Tagged Template Deformation

Raphael Prevost[1,2,⋆], Benoit Mory[1], Remi Cuingnet[1], Laurent D. Cohen[2], Roberto Ardon[1]

[1] Philips Research Medisys, Suresnes, France
[2] CEREMADE UMR 7534, Universite Paris Dauphine, Paris, France

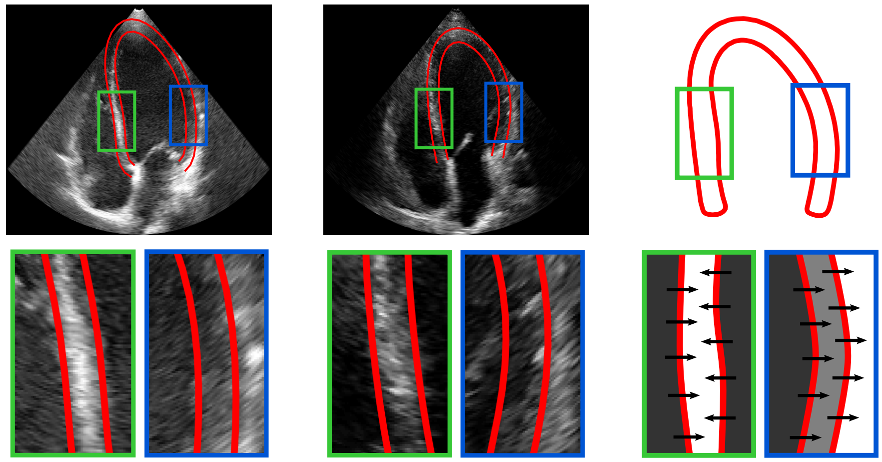
Abstract. Model-based approaches are very popular for medical image segmentation as they carry useful prior information on the target structure. Among them, the implicit template deformation framework recently bridged the gap between the efficiency and flexibility of level-set region competition and the robustness of atlas deformation approaches. This paper generalizes this method by introducing the notion of tagged templates. A tagged template is an implicit model in which different subregions are defined. In each of these subregions, specific image features can be used with various confidence levels. The tags can be either set manually or automatically learnt via a process also hereby described. This generalization therefore greatly widens the scope of potential clinical application of implicit template deformation while maintaining its appealing algorithmic efficiency. We show the great potential of our approach in myocardium segmentation of ultrasound images.

1 Introduction

Segmentation of medical images is an important part of clinical work-flow, typically used to assess anatomical information such as the volume or the shape of an organ. Leveraging the strong anatomical priors available for medical images, model-based approaches are particularly effective and popular. In a number of clinical applications, methods based on atlas deformation achieve state-of-the-art performance results [16]. Yet, they suffer from a high computational burden and can only be employed for images with a standardized acquisition protocol (such as CT or MR images). Very recently though, the *implicit template deformation* framework [15, 12] bridged the gap between level-set region competition [4] and atlas deformation approaches. Its unique properties (computational efficiency, topology preservation, compatibility with user interactions) were employed to achieve fast and reliable segmentation of different kinds of medical images [14, 6, 8]. Yet even if an advanced shape prior can be embedded within this framework [13], it still assumes that appearance can be globally defined, which is a strong constraint. In this paper, we present a method to use an enriched model that couples the prior information on the object's appearance with its shape.

Elaborated appearance models have been already proposed in other frameworks. One of the earliest was the well-known *active appearance model* [3] that

⋆ Raphael Prevost is now affiliated to ImFusion GmbH (Munich, Germany).



(a) (b) (c)

Fig. 1. The appearance of an object (here the myocardium) may vary (a) between different regions (*e.g* green and blue). However, this variation is consistent across subjects (b) so it can be embedded as a prior information within the model (c). For each edge of the model, we know what direction the image gradient (black arrows) should have.

however suffers from the usual problems of explicit methods. Level-set approaches were also extended to take into account variations in the appearance prior. In [11], the appearance model is a function of the distance to the shape boundary and in [5] a labeling function balancing the contribution of the data-fidelity term with respect to the regularization term is introduced. While this allows a better representation than a global appearance prior, it is still too constrained for many medical applications. Our work can be thought of as a generalization and extension of both ideas. Finally, an extension of the *MetaMorph* framework was proposed in [9]. This work aimed at learning jointly the shape and the appearance of the organ to be segmented. However, this results in a mutual information registration problem that needs to be solved at test time, which would be too time-consuming in 3D for our target applications.

