DYNAMIC TEXTURE BASED HEART LOCALIZATION AND SEGMENTATION IN 4-D CARDIAC IMAGES

Junzhou Huang, Xiaolei Huang, Dimitris Metaxas

CBIM, DCIS,
Rutgers University, NJ, USA

Leon Axel

Department of Radiology,
New York University, New York, USA

ABSTRACT

In this paper we present a dynamic texture based motion segmentation approach to address the challenging problem of heart localization and segmentation in 4D Spatio-temporal cardiac images. Our approach introduces time-dependent dynamic constraints into model-based segmentation, and it has the advantage of producing segmentation results that are both spatially and temporally consistent. Compared with previous methods that segment cardiac contours, our method offers the following advantages: 1) with distinct dynamic signatures, the heart can be quickly localized in 4D cardiac images; 2) heart dynamics are learned online and adaptively by analyzing the dynamic texture from the video sequence of a cardiac cycle and then incorporated in the segmentation process; and 3) the proposed dynamic features can be easily integrated with model-based segmentation methods. We illustrate our framework through combining the new dynamic constraints with active contour models, and demonstrate its performance on sequences of 4D MRI and tagged MRI images of the heart. We also validate the accuracy of the segmentation results by comparing with ground truth marked by experts.

1. INTRODUCTION

Cross-sectional imaging techniques such as ultrasound, CT, conventional MRI, and tagged MRI can provide non-invasively tomographic images of the heart at different times in the cardiac cycle. For instance, a typical 4D spatial-temporal MRI data set has cross-sectional images of the heart at 18 locations over 24 time points in the cardiac cycle. Given such large amount of data, quantitative analysis of the heart wall motion using these images becomes attractive. Systematic quantitative analysis of heart-wall motion and blood flow often requires accurate localization and segmentation of the epicardial and endocardial surfaces on 4D images from large groups of population with high efficiency. Since manual drawing of cardiac contours is very time-consuming, there has been vigorous research on methods to automate the cardiac segmentation [1, 2, 3, 4, 5] and motion estimation [6, 7] process.

These methods largely fall into four categories: Morphological [1, 2], Deformable Model based [3, 6, 7], Statistical Shape and Appearance Model based [4, 8], and Algebraic based [9]. The morphological approaches [1, 2] apply a hierarchical series of image processing steps to locate, segment and track the Left Ventricle (LV). It is more difficult for them to handle the Right Ventricle (RV) since some RV cross sections are too thin to be reliably detected by morphological operations. Model-based methods for cardiac segmentation are gaining popularity because they comine high-level knowledge with low-level image features to achieve robustness against noise, image artifacts, cluttered objects and complext heart geometry and motion. The models being used are either active contour (or deformable) models [3, 6, 7] that deform under the influence of internal smoothness and external image forces toward heart boundary, or statistical shape and appearance models [10, 4, 8] that are learned *a priori* from examples to capture variations in shape and appearance of the heart in images of a particular modality. The deformable models do not require learning and can estimate boundary with smooth curves/surfaces that bridge over boundary gaps, but their accuracy highly depends on initialization. Some recent works have introduced region and appearance constraints [11, 12] into traditional deformable models to alleviate the requirement of manual initializations. Statistical shape and appearance models [10, 4, 8] can encode high-level knowledge about shape and appearance of the heart in a more specific manner and are hence more robust in boundary segmentation. The main concerns related to these models are that the data collection, annotation and training processes are often time-consuming, and prior models learned on images from one imaging facility may not be useful on images from other facilities with different imaging parameters. Algebraic based method [9] exploits Polysegment and spatial Generalized Principal Component Analysis (GPCA) to segment beating hearts based on techniques of dynamic texture analysis developed in [13]. This algebraic technique is less computationally intensive, and good qualitative results in [9] demonstrate the advantage of dynamic texture analysis. However, the algorithm does not output the final cardiac contours and there were no quantitative results to confirm its accuracy on large datasets.

