{n}, E)$ is obtained by filtered backprojection (FBP) of the projections on $\mathcal{C}(\underline{n}, E)$:

$$f_0(\underline{x}, \underline{n}, E) = \int_{\mathcal{C}(\underline{n},E)} d\underline{\theta}\; e^{-\mu_0\, \underline{x}\cdot\underline{\theta}}\; g^F\Big(\underline{\theta}, \underline{x} - (\underline{x}\cdot\underline{\theta})\,\underline{\theta}\Big) \tag{5}$$

with

$$g^F(\underline{\theta}, \underline{s}) = \int_{\underline{s}'\cdot\underline{\theta}=0} d\underline{s}'\, g(\underline{\theta}, \underline{s} - \underline{s}')\, \delta(\underline{s}'\cdot\underline{n})\, k_\mu\Big(\underline{s}'\cdot(\underline{n}\times\underline{\theta})\Big) \tag{6}$$

where k_μ is the notch filter used in the FBP formula of Tretiak and Metz for the 2-D exponential Radon transform [4], i.e.

$$k_\mu(r) = 1/2 \int_{|\nu|>\mu_0/2\pi} d\nu\, |\nu|\, e^{j2\pi r\nu}\,. \tag{7}$$

The convolution kernel $w(\underline{x}, \underline{n}, E) = \delta(\underline{x}\cdot\underline{n})\, \hat{w}(\underline{x}\cdot\underline{a}, \underline{x}\cdot\underline{b})$ where δ is the 1D Dirac function and

$$\hat{w}(u,v) = \frac{\sinh \mu_0 v}{\pi v}\, q(u) + \frac{\overline{\mu}_0}{\pi\, u}\Big\{\frac{2\sinh\mu_0 v}{\mu_0 v} - \frac{\sinh\mu_0(v+i\,u)}{\mu_0(v+i\,u)} - \frac{\sinh\mu_0(v-i\,u)}{\mu_0(v-i\,u)}\Big\} \tag{8}$$

In the above definition,

$$q(u) = \int_{\mathbb{R}} d\sigma\, i\,\mathrm{sign}(\sigma)\, e^{i2\pi u\sigma} \tag{9}$$

is the convolution kernel of the Hilbert transform. Note that $w(\underline{x}, \underline{n}, E)$ is an odd function in $\underline{x}$, i.e $w(-\underline{x}, \underline{n}, E) = -w(\underline{x}, \underline{n}, E)$.

Basically, for exponential projections $g(\underline{\theta}, \underline{s})$ measured on the (360°) great circle $\mathcal{C}(\underline{n}, E) \cap \mathcal{C}(\underline{n}, -E)$ the Tretiak and Metz FBP gives an accurate reconstruction f. However, on the (180°) half circle $\underline{\theta} \in \mathcal{C}(\underline{n}, E)$, the reconstruction gives an incorrect f_0. The images f_0 and f are linked by w according to equation (4).

3 An equatorial band

The discussion is now focussed on the half equatorial band of figure 1a. Let A be the set of unit vectors corresponding to all the half great circles in Ω. See figure 1b. For each vector $\underline{n}$, the starting point of the half great circle is at $\underline{OE} = \cos\eta\, \underline{e}_y + \sin\eta\, \underline{e}_z$ with $\tan\eta = -\,\underline{n}\cdot\underline{e}_y/\underline{n}\cdot\underline{e}_z$.

For each half great circle in Ω, an integral equation similar to (4) can be written. Integrating equation (4) over all $\underline{n} \in A$, one obtains an integral equation for f which involves all the data from Ω:

$$f(\underline{x}) = f_0(\underline{x}) + W(\underline{x}) * f(\underline{x}) \tag{10}$$

with

$$f_0(\underline{x}) = c \int_A d\underline{n}\, f_0(\underline{x}, \underline{n}, E)\,, \tag{11}$$

and

$$W(\underline{x}) = c \int_A d\underline{n}\, w(\underline{x}, \underline{n}, E)\,. \tag{12}$$

In these equations, c is a normalization constant:

$$c = 1/\int_A d\underline{n} = 2\pi\,(1 - \cos\theta_0) \tag{13}$$

where θ_0 is the half aperture of the band.

---

[1] A parallel submission to this conference by Wagner, Noo and Clackdoyle addresses geometries consisting of a union of circles on the sphere. This is a distinct class of geometries.

Using the same argument as in [9], it can be shown that $f_0(\underline{x})$ can be calculated in a fully 3-D FBP way. The expression is

$$f_0(\underline{x}) = \int_\Omega d\underline{\theta}\; e^{-\mu_0\, \underline{x}\cdot\underline{\theta}}\, g^F\big(\underline{\theta}, \underline{x} - (\underline{x}\cdot\underline{\theta})\,\underline{\theta}\big) \tag{14}$$

where $g^F(\underline{\theta}, \underline{s})$ is obtained from $g(\underline{\theta}, \underline{s})$ by 2-D convolution:

$$g^F(\underline{\theta}, \underline{s}) = \int_{\underline{s}'\cdot\underline{\theta}=0} d\underline{s}'\, g(\underline{\theta}, \underline{s} - \underline{s}')\, h_\mu(\underline{\theta}, \underline{s}')\,. \tag{15}$$

The convolution filter h_μ is given by

$$h_\mu(\underline{\theta}, \underline{s}) = \begin{cases} 2\,c\,\dfrac{k_\mu(||\underline{s}||)}{||\underline{s}||} & \text{if} \quad \underline{\theta} \times \dfrac{\underline{s}}{||\underline{s}||} \in A \\[2ex] 0 & \text{otherwise} \end{cases} \tag{16}$$

As readily observed from its definition, the filter h_μ is a generalized function with singularities at $\underline{s} = 0$. The implementation of equation (15) therefore requires the use of some regularization technique. For an accurate computation of $g^F(\underline{\theta}, \underline{s})$ from samples of $g(\underline{\theta}, \underline{s})$ on a Cartesian grid, we recommend the implementation of (15) in the Fourier domain with some apodizing frequency window, such as the Hanning window. Such an implementation requires the knowledge of the Fourier transform of the filter h_μ. It is shown in [9] that this transform is

$$\begin{aligned} H_\mu(\underline{\theta}, \underline{\nu}) &= \int_{\underline{s}\cdot\underline{\theta}=0} d\underline{s}\, e^{-j2\pi \underline{s}\cdot\underline{\nu}}\, h_\mu(\underline{\theta}, \underline{s}) \\ &= \frac{c}{2} \int_{\mathcal{C}_\mu(\underline{\theta})\cap A} d\underline{n}\, w(\underline{n})\, |\underline{\nu}\cdot(\underline{n}\times\underline{\theta})|\,, \qquad \underline{\nu}\cdot\underline{\theta} = 0 \end{aligned} \tag{17}$$

where $\mathcal{C}_\mu(\underline{\theta})$ is a subset of the great circle $\mathcal{C}(\underline{\theta})$ of unit vectors orthogonal to $\underline{\theta}$:

$$\mathcal{C}_\mu(\underline{\theta}) = \mathcal{C}(\underline{\theta}) \setminus \big\{\underline{n} \in \mathcal{C}(\underline{\theta}) \;:\; |\underline{\nu}\cdot(\underline{n}\times\underline{\theta})| < \mu_0/2\pi\big\}\,. \tag{18}$$

Note, in particular, that $\mathcal{C}_\mu(\underline{\theta})$ is empty when $||\underline{\nu}|| < \mu_0/2\pi$ because $|\underline{\nu}\cdot(\underline{n}\times\underline{\theta})| < \mu_0/2\pi$ for any $\underline{n}$ in this case. Therefore, $H_\mu(\underline{\theta}, \underline{\nu}) = 0$ if $||\underline{\nu}|| < \mu_0/2\pi$.

