CONTENT-ADAPTIVE MESH MODELING FOR FULLY-3D TOMOGRAPHIC IMAGE RECONSTRUCTION[1]

Yongyi Yang, Jovan G. Brankov, and Miles N. Wernick

Dept. of Electrical and Computer Engineering
Illinois Institute of Technology
3301 S. Dearborn St., Chicago, IL 60616

ABSTRACT

In this paper we propose the use of a content-adaptive volumetric mesh model for fully three-dimensional (3D) 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 mesh representation can 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, resulting in efficient algorithms. The proposed methods are tested using gated cardiac-perfusion images. Initial results demonstrate that the proposed approach achieves good performance when compared to several commonly used methods for image reconstruction, and produces results very rapidly.

1. INTRODUCTION

In recent years there has been growing interest in fully-3D tomographic image reconstruction. A major challenge in fully-3D reconstruction lies in its memory requirement and demanding computation time. Like their 2D counterpart, most 3D reconstruction methods have traditionally been developed based on voxel image representations [1]. Bayesian priors (e.g., [2]) or regularization terms (e.g., [3]) are often used to combat the effect of noise.

Alternative model-based reconstruction approaches have also been proposed. For example, cylindrical models were proposed in [4] and surface models were used in [5,6].

In our previous work in [7], a content-adaptive mesh modeling approach was proposed for 2D image reconstruction. It was demonstrated that such an approach can outperform several well-known reconstruction algorithms in terms of both reconstructed image quality and computation time. In this study, we extend this approach to fully-3D image reconstruction. In this new approach, the image is first modeled by a volumetric mesh model, on the basis of which a customized basis representation is obtained for the image. The parameters of this representation are then 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 [8]. 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. To demonstrate the idea, we show in Figure 1 (left) a mesh model obtained for a 2D cardiac perfusion image in our previous work [7]; also shown in Figure 1 (right) is a 3D mesh model where, for clarity, only the distribution of the mesh nodes on the surface of organs is shown.

The potential benefits of using a CAMM for image reconstruction include: 1) it provides an compact representation of the image in that the number of parameters (i.e., mesh nodes) is typically much less than the number of required voxels in a voxel image representation, thus reducing the number of unknowns. This can help alleviate both the underdetermined nature of the reconstruction problem and the data storage

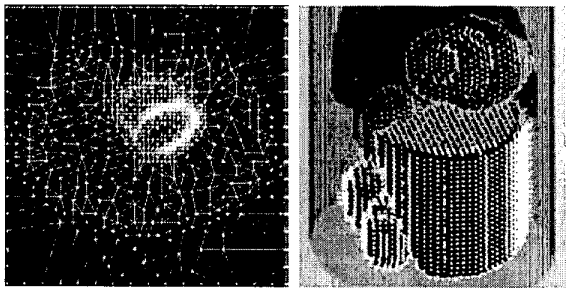

Figure 1. *Left*: a 2D content-adaptive mesh model of the torso, including the heart; *Right*: a 3D mesh model where, for clarity, only the distribution of the mesh nodes on the surface of organs is shown.

[1] This research was supported by the Whitaker Foundation and by NIH/NHLBI grant HL65425.

requirement, particularly for the case of 3D reconstruction; 2) this reduction in the number of unknowns can also lead to a fast computation; 3) a CAMM provides a natural spatially-adaptive smoothness mechanism; and 4) a mesh model can also be used for motion tracking in an image sequence, by allowing the mesh to deform over time [9]. Thus, a CAMM also 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 is 3D in this study. In a mesh model, the domain D is partitioned into M non-overlapping mesh elements, denoted by D_m, $m=1,2,\cdots,M$. The image function is represented as

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

where $\mathbf{x}_n$ is the nth mesh node, $\varphi_n(\mathbf{x})$ is the interpolation basis function associated with $\mathbf{x}_n$, N is the total number of mesh nodes used, and $e(\mathbf{x})$ is the modeling error. 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 this study, tetrahedrons are used for D_m, and linear interpolation functions are used for $\varphi_n(\mathbf{x})$.

