



HAL
open science

Shape prior based image segmentation using manifold learning

Arturo Mendoza Quispe, Caroline Petitjean

► **To cite this version:**

Arturo Mendoza Quispe, Caroline Petitjean. Shape prior based image segmentation using manifold learning. IEEE Image Processing, Tools and Applications, Nov 2015, Orléans, France. hal-01216823

HAL Id: hal-01216823

<https://hal.science/hal-01216823>

Submitted on 17 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Shape prior based image segmentation using manifold learning

Arturo Mendoza Quispe and Caroline Petitjean

Université de Rouen, LITIS EA 4108

76801 Saint-Etienne-du-Rouvray, France

e-mail: arturo.mendoza.quispe@gmail.com, caroline.petitjean@univ-rouen.fr

Abstract—In image segmentation, the shape knowledge of the object may be used to guide the segmentation process. From a training set of representative shapes, a statistical model can be constructed and used to constrain the segmentation results. The shape space is usually constructed with tools such as principal component analysis (PCA). However the main assumption of PCA that shapes lie a linear space might not hold for real world shape sets. Thus manifold learning techniques have been developed, such as Laplacian Eigenmaps and Diffusion Maps. Recently a framework for image segmentation based on non linear shape modeling has been proposed; still some challenges remain, such as the so-called out-of-sample extension and the pre-image problems. This paper presents such a framework relying on Diffusion Maps to encode the shape variations of the training set, and graph cut for the segmentation part. Finally, some segmentation results are shown on a medical imaging application.

Keywords—Image segmentation, Shape prior based segmentation, Shape modeling, Manifold learning.

I. INTRODUCTION

In many segmentation applications, the shape of the object to be segmented is known a priori, up to some variability. Image segmentation can benefit from the use of information regarding the shape or the gray levels of the object to be segmented, to increase robustness against noise and occlusions. The shape knowledge may be incorporated by first constructing a statistical shape model from training cases, and then constraining the segmentation results to be within the learned shape space. The shape space can be constructed with statistical analysis, e.g. with tools such as principal component analysis (PCA). The seminal work of Active Shape Models [1], based on point distribution modeling (PDM), established much of the framework in this area. However, the parameterized representation of the PDM and manual positioning of landmarks drastically limits its usage on some applications. In many subsequent research works, a different, implicit representation of shapes, namely the signed distance map, has been used and exploited successfully, in particular in variational frameworks, be it the level sets [2], [3] or the graph cuts [4]. The use of a statistical shape model often imposes an iterative process, which alternates between shape model registration (estimation of pose parameters) and graph-cut based segmentation. For example, in [4], the method consists in alternatively searching for the PCA, GMM and pose parameters using gradient descent (maximisation step) and segmenting by graph cut using the current shape from the

PCA (estimation step), as in an EM framework.

A limitation of these approaches is that statistical analysis of the shape dataset is made via a PCA. Yet the main assumption of PCA is that shapes lie in a linear space and follow a multivariate Gaussian distribution, an hypothesis which might not hold for real world shape sets. Thus non linear or manifold learning techniques have been developed, among which Multi-Dimensional Scaling (MDS) [5], Isomap [6], Local Linear Embedding (LLE) [7], Laplacian Eigenmaps [8], and Diffusion Maps [9].

In image analysis, these nonlinear methods have shown their potential in facial recognition, hyperspectral image classification, gait recognition, hand-written character recognition [10] and several medical imaging tasks such as segmentation and registration [11]. For the segmentation task, non linear shape statistics were first introduced with kernel PCA [12], [13] and pursued in the following years with Laplacian Eigenmaps [14] and Diffusion Maps [15], [16]. Etyngier et al. in particular were the first to introduce non-linear shape priors into a deformable model framework [15]. In their work, Diffusion Maps have been proven capable of representing the intrinsic non-linearity found in many datasets. Diffusion Maps are known to be robust to noise and computationally inexpensive [9].