In this paper, we observe that, while deformable-model based methods [3, 7] do not require *a priori* learning and tend to be more efficient, previous methods do not utilize fully the dynamic information in a video sequence to facilitate heart localization and segmentation. On the other hand, Algebraic

based method [9], shows the advantages of dynamic texture analysis [13], but its accuracy is not validated through quantitative results. In this paper, we adapt the technique of dynamic texture analysis [13] to make it easy to be coupled with deformable-model framework. After obtaining dynamic parameters learned from a single image sequence, we can automate heart localization in the sequence, then augment active contour models using the derived dynamic information to achieve robust and temporally-consistent segmentation and output smooth cardiac contours. Both qualitative and quantitative results show the power of our method in heart localization and segmentation.

2. BACKGROUND AND OUTLINE

2.1. Dynamic texture segmentation

Dynamic textures are sequences of images of moving scenes that exhibit certain stationary properties over time [13][14], which can be modelled by a linear dynamical system. The dynamics is represented as a time-dependent state process $x_t \in R^N$, and the appearance of the image frame at time t, $y_t \in R^M$, can be modelled as follows:

$$x(t+1) = A_i * x(t) + B_i * v(t) \quad (1)$$

$$y_i(t) = C_i * x(t) + D_i * w(t) \quad (2)$$

In the equation, $i = 1, ..., K$ represents that there are K partitions, $\{\Omega_i\}_{i=1,...,K}$, in the image sequence domain, where each region has distinct dynamic signatures. $A_i \in R^{N \times N}$ and $C_i \in R^{M_i \times N}$ are dynamic parameter matrices. As introduced in [13],with the obtained A_i and C_i, we can compute an observation matrix O_i:

$$O_i = [C_i^{'}, A_i^{'} * C_i^{'}, ..., (A_i^{'})^N * C_i^{'}]' \quad (3)$$

The observation matrix at every pixel, $\bar{x} \subset \Omega$, $O(\bar{x})$, can be derived from its neighboring pixels $Neigh(\bar{x}) \subset \Omega$. Then, the quantitative dynamic signature of each pixel can be computed as:

$$S(\bar{x}) = (\cos\gamma_1(\bar{x}), ..., \cos\gamma_N(\bar{x})) \quad (4)$$

where $\{\gamma_j(\bar{x})\}_{j=1,...,N}$ is the principle angle between the subspaces of the observation model $O(\bar{x})$ and a reference model $O(\bar{x_0})$. These dynamic signatures offer high-dimensional distance measures that are powerful. Thus, they can help robust image segmentation under challenging conditions.

2.2. Proposed improvements and outline of our approach

The dynamic textures introduced above is a recently-developed motion segmentation technique that is particularly useful in image sequences that capture moving scenes with underlying stationary properties. However, a linear dynamical system is assumed in the original technique. In this paper, we adapt this technique to cardiac motion analysis by extending its linear dynamics assumption to piece-wise linear since cardiac motion is nonlinear and periodic.

Our overall heart localization and segmentation framework is outlined by the flow-chart in Figure 1. The major algorithmic steps include: 1) subtracting background of the input image sequence; 2) computing dynamic signatures of the input using dynamic texture analysis; 3) fast K-mean clustering according to the dynamic signatures of each pixel [15]; 4) localizing the interested object; 5) combining the dynamic constraints with shape and appearance information in an active contour model-based framework to segment the interested object.

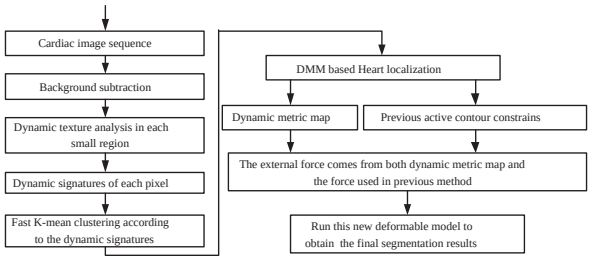

Fig. 1. The flow chart of our approach

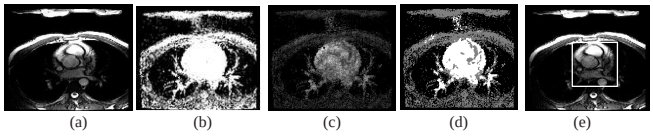

Fig. 2. Localization example. (a) one example frame, (b) background subtraction, (c) Dynamic Metric Map, (d) K-means clustering, (e) heart localization.