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

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

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

$$\mathbf{f}=\Phi\mathbf{n}+\mathbf{e}, \qquad (3)$$

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, and $\mathbf{e}$ is a vector denoting the error $e(\mathbf{x})$.

For tomographic image reconstruction, the imaging equation is typically written in terms of the voxel representation $\mathbf{f}$ as

$$E[\mathbf{g}]=\mathbf{H}\mathbf{f}, \qquad (4)$$

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:

$$E[\mathbf{g}]=\mathbf{H}[\Phi\mathbf{n}+\mathbf{e}]\equiv\mathbf{A}\mathbf{n}+\hat{\mathbf{e}}, \qquad (5)$$

where $\mathbf{A}=\mathbf{H}\Phi$, and $\hat{\mathbf{e}}=\mathbf{H}\mathbf{e}$.

The modeling error $\hat{\mathbf{e}}$ in (5) can be ignored when compared to the noise level in the imaging data, as a CAMM can provide a very accurate representation of the original image. Thus, we have

$$E[\mathbf{g}]\approx\mathbf{A}\mathbf{n}. \qquad (6)$$

The reconstruction problem becomes that of estimating $\mathbf{n}$ from the observed data $\mathbf{g}$. The image $\mathbf{f}$ can then be obtained from (3) (with $\mathbf{e}$ ignored).

2.2 Mesh domain reconstruction

In this paper we investigate maximum-likelihood (ML) estimate of the nodal values in $\mathbf{n}$. The ML estimate is obtained as

$$\hat{\mathbf{n}}_{ML}=\arg\max_{\mathbf{n}}\left\{\log\left[p(\mathbf{g};\mathbf{n})\right]\right\}, \qquad (7)$$

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 [10]:

$$\mathbf{n}_s^{(j+1)}=\frac{\mathbf{n}_s^{(j)}}{\sum_t \mathbf{A}_{ts}}\sum_t \mathbf{A}_{ts}\left(\frac{\mathbf{g}_t}{\sum_k \mathbf{A}_{tk}\mathbf{n}_k^{(j)}}\right), \qquad (8)$$

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 $\mathbf{A}$. This method is designated as MESH-EM in the following.

3. PRELIMINARY RESULTS

3.1 Evaluation image data

To demonstrate the proposed CAMM-based reconstruction approach, we used the 4D gated mathematical cardiac-torso (gMCAT) D1.01 phantom [12], which is a time sequence of 16 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 a clinical Tc^{99m} study. No attenuation correction was used.

3.2 Reconstruction methods considered

In addition to the two proposed reconstruction algorithms, we also considered in this preliminary study the following two well-known reconstruction procedures for comparison purposes: (1) filtered back projection (FBP); and (2) pixel-based ML. To help reduce the noise level, the results from these two methods were post-processed with a 3D low-pass filter of order 17 with a cutoff frequency of 0.65 (normalized by π). For consistency in the comparison, this same post-filtering was also applied to the proposed

mesh reconstruction method in the final results. Each of the iterative reconstruction algorithms was run for 30 iterations.

3.3 Volumetric mesh generation

The key to the proposed approach lies in how to construct a CAMM that is compact and accurate for representing the volumetric image to be reconstructed. For this purpose we extended our method in [13] to the 3D case. This method consists of the following three steps: 1) extract a feature map $\sigma(\mathbf{x})$ from the image $f(\mathbf{x})$ based on the largest magnitude of its second directional directives; 2) apply the well-known Floyd-Steinberg error-diffusion algorithm, a method originally designed for digital halftoning [14], to distribute mesh nodes non-uniformly in the 3D image domain, with density proportional to the feature map $\sigma(\mathbf{x})$; and 3) use a 3D Delaunay triangulation algorithm [15] to connect the mesh nodes.

