

# 4D Shape Priors for a Level Set Segmentation of the Left Myocardium in SPECT Sequences

Timo Kohlberger<sup>1</sup>, Daniel Cremers<sup>2</sup>, Mikael Rousson<sup>1</sup>, Ramamani Ramaraj<sup>1</sup>,  
and Gareth Funka-Lea<sup>1</sup>

<sup>1</sup>Siemens Corporate Research, Inc., Imaging and Visualization Department,  
Princeton, NJ, USA

`Timo.Kohlberger@siemens.com`

<sup>2</sup>Department of Computer Science, University of Bonn, Germany  
`dcremers@cs.uni-bonn.de`

**Abstract.** We develop a 4D (3D plus time) statistical shape model for implicit level set based shape representations. To this end, we represent hand segmented training sequences of the left ventricle by respective 4-dimensional embedding functions and approximate these by a principal component analysis. In contrast to recent 4D models on explicit shape representations, the implicit shape model developed in this work does not require the computation of point correspondences which is known to be quite challenging, especially in higher dimensions. Experimental results on the segmentation of SPECT sequences of the left myocardium confirm that the 4D shape model outperforms respective 3D models, because it takes into account a statistical model of the temporal shape evolution.

## 1 Introduction

Model-based imaging analysis of the left ventricle (LV) has gained an important role in diagnosis and treatment of heart diseases [8]. Segmentation in particular, either of the inner volume or the myocardium has found to be a significant prerequisite of further quantitative analysis, such as the estimation of the ejection fraction within one cardiac cycle. With standard 3D approaches, segmentation is carried out for each recorded image volume separately, i.e. by neglecting all temporal correlations. Recent approaches however aim at taking also the temporal dimension into account, arising from the fact that more and more 4D, i.e. 3D plus time, imaging techniques have become available. Relative to cardiac imaging, which provides a relatively high three-dimensional spatial resolution, isotope imaging methods, like gated perfusion SPECT, still have low resolution in both space and time and exhibit a relatively low signal to noise ratio. On the other hand, SPECT imaging inherently provides an excellent mean for cardiovascular diagnosis, because it allows direct visualization of physiology.

In terms of shape modeling, either explicit representations by landmarks or binary images/volumes are employed, or implicit ones, describing the separating three-dimensional contour as the zero-level set of a four dimensional function [16, 11]. Additionally, a statistical model is incorporated, in order to describe

inter-patient shape variabilities. In most cases, Principal Component Analysis (PCA) is the method of choice. In the case of explicit shape representations [4, 9], PCA is either applied directly to the landmark coordinates or to the components of a deformation field relative to a mean shape, or, for the implicit representation, to the components of the embedding level set functions [11, 17, 20].

The extension of shape models to the spatio-temporal case has been mainly pursued with explicit (parametric) shape models [18, 3, 12, 14]. While these methods have shown quite some success in modeling and estimating the temporal evolution of surfaces, one should point out that explicit shape representations are known to have certain inherent limitations. Firstly, when matching two parametric surfaces (in the shape learning phase), there is the fundamental correspondence issue of which points on one surface are associated with which points on the other surface. Explicit representations are typically based on a specific point correspondence, estimating this correspondence and thus solving this combinatorial problem is an enormous computational challenge [6], especially in the 4D case considered in this paper. Misaligned point correspondences will appear as artifacts in the subsequently estimated shape deformation modes. Secondly, in the shape estimation step, the evolution of parametric surfaces has certain limitations. The spatial resolution of the surface representation is determined by the local density of control points which may fluctuate due to the motion of points. This evolution may become unstable once control or marker points start overlapping. Splitting or remerging of contours in the segmentation of multiply-connected structures is by default not possible. Sophisticated regriding mechanisms are required to handle these problems [7]. At the same time, the required amendments to the control point evolution may lead to artifacts in the shape deformation that are not inherent to the geometric shape warping problem.