4 Solution of the integral equation

In this section, we show that the integral equation (10) admits a unique solution which can be expressed in the form of a Neumann series. First, note that the kernel $W(\underline{x})$ (equation (12)) is an odd function, i.e. $W(-\underline{x}) = -W(\underline{x})$ because $w(-\underline{x}, \underline{n}, E) = -w(\underline{x}, \underline{n}, E)$.

Let R be such that $f(\underline{x}) = 0$ for $|\underline{x}| > R$ and let

$$\chi(\underline{x}) = \begin{cases} 1 & \text{if } |\underline{x}| < R \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

In practice R is always finite since f is physically restricted to a finite region. Using χ, the integral equation (10) can be rewritten in the form

$$f = \chi\, f_0 + \chi\,(W * f) = \chi\, f_0 + K\, f \tag{20}$$

where K is an operator such that $K\,f = \chi(W * f)$. Since W is odd, we note that K is skew-symmetric.

Using the same arguments as those developed in [13], one can show that K is bounded. Let $\gamma = 1/(1 + ||K||^2)$. We introduce a modified operator $\hat{K} = (1-\gamma)\, I + \gamma\, K$ where I is the identity operator and rewrite (20) in the form

$$f = \gamma\,\chi\, f_0 + \hat{K}\, f\,. \tag{21}$$

By definition, $||\hat{K}|| < 1$. Therefore, the integral equation (21) admits a unique solution

$$f = \gamma \sum_{l=0}^{\infty} \hat{K}^l\, \chi\, f_0\,. \tag{22}$$

See [15] for mathematical details.

The reconstruction of f from formula (22) can be implemented in the following way:

- Step 1: Compute $\chi\, f_0$ from the data g using (14).
- Step 2: Compute $f_n = \hat{K}\,\chi\, f_{n-1}$ for $n = 1, \ldots, N$.
- Step 3: Compute $f_N \simeq \gamma\,\chi\, f_0 + \gamma \sum_{n=1}^{N} f_n\,.$

The function f_N represents the reconstructed image. The accuracy $||f_N - f||$ of the reconstruction depends on $||K||$. In the absence of noise, the smaller $||K||$, the smaller $||\hat{K}||$ and thus the smaller the number of terms N required for a given accuracy because the series converges faster.

Figure 2 shows results obtained from computer-simulated projections of a heart phantom made up of ellipsoids, with 20% activity in the ventricules. The attenuation coefficient was $\mu = 0.0152$ and the half aperture of the band was $\theta_0 = 25°$. There were 120×30 projections (each of 128×128 square pixels of side 1.2 mm) and the reconstruction was performed on a grid of 128^3 cubic voxels of side 1.2 mm.

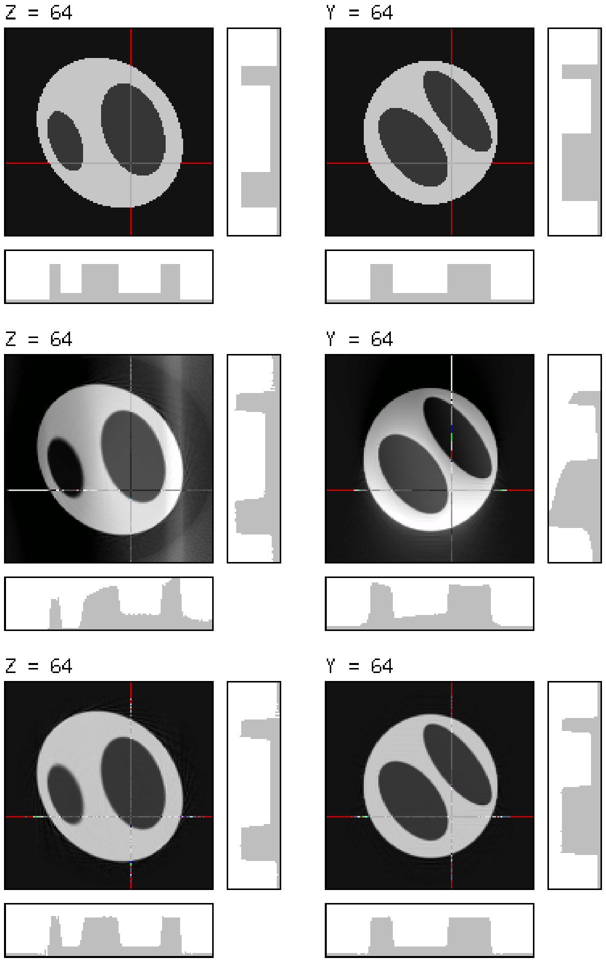


Figure 2: (top) Original phantom, (middle) FBP reconstruction f_0 using the data on Ω, (bottom) reconstruction using 10 terms of the Neumann series (equation 22). Reconstruction time: about 5 min. cpu per iteration on a SUN ULTRA 10.

References

[1] A. Markoe, "Fourier inversion of the attenuated X-ray transform", *SIAM J. Math. Anal.*, 15(4), 718-722, 1984.

[2] C. Mennessier, F. Noo, R. Clack, G. Bal and L. Desbat, "Attenuation correction in SPECT using consistency conditions for the exponential ray transform", *Phys. Med. Biol.*, 44, 2483-2510, 1999.

[3] R. Clack, P. E. Christian, M. Defrise and A. E. Welch, "Image reconstruction for a novel SPECT system with rotating slant-hole collimators". *In Conf. Rec. 1995 IEEE Med. Imag.* Conf., 1948-1952, 1996.

[4] O. Tretiak and C. Metz, "The exponential Radon transform", *SIAM J. Appl. Math.*, 39(2), 341-354, 1980.

[5] C. E. Metz and X. Pan, "A unified analysis of exact methods of inverting the 2D exponential Radon transform, with implications for Noise Control in SPECT", *IEEE Trans. Med. Imag.*, 14(4), 643-658, 1995.

[6] X. Pan and C. E. Metz, "Analysis of noise properties of a class of exact methods of inverting the 2D exponential Radon transform", *IEEE Trans. Med. Imag.*, 14(4), 659-668, 1995.

[7] I. A. Hazou, "Inversion of the exponential X-ray transform. I: Analysis", *Math. Methods in the Applied Sciences*, Vol. 10(10), 561-574, 1988.

[8] Y. Weng, G. L. Zeng and G. T. Gullberg, "Filtered backprojection algorithms for attenuated parallel and cone-beam projections sampled on a sphere", *in Three-dimensional Image Reconstruction in Radiation and Nuclear Medicine*, ed. P.Grangeat and J.-L. Amans (Dordrecht: Kluwer), 19-34, 1996.

[9] J.-M. Wagner and F. Noo, "Three-dimensional image reconstruction from exponential parallel-beam projections", *IEEE Transactions on Nuclear Sciences*, (to appear) June 2001.

[10] S. S. Orlov, "Theory of three dimensional reconstruction. 1. Conditions of a complete set of projections.", *Sov. Phys.-Crystallogr.*, 20, 312-314, 1975.