Of course, for tomographic image reconstruction the mesh structure has to be estimated from the observed data. The following procedure was demonstrated to work well in our studies. First, the projection data are summed over the 16 gated frames. From these summed projections an image is reconstructed using the filtered back projection (FBP) algorithm. The resulting image, denoted by $\bar{f}(\mathbf{x})$, provides a rough estimate of the heart summed over all 16 frames. The mesh structure is then created based on $\bar{f}(\mathbf{x})$ using the steps described above. The resulting 3D mesh was shown in Figure 1 (right), where, for clarity, only the distribution of the mesh nodes (instead of the tetrahedral elements) was shown. The number of mesh nodes used was 6,494 (comparing to 131,072 voxels). As can be seen, the mesh obtained by the proposed method is well adapted to the content of the 3D volumetric image. Specifically, mesh nodes had been placed densely in the important heart regions, and sparingly in the background. More results are provided at the following web site in animations for better visualization purposes: *http://www.ipl.iit.edu/brankov/Rotate.htm.*

The obtained mesh structure was then used as a basis on which each of the 16 3D image frames in the sequence was reconstructed. In future work, we will optimize the mesh to track motion from frame to frame.

3.4 Fully 3D CAMM reconstruction

For visual comparison, some representative 2D slices of frame #1, obtained by different reconstruction methods, are shown in Figure 2. The images in Fig.2(a) were from the original phantom, degraded by the intrinsic system blur. These images represent the ideal case of noise-free projection data. The images reconstructed using FBP are shown in Fig.2(b). The ML-EM results are shown in Fig.2(c), and the MESH-EM results are given in Fig.2(d). The MESH-EM algorithm appears to produce slightly better images, capturing the heart wall and achieving reasonable smoothness in the background. The ML-EM algorithm produced similar results, but slightly more noisy than the MESH-EM. As a preliminary assessment of the accuracy, the peak-signal-to-noise-ratio (PSNR) was computed for the reconstructed 3D images. The PSNRs of the reconstructed images for frame #1 by the FBP, ML-EM, and MESH-EM are 17.69 dB, 21.46 dB, and 22.22 dB, respectively.

As for the execution time, the MESH-EM takes about 2.8 sec for one 3D frame, while the ML-EM takes about 19.7 sec (implemented in MATLAB on a 2 GHz Pentium-4 PC). A note is that the MESH-EM requires an overhead of pre-computing the mesh-domain matrix $\mathbf{A} = \mathbf{H}\Phi$ in Eq.(6). In our implementation the total time for computing $\mathbf{A}$ was around 48 sec per frame (equivalent to about 2.5 iterations of ML-EM). This can be further reduced in a more efficient implementation. Nevertheless, the effect of this overhead will diminish after only 3 iterations as the MESH-EM is mush faster per iteration when compared to the ML-EM.

These results, though very preliminary, indicate that the use of a CAMM in fully 3D image reconstruction can achieve good image quality at low computational cost. We will use more comprehensive evaluation metrics (e.g. bias-variance plot) to better characterize the performance of the proposed technique in future work.

5. REFERENCES

[1] B. Bendriem and D. W. Townsend, *The Theory and Practice of 3D PET*, Kluwer Academic Publishers, 1998.

[2] 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.

[3] J. Fessler, "Penalized weighted least-squares reconstruction for positron emission tomography," *IEEE Trans. Med. Imaging*, vol. 13, pp. 290-300, 1994.

[4] 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.

[5] 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.

[6] G. R. Jennings and D. R. Wolf, "Tomographic reconstruction based on flexible geometric models," *IEEE Int. Conf. on Image Proc.*, 1994.

[7] J. G. Brankov, Y. Yang, and M. N. Wernick, "Tomographic image reconstruction using content-adaptive mesh modeling," *IEEE Inter. Conf. Image Proc.*, vol. 1, pp. 690-693, Thessaloniki, Greece, Oct., 2001.