Inspired by [15] and [16], we investigate the use of a non-linear shape prior in a graph cut framework. In our paper, we model a category of shapes as a smooth finite-dimensional sub-manifold of the infinite-dimensional shapes space, termed the shape prior manifold, using a manifold learning techniques, the Diffusion Maps [9], [17], [18], to approximate a mapping from the original shape space into a low-dimensional space. Advantageously this mapping is an isometry from the original shape space. We include the shape prior in the segmentation process, through a non-linear energy term designed to attract the shape towards its projection onto the manifold. Doing so requires to be able to project the segmentation result, i.e. a new shape (different from the training set) onto the manifold and, vice-versa, to extract the shape representation of any point in the manifold. This issues are known as the out-of-sample extension and the pre-image problem [19], [20], respectively; and like most of the manifold learning techniques, Diffusion Maps are not equipped with such capabilities.

The remainder of this document is organized as follows. Section II introduces the necessary background in manifold

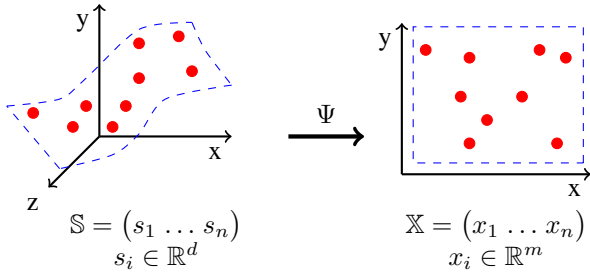


Fig. 1: Manifold learning framework exemplified with $d = 3$ and $m = 2$. Each d -dimensional shape s_i has an m -dimensional counterpart x_i in the manifold.

learning. Section III describes the method for this nonlinear shape-prior based segmentation framework. Results are reported in section IV, and conclusions in Section V.

II. LEARNING THE SHAPE PRIOR MANIFOLD

Manifold learning is the process of recovering the underlying low dimensional structure for data sets which lie on non-linear manifolds in a high-dimensional space [21]. It is closely related to the notion of dimensionality reduction. If the original data lies in a subspace of \mathbb{R}^d , then manifold learning aims at finding coordinates for each point in a smaller m -dimensional space, \mathbb{R}^m with $m \ll d$. The mapping from \mathbb{R}^d to \mathbb{R}^m is denoted by Ψ .

In our work, each data point is a shape, that is defined as a simple bounded, closed and non-intersecting surface. The shape is represented by its signed distance map to the contour [3]. The space of such shapes is denoted by $\mathbb{S} \subset \mathbb{R}^d$ (where d is the number of pixels of the shape binary image or distance map) and the shape prior manifold after the embedding is denoted $\mathbb{X} \subset \mathbb{R}^m$. Figure 1 is an illustration of manifold learning framework.

In this section, we will review how to construct an embedded space using Diffusion Maps, and how to handle the out-of-sample and the pre-image problems.

A. Diffusion maps

The first step is to view the data points $\mathbb{S} = \{s_1, \dots, s_N\}$ as the nodes of symmetric graph in which any two nodes s_i and s_j are connected by an edge. The strength of this connection is measured by a non-negative weight $w(s_i, s_j)$ that reflects their similarity. A common weighting function is the Gaussian kernel:

$$w(s_i, s_j) = \exp(-(s_i - s_j)^2 / (2\sigma^2))$$

with σ a parameter usually chosen so that it preserves at least one connection in each node. It is to be noted that this notion of similarity is completely application-driven.

Following the classical construction of the normalized graph Laplacian, and thanks to the properties of the induced Markov Chain P , we can create a random walk on the data set by

forming the following kernel of probabilities of transition in t time steps:

$$p_t(s_i, s_j) = \left(\frac{w(s_i, s_j)}{d(s_i)} \right)^t$$

where $d(s_i) = \sum_{j=1}^n w(s_i, s_j)$ is the degree of node s_i . In [8], it is shown that when data is uniformly sampled from a low-dimensional manifold, the first largest eigenvectors of the matrix of transition of this Markov chain are the discrete approximation of the eigenfunctions of the Laplace-Baltrami operator on the manifold, thus making it possible to perform the following eigendecomposition:

$$p_t(s_i, s_j) = \sum_{l \geq 0} \lambda_l^t \psi_l(s_i) \phi_l(s_j)$$

where $\{\lambda_l\}$ are eigenvalues of P , and ψ_l and ϕ_l are the corresponding left and right eigenvectors.