[11] M. Defrise, D. W. Townsend and R. Clack, "Three-dimensional image reconstruction from complete projections", *Phys. Med. Biol.*, 34(5), 573-587, 1989.

[12] M. Braunstein and R. Y. Levine, "Optimum beam configurations in tomographic intensity modulated radiation therapy", *Phys. Med. Biol.*, 45, 305-328, 2000.

[13] F. Noo and J.-M. Wagner, "Image reconstruction in 2D SPECT with 180-degree acquisition," *submitted to Inverse Problems.*

[14] J. B. Ra, C. B. Lim, Z. H. Cho, S. K. Hilal and J. Correll "A true three-dimensional reconstruction algorithm for the spherical positron emission tomograph", *Phys. Med. Biol.*, 27, 37-50, 1982.

[15] N. I. Akhiezer and I. M. Glazman, *Theory of linear operators in Hilbert space. Volume I*, Pitman Advanced Publishing Program (Boston. London. Melbourne.), 1981.

Accurate Iterative Reconstruction Algorithm for Sparse Objects: Application to 3-D Blood-Vessel Reconstruction from a Limited Number of Projections

Meihua Li[1], Haiquan Yang[2], and Hiroyuki Kudo[3]
[1]Doctoral Program in Engineering, University of Tsukuba, Tsukuba 305-8573, Japan
[2]Department of Mathematics, Jilin University, Changchun 130012, Jilin, China
[3]Institute of Information Sciences and Electronics, University of Tsukuba, Tsukuba 305-8573, Japan

Abstract

This paper proposes an accurate row-action type iterative method which is appropriate to reconstruct sparse objects from a limited number of projections. The main idea is to use the L_p norm with $p \approx 1.1$ to pick up a sparse solution from a set of feasible solutions to the measurement equation. We also show that this method works well in the 3-D blood-vessel reconstruction.

I. Introduction

Image reconstruction from a limited number of projections is a well-investigated subject in tomographic reconstruction fields. For example, this problem possesses applications in visualization of 3-D blood-vessel structure from angiographic projections and cardiac imaging with various modalities [1]. In spite of a lot of works, however, it is fair say that successful reconstruction algorithms which can be used in clinical routine do not exist. This paper proposes an iterative reconstruction algorithm which is very powerful for sparse objects. Here, the sparse objects refer to objects which have non-zero pixel values only on a relatively small number of pixels. Such objects appear in many instances of tomographic imaging. In particular, we have the following three applications in mind. The first application is the reconstruction of cerebral or cardiac blood-vessel structure from a limited number of angiographic projections measured with a C-Arm or rotational-angiographic devices (digital subtraction technique is normally used to eliminate unnecessary background objects). The second application is the cardiac SPECT or PET imaging where the cross section can be approximately regarded as a sparse object because the isotope normally concentrates only on the heart. The last interesting application is the reconstruction of tomographic dynamic sequences. This problem can be also formulated as the reconstruction of a sparse object in the following way. Let t denote time and assume that we have a good reconstruction f_t at time t. Then, $f_{t+1} - f_t$ becomes a sparse object if the motion in the cross section is not so large. Therefore, we can reconstruct $f_{t+1} - f_t$ from a limited number of projections g_{t+1} at time $t+1$ if an accurate reconstruction algorithm for sparse objects exists, which leads to an accurate reconstruction of f_{t+1}.

The proposed algorithm is outlined as follows. The tomographic reconstruction problem can be formulated as solving a linear equation $A\vec{x} = \vec{b}$ where $\vec{x}$ is an image vector, $\vec{b}$ is a set of measured line integrals, and A is an $m \times n$ matrix relating $\vec{x}$ to $\vec{b}$. When the number of projections is small, many feasible solutions to $A\vec{x} = \vec{b}$ exist because $m < n$. To pick up a good solution having sparsity from a set of feasible solutions, we formulate the problem as a bound-constrained minimum norm problem (PPB):

$$\text{minimize } \parallel \vec{x} \parallel_p^p /p \text{ subject to } A\vec{x} = \vec{b} \text{ and } 0 \leq \vec{x} \leq 1$$

where $\parallel \vec{x} \parallel_p$ denotes the L_p norm of $\vec{x}$ and the inequality $0 \leq \vec{x} \leq 1$ can be understood componentwise. The value of norm parameter p has a large effect on the solution. The best value of p for sparse objects is $p = 0$ because $\lim_{p \to 0} \parallel \vec{x} \parallel_p^p$ is equivalent to the number of non-zero pixels. This choice allows to pick up an object which has minimum number of non-zero pixels. However, $\parallel \vec{x} \parallel_p^p /p$ is non-convex and is non-differentiable when $0 < p < 1$ which makes it impossible to use well-known convex optimization techniques to construct an iterative algorithm. To overcome this drawback, we use $p \approx 1.1$ (a slightly larger value than 1). For this choice, $\parallel \vec{x} \parallel_p^p /p$ is both convex and differentiable so that we can use standard convex optimization techniques. Furthermore, $p \approx 1.1$ still allows to pick up a sparse solution to $A\vec{x} = \vec{b}$ compared to the use of ordinary Euclidean norm $p = 2$. We construct an iterative method for the above constrained minimization problem by using the dual coordinate ascent method [2]. This method converts the original problem (PPB) into a dual problem by using the so-called Lagrangian duality. Since the dual problem corresponding to (PPB) becomes a simple unconstrained maximization problem, we solve it by using the coordinate ascent method. In the primal space, this method can be considered as a variant of Bregman's method for convex programming excepting the bound constraint $0 \leq \vec{x} \leq 1$ [3]. The resulting algorithm is of row-action type similarly to the Algebraic Reconstruction Technique (ART). The iteration converges fast by using a projection access order proposed in the literature [4].

We believe that this method possesses some important applications in tomography. As a first step, we have applied this method to the reconstruction of 3-D blood-vessel structure from a limited number of cone-beam projections. The simulation results demonstrate that an accurate reconstruction is possible from only 8 projections for which the Feldkamp method and the ART method are not valid anymore. Furthermore, we apply the method to real-data measured with the cone-beam tomographic imaging system using a synchrotron radiation x-ray source [5].

II. PROPOSED METHOD

A. Problem Formulation

In tomography, the goal is to reconstruct an object from line-integral projection data. A discrete version of the projection process can be represented as

$$A\vec{x} = \vec{b}$$

where $A = (a_{ij})$ is a real $m \times n$ matrix representing the projection operator, $\vec{x} = (x_1, \cdots, x_n)^\top$ is a real vector representing the object, and $\vec{b} = (b_1, \cdots, b_m)^\top$ is the corresponding projection data. Let $\vec{c}_j$ denote the j-th column of matrix A.

For convenience to explain the method, we first consider the following minimum norm problem (PP):

$$\text{minimize } \parallel \vec{x} \parallel_p^p /p \text{ subject to } A\vec{x} = \vec{b} \tag{1}$$

where $1 < p < 2$. We use $p \approx 1.1$ for sparse objects as mentioned in Section I. Though the problem (PP) is simpler than the problem (PPB), it is still a nonlinear constrained optimization, and is difficult to solve directly. However, its dual will be an unconstrained optimization.

