

Incorporating Shape Variability in Image Segmentation via Implicit Template Deformation

Raphael Prevost^{1,2}, Remi Cuingnet¹, Benoit Mory¹,
Laurent D. Cohen², and Roberto Ardon¹

¹ Philips Research Medisys, Suresnes, France

² CEREMADE UMR 7534, CNRS, Université Paris Dauphine, Paris, France

Abstract. Implicit template deformation is a model-based segmentation framework that was successfully applied in several medical applications. In this paper, we propose a method to learn and use prior knowledge on shape variability in such framework. This shape prior is learnt via an original and dedicated process in which both an optimal template and principal modes of variations are estimated from a collection of shapes. This learning strategy requires neither a pre-alignment of the training shapes nor one-to-one correspondences between shape sample points. We then generalize the implicit template deformation formulation to automatically select the most plausible deformation as a shape prior. This novel framework maintains the two main properties of implicit template deformation: topology preservation and computational efficiency. Our approach can be applied to any organ with a possibly complex shape but fixed topology. We validate our method on myocardium segmentation from cardiac magnetic resonance short-axis images and demonstrate segmentation improvement over standard template deformation.

1 Introduction

Model-based methods are particularly effective and popular in medical image segmentation. Among them, template deformation has recently been used in various applications [1–4] for its interesting properties (computational efficiency, topology preservation, compatibility with user interactions). This variational method consists in seeking a segmenting implicit function as a deformed implicit template. This template, acting as a shape prior, is therefore of paramount importance. However, in previous works the initial template was either set as a synthetic model (*e.g.* ellipsoid for a kidney [2–4]) or as a segmented organ from a single arbitrary image [1]. Despite the consensus that learning shape priors is a powerful approach to improve robustness [5, 6], this has never been proposed in the context of segmentation by implicit template deformation.

In this paper, our goal is thus to improve implicit template deformation by taking into account learnt shape information while keeping both topology preservation and computational efficiency. Shape learning is often performed through statistical analysis of boundary vertices [6, 7] or implicit functions [8, 9]. However

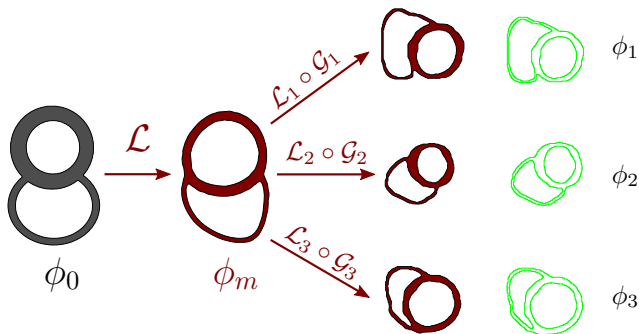


Fig. 1. Given an initial synthetic shape ϕ_0 , a set of shapes $\{\phi_i\}_i$ is simultaneously segmented via implicit template deformation while an intermediate mean shape $\phi_m = \phi_0 \circ \mathcal{L}$ (with the same topology as ϕ_0) is estimated

neither of these two representations guarantees topology preservation after averaging. To do so, other methods represent shapes via diffeomorphisms [10, 11], but they come with a high computational cost that is not compatible with real-time segmentation. We therefore propose a dedicated learning approach by using the template deformation energy as a pre-metric in the shapes space. This yields a co-segmentation process (similar to [12] for registration), within which an optimal template is estimated (see Figure 1). We also capture further information by building a space of main deformations around this template. Finally, we introduce a generalization of the template deformation formulation by using the computed statistics in the regularization term. The proposed framework is generic and can be applied to any organ with a possibly complex and variable shape but a fixed topology. We demonstrate its efficacy and interest by addressing the problem of myocardium segmentation in 2D cine-MR images.

In the following, Section 2 introduces the main notations and recalls the implicit template deformation framework. In Section 3, we develop the learning framework to estimate statistics that will be used to improve segmentation in Section 4. Optimization details are provided in Section 5. Validation results on clinical data are presented in Section 6 and discussion concludes the paper.