Finally, unifying the notions from Markov chains and potential theory, the diffusion distance is defined as:

$$D_t^2(s_i, s_j) = \sum_{l \geq 0} \lambda_l^{2t} (\psi_l(s_i) - \psi_l(s_j))^2$$

Let us retain only m terms, we can now construct the diffusion map:

$$x = \Psi_t(s) = (\lambda_1^t \psi_1(s) \lambda_2^t \psi_2(s) \dots \lambda_m^t \psi_m(s))^T$$

The eigenvectors are thus used as the new coordinates for the data, in the reduced m -dimensional space. It is interesting to note that the scaling parameter σ of the Gaussian kernel and the number of random walks t allow the Diffusion Maps to act either as a local or a global method by trying to preserve the respective parameterization of the data at each scale, unlike many other manifold learning algorithms. Here we use $t = 1$. For a more algorithmic description of the framework, please refer to [18].

B. Out-of-sample extension

The embedding is computed with the points of the training set, but does not allow to project points which are not part the training set, i.e. which are "out-of-sample". This case arises frequently; in particular, for our situation, the segmentation result will have to be projected in the reduced space, and as it does not belong to the original shape set, an out-of-sample approximation is needed.

Various solutions have been proposed for out-of-sample [9], [22]. The Nyström extension, a method used to speed up kernel methods computations, is one of the most frequently used [19]. It is based on the assumption that the similarity measure used to compute the embedding may be expressed as a kernel function, which is the case of Diffusion Maps. The Nyström extension is based on the principle of convergence of eigenvectors and eigenvalues as the number of examples increases. In [19], Bengio et al. show that for any point

$s^* \in \mathbb{R}^d$, the j -th coordinate of the mapped point $x^* \in \mathbb{R}^m$ can be approximated by [19]:

$$x_j^* = \hat{\Psi}_j(s^*) = \frac{1}{\lambda_j} \sum_{i=1}^n \psi_j(s_i) k(s^*, s_i)$$

Thus, the new point s^* is mapped as a weighted linear combination of the corresponding maps for the training points.

C. Pre-image problem

The pre-image of any point $\hat{x} \in \mathbb{R}^m$ is a point $\hat{s} \in \mathbb{R}^d$ in the original data space so that $\Psi(\hat{s}) = \hat{x}$ (i.e. $\hat{s} = \Psi^{-1}(\hat{x})$). This is an ill-posed problem [20] since such a pre-image might not exist; therefore, we settle for an approximate solution. Following [15], we construct the pre-image based on the neighboring training samples and assume that the point falls into the convex-hull of training samples in the reduced space.

We compute a Delaunay triangulation in the reduced space and find the neighbors of \hat{x} as the $m + 1$ vertices of the m -dimensional triangle to which \hat{x} belongs to. Next, we define the pre-image of \hat{s} as the weighted mean shape of its neighbors $\mathcal{N} = \{s_0, \dots, s_m\}$ so that:

$$\hat{s} = \operatorname{argmin}_{\theta_i, s} \sum_{i=0}^m \theta_i d(s^*, s_i) \quad (1)$$

with $\sum_{i=0}^m \theta_i = 1$ and $\theta_i \geq 0, \forall i = 0, \dots, m$ and the distance between shapes is defined as:

$$d(s_i, s_j) = \sum_{\text{pixels } p} (H(s_i(p)) - H(s_j(p)))^2 \quad (2)$$

where $H(s)$ is the Heaviside function [16]. The pre-image is thus computed as:

$$\hat{s} = \sum_{i=0}^m \theta_i s_i \quad (3)$$

where the coefficients $\{\theta_i\}$ are initialized as the barycentric coefficients of \hat{s} with respect to its neighbors and updated with gradient descent: $\theta_i = \theta_i + \lambda d(s^*, s_i)$.

