LEFT ENDOCARDIUM SEGMENTATION USING SPATIO-TEMPORAL METAMORPHS

Xinyi Cui[1], *Shaoting Zhang*[1], *Junzhou Huang*[2], *Xiaolei Huang*[3], *Dimitris N. Metaxas*[1], *Leon Axel*[4]

[1]Department of Computer Science, Rutgers University, Piscataway, NJ, USA
[2]Department of Computer Science & Engineering, University of Texas at Arlington, TX, USA
[3]Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA, USA
[4]Radiology Department, New York University, New York, NY, USA

ABSTRACT

The Metamorphs model is a robust segmentation method which integrates both shape and appearance in a unified space. The standard Metamorphs model does not encode temporal information. Thus it is not effective in segmenting time series data, such as a cardiac cycle from MRI. Furthermore, it needs manual interaction to initialize the model, which is time consuming for temporal data. In this paper, we proposed a model to seamlessly couple both spatial and temporal information together in the Metamorphs method. It is also able to automatically initialize the model instead of manual initialization. We model energy terms as probability maps, then different energy terms can be easily fused by multiplying them together. Temporal Spectral Residual (TSR) is employed to rapidly generate a probability map in temporal data. Compared to traditional Metamorphs, the computational overhead of our model is very light due to the efficiency of the TSR method and the ease of coupling different energy functions by using probability maps. We validate this algorithm in a task of segmenting the left ventricle endocardium from 2D MR sequences, and our method shows performance superior to the traditional Metamorphs.

Index Terms— Segmentation, spatio-temporal, Metamorphs, cardiac MRI

1. INTRODUCTION

Automated object segmentation is a fundamental problem in medical image analysis. It is challenging to robustly segment objects because of the common presence of cluttered objects, object texture, image noise, and various other artifacts in medical images. In recent decades, deformable model-based segmentation methods have been extensively studied and achieved considerable success [7, 2, 9], because of their ability to integrate high-level knowledge with low-level image processing. One of the seminal works in this area, the "Snakes" [7], models the segmentation task as an energy-minimizing framework driven by both external forces derived from image appearance cues and internal forces from shape smoothness constraint. Many variations have been proposed to improve the robustness of the Snakes model. Most of them focus on handling noise and spurious image edges, since the traditional Snakes, solely relying on the image gradient information, can easily get stuck at local minima. Methods such as Region analysis strategies [13] have been incorporated in Snake-like models to improve their robustness to noise. In particular, instead of modeling traditional "active contours", Metamorphs [6] have been introduced to model "deforming disks or volumes". It is able to integrate both shape and appearance in a unified space, where the estimated object boundary by shape and appearance are encoded as probability map. Since the model has incorporated not only boundary shape but also interior appearance, it is more robust to ambiguous boundaries and complex internal textures. The deformation of this parametric model is based on Free Form Deformation, which is very efficient, as it reduces the deformation space from all pixels to a few parameters. The Metamorphs method is effective on diverse tasks, such as segmenting the heart and liver from MRI, and prostate from ultrasound images. Several ways have been proposed to extend and further improve the original Metamorphs method. Shen et al. proposed Active Volume Model (AVM) [8] to perform volume segmentation in 3D images. The AVM model's shape is represented by a simplex mesh and its volumetric interior carries the various visual appearance feature statistics. A shape prior constraint based on Active Shape Model (ASM) [2] has also been incorporated into 3D Metamorphs [5, 10]. It constrains the intermediate shape by following the shape pattern from existing data, which makes it able to recover or preserve local shape details. Despite the superior performance of the Metamorphs method and these extensions and improvements, it is still not clear how to effectively integrate temporal information into this model for spatio-temporal segmentation, such as segmenting the left ventricle endocardium in a cardiac cycle. A possible solution is to use Metamorphs in each time frame independently, or use the result from the previous frame as the initialization for the next. However, neither of these truly employ the temporal information. Furthermore, it is also desirable to automatically initialize the model instead of manually locating the organ as is done in standard Metamorphs, since manual initialization is very time-consuming, especially for temporal data.

In this paper, we propose a unified framework to seamlessly couple spatial and temporal information together in the Metamorphs model. It is also able to automatically initialize the model in temporal data. As discussed above, boundary estimation from information sources such as gradient and region are encoded as probability maps in a standard Metamorphs framework, since different probability maps can be easily integrated by multiplying them together. Thus, we design our temporal energy as probability maps as well. Temporal Spectral Residual (TSR) [3] is employed to rapidly generate such maps in temporal data. TSR is an automatic saliency detection method which only needs a Fourier spectrum analysis, so it can rapidly predict the foreground and background. This temporal energy is used as both the initial energy of the model and a constraint during runtime. The weight of this temporal energy is adaptively decreased during deformation, since the region energy from texture information becomes more and more reliable when it is close to the boundary. Compared to traditional Metamorphs, the computational overhead of our model is very light due to the efficiency of this TSR method and the ease of coupling different energy functions by using probability maps. We validate this algorithm in a task of segmenting the left ventricle endocardium from 2D MR sequences. Compared to the standard Metamorphs, our method is fully automatic, and is able to effectively segment regions-of-interest in the time series data.