2 Segmentation by Implicit Template Deformation

Implicit template deformation [1, 2] is a variational framework for image segmentation. The segmentation is defined through the zero level-set of an implicit function $\phi : \Omega \rightarrow \mathbb{R}$, and ϕ is positive (resp. negative) inside (resp. outside) the segmentation. In this framework, the set of admissible segmentations \mathbb{S} is defined via an implicit template $\phi_0 : \Omega \rightarrow \mathbb{R}$ as the set of all implicit functions with the same topology as ϕ_0 , i.e. $\mathbb{S} = \{\phi : \Omega \rightarrow \mathbb{R} \text{ s.t. } \phi = \phi_0 \circ \psi, \psi \text{ is diffeomorphic}\}$. The unknown is thus the transformation $\psi : \Omega \rightarrow \Omega$ which is sought as a minimum of a region competition energy:

$$\min_{\psi} \left\{ \int_{\Omega} H(\phi_0 \circ \psi) r_{int} + \int_{\Omega} (1 - H(\phi_0 \circ \psi)) r_{ext} + \lambda \mathcal{R}(\psi) \right\}, \quad (1)$$

where H denotes the Heaviside function ($H(a) = 1$ if $a > 0$, 0 otherwise) while r_{int} and r_{ext} are image-based functions such as $r_{int}(\mathbf{x})$ is lower (resp. higher) than $r_{ext}(\mathbf{x})$ if voxel \mathbf{x} seems to belong to the target object (resp. background). $\mathcal{R}(\psi)$ is a constraint term on ψ that prevents the segmentation $\phi = \phi_0 \circ \psi$ to deviate too much from the initial template ϕ_0 ; it is weighted by a parameter λ . In [2], ψ is decomposed into (i) a global transformation $\mathcal{G} \in \mathbb{G}$ (e.g. a similarity) accounting for the pose of the template in the image, and (ii) a diffeomorphism $\mathcal{L} \in \mathbb{D}(\Omega)$ that yields local deformation and does change the shape of the template. This decomposition allows to define the regularization as a function of the deformation only $\mathcal{R}(\psi) = \mathcal{R}(\mathcal{L}) = \frac{\lambda}{2} \|\mathcal{L} - Id\|_2^2$. The problem finally reads

$$\min_{\mathcal{L}, \mathcal{G}} \left\{ \int_{\Omega} H(\phi_0 \circ \mathcal{L} \circ \mathcal{G}) r_{int} + \int_{\Omega} (1 - H(\phi_0 \circ \mathcal{L} \circ \mathcal{G})) r_{ext} + \frac{\lambda}{2} \|\mathcal{L} - Id\|_2^2 \right\}. \quad (2)$$

In such a setting, ϕ_0 not only fixes the topology of the segmentation but also acts as a shape prior, which makes its choice of paramount importance. Moreover, the term \mathcal{R} could be improved by taking into account shape variability of the considered organ. In the next section, we develop a framework to tackle both problems by estimating statistics on a collection of shapes.

3 A Dedicated Learning of Shape Variability

Consider N shapes of a given organ (coming for example from manual expert segmentations) implicitly represented by $\{\phi_i\}_{i=1..N} \subset \mathbb{S}$. From this set we aim to extract useful statistical information in terms of segmentation, that is to say a mean shape and a deformation model.

In order to estimate statistics in \mathbb{S} , we first define an adapted pre-metric in this space. Any shape $\phi_1 \in \mathbb{S}$ can be warped to another shape $\phi_2 \in \mathbb{S}$ via implicit template deformation by solving Problem (2) with $\phi_0 := \phi_1$, and for example $r_{int}^{\phi_2} := \max(-\phi_2, 0)$ and $r_{ext}^{\phi_2} := \max(\phi_2, 0)$. This leads to a dedicated definition of shape dissimilarity:

$$C^2(\phi_1, \phi_2) = \min_{\substack{\mathcal{L} \in \mathbb{D} \\ \mathcal{G} \in \mathbb{G}}} \left\{ \int_{\Omega} H(\phi_1 \circ \mathcal{L} \circ \mathcal{G}) r_{int}^{\phi_2} + \int_{\Omega} (1 - H(\phi_1 \circ \mathcal{L} \circ \mathcal{G})) r_{ext}^{\phi_2} + \frac{\lambda}{2} \|\mathcal{L} - Id\|_2^2 \right\}$$

