Improved Sparse Shape Composition Model for Multi-shape Prior

Bing Wang[1], Chonghao Fan[2], Hongzhi Xie[3], Lixu Gu[2,4]

[1] College of Mathematics and Computer Science, Hebei University, China
[2] Multi-disciplinary Research Center, Hebei University, China
[3] Department of Cardiovascular, Peking Union Medical College Hospital, China
[4] School of Biomedical Engineering, Shanghai Jiao Tong University, China

Abstract. Shape prior modeling is a challenging and crucial component in various image segmentation applications. Most existing methods aim at dealing with single object's shape variation, which are not directly applicable for multi-shape prior modeling. In this paper, we present an extension of recently proposed Spare Shape Composition model (SSC) for multi-shape prior modeling. In this extension, multiple shapes of one patient are regarded as a group. A sparse linear composition of training groups is computed iteratively to infer/refine the input group. Thus, not only the a-priori information of each shape but also the a-priori codependency information among different shapes is implicitly incorporated on-the-fly. To validate the efficacy of our method, a 2D left ventricular endocardium and epicardium localization experiment was conducted. The localization result demonstrates that the utilization of our method can achieve more accurate and stable localization compared with SSC.

1 Introduction

In various image segmentation applications, especially in the field of medical image segmentation, methods solely relying on image appearance cues usually tend to achieve unsatisfactory result. The fact that objects usually contain strong shape priors, gives rise to various shape model based segmentation methods. Leventon et al. [1] defined a probability distribution over the variances of training shapes, and utilized it to restrict the flow of the geodesic active contour. In [2], Cremers et al. incorporated statistical shape knowledge in the evolution process of a Mumford-Shah based segmentation [3]. Ali et al. [4] integrated prior shape constraints into a graph cuts framework for kidney segmentation. Such segmentation approaches have been proven to be one of the most successful methods in practice, and outperform the conventional methods in both robustness and accuracy owing to the integration of a-priori information.

Shape prior modeling plays a significant role in these methods, and is very crucial for the final accuracy and robustness. A straightforward approach is to learn from a number of training shapes by statistical means, leading to statistical shape models (SSMs) [5]. Active Shape models [6] and Active Appearance models [7] proposed by

Cootes et al. in 1995 and 2001, probably are two of the most popular methods in this area. Another widely-used method is level set shape prior model [8]. Subsequently, plenty of adaptations of these algorithms were proposed.

Recently, sparsity theory was introduced into shape prior modeling by [9, 10]. In their model, a sparse composition of training shapes is computed adaptively to infer/refine an input shape. Thus, it was named as Sparse Shape Composition model. With such a setting, it is capable of modeling complex shape variations, and preserving local details very well. Furthermore, when modeling a sparse error vector, it becomes quite robust to sparse non-Gaussian errors.

However, most of existing models are focusing on single shape prior modeling, and not directly applicable to deal with multiple shapes due to the lack of prior co-dependency information utilization. Such co-dependency among different shapes is of great value for various medical image analysis tasks. For instance, the implicit spatial relationship between endocardium and epicardium of left ventricle can be utilized as a supplementary to their shape priors for accurate localization or segmentation purpose. Endocardium and epicardium of left ventricle are very important for quantitative analysis of global and regional cardiac function, such as ejection fraction (EF), left ventricle myocardium mass (MM), and stroke volume (SV) [11]. In Fig.1 a 2D instance from cardiac cine-MR short axis images is shown.

Inspired by [9], we present an extension of SSC which aims at dealing with multi-shape prior modeling. In our method, multiple interested shapes are regarded as a group, and modeled together by a sparse linear combination of training groups. With such a mechanism, the a-priori spatial constraint among different shapes is also implicitly applied. It has the same advantages of SSC, due to the utilization of the same basic idea and optimization framework proposed in SSC.