2. METHODOLOGY

Online Robust Deformable Model. In order to find the boundary robustly, we employ an online deformation framework based on the Metamorphs. The standard Metamorphs model is driven by both gradient and region information derived from the image, as region information alleviates the problems caused by unclear boundaries and complex textures. However, Shen et al. [8] have shown that, in many cases, including the gradient-based boundary information in this framework does not improve the segmentation performance. Therefore, we just use the image intensity feature and its predicted object regions to derive image forces. A new energy derived from temporal information is also modeled, which is introduced in Sec. 2. These two energies are coupled using probability maps. The overall energy function is defined as:

$$E = \sum_{i=1}^{n} E_{int} + \left(\sum_{i=1}^{n} E_R + k_T E_T \right) \qquad (1)$$

where E_{int} is the internal (smoothness) energy [6]. The difference of the internal energy from [6] is that we take the sum over the whole sequence. E_R is the external region energy term, n is the number of time frames, E_T is the external temporal energy term, and k_T is the weight to balance the contributions of the two external energy terms E_R and E_T. The details of the region term are introduced in Sec. 2, and the temporal term is discussed in Sec. 2. The balance between the internal and external energies is naturally controlled by the smoothness constraint of the shape model [8]. Different from standard Metamorphs, our objective function is designed for the whole sequence of temporal data instead of each time frame separatively.

To initialize the model, we use a TSR-based saliency detection method to rapidly and automatically generate probability maps for all time frames, which denote the pixel probabilities of being foreground or background. This energy drives the model to deform. After this initialization, an intermediate result is generated and energy from the region term can be calculated and iteratively updated. Although temporal energy continues to serve as a constraint during the evolution of the model, its importance is adaptively decreased, since the region term becomes more and more reliable when the model is close to the boundary. We define the weight k_T as $k_T = 1 - e^{-|\Delta \mathcal{M}|}$, where M is the current shape model. Thus $|\Delta \mathcal{M}|$ is the magnitude of deformation change in model shape. In the beginning, the shape deforms quickly, so k_T is relatively large. It means that we trust more in the temporal energy, and put larger weights on it. After several iterations, the region term should be more important since the model is close to the boundary. At this time the shape deforms less, so k_T is smaller. The model converges when k_T approaches zero as the shape stops deforming. In the following we introduce the two external energy terms in our model.

Region Energy using Nonparametric Kernel Method. The region energy E_R is based on the intensity distribution of the model interior. A nonparametric kernel-based method is employed to model it. Such a nonparametric approximation is differentiable, generic, and can represent complex multi-modal intensity distributions. Suppose the model is placed on an image (denote its intensity as I), where the image region bounded by the current model $\Phi_{\mathcal{M}}$ is $\mathcal{R}_{\mathcal{M}}$, then the probability of a pixel's intensity value i being consistent with the model interior intensity can be derived using a Gaussian kernel as:

$$\mathbf{P}(i \big| \Phi_{\mathcal{M}}) = \frac{1}{V(\mathcal{R}_{\mathcal{M}})} \iint_{\mathcal{R}_{\mathcal{M}}} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(i-I(\mathbf{y}))^2}{2\sigma^2}} d\mathbf{y} \qquad (2)$$

where $V(\mathcal{R}_{\mathcal{M}})$ denotes the volume of $\mathcal{R}_{\mathcal{M}}$, and σ is a constant specifying the width of the gaussian kernel.

Using this nonparametric approximation, the intensity distribution of the model interior gets updated automatically while the model deforms. The initialization of the model texture is flexible, e.g., starting with a small model inside the texture region to be segmented.