B. Lagrangian Duality

From the reference [6], the Lagrangian dual of the problem (PP) is the following maximization problem (DP):

$$\text{maximize } D(\vec{y}) = \vec{b}^\top \vec{y} - \parallel A^\top \vec{y} \parallel_q^q /q \tag{2}$$

where $\vec{y} = (y_1, \cdots, y_m)^\top$ is a real vector, and $q = p/(p-1)$. Let $\vec{z}(\vec{y}) = (z_1(\vec{y}), \cdots, z_n(\vec{y}))^\top$ be a vector function whose j-th component is $z_j(\vec{y}) = |\vec{c}_j^\top \vec{y}|^{q-1}\text{sign}(\vec{c}_j^\top \vec{y})$. If $\vec{x}^*$ solves the problem (PP), then there exists $\vec{y}^*$ such that $\vec{x}^* = \vec{z}(\vec{y}^*)$ and $\vec{y}^*$ solves the problem (DP). Conversely, let $\vec{y}^*$ solve the problem (DP), then the vector $\vec{x}^* = \vec{z}(\vec{y}^*)$ solves the problem (PP).

C. Solving Maximization Problem (DP)

Obviously, the number of the unknowns in the problem (DP) is less than that in the problem (PP). Furthermore, the problem (DP) is an unconstrained maximization which is easier to solve than the problem (PP).

Differentiating (2), we get the following equations:

$$D'_{y_i}(\vec{y}) = 0, \quad i = 1, \cdots, m. \tag{3}$$

Equation (3) is a system of nonlinear equations which yields a solution of (2). Using the newton-like iteration to each equation in parallel, we get

$$y_i^{(k+1)} = y_i^{(k)} - \beta \frac{D'_{y_i}(\vec{y}^{(k)})}{D''_{y_i y_i}(\vec{y}^{(k)})}, \quad i = 1, \cdots, m. \tag{4}$$

where β is the relaxation parameter. We define

$$\Delta_i(\omega_1, \cdots, \omega_n) = \frac{b_i - \displaystyle\sum_{j=1}^{n} a_{ij} |\omega_j|^{q-1} \text{sign}(\omega_j)}{(q-1) \displaystyle\sum_{j=1}^{n} (a_{ij})^2 |\omega_j|^{q-2}},$$

then for $i = 1, \cdots, m$, (4) can be written as

$$y_i^{(k+1)} = y_i^{(k)} + \beta \Delta_i(\vec{c}_1^\top \vec{y}^{(k)}, \cdots, \vec{c}_n^\top \vec{y}^{(k)}), \tag{4'}$$

In fact, for convenience to implement, we use the sequential Gauss-Seidel type iteration scheme (5) in our program:

$$y_i^{(k+1)} = y_i^{(k)} + \beta \Delta_i(\omega_1^{(k,i)}, \cdots, \omega_n^{(k,i)}), \quad i = 1, \cdots, m, \tag{5}$$

where $\omega_j^{(k,i)} = \sum_{l=1}^{i-1} a_{lj} y_l^{(k+1)} + \sum_{l=i}^{m} a_{lj} y_l^{(k)}$ is another expression of $\vec{c}_j^\top \vec{y}$ in the iteration. Equation (5) can be regarded as the coordinate ascent method applied to maximize $D(\vec{y})$.

D. Algorithm

According to the duality, if $\vec{x}^*$ and $\vec{y}^*$ are solutions of the problems (PP) and (DP) respectively, then $x_j^* = |\vec{c}_j^\top \vec{y}^*|^{q-1}\text{sign}(\vec{c}_j^\top \vec{y}^*)$, $j = 1, \cdots, n$. Though it is difficult to get an explicit iteration scheme for $\vec{x}$, we can get one for $\omega_j^{(k,i)}$. We define $\vec{\mu}^{(k,i)} = (y_1^{(k+1)}, \cdots, y_{i-1}^{(k+1)}, y_i^{(k)}, y_{i+1}^{(k)}, \cdots, y_m^{(k)})^\top$, $\vec{\mu}^{(k,i+1)} = (y_1^{(k+1)}, \cdots, y_{i-1}^{(k+1)}, y_i^{(k+1)}, y_{i+1}^{(k)}, \cdots, y_m^{(k)})^\top$, $\vec{\mu}^{(k,m+1)} = \vec{\mu}^{(k+1,1)}$, $\omega_j^{(k,i)} = \vec{c}_j^\top \vec{\mu}^{(k,i)}$, and $\omega_j^{(k,m+1)} = \omega_j^{(k+1,1)}$. Let $\vec{\Omega}^{(k,i)} = (\overbrace{0, \cdots, 0}^{i-1}, \Delta_i(\omega_1^{(k,i)}, \cdots, \omega_n^{(k,i)}), \overbrace{0, \cdots, 0}^{m-i})^\top$. Then from (5), we obtain

$$\vec{\mu}^{(k,i+1)} = \vec{\mu}^{(k,i)} + \beta \vec{\Omega}^{(k,i)}. \tag{5'}$$

Taking an inner product with $\vec{c}_j$, we finally get

$$\omega_j^{(k,i+1)} = \omega_j^{(k,i)} + \beta a_{ij} \Delta_i(\omega_1^{(k,i)}, \cdots, \omega_n^{(k,i)}), \tag{6}$$

$$j = 1, \cdots, n.$$

Note that (6) is a row-action type iteration. Note also that (6) is exactly the ART method when $p = 2$. Unfortunately, the iteration is instable when the denominator of Δ_i in (6) is near 0. In our program, we set the denominator to a real constant MIN if its value is less than the constant MIN.

The algorithm is summarized as follows.

[STEP 1] Give an initial vector $\vec{\omega}^{(0)} = (\omega_1^{(0)}, \cdots, \omega_n^{(0)})^\top$ such that $\vec{\omega}^{(0)} = A^\top \vec{y}^{(0)}$ for some $\vec{y}^{(0)}$. ■

[STEP 2] For $k = 0, 1, \cdots$, do the following iteration until k is large enough or $\parallel \vec{\omega}^{(k)} - \vec{\omega}^{(k+1)} \parallel$ is small enough. ■

[STEP 2.1] Let $\vec{\omega}^{(k,1)} = \vec{\omega}^{(k)}$. ■

[STEP 2.2] For $i = 1, \cdots, m$, do the iteration (6). ■

[STEP 2.3] Let $\vec{\omega}^{(k+1)} = \vec{\omega}^{(k,m+1)}$. ■

[STEP 3] Suppose $\vec{\omega}^*$ is the final result of [STEP 2], then $\vec{x}^*$, whose j-th component is $x_j^* = |\omega_j^*|^{q-1}\text{sign}(\omega_j^*)$, is the required result. ■

E. Dealing with Bound Constraint

In this subsection, we consider the bound-constrained minimum norm problem (PPB). This problem is equivalent to the following problem:

$$\text{minimize } F(\vec{x}) = \sum_{j=1}^{n} f(x_j) \text{ subject to } A\vec{x} = \vec{b}$$

where $f(t) = \begin{cases} t^p/p & 0 \leq t \leq 1 \\ \infty & \text{otherwise} \end{cases}$. The Lagrangian dual of the above problem is

$$\text{maximize } \vec{b}^\top \vec{y} - G(A^\top \vec{y})$$

where $G(\vec{x}) = \sum_{j=1}^{n} g(x_j)$, $g(t) = \begin{cases} 0 & t < 0 \\ t^q/q & 0 \leq t \leq 1 \\ t - 1/p & t > 1 \end{cases}$. Similarly to the previous (PP) case, we can get the following iteration scheme which is like the iteration (6).

$$\omega_j^{(k,i+1)} = \omega_j^{(k,i)} + \beta a_{ij} \frac{b_i - \sum_{l=1}^{n} a_{il} g'(\omega_l^{(k,i)})}{\sum_{l=1}^{n} a_{il}^2 g''(\omega_l^{(k,i)})}.$$

Unfortunately, $g(t)$ does not have the second derivative at $t = 1$. Therefore, we define $g''(1) = q - 1$ for implementation. And we set the denominator of the fraction in the above iteration scheme to a real constant MIN if its value is less than MIN. Suppose $\vec{\omega}^*$ is the final result of the above iteration, then $\vec{x}^*$, whose j-th component is $x_j^* = g'(\omega_j^*)$, is the solution of the problem (PPB).

