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.

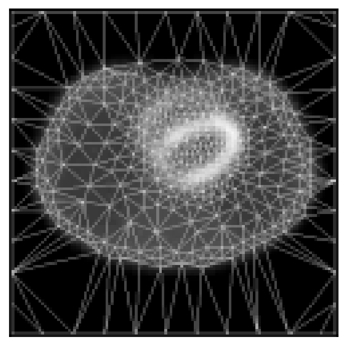

Figure 1. Mesh structure used in our experiment.

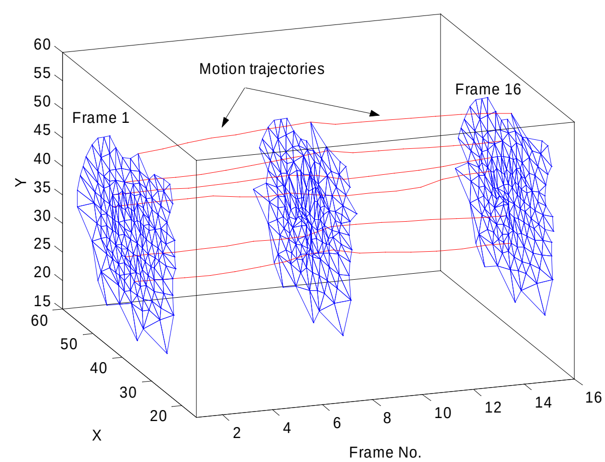

Figure 2. Deformable mesh (shown for frames 1, 8, and 16), and motion trajectories for some selected mesh nodes throughout the sequence.

[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]:

$$J = \frac{1}{2} W_m \sum_{m=1}^{M} \left[\int_{D_m} \left(f_t(\mathbf{x} + \mathbf{d}(\mathbf{x})) - f_r(\mathbf{x}) \right)^2 d\mathbf{x} \right] + \frac{1}{2}(1 - W_m) E_d \quad , \tag{2}$$

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:

$$h_t(l) = \frac{1}{C}\left(1 - \frac{2|l|}{K}\right)^{\gamma}, l = -K/2, \ldots, K/2, \tag{6}$$

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.

[2] Image sequences available for cine viewing at *http://www.iit.edu/~branjov/3D01.htm*

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.

[12] K. Lange and R. E. Carson, "EM reconstruction algorithms for emission and transmission tomography," *J. Comput. Assist. Tomogr.*, vol. 8, no.2, pp. 306–316, 1984.

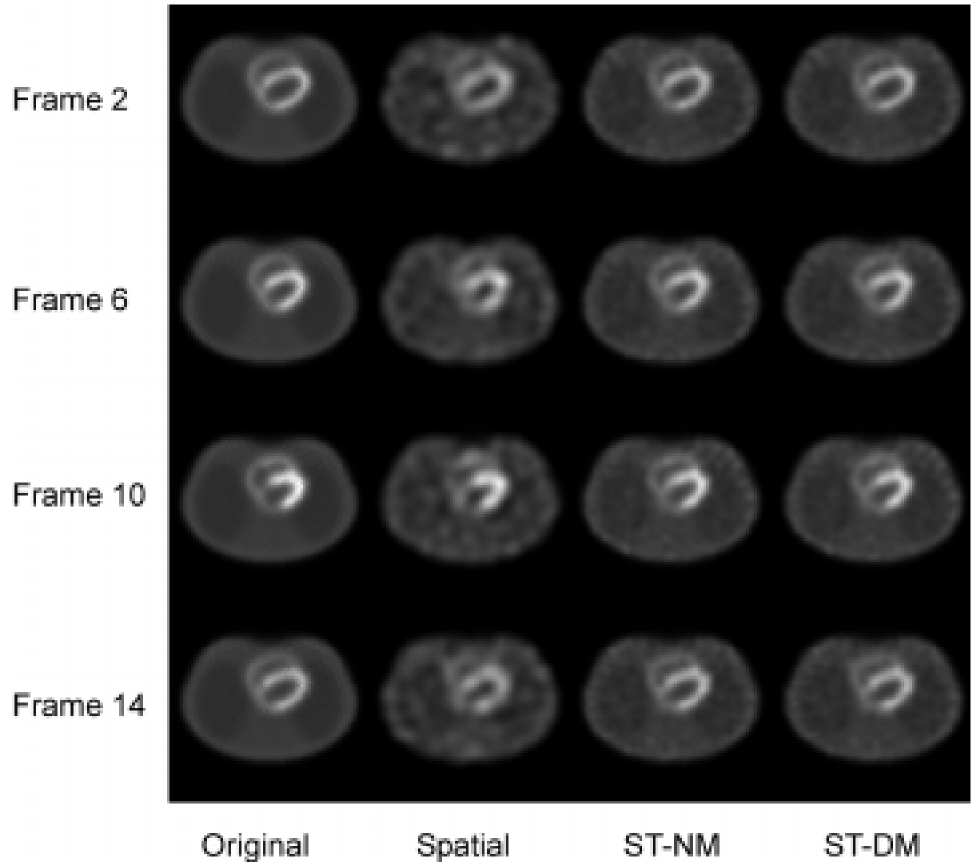

Figure 3. Results obtained by maximum-likelihood expectation-maximization. "Spatial" denotes spatial smoothing only. "ST-NM" denotes spatial-temporal smoothing without motion compensation. "ST-DM" denotes the proposed spatial-temporal smoothing with motion compensation achieved using a deformable mesh. "Original" denotes the phantom degraded by the system blur to represent an approximate best-case image for comparison.

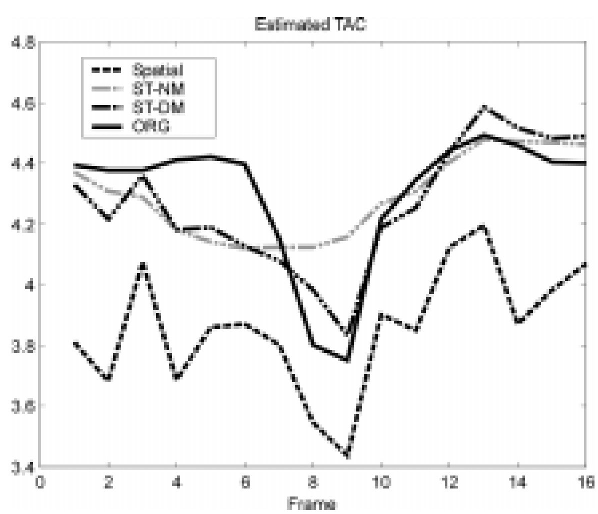

Figure 4. Time activity curves (TAC) for a small region in the left ventricular wall vs. the frame number. Note the failure of ST-NM to capture the motion at frame 8 and 9.