The main contributions of this article are essentially methodological and consist in (*i*) an interpretation of template deformation as a flux minimization problem, (*ii*) an enrichment of the shape template with an appearance model that is able to exploit specific image features with various confidence levels, (*iii*) a method to automatically learn such enhanced models.

2 Segmentation by implicit template deformation

Implicit template deformation [12, 13] is a variational framework for image segmentation that consists in finding an implicit function $\phi : \Omega \to \mathbb{R}$ whose zero level-set will be the segmentation boundary. The key particularity is that the set of admissible implicit functions is defined with respect to an initial *implicit* *template* $\phi_0 : \Omega_0 \rightarrow \mathbb{R}$ as the set of functions obtained by deforming ϕ_0 with a geometric transformation ψ. Provided some constraints on ψ, the template ϕ_0 can be considered as a shape model: if ψ is a diffeomorphism, then ϕ and ϕ_0 will have the same topology. The transformation ψ is sought as the minimum of a region competition energy and a regularization term:

$$\min_{\psi} \mathcal{R}(\psi) + \int_{\Omega} H(\phi_0 \circ \psi(\mathbf{x}))\ r_{int}(\mathbf{x})\ d\mathbf{x} + \int_{\Omega} (1 - H(\phi_0 \circ \psi(\mathbf{x})))\ r_{ext}(\mathbf{x})\ d\mathbf{x} \quad (1)$$

where H is the Heaviside step function, and $r_{int} : \Omega \rightarrow \mathbb{R}^+$ and $r_{ext} : \Omega \rightarrow \mathbb{R}^+$ are pointwise classification error functions in the foreground (resp. background) region. $\mathcal{R}$ is a regularization term that penalizes the magnitude and irregularities of local deformations induced by ψ, typically set to a Free-Form deformation.

Choice of the implicit template In most previous work, the implicit template ϕ_0 was either set to the signed distance function of a simple geometric shape [14] or to a given pre-segmented shape [15]. However, another (and better) strategy is to learn ϕ_0 as the mean of a training set [13].

Regularization prior The regularization term can be defined as $\mathcal{R}(\psi) = \frac{1}{2}\|\psi - \mathbf{Id}\|_\sigma^2$, where $\|x\|_\sigma^2 = \langle x, K_\sigma^{-1} * x\rangle$ and K_σ a Gaussian kernel. When a learning database is available, this term can be generalized to $\mathcal{R}(\psi) = \frac{1}{2}\|\psi - P_{\mathbb{L}}\psi\|_U^2$ where $P_{\mathbb{L}}$ is the operator that projects a deformation onto the space $\mathbb{L}$ of the first variation modes of a PCA analysis [13].

Appearance prior The image-based functions r_{int} and r_{ext} are usually defined as logarithms of probabilities, following maximum likelihood principles. Such probabilities can be learnt via classifiers, such as random forests [6]. However they are estimated using only local image features and are therefore *independent from the shape model*. In Figure 1, we show a clinical application in which this approach is too restrictive. Indeed, the appearance of the target organ may vary (depending on the location with respect to the model).

3 Tagged template deformation

A flux minimization problem The image-based term in (1) is a volume integral and can be written as $E(\psi) = \int_{\Omega} H(\phi_0 \circ \psi(\mathbf{x}))\ (r_{int} - r_{ext})(\mathbf{x})\ d\mathbf{x}$ plus a constant term that is neglected in the minimization. However, under simple regularity assumptions, this energy can be reformulated as a surface integral:

$$E(\psi) = \int_{(\phi_0 \circ \psi)^{-1}(0)} \langle \mathbf{f}(\mathbf{x})\ , \mathbf{n}(\mathbf{x}) \rangle\ d\mathbf{x} \quad (2)$$

with $\mathbf{n}$ the normal of the current segmentation and $\mathbf{f} = \mathcal{G} * (r_{int} - r_{ext})$ the convolution of the image term with the Green function $\mathcal{G}$ of Poisson equation [1]. E is therefore the flux of the vector field $\mathbf{f}$ across the segmentation boundary.