Although the region energy is already able to generate good results for static images, it still has two problems for temporal data. First, the results may not be consistent between two neighboring frames since there is no temporal con-

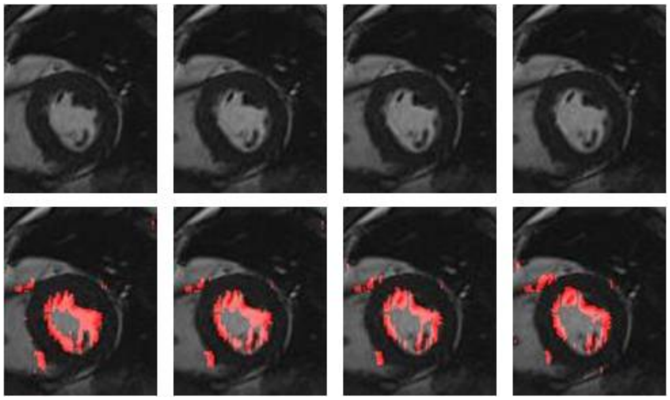

Fig. 1. Salient motion regions detected by TSR algorithm on a cardiac MR sequence. (marked in red color in the bottom row).

straint. Second, we may need to perform manual initialization for each time frame. This is not efficient, as there may be more than 20 frames in each temporal dataset. This motivates us to design the temporal energy E_T. In our approach, instead of manual initialization of a model inside the texture region, we automatically initialize the model according to the object boundary estimated from the temporal energy term E_T, which is discussed in the following section.

Temporal Energy using Temporal Spectral Residual. To solve the above mentioned problems, we propose a new energy, defined in temporal space. The goal is to seamlessly couple both region and temporal terms. To achieve this, it is preferable that the temporal term is also defined as a probability map, which denotes the probability of being foreground or background. Then it is consistent with the Metamorphs framework, and two terms can be seamlessly coupled in the same framework and be balanced with weight k_T as in Eq. 1. The computational overhead should also be small. We use TSR to efficiently generate this temporal energy term as a probability map.

TSR is an efficient method to find salient motion regions in video sequences. The main idea is to roughly remove the redundant part of a volume data (the static part of temporal slices) and keep the salient motion regions. This algorithm is able to provide reliable motion regions without needing initial labeling or any training data. It uses the Spectral Residual algorithm (SR) [4], which is a saliency detection algorithm on 2D images using statistics of Fourier Transformation. TSR uses SR to find the salient regions along the *temporal axis in a video sequence*.

Treating a sequence of 2D images as a 3D volume, TSR is able to roughly locate the cardiac region in temporal space, as shown in Figure 1. A cardiac motion cycle contains cardiac motion regions (the motion saliency) and static regions (regions which do not change much during the cardiac motion cycle). The motion salient regions are used to generate the E_T energy term in the Metamorphs, for both initialization and deformation constraint.

Denote the temporal axis of the sequence as T, then the temporal slices are represented by slabs of XT and YT, where X and Y are the axes of each image frame. To find the salient regions along the temporal axis, the TSR algorithm employs SR on XT and YT separately. Static regions do not change much over time in the cardiac MR sequence. When viewed in the temporal domain, pixel intensity stays almost the same over temporal slices. Thus given a temporal slice (i.e., XT or YT planes), the SR algorithm is able to remove the static part and keep the saliency region. Merging the results from both temporal planes by majority voting generates a reliable saliency map for the motion region, i.e., left ventricle endocardium in this case. Figure 1 shows a representative result. The probability maps detected by TSR are mostly concentrated around the boundary of the left ventricle endocardium, which provides a rough initialization and temporal energy term E_T for the constraint of the deformable model. In addition, as this algorithm only needs the Fourier transform, the computational cost is very low. It does not significantly increase the computation time of the Metamorphs.

3. EXPERIMENTS

In this section, we validate this proposed Spatio-temporal Metamorphs by segmenting the left ventricle endocardium from MRI. Segmentation experiments were performed on a set of 10 MRI cine sequences. Each sequence has 25 heart phases (frames) and a total duration of 1 sec (approximately one heart-beat). General cardiac segmentation has been previously extensively investigated [12, 14, 1]. Our focus here is to demonstrate the improved performance relative to standard Metamorphs on this temporal data.

This segmentation task is challenging because of two facts. The papillary muscle and trabeculae inside the left ventricle endocardium has large frequent movement. Such movement causes high gradient, which can adversely affect the accuracy of segmentation algorithms. Furthermore, the morphology of the papillary muscle changes along time series. It is hard to obtain a consistent result by performing a segmentation algorithm on a single frame. The bottom row in Fig. 2 shows a segmentation result of the Metamorphs that fails to maintain the consistency over time. The inferior papillary muscle is captured in the first frame, but then lost in the next. The Spatio-temporal Metamorphs uses the temporal constraint, so it considers the papillary muscle movement over the whole MR sequences, and is able to keep the results more consistent. Our result is shown in the top row in Fig. 2, where the papillary muscle area is successfully captured along all frames, due to the constraint from temporal energy.