III. Experimental Results

A. Simulation Studies

We have applied the proposed method to reconstruct 3-D blood-vessel structure from a limited number of cone-beam projections. We have used the 3-D blood-vessel phantom developed by the phantom group of Siemens (http://www.imp.uni-erlangen.de/forbild/english/results/index.htm). The x-ray source positions are located on the circle with uniform angular interval over 180°. The number of source positions is 8 or 4 and each projection consists of 256 × 256 pixels. The reconstructed image has 256 × 256 × 256 pixels. We have compared the proposed method with the conventional ART method which has been often used to this kind of limited-data problem in the literature. We used a projection data access order proposed in [4] which allows a fast convergence. By using this data access order, ten iterations were enough to obtain satisfactory images. The computations were performed with a PC with a Pentium III 700 MHz processer and the required computational time for ten iterations was about 30 minutes which is reasonable in practice. Five transaxial slices of reconstructed images after ten iterations are shown in Fig. 1 and Fig. 2 together with the corresponding slices of the phantom. In Fig. 3, we also show 3-D graphic display corresponding to the reconstructed images which is generated by using the volume rendering software. The threshold value to pick up the blood vessels from the reconstructed images is manually optimized dependent on each method. The proposed method succeeds in accurately reconstructing the fine blood-vessel structure whereas the ART method produces severe artifacts which make it impossible to recognize thin blood vessels. These results strongly demonstrate that the use of L_p norm with $p \approx 1.1$ is very powerful for sparse objects compared to $p = 2$ corrsponding to the ART method. Furthermore, the use of bound constraint $0 \leq \vec{x} \leq 1$ could improve the reconstructed images.

B. Real Data

We have also applied the proposed method to real data measured with our cone-beam tomographic imaging system using a synchrotron radiation x-ray source [5]. The cardiac blood-vessel phantom is used as a test object. The result will be presented at the conference. We are also applying the proposed method to cardiac SPECT data. The current result shows that the proposed method more accurately recovers the sharp boundary of blood pool compared with the ART method.

IV. Conclusions

We have proposed an accurate iterative method which is appropriate to reconstruct sparse objects from a limited number of projections. The main idea of the proposed method is to use the L_p norm with $p \approx 1.1$ to pick up a sparse solution from a set of feasible solutions. The algorithm is of row-action type and can be efficiently implemented similarly to the ART method. We have also shown that this method works well in the 3-D blood-vessel reconstruction from a limited number of cone-beam projections.

V. References

[1] D.Saint-Felix, Y.Trousset, C.Picard, C.Ponchut, R.Romeas, and A.Rougee," In vivo evaluation of a new system for 3D computerized angiography," Phys.Med.Biol., 39, pp.583-595, 1994

[2] D.P.Bertsekas," Non-linear programming," Athena Scientific, 1995

[3] Y.Censor and A.Lent," An iterative row-action method for interval convex programming," J.Opt.Theory Appl., 34, pp.321-353, 1981

[4] G.T.Herman and L.B.Meyer," Algebraic reconstruction techniques can be made computationally efficient," IEEE Trans.Med.Imaging, 12, pp.600-609, 1993

[5] T.Saito, H.Kudo, T.Takeda *et al.*," Three-dimensional monochromatic x-ray computed tomography using synchrotron radiation," Opt.Eng., 37, pp.2258-2268, 1998

[6] A.Dax," On minimum norm solutions," J.Opt.Theory Appl., 76, pp.183-193, 1993

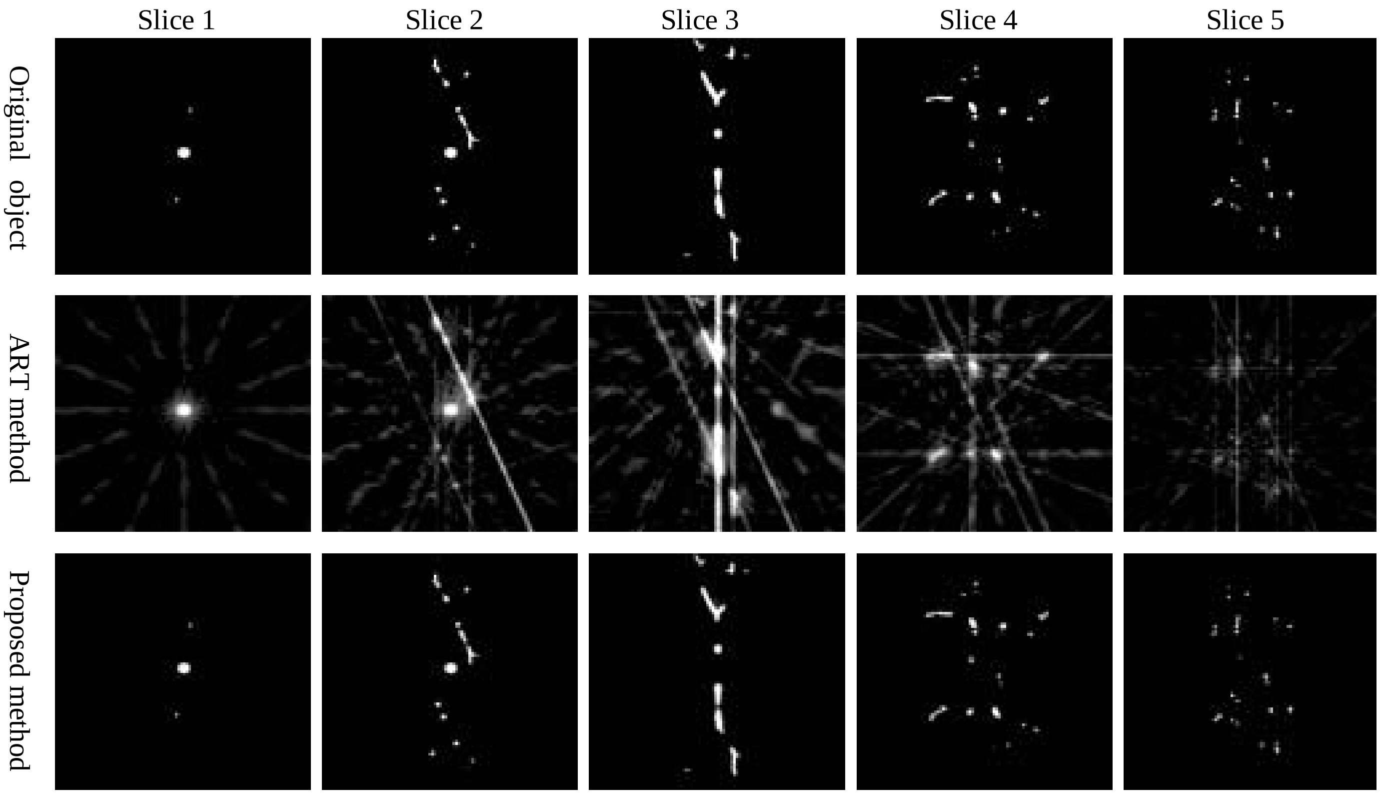


Fig. 1: Reconstructed transaxial slices from 8 projections.