As in [12], our dedicated notion of mean is defined via a minimization problem :

$$\phi_m = \operatorname{argmin}_{\phi \in \mathbb{S}} \sum_{i=1}^N C^2(\phi, \phi_i) \Leftrightarrow \phi_m = \phi_0 \circ \left\{ \operatorname{argmin}_{\mathcal{L} \in \mathbb{D}(\Omega)} \sum_{i=1}^N C^2(\phi_0 \circ \mathcal{L}, \phi_i) \right\} \quad (3)$$

The right-handside equivalence comes from the constraint $\phi_m \in \mathbb{S}$. Indeed, the mean template has to preserve the topology of the training shapes. Expanding the segmentation costs and neglecting constant terms in Equation (3) yields the following optimization problem to solve:

$$\min_{\substack{\mathcal{L} \in \mathbb{D}(\Omega) \\ (\mathcal{L}_i)_{i \in \mathbb{D}(\Omega)}^N, (\mathcal{G}_i)_{i \in \mathbb{G}^N}}} E_{learn} = - \sum_{i=1}^N \int_{\Omega} H(\phi_0 \circ \mathcal{L} \circ \mathcal{L}_i \circ \mathcal{G}_i) \phi_i + \frac{\lambda}{2} \|\mathcal{L}_i - Id\|_2^2. \quad (4)$$

This can be interpreted as segmenting simultaneously all training shapes $\{\phi_i\}_i$ starting from ϕ_0 while estimating an optimal common intermediate shape $\phi_0 \circ \mathcal{L}$ (see Figure 1). In Eq (4), the energy is minimized (see Section 5) with respect to three kinds of variables

- the global transformations $(\mathcal{G}_i)_i$, called the **poses**, that register all shapes to ϕ_0 with translation, rotation and scaling. As they are part of the optimization process (see Section 5), they do not bias the learning, as a fixed pre-alignment (*e.g.* [8, 9]) would do.
- the **common deformation** \mathcal{L} , which includes the common parts of the deformations from ϕ_0 to all the training shapes.
- the local deformations $(\mathcal{L}_i)_i$, called the **residual deformations**, are the residual components of the deformations from $\phi_0 \circ \mathcal{L}$ to ϕ_i . Unlike \mathcal{L} , their magnitude is penalized so that any deformation which is common to all the training set will be preferably included in \mathcal{L} .

The optimal common deformation \mathcal{L}^* can be used to define the optimal template as $\phi_m = \phi_0 \circ \mathcal{L}^*$. This shape globally minimizes the magnitude of residual deformations to each shape of the dataset. Note that \mathcal{L} is not penalized so the choice of ϕ_0 defines the topology of ϕ_m but does not affects it further.

The optimal residual deformations $(\mathcal{L}_i^*)_i$ are also available and can be used to capture further information on the variability of the training shapes. We build a space of principal deformations \mathbb{L} to constrain future segmentation of new images. As in [13], a principal component analysis (PCA) is applied to the residual deformations $(\mathcal{L}_i^*)_i$ to find a suitable parametrization of such a space. Any deformation $\ell \in \mathbb{L}$ can then be written as a linear combination of the offset $\bar{\ell}$ and $(\ell_k)_{k=1..M}$ the first M modes of variation:

$$\ell[w] = \bar{\ell} + \sum_{k=1}^M w_k \ell_k \quad , \quad w \in \mathbb{R}^M. \quad (5)$$

The space of diffeomorphisms is not stable under linear combinations so elements of \mathbb{L} are not necessarily diffeomorphisms. Nevertheless we show hereafter how this space indirectly in a topology-preserving segmentation framework.

4 Generalized Implicit Template Deformation

The previously estimated statistical information can be used to improve the implicit template deformation described in Eq. (2). A first improvement is achieved by replacing the original template ϕ_0 by the mean template $\phi_m = \phi_0 \circ \mathcal{L}^*$. Secondly, the estimation of the deformation can also be enhanced by using the space of principal deformations \mathbb{L} . In most previous work [7, 9, 14], the learnt variable is directly expressed as a linear combination of modes. Here we rather modify the regularization term \mathcal{R} so that diffeomorphism \mathcal{L} is constrained with respect to the set \mathbb{L} instead of the identity. Thus, only deformations that cannot be explained through the learnt space \mathbb{L} are penalized. The rationale is to use linear

combinations of diffeomorphisms indirectly to preserve both the topology and the computational efficiency. The new segmentation energy therefore reads

$$E_{seg}(\mathcal{L}, \mathcal{G}, w) = \int_{\Omega} H(\phi_m \circ \mathcal{L} \circ \mathcal{G}) r_{int} + (1 - H(\phi_m \circ \mathcal{L} \circ \mathcal{G})) r_{ext} + \frac{\lambda}{2} \|\mathcal{L} - \ell[w]\|_2^2. \quad (6)$$