Table 1 quantitatively compares the accuracy of these two algorithms compared with expert segmentation. The proposed method achieves higher mean values of sensitivity and

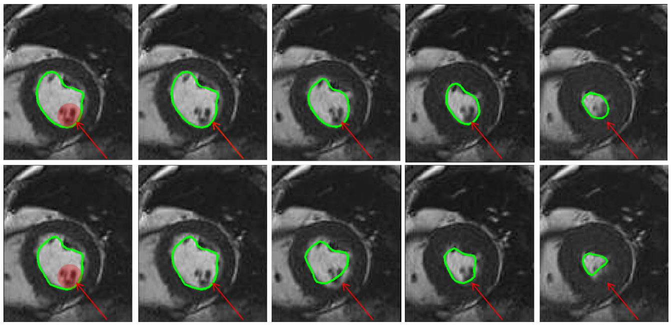

Fig. 2. Comparison between the proposed method (top row) and the standard Metamorphs (bottom row). The light red area in the first column is the inferior papillary muscle in the left ventricle endocardium. The high frequent movement makes it hard to capture the papillary muscle consistently.

specificity with lower standard deviations. It demonstrates that our method is more accurate and stable. We also conducted running time analysis in Table 1. We implemented this method using Matlab on a Quad CPU 2.4GHZ PC, with a volume data with 25 frames of size 256 by 256. Standard Metamorphs takes 30.1 seconds. Standard Metamorphs with TSR takes 30.65 seconds. TSR takes only additional 1.5% of running time, which is a small portion of the overall computation. Furthermore, the use of the motion information from the whole series with TSR can further boost the calculation speed by enabling parallel computing on a multi-core platform. The standard Metamorphs method is calculated sequentially, as the current frame depends on the result from the previous frame as the initialization. The Spatio-temporal Metamorphs does not have this constraint. Each frame can be updated in parallel, thus the computational time here is further reduced to 4.8s, which is only 15.95% of the standard Metamorphs. Thus the proposed method can greatly speed up the calculation of the standard Metamorphs.

4. CONCLUSIONS

We proposed a robust segmentation method which employs both region and temporal information as image forces. The model is automatically initialized based on predicted saliency regions from the Temporal Spectral Residual algorithm. Such temporal energy is further used as a constraint during deformation with adaptively changing weights. Its computational overhead is very light compared to standard Metamorphs because of the efficiency of this TSR method and the ease of coupling different energy functions by using probability maps. This algorithm was validated in a task of segmenting the left ventricle endocardium from 2D MR sequences, and shows improved performance. In the future, we would like to extend this framework to 3D. Furthermore, we will incorporate shape priors [11] to improve the robustness of our method.

	Sensitivity	Specificity	Time
Metamorphs	0.79 ± 0.14	0.94 ± 0.05	30.1s
Our method	0.93 ± 0.05	0.96 ± 0.03	4.8s

Table 1. Quantitative comparisons of sensitivity, specificity and computational time. Mean values and standard deviations are reported.

5. REFERENCES

[1] R. Chandrashekara, R. Mohiaddin, and D. Rueckert. Cardiac motion tracking in tagged mr images using a 4D b-spline motion model and nonrigid image registration. In *ISBI*, pages 468–471, 2004.

[2] T. Cootes, C. Taylor, D. Cooper, and J. Graham. Active shape model - their training and application. *CVIU*, 61:38–59, 1995.

[3] X. Cui, Q. Liu, and D. Metaxas. Temporal spectral residual: fast motion saliency detection. In *MM*, pages 617–620, 2009.

[4] X. Hou and L. Zhang. Saliency detection: A spectral residual approach. In *CVPR*, 2007.

[5] J. Huang, X. Huang, D. Metaxas, and L. Axel. Adaptive metamorphs model for 3d medical image segmentation. In *MICCAI*, pages 302–310, 2007.

[6] X. Huang and D. Metaxas. Metamorphs: Deformable shape and appearance models. *TPAMI*, 2008.

[7] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *IJCV*, 1:321–331, 1987.

[8] T. Shen, H. Li, and X. Huang. Active volume models for medical image segmentation. *TMI*, 2011.

[9] L. Staib and J. Duncan. Boundary finding with parametrically deformable models. *TPAMI*, 14(11):1061–1075, 1992.

[10] S. Zhang and et.al. 3D segmentation of rodent brain structures using hierarchical shape priors and deformable models. In *MICCAI*, volume 6893, pages 611–618, 2011.

[11] S. Zhang, Y. Zhan, M. Dewan, J. Huang, D. N. Metaxas, and X. S. Zhou. Towards robust and effective shape modeling: Sparse shape composition. *MedIA*, 2012.

[12] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu. Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features. *TMI*, 27(11):1668–1681, 2008.

[13] S. Zhu and A. Yuille. Region Competition: Unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation. *TPAMI*, 18(9):884–900, 1996.

[14] X. Zhuang, K. Rhode, R. Razavi, D. Hawkes, and S. Ourselin. A registration-based propagation framework for automatic whole heart segmentation of cardiac MRI. *TMI*, 29(9):1612–1625, 2010.