ICIP-2001 paper proposal:

Tomographic Image Reconstruction Using Content-Adaptive Mesh Modeling

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 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 can also lead to development of efficient numerical reconstruction algorithms. The proposed methods are tested using gated cardiac-perfusion images. 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 improving the quality of reconstructed images in tomography. Most of these methods are pixel-based, i.e., the image is represented and computed directly in a pixel basis. Bayesian priors (e.g., [1]) or regularization terms (e.g., [2]) are typically used to combat the effect of noise.

Alternative model-based 2D and 3D 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. 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 the underdetermined nature of the reconstruction problem; 2) this reduction in the number of unknowns can lead to a fast computation; and 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. 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

$$f(\mathbf{x}) = \sum_{n=1}^{N} \varphi_n(x) f(\mathbf{x}_n), \tag{1}$$

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.

Now let $\mathbf{n}$ denote a vector formed by the nodal values of the mesh model, i.e.,

$$\mathbf{n} \equiv \left[f(\mathbf{x}_1), f(\mathbf{x}_2), \cdots f(\mathbf{x}_n) \right]^{\mathrm{T}}. \tag{2}$$

If $\mathbf{f}$ denotes the pixel representation of the image function $f(\mathbf{x})$ over D, then from (1) and (2) one can obtain

$$\mathbf{f} = \Phi \mathbf{n}, \tag{3}$$

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 pixel representation $\mathbf{f}$ as

$$E[\mathbf{g}] = \mathbf{Hf}, \quad (4)$$

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:

$$E[\mathbf{g}] = \mathbf{H}\Phi\mathbf{n} \equiv \mathbf{An}, \quad (5)$$

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 investigated maximum-likelihood and least-squares estimates of $\mathbf{n}$.

A. Maximum-Likelihood Estimate

The maximum-likelihood (ML) estimate is obtained as

$$\hat{\mathbf{n}}_{ML} = \arg\max_{\mathbf{n}} \left\{ \log\left[p(\mathbf{g};\mathbf{n}) \right] \right\}, \quad (6)$$

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]:

$$\mathbf{n}_s^{(j+1)} = \frac{\mathbf{n}_s^{(j)}}{\sum_t \mathbf{A}_{ts}} \sum_t \mathbf{A}_{ts} \left(\frac{\mathbf{g}_t}{\sum_k \mathbf{A}_{tk} \mathbf{n}_k^{(j)}} \right), \quad (7)$$

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:

$$\hat{\mathbf{n}}_{LS} = \arg\min_{\mathbf{n}} \|\mathbf{g} - \mathbf{An}\|^2, \quad (8)$$

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 four-dimensional (4D) gated mathematical cardiac-torso gMCAT D1.01 phantom [10]. This is a time sequence of 16 three-dimensional (3D) 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 conditions observed in a typical clinical Tc^{99m} study. In our experiments, a single slice (No.70) was chosen, which has 55,000 counts per frame (a total of 16 frames). No attenuation map was used. Each image frame was reconstructed separately, and a single mesh structure is 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 $\alpha = 1$, $\beta = 1.2$, $\delta = 3$, $\gamma = 0.35$. 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 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:

$$\vartheta(\mathbf{x}) = \max\left(\left|\nabla_{xx}^2 \bar{f}(\mathbf{x})\right|, \left|\nabla_{xy}^2 \bar{f}(\mathbf{x})\right|, \left|\nabla_{yy}^2 \bar{f}(\mathbf{x})\right| \right). \quad (9)$$

From this image, we compute a feature map as follows:

$$\sigma(\mathbf{x}) = \begin{cases} \vartheta(\mathbf{x})^{0.475} & \mathbf{x} \in \textit{Heart region} \\ \vartheta(\mathbf{x})^{0.95} & \mathbf{x} \in \textit{Background} \end{cases}. \qquad (10)$$

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 1/7 the number of pixels. Note that the algorithm places mesh nodes densely in the important heart regions, and sparingly in the background. This mesh was used as a basis on which to reconstruct each of the image frames in the sequence. In future work, we will optimize the mesh to track motion from frame to frame.

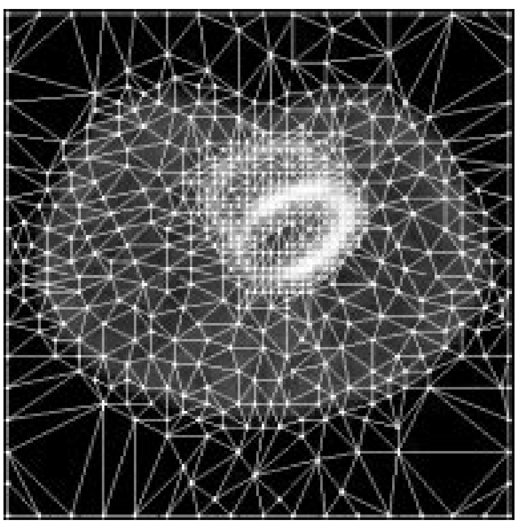

Figure 1. Content-adaptive mesh model of the torso, including the heart, using 609 mesh nodes.