3. METHODOLOGY

3.1. Dynamic metric map

Suppose $P(X,t)$ represents a cardiac image sequence. X denotes the position (x,y) and $t \in T$ denotes the time axis. Before dynamic analysis, we first exclude those non-moving pixels by subtracting the background of the image sequence :

$$Var(X) = Var(P(X,t),2) \quad (5)$$

$$Var(X) > TH \quad (6)$$

Here, $Var(P(X,t),2)$ represents the difference or subtraction operation between two consecutive frames, and TH is a threshold value below which we consider as non-moving background. This step is similar to ground-figure separation based on motion differences. Figure 2 (a) shows one of the frames from a cardiac image sequence. In this sequence, the heart is deforming due to heart beating; the chest cavity moves due to breathing and image noise; and there also exists the nearly non-moving image background. Figure 2 (b) shows the separation result after subtracting the non-moving background.

To perform dynamic texture analysis, we retrieve an $l \times l$ window centered at $P(X)$ on each frame. Let us denote the

retrieved $l^2 \times T$ volume as $V(X)$, where T is the number of the image frames. With dynamic analysis, we can generate a local spatio-temporal signature $S(X)$ by Eq. 4 using pixel dynamics in the region $V(X)$. Now having computed the dynamic signature for each pixel, we can compute the dynamic metric map for the whole image sequence. For each pixel X, we compute the Martin's distance $D(X)$ to combine together all elements in a dynamic signature $S(X)$ [13].

Now, each pixel has one dynamic value based on $D(X)$. We can then build the Dynamic Metric Map (DMM) for each image sequence using the pixels' dynamic metric values:

$$DMM(X) = D(X) \tag{7}$$

Figure 2 (c) shows the derived Dynamic Metric Map for the example cardiac sequence. From the result, we can see that the dynamic metric value within the heart region is very different from those in the chest wall region in the Dynamic Metric Map. The difference is much more distinctive compared with the intensity difference between these regions as shown in Figure 2 (a). This indicates that it is more advantageous to use dynamic information for heart localization and segmentation than using appearance information alone.

3.2. DMM based heart localization

Deformable Model based segmentation methods are usually sensitive to the initialization condition to some extent. Hence fast and robust heart localization and initialization is very important for achieving good final segmentation results.

Based on the dynamic metric map introduced above, a robust heart localization method is developed as follows. Suppose we have computed the Dynamic Metric Map $DMM(X)$ of a cardiac image sequence $P(X,t)$.

We also acquire the result of background subtraction $Var(X)$ for each pixel X using Eq.6. Then, each pixel X has two kinds of feature values, $Var(X)$ and $DMM(X)$ respectively. They can be combined together to form a feature vector for heart localization: $Feature(X) = [Var(X), DMM(X)]$. Based on these features, the fast K-means clustering method [15] can be used to separate pixels in the heart region from those in other regions. The algorithm avoids unnecessary distance calculations by applying the triangle inequality and accelerated k-means method. Figure 2 (d) shows the clustering result in the example images. We can see that the heart region consists of mostly white pixels while the chest cavity region consists of mostly gray pixels. This results from the very different dynamic signatures in these two regions. After we obtained the clustered pixels in the heart region, we get the final result of heart localization as shown in Figure 2 (e).

3.3. DMM based segmentation model

During heart localization, we use T number of consecutive image frames to build the Dynamic Metric Map (DMM) for final pixel clustering. It is robust for localizing the heart region. However, when T is large, the dynamic signature $S(X)$ can not accurately represent the dynamics in every individual frame. With the Gauss-Markov model theory, fortunately, we can assume the dynamics of each pixel $P(X,t_0)$ is smooth in a piecewise $l \times l \times t$ volume centered at $P(X,t_0)$, where $t \ll T$. Thus, we can acquire $DMM(X,t_0)$ of each frame by computing dynamic metric values of all pixels in a $l \times l \times t$ volume, then applying DMM computation (see Sec. 3.1) on this volume.