[8] 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.

[9] 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.

[10] 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.

[11] E. K. P. Chong and S. H. Zak, *An Introduction to Optimization.* New York: John Wiley & Sons, Inc., 1996.

[12] 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.

[13] Y. Yang, J. Brankov, and M. N. Wernick, "A fast algorithm for accurate content-adaptive mesh generation," *IEEE Inter. Conf. Image Proc.*, vol. 1, pp. 868-871, Thessaloniki, Greece, Oct., 2001.

[14] R. Floyd and L. Steinberg, "An adaptive algorithm for spatial gray scale," SID Int. Sym. Digest of Tech. Papers, 1975.

[15] F. Preparata and M. Shamos, *Computational Geometry--An Introduction*, Springer-Verlag, New York, 1985.

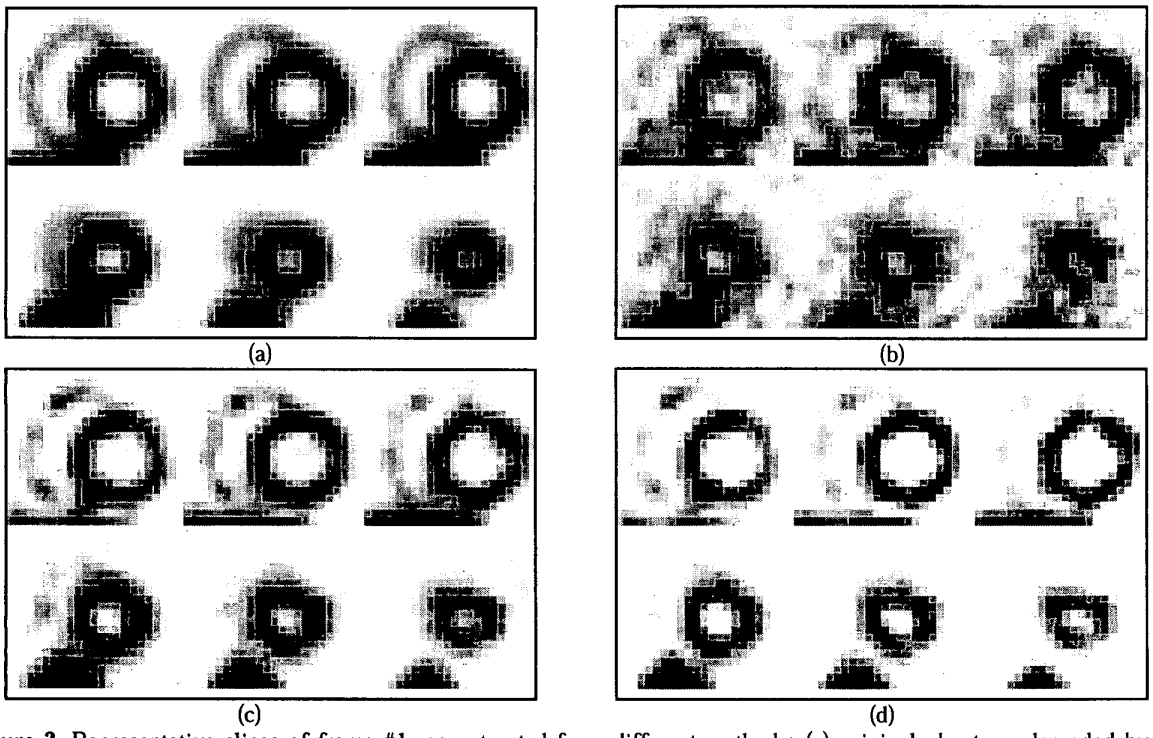

Figure 2. Representative slices of frame #1 reconstructed from different methods: (a) original phantom, degraded by the intrinsic system blur, (b) FBP with post-filtering, (c) ML-EM, and (d) proposed MESH-EM. For consistency, the same post-filtering was also applied to the images in (c) and (d).