While the level set method overcomes the above drawbacks of explicit representations, there have only been very few recently published works on extending the shape modeling to the spatio-temporal domain [5, 19]. While the latter approaches treat the temporal dimension separately and employ dynamical models, we propose to treat time as an ordinary fourth dimension and thus implementing a segmentation of the volume sequence as a whole. Unlike Chandrashekhara et al. [3], who developed this idea for deformation fields relative to an explicit (parametric) mean sequence, we propose to apply a 4D PCA subspace estimation directly to the level set functions. As a consequence, we do not require a sophisticated non-rigid registration of training sequences, or the challenging combinatorial correspondence searches required by explicit landmark-based approaches. Moreover, due to considering all volumes of a sequence at the same time, and thereby also learning the temporal evolution of the shape model, segmentation quality can be shown to improve over treating each volume separately [11]. In addition, temporal alignment can be incorporated into the segmentation process straightforwardly, avoiding the need for any pre-alignment of the input sequence.

## 2 Approach

### 2.1 Level Set Representation

Let the recorded image volume sequence be denoted by the continuous intensity function  $I : \Omega \times [1, T] \rightarrow \mathbb{R}$  defined on the 3D volume  $\Omega \in \mathbb{R}^3$  and the time period  $[1, T]$ . In contrast to other approaches treating intensity volumes at different times points separately, in the following we will always consider  $I$  as a function in the four-dimensional spatio-temporal space  $\Omega \times [1, T]$ .

In order to be independent of contour parametrization, support topological changes as well as avoiding correspondence problems when introducing a prior shape model later on, we pursue the approach in [16] in application to the four-dimensional space and embed any contour  $C : [0, 1]^2 \times [1, T] \Rightarrow \mathbb{R}^2$  as the zero level set of a function  $\phi : \Omega \times [1, T] \Rightarrow \mathbb{R}$ , such that

$$C(\mathbf{x}, t) = \{(\mathbf{x}, t) \mid \phi(\mathbf{x}, t) = 0\},$$

which has been successfully applied to 2D [13] and 3D [10] image segmentation. In aiming at having a unique correspondence between a shape  $C$  and its embedding function, we furthermore assume  $\phi$  to be a signed distance function, i.e.  $|\nabla\phi| = 1$  almost everywhere.

However, since we want to estimate spatial translation and scale as well as temporal translation and scale separately, we introduce the intermediate transformation

$$T_\theta(\mathbf{x}, t) := \begin{pmatrix} \theta_{m_s} \mathbf{x} - \boldsymbol{\theta}_{d_s} \\ \theta_{m_t} t - \theta_{d_t} \end{pmatrix},$$

with  $\theta_{d_s} \in \mathbb{R}^3$  and  $\theta_{d_t} \in \mathbb{R}$  referring to translation parameters in space and time, respectively, and  $\theta_{m_s}^{-1} \in \mathbb{R}$  and  $\theta_{m_t}^{-1} \in \mathbb{R}$  to those in scale<sup>1</sup>, respectively. By this, the embedding function will, in the following, be of the form:

$$\phi_\theta(\mathbf{x}, t) := \theta_{m_s}^{-1} \phi(T_\theta(\mathbf{x}, t)).$$

### 2.2 Segmentation as Variational Problem

Subsequently, the problem of segmenting the image sequence  $I$  implicitly by the 4D level set function  $\phi$  is stated as the joined minimization of the energies

$$E(\phi, \theta) = E_{data}(\phi, \theta) + \nu E_{shape}(\phi), \quad (1)$$

where the data-driven energy term  $E_{data}(\phi, \theta) = -\log \mathcal{P}(I \mid \phi, \theta)$  measures the probability of observing the image sequence  $I$  given the segmentation  $\phi$  (at place  $\theta_d$  and scale  $\theta_m$ ), and  $E_{shape}(\phi) = -\log \mathcal{P}(\phi)$ , which reflects the probability of the occurrence of a segmentation  $\phi$  in relation to a set of 4D training segmentations  $\{\hat{\phi}_1, \dots, \hat{\phi}_M\}$ .

---

<sup>1</sup> Here, we consider only isotropic scale in space.