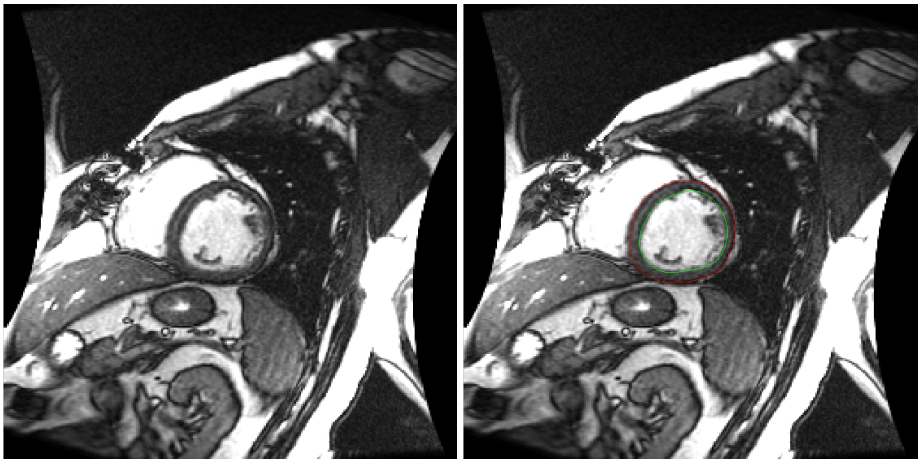

Fig. 1. Left: 2D cardiac cine-MR short axis image. Right: Manually delineated endocardial contour (drawn as green) and epicardial contour (drawn as red) of left ventricle.

2 Improved Sparse Shape Composition model

Our model aims at modeling any number of complicated shapes simultaneously with a pre-defined training repository. Shapes may refer to shapes of different objects (e.g., shapes of femur bone, femur cartilage, tibia bone and tibia cartilage of one patient) or shapes of a time-varying object (e.g., shapes of heart), and regarded as a group in our method. The details of how we advanced SSC to our method will be presented in this section.

2.1 From shape representation to group representation

Following SSC, explicit parametric shape representation[1] is employed in this method. Specifically, shape instance is represented by a column vector concatenated by coordinates of all its vertices. For instance, column vector $shape$ of a 3D mesh which contains 100 vertices is concatenated as Eq. (1).

Let m represents the number of shapes required to be modeled. Vertex number and column vector of the ith shape are notated by k_i and $shape^i \in \mathbb{R}^{k_i \times d}$ respectively, where d stands for the dimension of shapes. Then, a column vector g which represents the group consisted of these m shapes can be constructed by concatenating $shape^i$ for $i = 1,2,\dots,m$, as shown in Eq. (2).

$$shape \triangleq [x_1\ y_1\ z_1\ x_2\ y_2\ z_2 \dots x_{100}\ y_{100}\ z_{100}]^T \tag{1}$$

$$g \triangleq \left[shape^{1^T}\ \ shape^{2^T}\ \dots\ \ shape^{i^T}\right]^T \in R^{\sum_i k_i \times d} \tag{2}$$

2.2 Matrix of training repository

Assume there are n group samples with manual delineation in the training repository, which are quite sufficient to model variations of shapes and the spatial relationship among them. A matrix which represents the training repository can be constructed based on these samples. An illustration can be found in Fig. 2. It should be noticed that there should be a consistency among columns through this matrix. Specifically, vertex numbers of m shapes should keep consistent, and shape vertices should be one-to-one corresponding through these groups. Two methods are introduced to acquire consistency of shapes in [9]. Both of them can be extended to groups quite intuitionally. We assume this consistency is already achieved here.

After the conversion of all these groups into column vectors, apply pre-alignment to eliminate the position and orientation difference and transform them into a standard coordinate system. Pre-alignment is a two-step procedure based on the generalized Procrustes analysis [12]: first, select a group vector as reference and align others to it, take the transformed vectors and the reference vector as initial aligned groups; second, in order to remove the bias caused by the selection of reference, compute the mean vector of these initial aligned groups, take it as the new reference and align

[1] 2D and 3D shapes are represented by curves or meshes composed of a number of vertices.

others to it to get the final aligned groups. These final aligned groups vectors are notated by $\tilde{g}_j$ for $j = 1,2,\ldots,n$. Shape vectors of $\tilde{g}_j$ are notated as $\tilde{s}_j^i$ for $i = 1,2,\ldots,m$.

Finally, aligned group vectors are assembled together parallelly to concrete the matrix of training repository which is denoted as $D = [\tilde{g}_1\ \tilde{g}_2\ \ldots\ \tilde{g}_n]$.