Minimization of (6) can be performed with an alternate scheme:

Update of the Segmentation. With $\ell[w]$ fixed, the energy is minimized through a gradient descent-like scheme on \mathcal{L} and \mathcal{G} (see Section 5).

Update of the Modes Weights. With \mathcal{L} and \mathcal{G} fixed, the update of $L[w]$ can be seen as a projection of \mathcal{L} onto \mathbb{L} . Indeed the energy minimization comes down to a simple quadratic problem, whose closed-form solution is

$$\forall k \in \{1, \dots, M\}, w_k = \langle \mathcal{L} - \bar{\ell}, \ell_k \rangle / \langle \ell_k, \ell_k \rangle. \quad (7)$$

The first step is similar to standard template deformation [2] and the second one is straightforward. Therefore, the proposed algorithm maintains the efficiency of the original algorithm. Further details on optimization are provided herebelow.

5 Optimization Schemes

Both learning (4) and segmentation (6) energies involve variables, either in \mathbb{G} for poses or in $\mathbb{D}(\Omega)$ for deformations, that are simultaneously updated. Variables in \mathbb{G} can be parameterized by a vector \mathbf{p} in \mathbb{R}^P (translation, rotation angles, scales). Minimization of a given energy E is done through a gradient descent on this vector: $\mathbf{p}^{(n+1)} \leftarrow \mathbf{p}^{(n)} - \Delta t \nabla_{\mathbf{p}} E$, where Δt is the time step. On the other hand, such a process is not suitable in $\mathbb{D}(\Omega)$ as this space is not stable under linear combinations. A more appropriate way is to combine diffeomorphisms via composition since $(\mathbb{D}(\Omega), \circ)$ is a group. Following [1], we therefore update any diffeomorphism \mathcal{L} in the following way: $\mathcal{L}^{(n+1)} \leftarrow (Id - \Delta t \nabla_{\mathcal{L}} E) \circ \mathcal{L}^{(n)}$. The regularity is enforced by a Gaussian filtering of the gradient as in [2].

6 Validation

We validated our method in the context of myocardium analysis and segmentation in cardiac short-axis 2D cine-MR images. This task is particularly challenging for model-based approaches because of the complex topology of the target object, i.e. a band around left and right ventricles.

Our dataset is composed of 245 MR images coming from 61 different patients, which was randomly split into a training set (120 images, 30 patients) and a testing set (125 images, 31 patients). The acquisitions have been synchronised so that each heart is in the same cardiac phase. In every image, a myocardium segmentation has been manually performed by a radiologist. The initial synthetic template ϕ_0 used is shown in Figure 1.

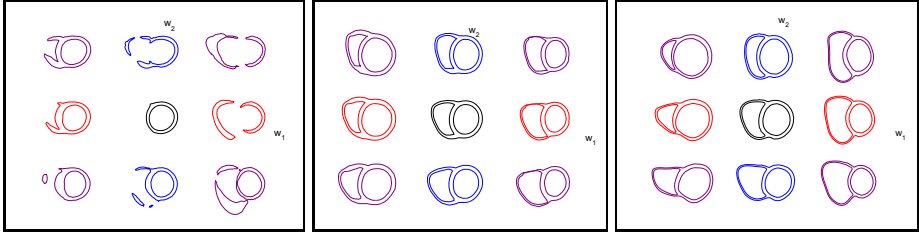


Fig. 2. Mean model and first two modes of the variation of the myocardium learnt on the training dataset using implicit shape model [9] (left), active shape model [7] (middle) and the proposed method (right). For our approach, the visualized shapes are the zero level-sets of $\phi_m \circ (\bar{\ell} + w_1 \ell_1 + w_2 \ell_2)$.