In particular, for the data energy term we make use of the global approach by Mumford and Shah [15] given as a level set formulation [2] in the spatio-temporal space  $\Omega \times [0, T]$ :

$$E_{data}(\phi, \theta) := \int_{[1, T]} \int_{\Omega} \left( I(T_{\theta}(\mathbf{x}, t)) - \mu_1(t) \right)^2 H\phi_{\theta}(\mathbf{x}, t) + \left( I(T_{\theta}(\mathbf{x}, t)) - \mu_2(t) \right)^2 \left( 1 - H\phi_{\theta}(\mathbf{x}, t) \right) d\mathbf{x} dt \quad (2)$$

with  $H\phi := H(\phi)$  denoting the Heaviside step function, which is one at coordinates  $(\mathbf{x}, t)$  where  $\phi$  is positive, and zero else, as well as  $\mu_1 = \frac{1}{\alpha_1} \int_{[1, T]} \int_{\Omega} I H\phi_{\theta} d\mathbf{x} dt$ ,  $\mu_2 = \frac{1}{\alpha_2} \int_{[1, T]} \int_{\Omega} I(1 - H\phi_{\theta}) d\mathbf{x} dt$ , and  $a_k = \int_{[1, T]} \int_{\Omega} H\phi_{\theta} d\mathbf{x} dt$ ,  $k = 1, 2$ .

With respect to the shape prior, we propose to conduct a PCA directly on the set of vectorized training *sequences* thereby reaching a low-dimensional parametrization of the subspace [11, 20]:

$$\phi_{\alpha}(\mathbf{x}, t) := \phi_0(\mathbf{x}, t) + \alpha^{\top} \psi(\mathbf{x}, t) \quad (3)$$

where  $\psi = (\psi_1, \dots, \psi_K)^{\top}$  denotes the eigenmode vector of the 4D levelset functions and  $\phi_0 = \frac{1}{M} \sum_{i=1}^M \hat{\phi}_i$  the mean function<sup>2</sup>. Note that in contrast to [20], the eigenmodes here describe *spatio-temporal* variations across the whole volume sequence. Within this subspace coordinate system, we assume the prior probability  $\mathcal{P}(\phi, \theta)$  as multivariate Gaussian distribution, and infer for an arbitrary function  $\phi_{\alpha}$  the energy

$$E_{shape}(\phi) := \alpha^{\top} \Sigma^{-1} \alpha$$

with  $\Sigma^{-1}$  denoting the inverse eigenvalue matrix resulting from the PCA.

Moreover, following the approach suggested in [20], we represent  $\phi$  by means of the parametrization given in (3) throughout the *whole* segmentation problem, i.e. also in the data term, and consequently substitute  $\phi$  by  $\phi_{\alpha}$  in (2). Thereby, we restrict the minimization of (1) to the much lower-dimensional subspace induced by the training shapes.

Due to the prior shape model being invariant towards translation and scale both in space and time, it is required for the training set to be free of such transformations prior to applying the PCA. Here, this is implemented by a prior alignment of each training shape  $\hat{\phi}_i$  to an arbitrarily chosen reference by minimizing the energy

$$\int_{[1, T]} \int_{\Omega} \left( \theta_{m_s}^{-1} \phi_{ref}(T_{\theta}(\mathbf{x}, t)) - \hat{\phi}_i(\mathbf{x}, t) \right)^2 d\mathbf{x} dt, \quad i = 1, \dots, M.$$

### 2.3 Gradient Descent Minimization

Determining a minimum of (1) with respect to  $\alpha$  and  $\theta$  is implemented by gradient descent with respect to both vectors, that is

<sup>2</sup> Note that  $\phi_{\alpha}$  is not a distance function, since the space of the latter is nonlinear.