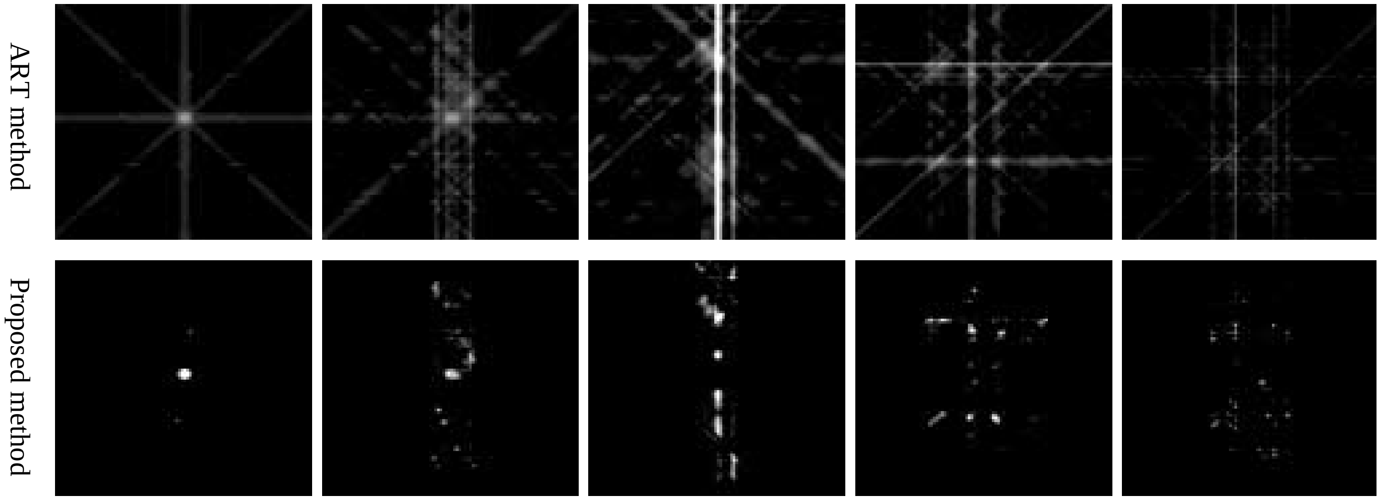


Fig. 2: Reconstructed transaxial slices from 4 projections.

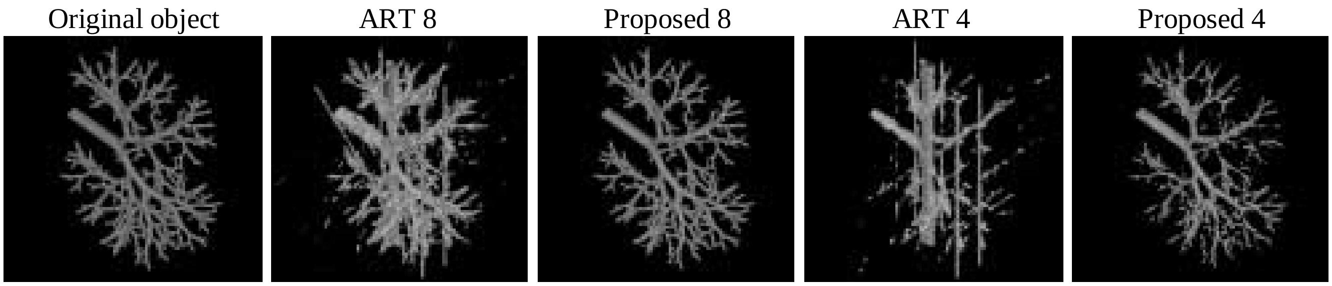


Fig. 3: 3-D graphic display of reconstructed images.

ART for helical cone-beam CT reconstruction

B. M. Carvalho[1,2], G. T. Herman[2], S. Matej[3]

[1]Department of Computer and Information Science, University of Pennsylvania.

[2]Center for Computer Science and Applied Mathematics, Temple University.

[3]Medical Image Processing Group, Department of Radiology, University of Pennsylvania.

Abstract

We report on our first results on the use of Algebraic Reconstruction Techniques (ART) on helical cone-beam Computerized Tomography (CT) data. Two variants of ART have been implemented: a standard one which considers a single ray in an iterative step and a block version which groups several cone-beam projections in calculating an iterative update. Both seem to produce high-quality reconstructions, although the number of cycles through the data to achieve those (between 15 and 20), while not huge, is larger than the number of cycles through the data needed for reconstructing volumes from data acquired from different modalities (1 iteration for PET data and 1 to 4 iterations for EM data). The reason for that maybe due to the uneven coverage of points by the data collection geometry, resulting in a slower rate of convergence.

I. Introduction

Algorithms for image reconstruction from projections form the foundations of modern methods of tomographic imaging in radiology, such as helical cone-beam X-ray computerized tomography (CT). Helical cone-beam CT is an image modality in which the cone-beam data acquisition is performed with a helical motion of the X-ray source-detector relative to the patient. The value of the helix pitch determines the speed of data acquisition, the bigger the pitch value, the faster the acquisition is. In [1] we showed that it is possible to obtain high-quality reconstructions from helical cone-beam CT data using ART (Algebraic Reconstruction Technique) even when applied to data acquired when using a considerably big pitch value.

An image modeling tool, which was described in a general context in [2, 3] and utilized in image reconstruction algorithms in [4, 5], is the representation of images and volumes using *blobs*, which are radially symmetric bell-shaped functions whose value at a distance r from the origin is

$$b_{m,a,\alpha}(r) = \frac{1}{I_m(\alpha)}\left[\sqrt{1-(r/a)^2}\right]^m I_m\left[\alpha\sqrt{1-(r/a)^2}\right], \qquad (1)$$

for $0 \leq r \leq a$ and is zero for $r > a$. In this equation I_m denotes the modified Bessel function of order m, a is the radius of the support of the blob and α is a parameter controlling the blob shape. A volume is represented as a superposition of N scaled and shifted versions of the same blob; i.e., as

$$\overline{f}(x,y,z) = \sum_{j=1}^{N} c_j b_{m,a,\alpha}\left(\sqrt{(x-x_j)^2+(y-y_j)^2+(z-z_j)^2}\right), \qquad (2)$$

where $\{(x_j,y_j,z_j)\}_{j=1}^{N}$ is the set of grid points in the three-dimensional (3D) Euclidean space to which the blob centers are shifted. Once we have chosen these grid points and the specific values of m, a and α, the volume is determined by the finite set $\{c_j\}_{j=1}^{N}$ of real coefficients; the task of the reconstruction algorithm in this context is to estimate this set of coefficients from the projection data.

The aim of [4, 5] was to study the choices of the grid points and of the parameters m, a and α, combined with implementation of the algorithm to estimate the coefficients, from the point of view of obtaining high-quality reconstructions in a reasonable time.