III. SEGMENTATION FRAMEWORK

The segmentation framework is an iterative process that consists of two stages applied consecutively. First, a segmentation step with the shape prior information embedded in a non-linear energy term designed to attract the shape towards its projection onto the manifold; followed by a shape prior update stage where the current estimate of the segmentation is taken into account to update the shape prior through a process of manifold traversal. The method is described in Figure 2 and detailed in Algorithm 1.

A. Segmentation with the shape prior

This phase incorporates the shape prior term into an image segmentation algorithm. As specified in the Introduction, variational approaches are flexible enough to allow to take into

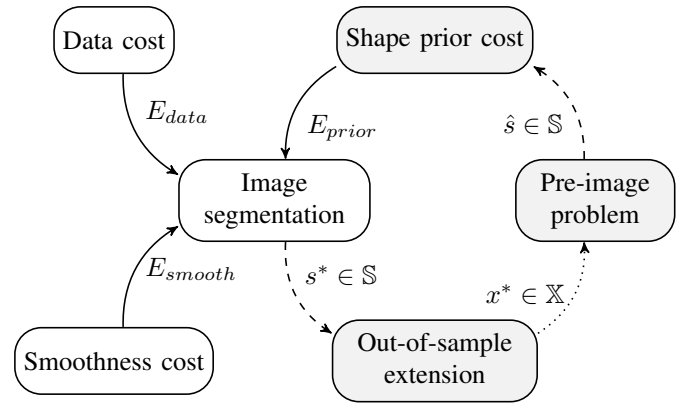


Fig. 2: Overview of the manifold-based shape prior segmentation framework

Algorithm 1 Prior based segmentation framework

Require: A set of n aligned training shapes s_i , an image I to be segmented

Ensure: A shape that segments image I

- 1: (A) **Shape manifold**
 - 2: Compute the shape manifold with the n training shapes
 - 3: (B) **Image segmentation**
 - 4: Initial segmentation yields shape s_1^*
 - 5: **for** $j = 1$ to $NbIterations$ **do**
 - 6: (i) Out-of-sample projection $x^* = \hat{\Psi}(s_j^*)$
 - 7: (ii) Find $m + 1$ nearest neighbours of x^* with Delaunay triangulation
 - 8: (ii) Define new shape as $\hat{s} = \sum_{i=0}^m \theta_i s_i$ with θ_i computed with gradient descent: $\theta_i = \theta_i + \lambda d(s^*, s_i)$
 - 9: (iv) Rigidly align \hat{s} on s_j^*
 - 10: (v) Image segmentation by graph cut minimization of $E(I, f)$ provides a segmentation s_{j+1}^*
 - 11: (vi) $s_j^* \leftarrow s_{j+1}^*$
 - 12: **end for**
 - 13: **return** shape s_j^*
-

account a shape prior. We thus use the Markov Random Fields framework with an energy minimization via Graph Cuts [23].

Let I denote the image to be segmented. The segmentation problem is formulated in terms of energy minimization; thus, we seek the labeling f with binary values O (object) or B (background) that minimizes the energy:

$$E(I, f) = E_{smooth}(I, f) + (1 - \lambda)E_{data}(I, f) + \lambda E_{prior}(f, \hat{s}) \quad (4)$$

where E_{smooth} measures the extent to which f is not piecewise smooth, E_{data} measures the disagreement between f and the observed data and E_{prior} is the shape prior term. Let us now review the design of each term. As smoothness and data terms are standard choices and not our main focus we will simply describe them. We used a low-pass filtering on the

image for E_{smooth} :

$$E_{smooth}(I, f) = \sum_{f_p \neq f_q, p, q} \exp\left(-\frac{(I_p - I_q)^2}{2\sigma_n^2}\right) \cdot \frac{1}{dist(p, q)} \quad (5)$$

where I_p and I_q are the gray levels of pixels p and q , $dist(p, q)$ the Euclidean distance between p and q , and σ_n a constant usually related to the image noise. For the data-driven energy term E_{data} , we use models of Gaussian, one for the object and one for the background.