$$\frac{d\alpha(\tau)}{d\tau} = -\frac{\partial E_{data}(\alpha, \theta)}{\partial \alpha} - \nu \frac{\partial E_{shape}}{\partial \alpha}, \quad \text{and} \quad \frac{d\theta(\tau)}{d\tau} = -\frac{\partial E_{data}(\alpha, \theta)}{\partial \theta}$$

with  $\tau$  denoting the artificial evolution time, as opposed to the physical time  $t$ , and  $d\tau$  the gradient step size. Thereby, the data term gradient reads

$$\frac{\partial E_{data}}{\partial \alpha_i} = \int_{[1, T]} \int_{\Omega} \left( (I - \mu_1)^2 - (I - \mu_2)^2 \right) \delta(\phi_{\alpha, \theta}) \theta_{m_s}^{-1} \psi_i(T_{\theta}(\mathbf{x}, t)) \, d\mathbf{x} \, dt,$$

with  $\delta$  referring to the regularized Dirac measure, and the prior gradient  $\frac{\partial E_{shape}}{\partial \alpha} = \Sigma^{-1} \alpha$ . Furthermore, the gradient of the scaling and translating transformation is given by

$$\frac{\partial E_{shape}}{\partial \theta_{\xi}} = \int_{[1, T]} \int_{\Omega} \left( (I - \mu_1)^2 - (I - \mu_2)^2 \right) \delta(\phi_{\alpha, \theta}) \nabla \phi_{\alpha, \theta} \frac{dT_{\theta}}{d\theta_{\xi}} \, d\mathbf{x} \, dt$$

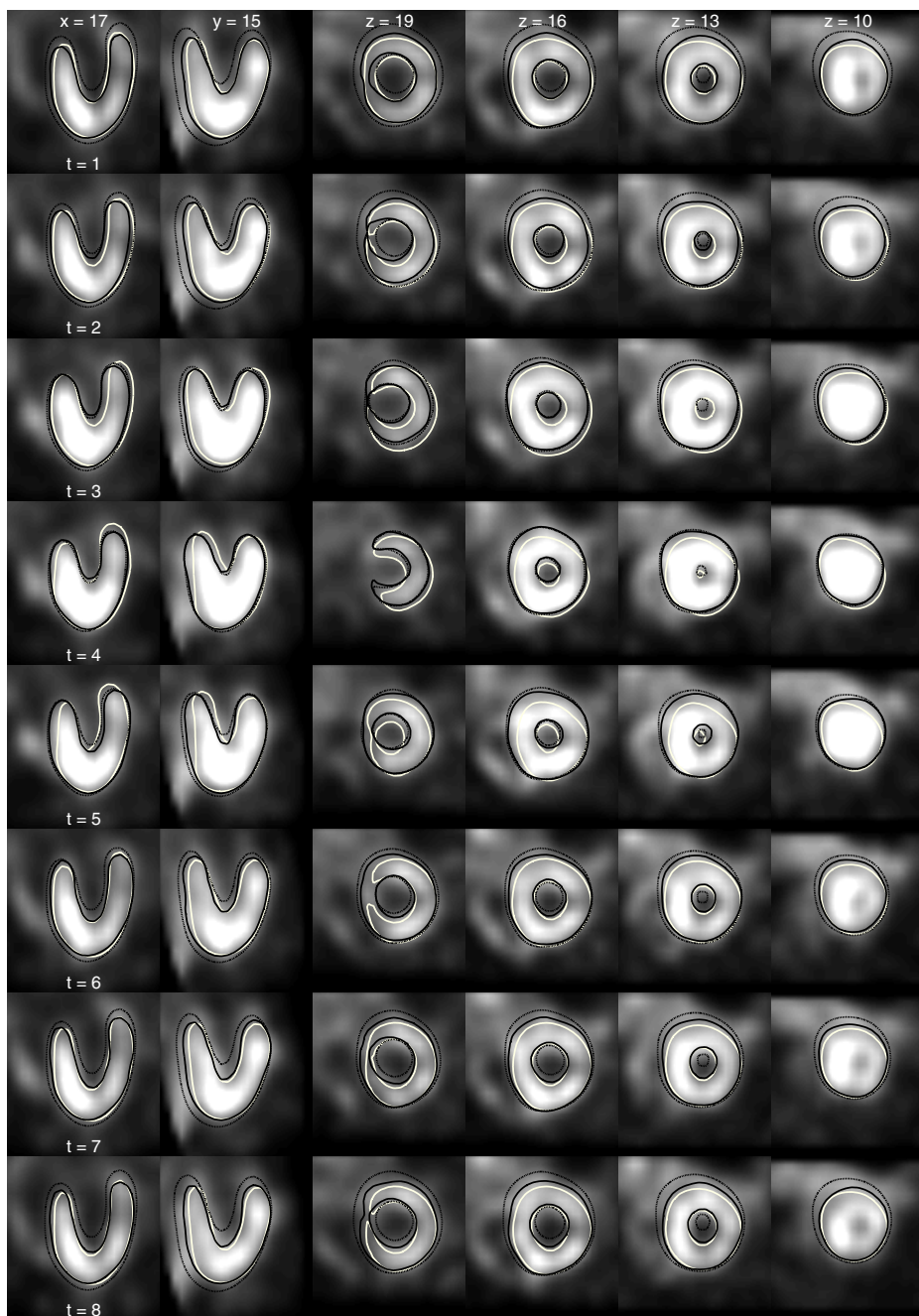
with  $\xi \in \{d, m\}$ , and  $\frac{dT_{\theta}}{d\theta_d} = -\begin{pmatrix} \mathbf{1}_3 \\ 1 \end{pmatrix}$ ,  $\frac{dT_{\theta}}{d\theta_m} = \begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}$ , respectively.