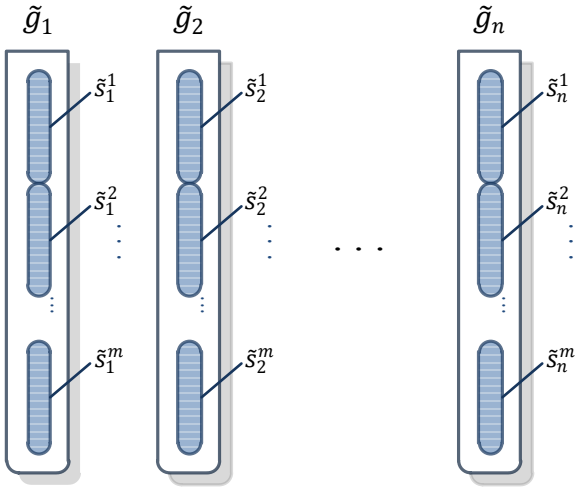

Fig. 2. Diagram illustrating the matrix of training repository.

2.3 Problem formulation and optimization framework

The basic idea of SSC can be intuitionally extended to groups: for any input group $y_g = [{y_1}^T\ {y_2}^T \ldots {y_m}^T]^T$, an optimal sparse linear combination of existing training groups can be found to approximately represent it. The weights or coefficients are denoted as $x \in R^n$, and the optimal value of it is denoted as $x_{opt} \in R^n$. x_{opt} is found by solving Eq. 3 utilizing the optimization method proposed in [9].

$$arg\ \min_{x,e,\beta} \left\| T(y_g, \beta) - Dx - e \right\|_2^2,\ s.t.\ \|x\|_0 \le k_1,\ \|e\|_0 \le k_2 \tag{3}$$

2.4 Difference from SSC

When modeling m shapes from a group, SSC required to be conducted for m times. Each time an optimal sparse weight vector is computed for a single input shape. Compared with the unique optimal weight vector x_{opt} calculated based on our method, these m vectors tend to differ from each other in practice. Thus, group constraint or co-dependency among different shapes is not utilized in the modeling process.

3 Endocardium and epicardium localization

Following the organ localization framework proposed in [9], we conducted an experiment to verify our method, i.e., left ventricular endocardium and epicardium localization from 2D cardiac cine-MR images [12], and compared it with the original SSC. In this experiment, endocardium and epicardium are regarded as a group and localized simultaneously; training repository is consisted of 15 groups from different patients; 91 images from 28 patients are tested.

3.1 Details of localization experiment

This experiment is focusing on the 2D images with papillary muscles, for the convenience of achieving consistency of group vectors. These images are capable of defining 8 landmarks through centers of the two largest papillary muscles. Two intersection points of the line which passes through these centers and the manual endocardial contour, are considered as endocardium landmarks; then, the midperpendicular of these landmarks can be found, which intersect with endocardium at the last two endocardium landmarks. The epicardium landmarks are defined in the same method. After the determination of landmarks, a certain number of vertices are interpolated between two neighboring landmarks along both manual contours to form the group vectors in the training matrix. In our experiment, the vertex number of endocardial contour in a group vector is 50, and 70 for epicardial contour. Two training samples from the repository are shown in Fig. 3. After the determination of landmarks and group vertices, the matrices of training landmarks and groups can be formed based on the method described in section 2.1 and 2.2, and notated as D_L and D_G.

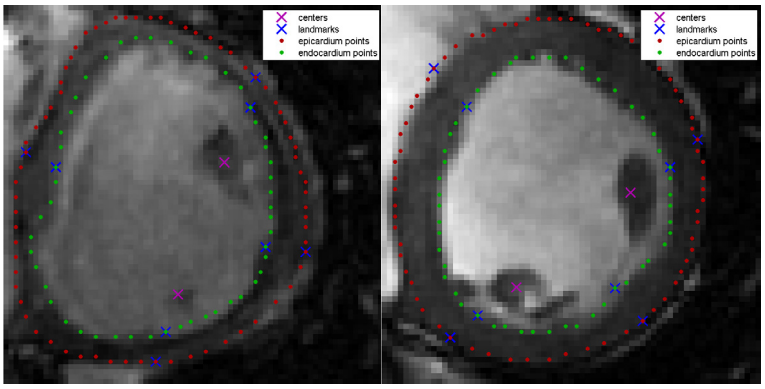

Fig. 3. Two images with papillary muscle centers, landmarks, endocardium and epicardium contour vertices.

Given a testing image, the localization procedure is as follows: first, manually labeling its landmarks, notated as y_L; then, x_{opt} and β_{opt} are computed by optimizing

Eq. (3) with D_L and y_L; at last, transforming $D_G x_{opt}$ back to the coordinate system of testing image as the group location.