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.

A final note is that 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).

4. REMARKS

In this paper we showed that the use of a CAMM in image reconstruction can achieve improved image quality at low computational cost. By the time of the conference, we hope to further develop this approach to achieve motion-compensated reconstruction of dynamic or gated image sequences by deforming the mesh model.

5. REFERENCE

[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," *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, "Trhee 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.

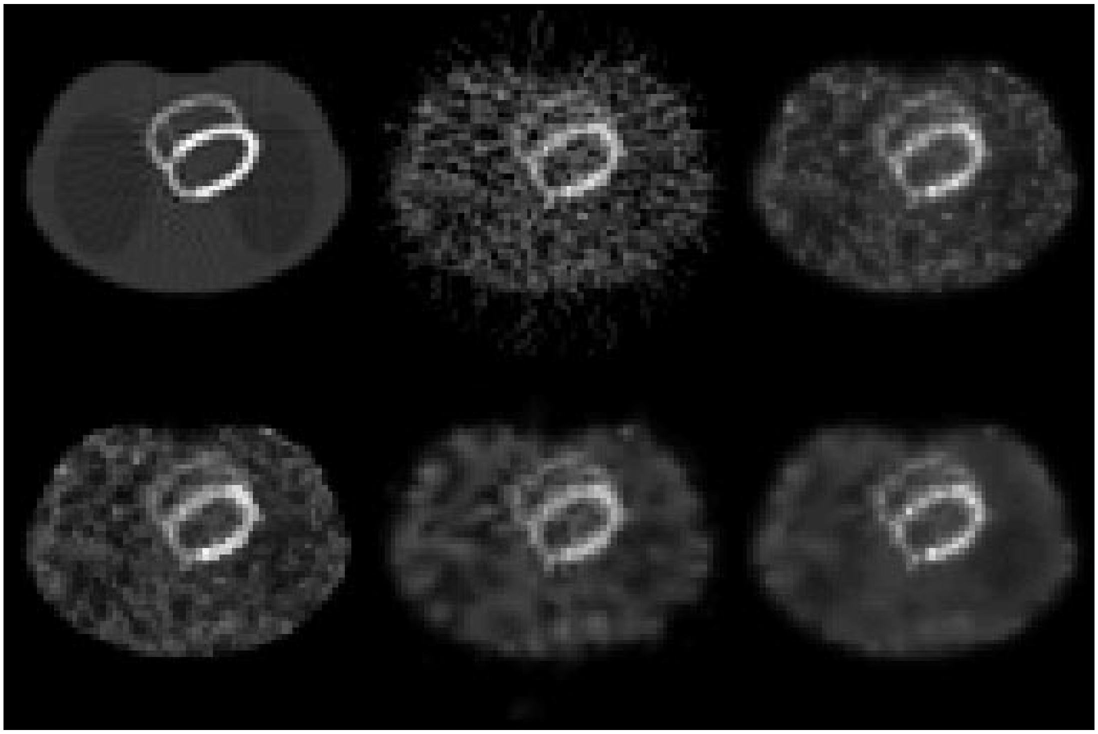

Figure 2. From left to right in top row: Original phantom, Filtered backprojection reconstruction, ML-EM reconstruction. Bottom row: MAP reconstruction, MESH-LS reconstruction, and MESH-EM reconstruction.

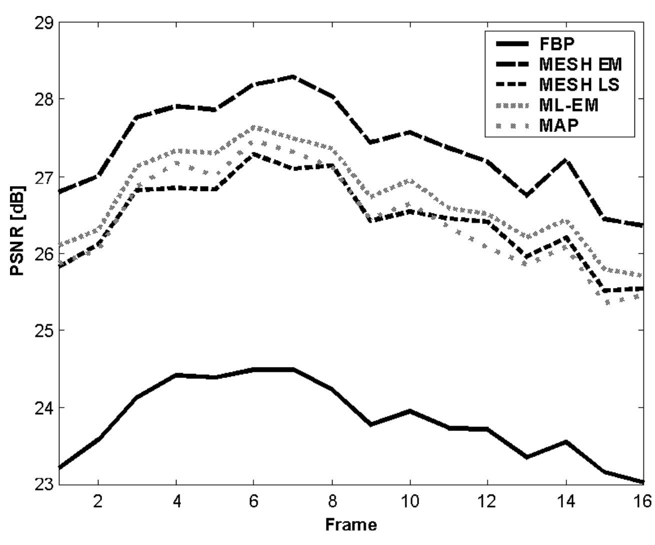

Figure 3. PSNR vs. frame number for various reconstruction methods.

TABLE 1. RECONSTRUCTION METHODS COMPARISON FOR EXECUTION TIME, ADDITIONAL MEMORY REQURAMENT AND AVREG PSNR.

	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