### 3 Experimental Evaluation and Discussion

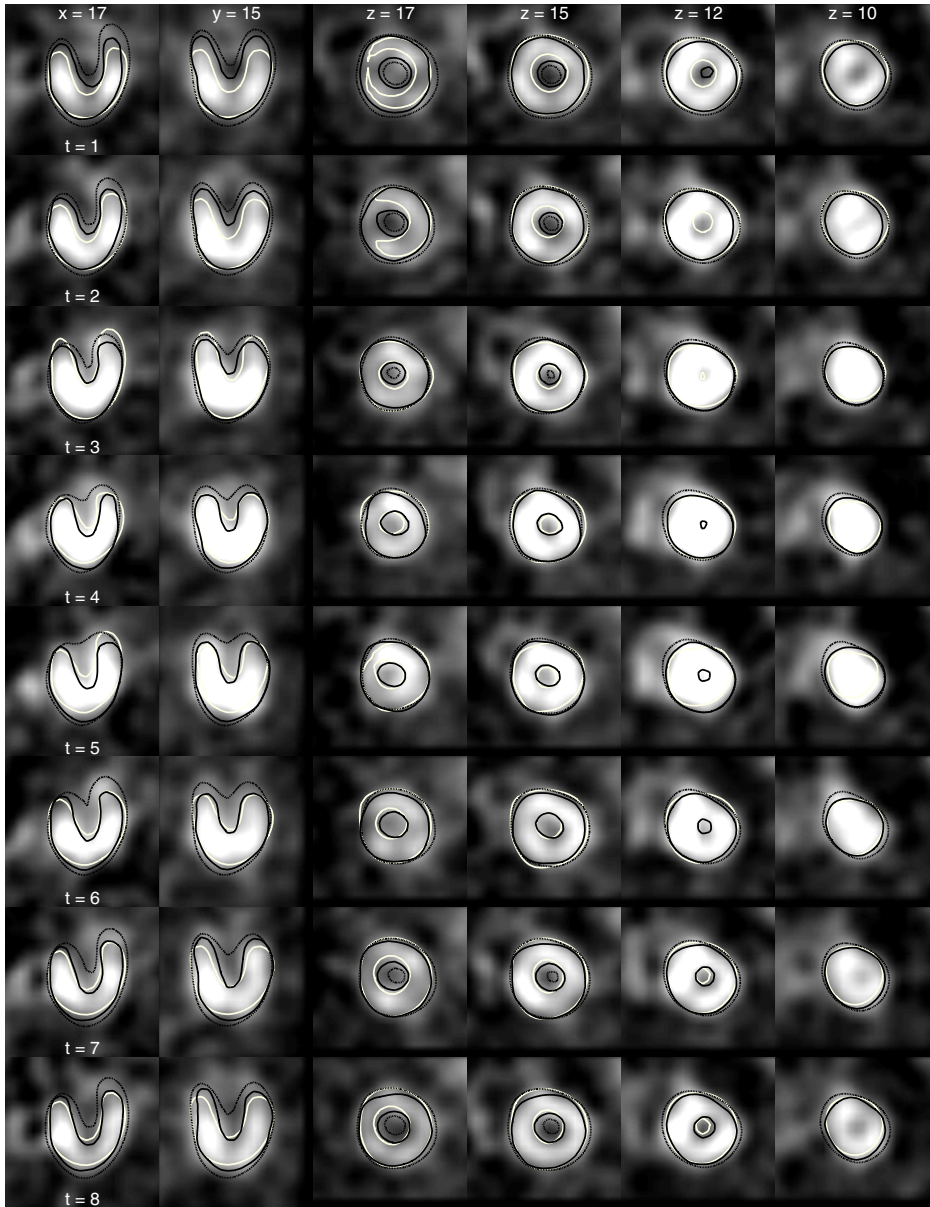
We compared the segmentation results of our proposed approach to that of Tsai et. al. [20] (3D, similarity transforms only, Chan-Vese) while measuring the *relative symmetric voxel error*:  $|b(\phi) \otimes b(\phi_{ref})| / |b(\phi_{ref})|$  with  $b(\phi)$  giving the binarized volume of  $\phi$ ,  $\otimes$  denoting the exclusive or operator and  $\phi_{ref}$  the ground truth distance function. As training and ground truth data sets we had 15 hand-segmented gated perfusion SPECT recordings of the left myocardium with a volume of  $30 \times 30 \times 27$  voxels (each of size  $6.59 \text{ mm}^3$ ) times eight frames over the cardiac cycle, and possible pathologies being removed. After alignment due to translation and scale, as explained above, 3D PCA was conducted over the vectorized  $15 \times 8$  volumes alike, whereas 4D PCA on the 15 vectorized sequences only. Ten gradient descent iterations were taken for 3D or 4D. Scaling and translation parameters  $\theta$  were performed prior to the interleaved descent for both  $\alpha$  and  $\theta$  in order to avoid local minima with respect to  $(\alpha, \theta)$  due to spatial or temporal misalignments. Different weightings  $\nu$  of the prior gradient were tried in the range from  $1/100$  to  $1/1$  of the data gradient's  $L^2$  norm and we chose the one yielding the best results w.r.t. the quality measure.