To integrate the DMM in model-based segmentation, we compute the Gradient Vector Flow (GVF) of the DMM, and add it into traditional active contour models as a new external image force. Then, the energy functional to minimize for the new dynamic contour model becomes:

$$E = \int_0^1 (E_{int}(C(s)) + E_{img}(C(s)) + E_{DMM}(C(s)))ds \tag{8}$$

Figure 3 shows improvements in the gradient vector flow forces based on DMM.

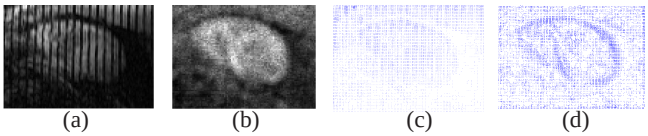

Fig. 3. (a) an image, (b) DMM, (c) Gradient based external forces, (d) DMM based external forces

The initialization for this new deformable model can be obtained from the heart localization result. And the added "dynamic" image forces increase the model's attraction range and make it insensitive to noise that are only present in image appearance but not in dynamic motion.

4. EXPERIMENTAL RESULTS

We conduct experiments on both tagged and un-tagged MRI 4D cardiac images to verify the effectiveness of the proposed segmentation algorithm. We present our validation results using five sequences of 4D spatio-temporal short-axis tagged MR images. Each sequence consists of 24 phases, with 16 slices (images) per phase. Since we do not use the first and last four phases in the sequences, we have 256 images for testing from each sequence. An expert is also asked to draw the epicardium (Epi) contours in these testing images.

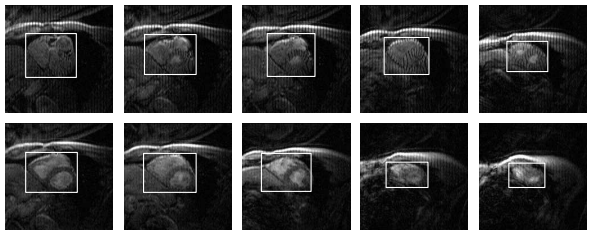

Fig. 4. DMM based localization results

In order to test the performance of the proposed DMM based localization method, we let it run on the validation im-

age sequences. We then manually check all localization results in the data set including 256×5 images. The correct-localization rate is 100%. Moreover, most localization results bound tightly around the heart regions. We show some example localization results in Figure 4. They demonstrate that the proposed method can robustly localize the heart even in cardiac images with low contrast, high noise and tagging lines. The DMM based heart localization process takes 0.288 seconds for each slice sequence including 24 images on a $1.5GHz$ laptop PC in MATLAB environment.

After localization, the final segmentation can be obtained using DMM-augmented deformable models introduced in section 3.3. The segmentation process takes less than 4 seconds for each image on a $1.5GHz$ laptop PC in MATLAB environment. Some example segmentation results are shown in Figure 5.

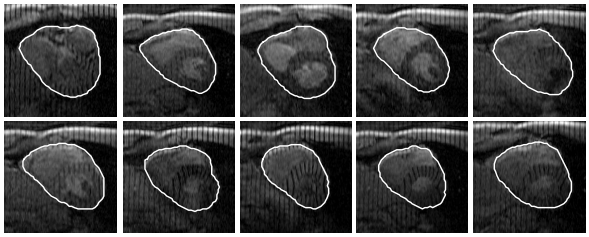

Fig. 5. DMM based segmentation results

Quantitative validation is performed by comparing the automated epicardiac segmentation results with expert solutions. Here, we will not deal with endocardiac segmentation. After good epicardiac segmentation, it can be easily done with the existing method [3] as there exist distinct appearance due to blood flow and less tagged line noises. Comparatively, epicardiac segmentation is very hard due to complex appearance, low contrast, tagged lines and noises.