In a number of clinical applications such as ultrasound, we may only assume what contrast the target object should have (*e.g* it is brighter than its neighborhood). A convenient choice for $\mathbf{f}$ is then the gradient of the image $\boldsymbol{\nabla} I$ (or its opposite, depending on the contrast of the target object). As pointed out in [10], the corresponding image-based functions are the image Laplacian $r_{int}(\mathbf{x}) = \pm \Delta I(\mathbf{x})$ and $r_{ext}(\mathbf{x}) = 0$. Our proposed model encompasses this case and even goes beyond by considering any surface or region-based feature.

Flux minimization with tagged model Let us assume that we have a set of K such vector fields $(\mathbf{f}_k)_{k=1 \ldots K}$. Instead of encoding the appearance of the whole target object, each $\mathbf{f}_k$ can be specialized to describe a particular region of the structure to be segmented. This region is defined via a *tag* function $T_k : \Omega_0 \to [-1, 1]$. The absolute value of T_k is a fuzzy indicator of the region, while its sign indicates whether the flux of $\mathbf{f}_k$ should be minimized or maximized. Note that T_k, as it is defined in the template referential Ω_0 rather than the image referential Ω, will have to be warped by ψ as well. The tagged implicit template deformation energy therefore reads

$$E(\psi) = \sum_{k=1}^{K} \int_{(\phi_0 \circ \psi)^{-1}(0)} T_k \circ \psi(\mathbf{x}) \; \langle \mathbf{f}_k(\mathbf{x}) \; , \mathbf{n}(\mathbf{x}) \rangle \; d\mathbf{x} = \sum_{k=1}^{K} E_k(\psi) \tag{3}$$

Numerical optimization After applying the divergence theorem, the derivative of E with respect to a parameter $\mathbf{p}_i$ of the transformation ψ reads:

$$\begin{aligned} \nabla_{\mathbf{p}_i} E_k(\psi) = & \int_{\Omega} \delta(\phi_0 \circ \psi) \left\langle \nabla \phi_0 \circ \psi \; , \; \frac{\partial \psi}{\partial p_i} \right\rangle \quad (T_k \circ \psi \; . \; div(\mathbf{f}_k) \; + \; \langle \boldsymbol{\nabla}(T_k \circ \psi), \mathbf{f}_k \rangle) \\ & + \int_{\Omega} H(\phi_0 \circ \psi) \left(\left\langle \boldsymbol{\nabla}(T_k \circ \psi) \; , \frac{\partial \psi}{\partial \mathbf{p}_i} \right\rangle div(\mathbf{f}_k) + \left\langle D^2(T_k \circ \psi) \; . \frac{\partial \psi}{\partial \mathbf{p}_i}, \mathbf{f}_k \right\rangle \right) \end{aligned}$$

Here the factor $\delta(\phi_0 \circ \psi)$ gives a small support to the first integrand, therefore computations are done only on the zero level-set of $\phi_0 \circ \psi$ instead of the whole volume (we refer the reader to [12] for details on implementation). The second integrand is however defined over the whole volume represented by $H(\phi_0 \circ \psi)$ and thus potentially represent a computational burden. Yet we point out that its dependence on the derivatives of T allows to reduce this overhead: If T has a sparse gradient (*e.g* is a piecewise-continuous function), these terms can also be computed from a small number of contributions.

Relationship with atlas deformation methods Tagged template deformation is closely related to atlas deformation method since the segmentation is obtained by registering a template. The great benefits of such methods are their robustness as well as topology preservation [16]. Our proposed framework yields the same advantages but the difference is that our tagged template is defined in a piecewise manner. This yields a much more efficient algorithm since image-based forces within the gradient computation are only computed on surfaces instead of the whole volume: our C++ implementation currently allows a segmentation of a 3D image in only a few seconds on a standard laptop. It is also more flexible due to the possibility of choosing the image features $(\mathbf{f}_k)_k$ which makes it applicable to non-standardized modalities such as ultrasound.