Regarding the shape prior term, the shape prior \hat{s} is first transformed into a probability atlas M [16] over all pixels p :

$$M(p) = \begin{cases} 1 & \text{if } \hat{s}(p) \leq 0 \\ e^{-\hat{s}(p)/\gamma_f(p)} & \text{if } \hat{s}(p) > 0 \end{cases} \quad (6)$$

where γ_f controls the spread of the influence of the prior; we propose to use different values for each one of the labellings $f = \{O, B\}$. Then, the prior energy term is defined as:

$$E_{prior}(f, \hat{s}) = - \sum_p \begin{cases} \log M(T(p)) & \text{if } f(p) = O \\ \log(1 - M(T(p))) & \text{if } f(p) = B \end{cases} \quad (7)$$

where T is an affine transform that aligns the shape prior model \hat{s} with the current segmentation estimate s^* .

B. Update of the shape prior

In this stage, the current shape prior \hat{s} based on the current estimation s^* is updated so that the prior resembles the target. This traversal through the manifold consists also in two phases. First, the current estimation s^* is embedded onto the feature space using the out-of-sample extension in order to obtain $x^* = \hat{\Psi}(s^*)$. Then, the pre-image is computed from this point in order to get the new shape prior $\hat{s} = \Psi^{-1}(x^*)$. This process should provide a new shape that increasingly resembles our best guess while also compensating for its inaccuracies.

IV. IMPLEMENTATION AND EXPERIMENTAL RESULTS

Organ segmentation in medical imaging is a field that can benefit from the use of shape prior [24]. In this paper, we use cardiac magnetic resonance images, where the aim is to segment the right ventricle (RV), a crescent shape object illustrated in Figure 3. For each of these images, the manual segmentation has been drawn by an expert. The binary maps obtained from the manually drawn contours are transformed into signed distance maps. We have a total of $n = 19$ images from different patients and their associated distance maps of dimension $d = 256 \times 216 = 55296$. All images are rigidly registered beforehand. $n - 1$ distance maps are used to construct the manifold. The remaining image is the one that will be segmented.

Before computing the embedding, one must fix the (reduced) dimension of the embedded space. In this paper, and as is commonly done, we fix it to the intrinsic dimension of the data. There are several ways to estimate the intrinsic dimension

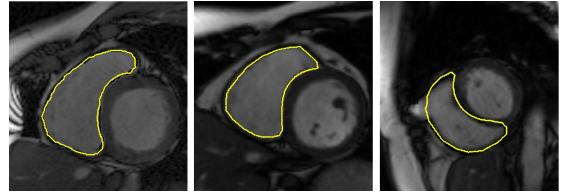


Fig. 3: Cardiac magnetic resonance images and manual segmentation of the RV in yellow

of the data [21]. The maximum likelihood estimator (MLE), which exploits neighborhood information, is one common way to do so [25]. For our data, the intrinsic dimension has been found to be equal to 3.

For the manifold construction using Diffusion Maps, we employed the MATLAB Toolbox for Dimensionality Reduction [26] and we implemented the out-of-sample extensions and pre-image calculations as described in Sections II-B and II-C since they are not available in the toolbox for Diffusion Maps. The kernel parameter for the smoothness cost (Eq. 5), σ , is fixed to the mean distance between all points in the dataset. Other parameters of the method include the λ weighting in Eq. 4. In [18], the segmentation framework is performed without any shape prior term during the first iterations of the algorithm since a robust alignment may not be possible. We propose a more dynamic approach by increasing the shape prior influence at each iteration: the value of λ is initialized at $1/8$, updated with a 5% increase at each step and with a maximum value of $3/8$, after which the weights remain unchanged. We use the MATLAB Wrapper for Graph Cuts [27], for optimizing Eq. 4 in an iterative manner as described previously. In an effort to foster research in manifold based shape prior segmentation, we make our implementation code available to the scientific community¹.

A preliminary result is the embedding of the data, that is obtained from Diffusion Maps, illustrated in Figure 4 using the two first embedded coordinates. This gives us an idea of the main variability in the training set. One can note that the horizontal axis reports information regarding the thickness of the RV shape, whereas the vertical axis focuses on the bottom left tail of the shape, which is more or less sharp.