Figures 1 and 2 depict two typical segmentation results for both methods in comparison to the hand-segmented ground truth. While the dataset in Fig. 1 was contained in the training set, that in Fig. 2 was not. In the first experiment, the 3D-based method yields a systematic over-segmenting especially at end diastole, which is mainly due to over-estimating the per-volume scales. The error for the 3D approach did range from 35% (end systole) up to 90% (end diastole) vs. 40% to 20% with the 4D approach (averages: 53% vs. 28%). In the second experiment, Fig. 2, the 3D prior subspace allows for too many variations<sup>3</sup>, due to ignoring

<sup>3</sup> A stronger weighting of the prior term inhibits most of the temporal variation.



**Fig. 1.** Segmentation result for a sequence included in the training set. White: Manually segmented ground truth. Gray: 3D-prior method. Black: 4D-prior method. Estimating one scale for the whole sequence yields better results than one for each volume, and in turn allows for a better fit with respect to shape.



**Fig. 2.** Segmentation result for a sequence not being in the training set. White: Manually segmented ground truth. Gray: 3D-prior method. Black: 4D-prior method. Due to taking the temporal correlations into account also, the 4D prior yields a more specific prior per time point. Despite a strong weighting of the 3D prior, better results for the latter were not achievable, while still allowing for temporal variations.

the temporal index. On the other hand, the fifteen training vectors of the 4D PCA already yield an acceptable generalization (the errors were: 35% to 80% vs. 40% to 30% and 57% and 27% in average). These and other experiments clearly show the advantage of the new approach of taking the whole temporal dimension into account — both in the data-dependent probability, as well as the statistical shape prior.

**Acknowledgements.** The authors would like to thank Günther Platsch and Darrell Burckhardt of Siemens Molecular Imaging.