4 Learning a tagged model from a database

In this section we infer the tags from an annotated database and a set of possible features $(\mathbf{f}_k)$. We assume that we have a set of images $(I_n)_{n=1\dots N}$ and their associated set of features $(\mathbf{f}_{n,k})_{k=1\dots K}$. As a pre-processing step, all images are registered to the model referential Ω_0; the features will be transported by the same transformations $(\psi_n)_{n=1\dots N}$. Such registrations do not need to be precise everywhere but in the neighborhood of the zero level-set of ϕ_0.

A feature $\mathbf{f}_k$ is significant at point $\mathbf{x}$ (*i.e.* has a high tag absolute value) if it is locally in agreement with the ground truth (*i.e.* it is aligned with the normal of the ground truth) across the whole training set. The function S_k defined below quantifies this significance:

$$S_k(\mathbf{x}) = \frac{1}{N}\sum_{n=1}^{N}\left\langle \frac{\boldsymbol{\nabla}\phi_0(\mathbf{x})}{|\boldsymbol{\nabla}\phi_0(\mathbf{x})|},\ \mathbf{f}_{n,k}\circ\psi_n(\mathbf{x})\right\rangle \tag{4}$$

We are mainly interested in the values of S on the zero level-set of ϕ_0. At such points, $\frac{\boldsymbol{\nabla}\phi_0(\mathbf{x})}{|\boldsymbol{\nabla}\phi_0(\mathbf{x})|}$ represents the inward unit normal of the hypersurface represented by ϕ_0. To better understand (4), let us consider the case $\mathbf{f}_{n,k} = \boldsymbol{\nabla} I_n$. If the point $\mathbf{x}$ belongs to an edge of the image I_n that follows the boundary of the model, then the image gradient $\boldsymbol{\nabla} I_n(\mathbf{x})$ will be collinear to the normal and their scalar product will be high (in absolute value). Therefore $S_k(\mathbf{x})$ will have a large magnitude where there is a consistent edge across the images of the database (positive for bright-to-dark edges and vice versa). Conversely, at points where the interior of the model is sometimes brighter and sometimes darker than its exterior (*i.e.* image edges that are not reliable), S_k will be close to zero. It will also vanish when the model boundary crosses a perpendicular edge. While S_k may seem as a good tag function candidates, their gradient has no reason to be sparse. As mentioned in Section 3, the efficiency of our method is directly dependent on this sparsity. We therefore rather define the tags as

$$T_k^* = \arg\min_T \int_{\Omega_0}\left(\frac{1}{2}\|T(\mathbf{x}) - S_k(\mathbf{x})\|^2 + \nu\|\boldsymbol{\nabla} T_k(\mathbf{x})\|\right)d\mathbf{x} \tag{5}$$

which is the usually called a *total-variation regularization* of S. Indeed the L_1-norm of the gradient has the interesting property of favoring piecewise-constant functions. Problem (5) is solved with the method described in [2]. Results of such a process will be given hereafter for myocardium in ultrasound images.

5 Application to myocardium segmentation in US images

Ultrasound imaging is widely used to diagnose and understand cardiac heart diseases. Myocardium segmentation in such images is thus an important field of research. However, because of its complex appearance, it is also very challenging and only interactive methods have been proposed so far, such as [7].

Our dataset is composed of 42 images coming from 14 subjects (both healthy volunteers and patients). The considered images are 2D long-axis, taken from a 4-chamber view of the heart, with a spatial resolution of 0.5 mm $\times$ 0.5 mm.