Denote the expert segmentation in the images as l_{true}, and the results from our method as l_{DMM}. We define the false negative fraction (FNF) to indicate the fraction of tissue that is included in the true segmentation but missed by our method: $FNF = \frac{|l_{true} - l_{DMM}|}{l_{true}}$. The false positive fraction (FPF) indicates the amount of tissue falsely identified by our method as a fraction of the total amount of tissue in the true segmentation: $\frac{|l_{DMM} - l_{true}|}{l_{true}}$. And the positive fraction (TPF) describes the fraction of the total amount of tissue in the true segmentation that is overlapped with our method: $\frac{|l_{DMM} \cap l_{true}|}{l_{true}}$. On the validation datasets, the proposed method achieved encouraging segmentation results: FNF, FPF and TPF are 2.7%, 5.6% and 97.3% respectively.

5. DISCUSSIONS AND CONCLUSIONS

In this paper, we have presented a new cardiac image segmentation approach that fully exploits dynamic analysis to achieve temporally-consistent segmentation results. Our main contributions are three folds. First, we introduce and improve the dynamic texture based motion segmentation method to compute dynamic signatures of image regions with different motion patterns. Second, the clustering of dynamic signatures gives us the distinctively-moving heart region, which is the heart localization result acquired without manual interaction or any prior atlas information. Third, we introduce the Dynamic Metric Map, from which we can derive dynamic information to be integrated into deformable models to achieve temporally-consistent segmentation results.

6. REFERENCES

[1] J. Goutsias and S. Batman, "Morphological methods for biomedical image analysis," *SPIE Handbook of Medical Imaging, M. Sonka, J. Fitzpatrick (Eds)*, vol. II, pp. 255–263, 2000.

[2] M.A. Guttman, J.L. Prince, and E.R. McVeigh, "Tag and contour detection in tagged mr images of the left ventricle," *IEEE Trans. on Medical Imaging*, vol. 13, no. 1, pp. 74–88, 1994.

[3] A. Montillo, D. Metaxas, and L. Axel, "Automated deformable model-based segmentation of the left and right ventricles in tagged cardiac MRI," in *Proc. of MICCAI*, 2003, pp. 507–515.

[4] M. Lorenzo-Valdes, G. I. Sanchez-Ortiz, R. Mohiaddin, and D. Rueckert, "Segmentation of 4D cardiac mr images using a probabilistic atlas and the em algorithm," in *Proc. of MICCAI*, 2003, pp. 440–450.

[5] M. Uzumcu, A. F. Frangi, M. Sonka, J.H.C. Reiber, and B.P.F. Lelieveldt, "ICA vs. PCA active appearance models: Applications to cardiac mr segmentation," in *Proc. of MICCAI*, 2003, pp. 451–458.

[6] Papademetris X., Sinusas A., Dione D. P., and Duncan J. S., "Estimation of 3d left ventricular deformation from echocardiography," *Medical Image Analysis*, vol. 5, pp. 17–28, 2001.

[7] Radeva P., Amini A., and Huang J., "Deformable b-solids and implicit snakes for 3d localization and tracking of spamm mri data," *Computer Vision and Image Understanding*, vol. 66, pp. 163–178, 1997.

[8] J. Lotjonen, J. Koikkalainen, D. Smutek, S. Kivisto, and K. Lauerma, "Four-chamber 3-D statistical shape model from cardiac short-axis and long-axis mr images," in *Proc. of MICCAI*, 2003, pp. 459–466.

[9] Ravichandran A., Vidal R., and Halperin H., "Segmenting a beating heart using polysegment and spatial gpca," in *Proc. of ISBI*, 2006, pp. 634–637.

[10] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," in *Proc. of ECCV*, 1998, vol. 2, pp. 484–498.

[11] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," in *International Journal of Computer Vision*, 2002, vol. 46, pp. 223–247.

[12] X. Huang, D. Metaxas, and T. Chen, "Metamorphs: Deformable shape and appearance models," in *Proc. of CVPR*, 2004, vol. 1, pp. 496–503.

[13] G. Doretto, D. Cremers, P. Favaro, and S. Soatto, "Dynamic texture segmentation," in *Proc. of ICCV*, 2003, pp. 1236–1242.

[14] A. Chan and N. Vasconcelos, "Layered dynamic textures," in *Proc. of NIPS*, 2005.

[15] C. Elkan, "Using the triangle inequality to accelerate k-means," in *Proc. of ICML*, 2003, pp. 147–153.