Figure 5 shows a segmentation performed with graph cuts where no shape prior is used, which emphasizes the need for such a prior.

Regarding the whole shape prior based segmentation framework, the initialization of the segmentation can be performed from a manual rough segmentation or a random positioning of the mean shape. From this first estimate, seeds – some for the background and some for the object – are randomly picked out in order to extract the relevant information needed to build the Gaussian models for the data cost energy term; at the same time, the smoothness cost is built. The shape prior is initialized to the mean of the shape set and placed over the

¹<http://www.litislabs.eu/Members/cpetitjean/shapepriorsegm/>

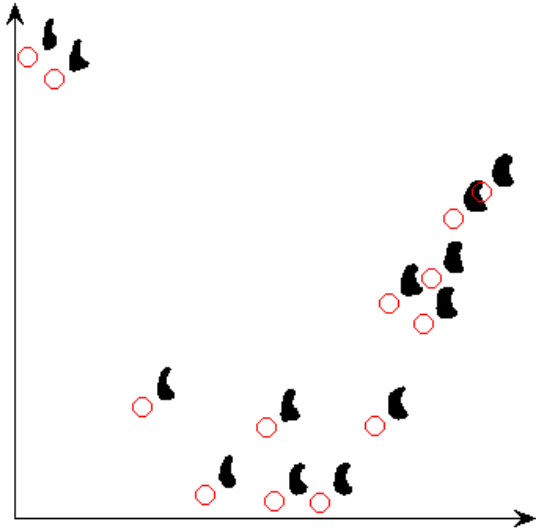


Fig. 4: Embedding example in 2D using Diffusion Maps

image accordingly. Finally, graph cuts are used to minimize the total energy of Eq. 4. Figure 6 reports an example of a traversal in the manifold restricted to 2 dimensions, with the evolution of the shape of the segmentation result. Segmentation results are provided in Figure 7, where the initial guess was manually and randomly placed; they are encouraging although room for improvement is left.

V. CONCLUSIONS

In this work we have shown the efficiency and the potential of an image segmentation framework with a nonlinear statistical shape prior, with some results obtained on a challenging image segmentation problem as a proof of concept. The use of a nonlinear shape prior requires the construction of a shape manifold and the design of specific tools to handle the out-of-sample extension, and its corollary problem, the pre-image computation. In the future we plan to pursue our investigations by considering other manifold learning techniques. We used Diffusion Maps here but some other algorithms such as Laplacian Eigenmaps could be considered as well. In the same spirit, other algorithms for out-of-sample apart from Nyström extension exist and their influence on the segmentation could be fruitfully investigated. Another point of improvement concerns the manifold traversal and the constrain on the resulting shape, in this work regarded as a simple linear combination of its neighbors.

ACKNOWLEDGEMENTS

This project is co-financed by the European Union with the European regional development fund (ERDF) and by the Haute-Normandie Regional Council.

REFERENCES

- [1] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer vision and image understanding*, vol. 61, no. 1, pp. 38–59, 1995.

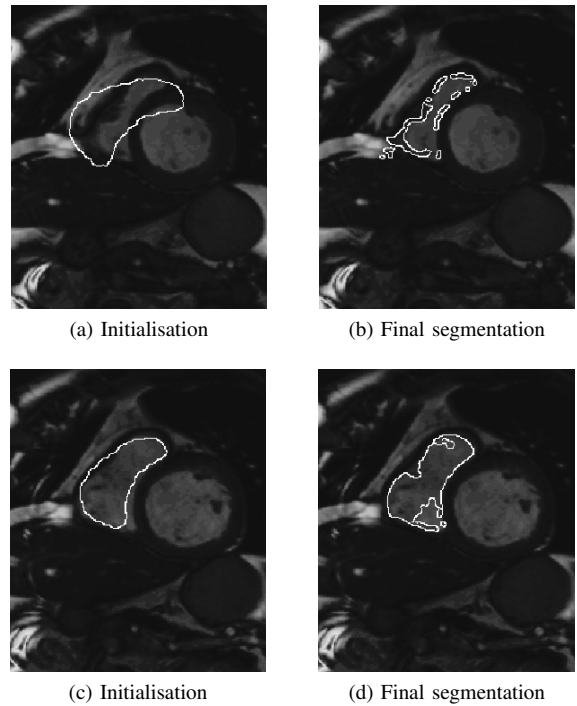


Fig. 5: Example of segmentation without using a shape prior term, i.e. $\lambda = 0$