Learned tags for myocardium in US images The result T^* of the tag learning process is presented in Figure 2. Here the mean model is defined as a closed curve including both the internal and external contours, and was estimated via the shape learning process described in [13]. For the sake of simplicity and clarity, we only used one feature which is the smoothed image gradient $\mathbf{f} = \boldsymbol{\nabla} I_\sigma$. The most significant and consistent edges (in the region of the septum for instance) are detected. The pixels at the apex are also clustered into a sub-region, but with lower confidence. Furthermore, we notice a tag inversion between the inner and outer boundary at the apex and the bottom-right of the model. Others areas (*e.g* at each part of the apex) are completely neglected: the segmentation will solely be interpolated by the shape prior without taking the image into account.

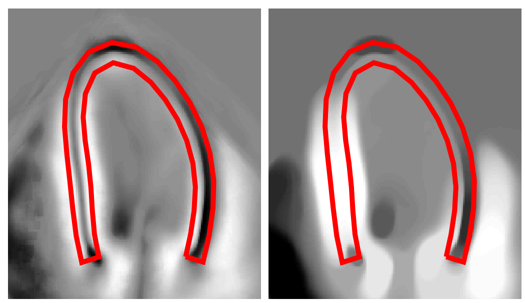


Fig. 2. Tags learning for myocardium segmentation in US in the referential of the mean model shown in red. Black represents -1, grey 0 and white 1. (Left) Mean scalar product map S. (Right) Tags T^* obtained after total-variation regularization of S. Different zones are detected, as expected from the images in Figure 1.

Evaluation of the myocardium segmentation A clinician clicked on 3 points in each image within the myocardium: one at the apex and one at each valve. The three points are used to initialize the position and size of the mean model (see left image in Figure 3). Besides, it is naturally possible with implicit template deformation to indicate some points that should lie inside or outside the segmentation (see [12]). We therefore also use these points as inner constraints.

The validation has been performed with a leave-one-patient-out strategy. We evaluate our approach by computing for each image (i) the mean absolute distance, (ii) the maximum distance and (iii) the Dice coefficient, between the segmentation and the ground truth. The results are summarized in Table 1. For comparison purposes, we also indicate the scores obtained with the initial contour (placed with 3 points) and with the baseline method (constant positive tags). All reported metrics are significantly better with the new tagged template deformation method (p-value ≤ 0.0001 with a Wilcoxon signed rank test).

The distance-based metrics (3.15 for the mean absolute distance and 9.76 for the maximum distance) are slightly higher than [7], namely 1.18 and 4.41.

	Mean distance	Max distance	Dice coeff.
Initialization	4.87 mm (1.53)	15.17 mm (3.65)	0.59 (0.17)
Standard temp. def.	5.10 mm (0.49)	16.65 mm (3.48)	0.59 (0.08)
Tagged temp. def.	**3.15 mm** (0.88)	**9.76 mm** (2.29)	**0.77** (0.06)

Table 1. Results for the myocardium segmentation averaged over the 42 images, reported in mm. Figures in brackets indicate standard-deviations.

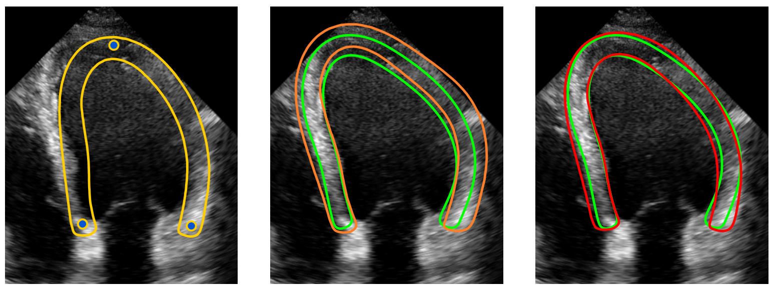


Fig. 3. Myocardium segmentation in US images compared to the ground truth in green. From left to right: Initialization of the mean model with 3 points, segmentation with the standard template deformation approach (orange) and the new model (red).

However, their method needed 6 points *on the contour*, while we only need 3 points *inside* the myocardium. Besides, their validation database was solely composed of healthy subjects. Images from patients with pathologies are more difficult to segment since the learning is less reliable.