II. Helical cone-beam reconstruction using ART

It has been pointed out in [6] that applying the simplest form of ART to cone-beam projection data acquired on a circular trajectory can result in substandard reconstructions, and it has been suggested that a certain alteration of ART leads to improvement. However, besides an illustration of its performance, no properties (such as limiting convergence) of the algorithm have been given. We still need a mathematically rigorous extension of the currently available theory of optimization procedures to include acceptable solutions of problems arising from cone-beam data collection. We discuss this phenomenon in the context of reconstruction using ART with blobs from helical cone-beam data collected according to the geometry of [7]. In [1] we showed that ART can indeed produce high-quality results when applied to helical cone-beam data. In this paper we will concentrate on how to improve the convergence rate of the reconstruction algorithm by making use of a block-ART algorithm.

A. Standard ART

For this discussion we adopt the notation of [8], because it is natural both for the assumed data collection and for the mathematics that follows. We let I denote the number of times the X-ray source is pulsed as it travels its helical path multiplied by the number of lines for which the attenuation line integrals are estimated in the cone-beam for a single pulse. Thus I is the total number of measurements and we use Y to denote the (column) vector of the individual measurements y_i, for $1 \le i \le I$. We let N denote the number of grid points at which blobs are centered; our desire is to estimate the coefficients $\left\{c_j\right\}_{j=1}^{N}$ and thereby define a volume using (2). For $1 \le i \le I$, we let a_{ij} be the integral of the values in the jth blob along the line of the ith measurement (note that these a_{ij} can be calculated analytically for the actual lines along which the data are collected) and we denote by A the matrix whose ijth entry is a_{ij}. Then, using c to denote the (column) vector whose jth component is c_j, this vector must satisfy the system of approximate equalities:

$$Ac \approx Y. \tag{3}$$

In the notation of [8] the traditional ART procedure for finding a solution of (3) is given by the iterations:

$$\begin{aligned} &c^{(0)} \quad \text{is arbitrary,} \\ &c_j^{(n+1)} = c_j^{(n)} + \omega^{(n)} \frac{y_i - \sum_{k=1}^{N} a_{ik} c_k^{(n)}}{\sum_{k=1}^{N} a_{ik}^2} a_{ij}, \\ &\qquad \text{for } 1 \le j \le N, \\ &n = 0, 1, \ldots, \quad i = n \bmod I + 1, \end{aligned} \tag{4}$$

where $\omega^{(n)}$ is a relaxation parameter. While this procedure has a mathematically well-defined limiting behavior (see, e.g., Theorem 1.1 of [8]), in practice we desire to stop the iterations early for reasons of computational costs. We have found that for the essentially parallel-beam data collection modes of fully 3D PET [9], Fourier rebinned PET [10] and Transmission Electron Microscopy [11], one cycle through the data (i.e., $n = I$) is sufficient to provide us with high-quality reconstructions. However, our preliminary experiments indicate that this does not happen with helical cone-beam data.

We conjecture that the reason for this is the following. Let us associate with the jth blob the value

$$s_j = \sum_{i=1}^{I} a_{ij}, \text{ for } 1 \le j \le N. \tag{5}$$

For the parallel mode of data collection the values of s_j are nearly the same for all the blobs. However, this is not the case for cone-beam data. If we use the data collection geometry of [7], the blob coefficients closer to the helical source trajectory will have higher s_j values than the blob coefficients on the opposite side of the trajectory and, as can be seen in (4), this results in some blob coefficients being updated more frequently than others, making it harder for the iterative algorithm to converge to an acceptable solution.

B. Block-ART

It is natural to consider instead of the row-action algorithmic scheme (4) its block-iterative version, in which all the measurements taken by a number of pulses of the X-ray source form a block. A powerful theory is developed for this in [8]. Let M be the number of blocks, Y_i be the L-dimensional vector of those measurements which form the ith block and let A_i be the corresponding submatrix of A (we assume that each block has the same number of measurements). Theorem 1.3 of [8] states that the following block-iterative algorithm has good convergence properties:

$$\begin{aligned} &c^{(0)} \quad \text{is arbitrary,} \\ &c^{(n+1)} = c^{(n)} + A_i^T \Sigma^{(n)} \left(Y_i - A_i c^{(n)}\right), \\ &n = 0, 1, \ldots, \quad i = n \bmod M + 1, \end{aligned} \tag{6}$$

where $\Sigma^{(n)}$ is an $L \times L$ relaxation matrix. This theory covers even fully-simultaneous algorithmic schemes (just put all the measurements into a single block). There are also generalizations of the theory which allow the block sizes and the measurement-allocation-to-blocks to change as the iterations proceed.

A variation on such a block-ART algorithm is to perform component-dependent weighting in the update of blob coefficients. The essence of this approach is to introduce in (6) a second $(N \times N)$ relaxation matrix $\Delta^{(n)}$ in front of the A_i^T. Then we need to answer the following: For what simple (in the sense of computationally easily implementable) pairs of relaxation matrices $\Sigma^{(n)}$ and $\Delta^{(n)}$ can we simultaneously obtain desirable limiting convergence behavior and good practical performance by the early iterates. Examples of the $\Delta^{(n)}$ to be studied are the diagonal matrix whose jth entry is the reciprocal of the s_j of (5) or, alternatively, the reciprocal of a similar sum taken over only those measurements i which are in the block used in the particular iterative step. A recently proposed simultaneous reconstruction algorithm which uses j-dependent weighting appears in [12], where it is shown that a certain choice of such weighting leads to substantial acceleration of the algorithm's initial convergence.

Here we define the weights to be used in the updates based on the following idea. Suppose that we have taken the projection data of an object for which all the blob coefficients c_j are 1. Then, it appears desirable to have a uniform assignment of the blob coefficients after a single step of a modified version of (6), assuming that the initial assignment of the blob coefficients is zero. Assuming that the $\Sigma^{(n)}$ is the identity matrix, we can achieve this aim by choosing $\Delta^{(n)}$ to be a diagonal matrix whose jth entry is inversely proportional to the sum over all lines in the block of the line integral through the jth blob multiplied by the sum of the line integrals through all the blobs. The mathematical expression for this is

$$\sum_{l=1}^{L} \left(a_{[(i-1)L+l]j} \sum_{k=1}^{N} a_{[(i-1)L+l]k} \right). \tag{7}$$

In order for this to work we have to ensure that the value of (7) is not zero. This is likely to demand the forming of blocks which correspond to more than one pulse of the X-ray source, since the rays forming a block should intersect all blobs in the reconstruction region.

III. Results

Both ART (4) and the block-ART (described by (6) and (7)) algorithms were used to reconstruct a modified 3D Shepp-Logan phantom [13] in which the values range from 0.00 to 2.00, using data collected from two helix turns, with 300 projections taken per turn and 64 rows and 128 channels per projection (i.e. $I = 2 \times 300 \times 64 \times 128$ =4,915,200). The cone and fan angles of the cone-beam were 9.46° and 21.00°, respectively. The reconstructed volumes consisted of a $95 \times 95 \times 191$ blob coefficients array (2) organized on a bcc grid (see [4]) that was interpolated to a cubic grid with $128 \times 128 \times 128$ voxels. Figure 1 shows a (x,z)-slice of the volume reconstructed using the standard ART algorithm (a) and the block-ART algorithm (b) and a (y,z)-slice using the the standard ART algorithm (c) and the block-ART algorithm (d). The grayscale window used to show the slices was [1.00,1.03]. Both algorithms were executed for 17 cycles using $\omega^{(n)} = 0.01$ for the standard ART algorithm and 0.1 as a relaxation parameter (multiplying the identity matrix $\Sigma^{(n)}$) for the block-ART algorithm. For the block-ART algorithm, the measured data was grouped into 75 blocks formed by 8 cone-beams each. As one can see, the visual quality of these reconstructions is similar, although the blob-ART reconstruction seems to produce a more uniform backgroung inside the skull of the phantom. The time needed for both reconstructions is similar since the block-ART algorithm only carries a small overhead for computing the weights for each particular (blob,block) pair. (This overhead can be eliminated by precomputation and storage of the weights.)