First we compare qualitatively the shape information learnt from the implicit shape model proposed in [9], the active shape framework [7] and our proposed approach. Figure 2 shows the mean shape and first two modes of variation for each method. The implicit method fails at recovering the true topology of the mean shape, as well as its first modes of variations. The explicit method performs better and provides a reasonable mean model. However, the modes of variation are less satisfying than the modes of deformations learnt with our approach (which tends to provide a better topology preservation and seem more realistic).

Then we evaluate how learnt information improves segmentation via implicit template deformation of unseen images. Myocardiums have been segmented in test images using (i) the synthetic model ϕ_0 as template, (ii) the estimated mean model ϕ_m as template, (iii) the new deformation model-based regularization term in addition to the mean model ϕ_m (with 5 modes). The image-based classification functions r_{int} and r_{ext} were negative log-likelihoods of intensity probability distributions inside and outside the myocardium. Performance of each algorithm is quantified using Dice coefficients between the segmentation and the expert ground truth. Results on the whole testing set are summarized in Figure 3. Both the replacement of the template ϕ_0 by ϕ_m and the introduction of the new regularization term improved the robustness of the segmentations (p -value < 0.0001 for a Wilcoxon signed-rank test). To illustrate this improvement, Figure 4 shows some results in three different cases, for the classical regularization term with two values of the shape constraint parameter $\lambda \in \{1, 2\}$ and the new model-based regularization term. In all settings, the template was the mean model ϕ_m . Consider Case #1: since the image term is reliable, a satisfying result is obtained with a small shape constraint. However, the myocardium deviates significantly from the mean shape: using a too strong constraint λ_2 prevents the algorithm to converge towards the right solution. Conversely in Case #2, the image information is much more ambiguous. This provokes some leaks with λ_1 , which shows there is no fixed value that allows a good segmentation in both cases. Yet by introducing the new regularization (fourth column), likely deformations are not penalized. This allows us to widen the capture range while still

avoiding unrealistic leaks. Finally, case #3 illustrates that our method may also improve the result even if no λ was originally successful.

Fig. 3. Boxplot of the Dice coefficients for myocardium segmentation in MR images via implicit template deformation with synthetic model ϕ_0 (left), mean model ϕ_m (middle), mean model ϕ_m and deformation model (right)

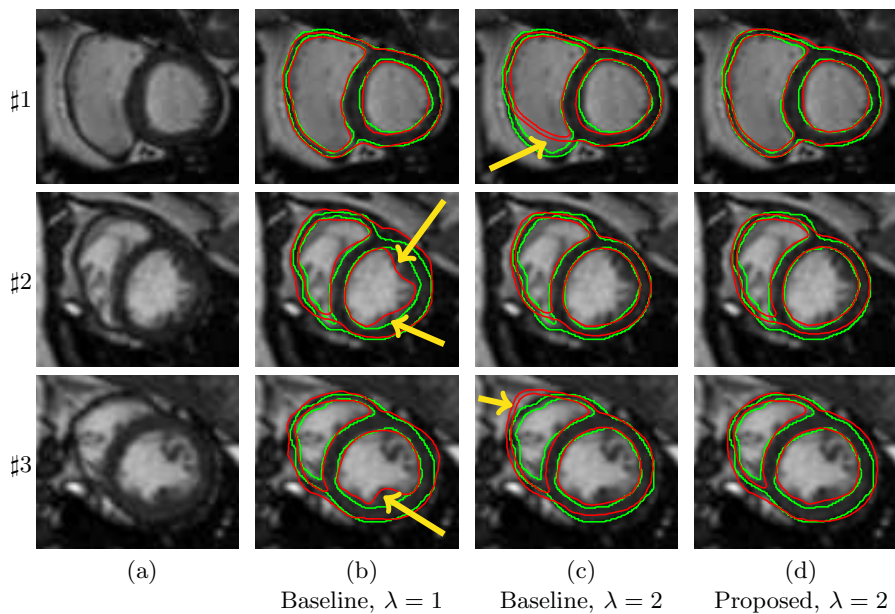
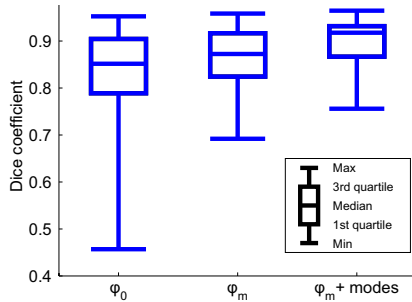


Fig. 4. Segmentation results (red) of different cases versus ground truths (green). Main failures are highlighted by yellow arrows. (a) Original images, (b,c) Baseline method [2] with small ($\lambda = 1$) and high ($\lambda = 2$) shape constraint, (d) Proposed method.