We also show visually the benefits of the tagged template over the baseline algorithm in Figure 3. As the standard template can only take into account gradient information in a single direction, it may segment correctly the septum but then cannot capture the correct boundary at the apex. Furthermore, it takes too much into account the image information at some points of the model (typically on both sides of the apex). The results might then be even worse than the initialization, as shows Table 1. Conversely, the segmentation obtained with the tagged model has a better behavior and is much closer to the ground truth.

6 Conclusion

By introducing tagged models, we have greatly enriched the prior information that is exploited in the promising framework of implicit template deformation. This extension widens the scope of potential clinical applications of this segmentation method; we indeed showed that major improvements were achieved over the standard approach in the context of myocardium segmentation in US images. Note however that this new framework is completely generic and valid both in 2D and 3D thanks to the implicit representation of shapes. It therefore paves the way for multi-organ segmentation: several organs can be represented by an implicit function, each of them being tagged in order to be attached to a dedicated image-based energy. It therefore represents a further step towards atlas-based methods, with a much more efficient and flexible approach though.

References

1. Arfken, G.B., Weber, H.J., Ruby, L.: Mathematical methods for physicists, vol. 6. Academic press NY (1985)
2. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. JMIV 40(1), 120–45 (2011)
3. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. Pattern Analysis and Machine Intelligence, IEEE Transactions on 23(6), 681–685 (2001)
4. Cremers, D., Rousson, M., Deriche, R.: A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. International journal of computer vision 72(2), 195–215 (2007)
5. Cremers, D., Sochen, N., Schnörr, C.: Towards recognition-based variational segmentation using shape priors and dynamic labeling. In: Scale Space Methods in Computer Vision. pp. 388–400. Springer (2003)
6. 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: MICCAI, LNCS, vol. 7512, pp. 66–74. Springer (2012)
7. Dietenbeck, T., Alessandrini, M., Barbosa, D., D'hooge, J., Friboulet, D., Bernard, O.: Detection of the whole myocardium in 2D-echocardiography for multiple orientations using a geometrically constrained level-set. MIA 16(2), 386–401 (2012)
8. Gauriau, R., et al.: A generic, robust and fully-automatic workflow for 3D CT liver segmentation. In: Abdominal Imaging. Computation and Clinical Applications, LNCS, vol. 8198, pp. 241–50. Springer (2013)
9. Huang, X., Li, Z., Metaxas, D.: Learning coupled prior shape and appearance models for segmentation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004, pp. 60–69. Springer (2004)
10. Kimmel, R., Bruckstein, A.M.: Regularized Laplacian zero crossings as optimal edge integrators. International Journal of Computer Vision 53(3), 225–243 (2003)
11. Leventon, M.E., Faugeras, O., Grimson, W.E.L., Wells III, W.M.: Level set based segmentation with intensity and curvature priors. In: Workshop on Mathematical Methods in Biomedical Image Analysis. pp. 4–11. IEEE (2000)
12. Mory, B., Somphone, O., Prevost, R., Ardon, R.: Real-time 3D image segmentation by user-constrained template deformation. In: MICCAI, LNCS, vol. 7510, pp. 561–8. Springer (2012)
13. Prevost, R., Cuingnet, R., Mory, B., Cohen, L.D., Ardon, R.: Incorporating Shape Variability in Image Segmentation by Implicit Template Deformation. In: MICCAI, LNCS, vol. 8151, pp. 82–89. Springer (2013)
14. Prevost, R., Mory, B., Correas, J.M., Cohen, L.D., Ardon, R.: Kidney detection and real-time segmentation in 3D contrast-enhanced ultrasound images. In: Proceedings of IEEE ISBI 2012. pp. 1559–62 (2012)
15. Saddi, K., Chefd'hotel, C., Rousson, M., Cheriet, F.: Region-based segmentation via non-rigid template matching. Proceedings of ICCV 0, 1–7 (2007)
16. Wolz, R., Chu, C., Misawa, K., Mori, K., Rueckert, D.: Multi-organ abdominal ct segmentation using hierarchically weighted subject-specific atlases. In: MICCAI, LNCS, vol. 7510, pp. 10–17. Springer (2012)