In order to compare our method with the original SSC, SSC was also employed in the same localization framework with same training data and input landmarks to separately localize the endocardium and epicardium of left ventricle.

3.2 Evaluation and comparison

Evaluation

In addition to visual evaluation of localization accuracy, three quantitative measures are employed: average perpendicular distance (APD) [12], standard deviation of perpendicular distances (SPD), and dice metric (DM) [13]. As the name implies, APD and SPD measures the perpendicular distances from points on the localization result to manual contour, and calculate the average and standard deviation of them. Higher APD or SPD implies that localization result doesn't match closely to the manual contour. DM measures the overlap rate of the areas surrounded by localization and manual contour. It ranges between 0 and 1, with higher DM indicating better match.

Comparison

Accuracy and robustness of endocardium and epicardium localization based on our method and the original SSC are compared in this section. Three localization cases from different patients are shown in Fig. 4. The left column shows a case that both SSC and our method achieve acceptable result. The middle column shows a case that SSC fails in the localization, whereas our method still performs well. The right column shows an extreme case that neither method locates the endocardium and epicardium accurately due to the insufficiency of training samples, which is quite rare in our experiment. Despite the failure in the third case, our method outperforms the original SSC in all three cases from both visual and quantitative point of view. Furthermore, in order to compare them from the big picture, global APD, SPD and DM of our method and SSC are calculated statistically on 91 testing images, as shown in Table 1.

Table 1. Global APD, SPD and DM

Measures	SSC		Improved SSC	
	Endocardium	Epicardium	Endocardium	Epicardium
Global APD	3.3554	2.7553	2.3493	2.2032
Global SPD	2.4982	1.9412	1.6836	1.5741
Global DM	0.8766	0.9224	0.9124	0.9363

From the evidence provided in Fig. 4 and Table 1, we can reasonably arrive at a conclusion that our method can achieve more accurate and stable localization in this study compared with SSC.