7 Conclusion

In this paper we have presented a method to include organ shape variability in the implicit template deformation framework. A variational approach was proposed to extract statistical information (mean and principal variations) from a collection of shapes. This training method is automatic, does not require landmarks correspondance and relies upon a definition of shape dissimilarity that is

directly derived from the implicit template deformation functional. We also proposed a generalization of the original segmentation algorithm in which the shape prior is automatically adapted to the current image during the deformation process with almost no additional cost (segmentation takes around one second on a standard computer). Quantitative results demonstrated the improvement over implicit template deformation for a 2D application. Our approach is very generic and can be used to segment any object with a complex shape but a fixed topology that shall be preserved. Furthermore, extension in 3D or to multiple objects (*e.g.* brain structures) is straightforward thanks to the implicit representation of shapes. Despite its paramount importance, the image-based term was not investigated as we focused on incorporating shape information on top of any pixelwise classifier. We plan to extend our framework to learn a dedicated appearance model as well.

References

1. Saddi, K., et al.: Region-Based Segmentation via Non-Rigid Template Matching. In: ICCV, vol. 1, pp. 1–7 (2007)
2. Mory, B., Somphone, O., Prevost, R., Ardon, R.: Real-Time 3D Image Segmentation by User-Constrained Template Deformation. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part I. LNCS, vol. 7510, pp. 561–568. Springer, Heidelberg (2012)
3. Prevost, R., et al.: Kidney Detection and Real-Time Segmentation in 3D Contrast-Enhanced Ultrasound Images. In: IEEE ISBI, pp. 1559–1562 (2012)
4. Cuingnet, R., Prevost, R., Lesage, D., Cohen, L.D., Mory, B., Ardon, R.: Automatic Detection and Segmentation of Kidneys in 3D CT Images Using Random Forests. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part III. LNCS, vol. 7512, pp. 66–74. Springer, Heidelberg (2012)
5. Cremers, D., et al.: A Review of Statistical Approaches to Level Set Segmentation: Integrating Color, Texture, Motion and Shape. IJCV 72(2), 195–215 (2007)
6. Heimann, T., Meinzer, H.P.: Statistical Shape Models for 3D Medical Image Segmentation: a Review. Medical Image Analysis 13(4), 543 (2009)
7. Cootes, T.F., et al.: Active Shape Models-Their Training and Application. Computer Vision and Image Understanding 61(1), 38–59 (1995)
8. Rousson, M., Paragios, N.: Shape Priors for Level Set Representations. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002, Part II. LNCS, vol. 2351, pp. 78–92. Springer, Heidelberg (2002)
9. Tsai, A., et al.: A Shape-Based Approach to the Segmentation of Medical Imagery using Level Sets. Trans. Med. Imaging 22(2), 137–154 (2003)
10. Vaillant, M., et al.: Statistics on Diffeomorphisms via Tangent Space Representations. NeuroImage 23(1), 161 (2004)
11. Arsigny, V., Commowick, O., Pennec, X., Ayache, N.: A Log-Euclidean Framework for Statistics on Diffeomorphisms. In: Larsen, R., Nielsen, M., Sparring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 924–931. Springer, Heidelberg (2006)
12. Joshi, S., et al.: Unbiased Diffeomorphic Atlas Construction for Computational Anatomy. NeuroImage 23(1), 151 (2004)
13. Rueckert, D., Frangi, A.F., Schnabel, J.A.: Automatic Construction of 3D Statistical Deformation Models using Non-Rigid Registration. In: Niessen, W.J., Viergever, M.A. (eds.) MICCAI 2001. LNCS, vol. 2208, pp. 77–84. Springer, Heidelberg (2001)
14. Leventon, M., et al.: Statistical Shape Influence in Geodesic Active Contours. In: CVPR, pp. 316–323 (2000)