Upon submission of the camera-ready version of this manuscript, the authors of [1] have pointed us to a recently submitted related approach.

## References

1. A. Abufadel, T. Yezzi, and R. Schafer. 4D segmentation of cardiac data using active surfaces with spatiotemporal shape priors. *Applied Pattern Recognition*. submitted.
2. T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Trans. Image Processing*, 10(2):266–277, 2001.
3. R. Chandrashekar, A. Rao, G.I. Sanchez-Ortiz, and D. Rueckert. Construction of a statistical model for cardiac motion analysis using nonrigid image registration. In *18th Int. Conf. on Information Processing in Medical Imaging (IPMI '03)*, 2003.
4. T.F. Cootes, C. Beeston, G.J. Edwards, and C.J. Taylor. A unified framework for atlas matching using active appearance models. In *16th Int. Conf. Information Processing in Medical Imaging*, volume 1613, pages 322–333. Springer Berlin, 1999.
5. D. Cremers and G. Funka-Lea. Dynamical statistical shape priors for level set based sequence segmentation. In *Third Workshop on Variational and Level Set Methods*, volume 3752 of *LNCS*, pages 210–221. Springer Berlin, 2005.
6. R. Davies, C. Twining, Cootes T., J. Waterton, and C. Taylor. A minimum description length approach to statistical shape modelling. *IEEE Trans. on Medical Imaging*, 21(5):525–537, 2002.
7. H. Delingette and J. Montagnat. New algorithms for controlling active contours shape and topology. In *ECCV 2000*, volume LNCS 1843, pages 381–395. Springer.
8. A.F. Frangi. Three-dimensional modeling for functional analysis of cardiac images: A review. *IEEE Trans. Medical Imaging*, 20(1):2–25, Jan. 2001.
9. A.F. Frangi, D. Rueckert, J.A. Schnabel, and W.J. Niessen. Automatic construction of multiple-object three-dimensional statistical shape models: Applications to cardiac modeling. *IEEE Trans. Medical Imaging*, 21(9):1151–1166, Sept. 2002.
10. S. Kichenassamy, A. Kumar, P.J. Olver, A. Tannenbaum, and A.J. Yezzi. Gradient flows and geometric active contour models. In *ICCV 1995*, pages 810–815, 1995.
11. M.E. Leventon, W.E. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *CVPR 2000*, volume 1, pages 1316–1323. IEEE, 2000.
12. M. Lorenzo-Valdéz, G.I. Sanchez-Ortiz, R. Mohiaddin, and D. Rueckert. Segmentation of 4d cardiac MR images using a probabilistic atlas and the EM algorithm. In *MICCAI 2005*, volume 2878, pages 440–450. Springer, Berlin, 2005.
13. R. Malladi, J. A. Sethian, and B.C. Vemuri. Shape modeling with front propagation: A level set approach. *PAMI*, 17(2):158–175, 1995.
14. J. Montagnat and H. Delingette. 4D deformable models with temporal constraints: Application to 4D cardiac image segmentation. *Med. Im. Anal.*, 9(1):87–100, 2005.



15. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42:577–685, 1989.
16. S.J. Osher and J.A. Sethian. Fronts propagation with curvature dependent speed: Algorithms based on hamilton-jacobi formulations. *J. of Comp. Ph.*, 79:12–49, 1988.
17. N. Paragios. A variational approach for the segmentation of the left ventricle in cardiac image analysis. *IJCV*, 50(3):345–362, July 2002.
18. D. Perperidis, R. Mohiaddin, and D. Rueckert. Construction of a 4d statistical atlas of the cardiac anatomy and its use in classification. In J. Duncan and G. Gerig, editors, *MICCAI 2005*, volume 3750 of *LNCIS*, pages 402–410. Springer, 2005.
19. W. Sun, M. Cetin, R. Chan, V. Reddy, G. Holmvang, V. Chandar, and A. Willsky. Segmenting and tracking the left ventricle by learning the dynamics in cardiac images. Technical Report LIDS 2642, MIT, Feb. 2005.
20. A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W.E. Grimson, and A. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Trans. Medical Imaging*, 22(2):137–154, Feb. 2003.