IV. Discussion

We presented here our first results on the use of ART techniques for the reconstruction of helical cone-beam CT data. Our future work will concentrate on how to optimize the algorithms, by both speeding up the execution of a single cycle and improving the rate of convergence of the algorithms, and the evaluation of these algotihms and other block-ART variants.

References

[1] G.T. Herman, S. Matej, and B.M. Carvalho. Algebraic reconstruction techniques using smooth basis functions for helical cone-beam tomography. In D. Butnariu, Y. Censor, and S. Reich, editors, *Inherently Parallel Algorithms in Feasibility and Optimization and their Applications*, Studies in Computational Mathematics. Elsevier Science, Netherlands, to appear.

[2] R.M. Lewitt. Multidimensional digital image representations using generalized Kaiser-Bessel window functions. *J. Opt. Soc. Amer. A*, 7:1834–1846, 1990.

[3] R.M. Lewitt. Alternatives to voxels for image representation in iterative reconstruction algorithms. *Phys. Med. Biol.*, 37:705–716, 1992.

[4] S. Matej and R.M. Lewitt. Efficient 3D grids for image reconstruction using spherically-symmetric volume elements. *IEEE Trans. Nucl. Sci.*, 42:1361–1370, 1995.

[5] S. Matej and R.M. Lewitt. Practical considerations for 3–D image reconstruction using spherically symmetric volume elements. *IEEE Trans. Med. Imag.*, 15:68–78, 1996.

[6] K. Mueller, R. Yagel, and J.J. Wheller. Anti-aliased three-dimensional cone-beam reconstruction of low-contrast objects with algebraic methods. *IEEE Trans. Med. Imag.*, 18:519–537, 1999.

[7] H. Turbell and P.-E. Danielsson. Helical cone-beam tomography. *Internat. J. Imag. Systems Tech.*, 11:91–100, 2000.

[8] P.P.B. Eggermont, G.T. Herman, and A. Lent. Iterative algorithms for for large partitioned linear systems with applications to image reconstruction. *Linear Algebra Appl.*, 40:37–67, 1981.

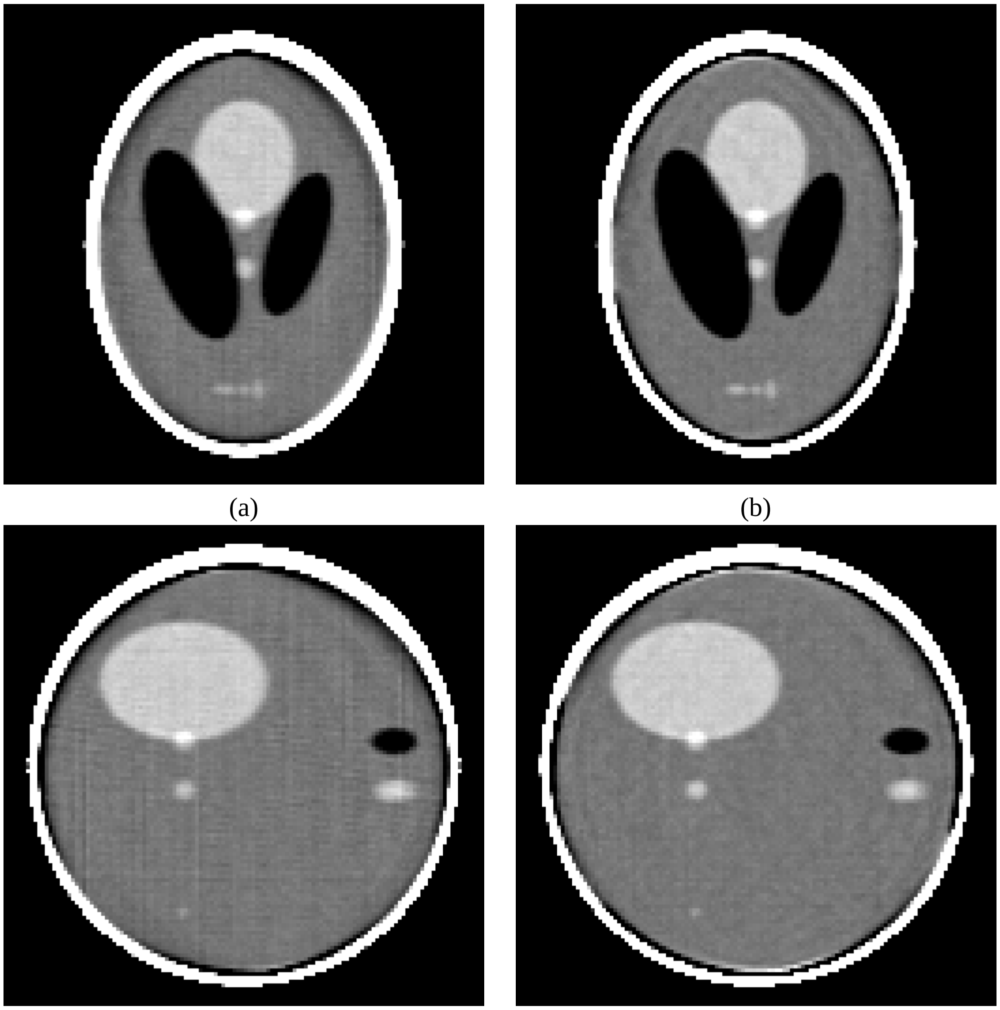
[9] S. Matej, G.T. Herman, T.K. Narayan, S.S. Furuie, R.M. Lewitt, and P. Kinahan. Evaluation of task-oriented performance of several fully 3–D PET reconstruction algorithms. *Phys. Med. Biol.*, 39:355–367, 1994.

[10] T. Obi, S. Matej, R.M. Lewitt, and G.T. Herman. 2.5D simultaneous multislice reconstruction by series expansion methods from Fourier-rebinned PET data. *IEEE Trans. on Med. Imag.*, 19:474–484, 2000.

[11] R. Marabini, G.T. Herman, and J.M. Carazo. 3D reconstruction in electron microscopy using ART with smooth spherically symmetric volume elements (blobs). *Ultramicrosc.*, 72:53–65, 1998.

[12] Y. Censor, D. Gordon, and R. Gordon. Component averaging: An efficient iterative parallel algorithm for large and sparse unstructured problems. *Parallel Comput.*, 27:777–808, 2001.

[13] C. Jacobson. *Fourier Methods in 3D-Reconstruction from Cone-Beam Data*. PhD thesis, Department of Electrical Engineering, Linkoping University, 1996.



(a) (b)

(c) (d)

Figure 1: Slices of reconstructed Shepp-Logan phantom using the standard ART algorithm (a) and (c), and the block-ART algorithm (b) and (d), showed using the gray-scale window settings of [1.00,1.03] (images (a) and (b) show a (x, z)-slice while images (c) and (d) show a (y, z)-slice). Both algorithms were executed for 17 cycles.