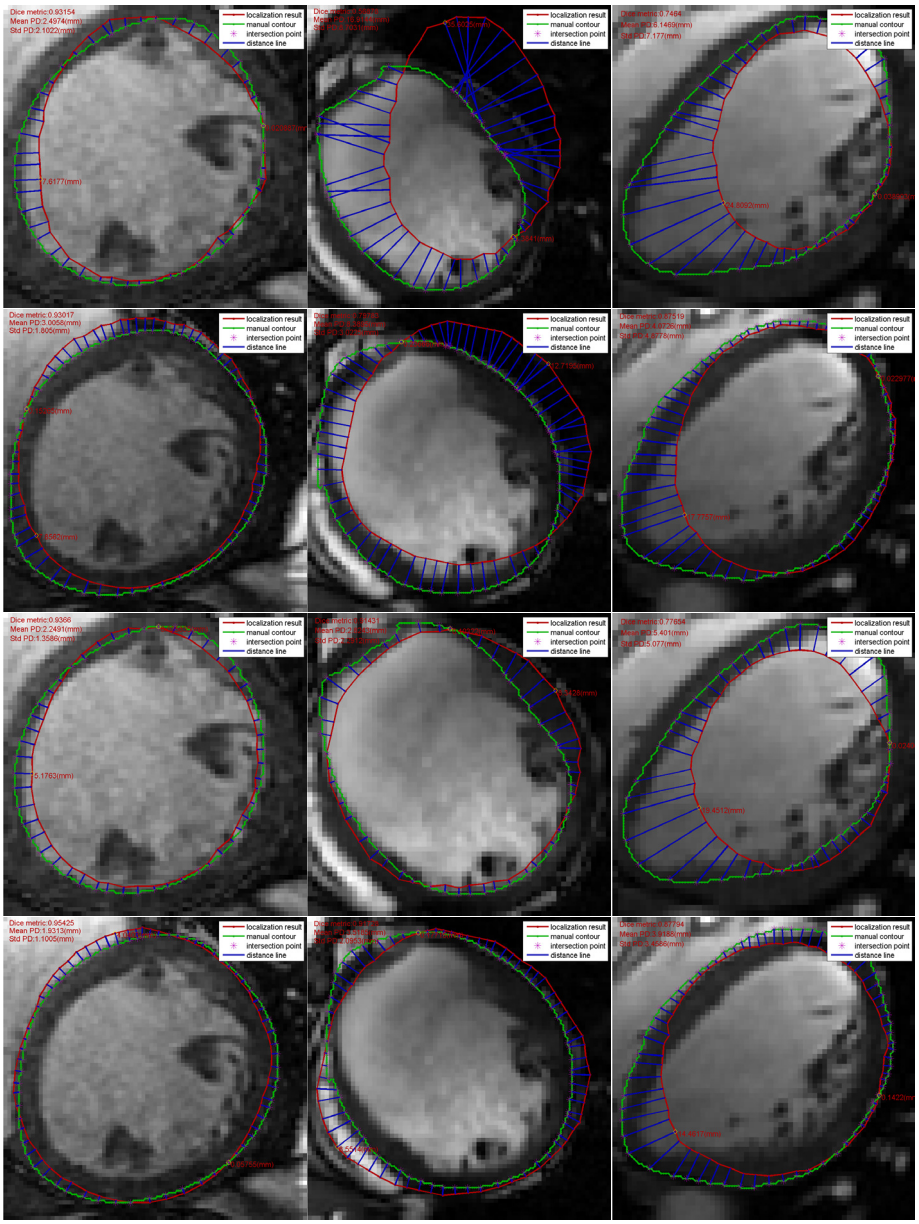

Fig. 4. Three localization cases. First and second row: endocardium and epicardium localization utilizing SSC. Third and fourth row: endocardium and epicardium localization utilizing our method.

4 Conclusion

In this paper, we proposed an extension of Sparse Shape Composition model for multi-shape prior modeling. In our method, multiple interested shapes from one patient are regarded as a group and modeled together to incorporate the co-dependency among different shapes. It is validated on a 2D endocardium and epicardium localization task, and exhibits more accurate and stable performance compared with original SSC. The success of our method is mainly relying on the incorporation of co-dependency among shapes.

In the future, we intend to apply this extension to various multi-shape segmentation tasks in clinical practices, especially to 3D shape sequences of time-varying organs.

References

1. Leventon, M. E., Grimson, W. E. L., & Faugeras, O. (2000). Statistical shape influence in geodesic active contours. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on* (Vol. 1, pp. 316-323). IEEE.
2. Cremers, D., Tischhäuser, F., Weickert, J., & Schnörr, C. (2002). Diffusion snakes: Introducing statistical shape knowledge into the Mumford-Shah functional. *International journal of computer vision*, *50*(3), 295-313.
3. Mumford, D., & Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on pure and applied mathematics*, *42*(5), 577-685.
4. Ali, A. M., Farag, A. A., & El-Baz, A. S. (2007). Graph cuts framework for kidney segmentation with prior shape constraints. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007* (pp. 384-392). Springer Berlin Heidelberg.
5. Heimann, T., & Meinzer, H. P. (2009). Statistical shape models for 3D medical image segmentation: A review. *Medical image analysis*, *13*(4), 543-563.
6. Cootes, T. F., Taylor, C. J., Cooper, D. H., & Graham, J. (1995). Active shape models-their training and application. *Computer vision and image understanding*, *61*(1), 38-59.
7. Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. *IEEE Transactions on pattern analysis and machine intelligence*, *23*(6), 681-685.
8. Rousson, M., & Paragios, N. (2002). Shape priors for level set representations. In *Computer Vision—ECCV 2002* (pp. 78-92). Springer Berlin Heidelberg.
9. Zhang, S., Zhan, Y., Dewan, M., Huang, J., Metaxas, D. N., & Zhou, X. S. (2012). Towards robust and effective shape modeling: Sparse shape composition. *Medical image analysis*, *16*(1), 265-277.
10. Zhang, S., Zhan, Y., & Metaxas, D. N. (2012). Deformable segmentation via sparse representation and dictionary learning. *Medical Image Analysis*, *16*(7), 1385-1396.
11. Radau P, Lu Y, Connelly K, Paul G, Dick AJ, Wright GA. "Evaluation Framework for Algorithms Segmenting Short Axis Cardiac MRI." The MIDAS Journal -Cardiac MR Left Ventricle Segmentation Challenge, http://hdl.handle.net/10380/3070
12. Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, 285-339.
13. Lynch, M., Ghita, O., & Whelan, P. F. (2008). Segmentation of the left ventricle of the heart in 3-D+ t MRI data using an optimized nonrigid temporal model. *Medical Imaging, IEEE Transactions on*, *27*(2), 195-203.