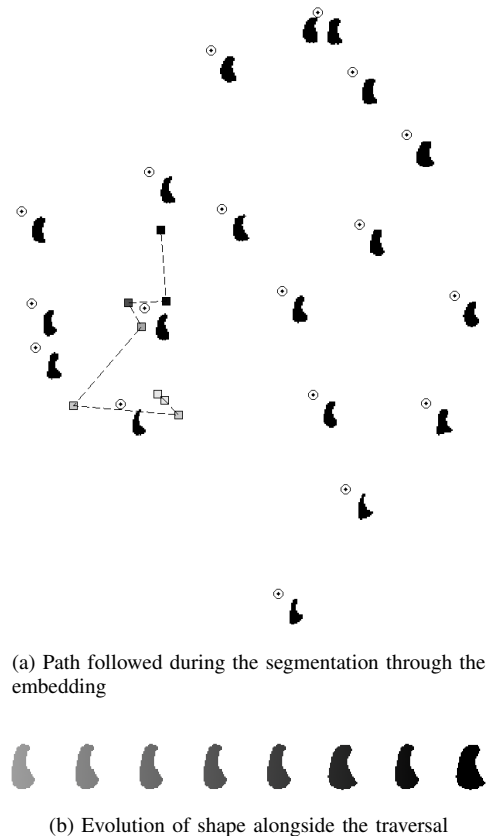


Fig. 6: Evolution of the segmented shape and path in the embedding along the segmentation process

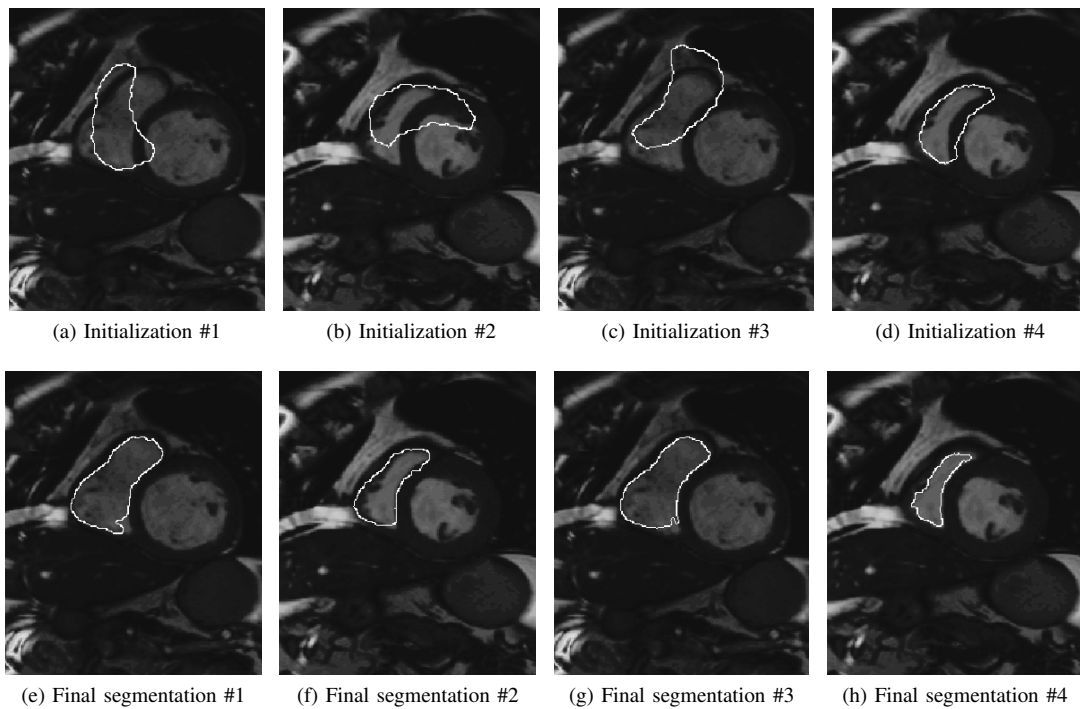


Fig. 7: Segmentation results on different images with varying initializations

- [2] M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *CVPR*, vol. 1. IEEE, 2000, pp. 316–323.
- [3] A. Tsai, A. Yezzi Jr, W. Wells, C. Tempny, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *Medical Imaging, IEEE Transactions on*, vol. 22, no. 2, pp. 137–154, 2003.
- [4] J. Zhu and R. Zabih, "Graph cuts segmentation with statistical shape priors for medical images," in *IEEE Southwest Symposium on Image Analysis and Interpretation*, 2008.
- [5] I. Borg and P. J. Groenen, *Modern multidimensional scaling: Theory and applications*. Springer Science & Business Media, 2005.
- [6] J. B. Tenenbaum, V. De Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.
- [7] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–26, 2000.
- [8] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, vol. 15, no. 6, pp. 1373–1396, 2003.
- [9] S. S. Lafon, "Diffusion maps and geometric harmonics," Ph.D. dissertation, Yale University, 2004.
- [10] R. Pless and R. Suenir, "A survey of manifold learning for images," *IPSI Transactions on Computer Vision and Applications*, vol. 1, pp. 83–94, 2009.
- [11] P. Aljabar, R. Wolz, and D. Rueckert, "Manifold learning for medical image registration, segmentation, and classification," *Machine Learning in Computer-Aided Diagnosis: Medical Imaging Intelligence and Analysis*, pp. 351–372, 2012.
- [12] D. Cremers, T. Kohlberger, and C. Schnörr, "Nonlinear shape statistics in mumpford-shah based segmentation," in *In European Conference on Computer Vision*. Springer, 2002, pp. 93–108.
- [13] Y. Rathi, S. Dambreville, and A. Tannenbaum, "Statistical shape analysis using kernel pca," in *Electronic Imaging 2006*. International Society for Optics and Photonics, 2006, pp. 60 641B–60 641B.
- [14] P. Yan, W. Zhang, B. Turkbey, P. L. Choyke, and X. Li, "Global structure constrained local shape prior estimation for medical image segmentation," *Computer Vision and Image Understanding*, vol. 117, no. 9, pp. 1017 – 1026, 2013.
- [15] P. Etyngier, F. Ségonne, and R. Keriven, "Shape priors using manifold learning techniques," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. IEEE, 2007, pp. 1–8.
- [16] O. Moolan-Feroze, M. Mirmehdi, M. Hamilton, and C. Bucciarelli-Ducci, "Segmentation of the right ventricle using diffusion maps and markov random fields," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*. Springer, 2014, pp. 682–689.
- [17] S. Lafon, Y. Keller, and R. R. Coifman, "Data fusion and multicue data matching by diffusion maps," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 11, pp. 1784–1797, 2006.
- [18] N. Thorstensen, P. Etyngier, F. Segonne, and R. Keriven, "Diffusion maps as a framework for shape modeling," *Computer Vision and Image Understanding*, vol. 115, no. 4, pp. 520–530, 2011.
- [19] Y. Bengio, J. Paiement, P. Vincent, O. Delalleau, N. Le Roux, and M. Ouimet, "Out-of-sample extensions for lle, isomap, mds, eigenmaps, and spectral clustering," *Advances in neural information processing systems*, vol. 16, pp. 177–184, 2004.
- [20] P. Arias, G. Randall, and G. Sapiro, "Connecting the out-of-sample and pre-image problems in kernel methods," in *Computer Vision and Pattern Recognition, 2007*, pp. 1–8.
- [21] H. Strange and R. Zwigglelaar, *Open Problems in Spectral Dimensionality Reduction*. SpringerBriefs in Computer Science, 2014.
- [22] K. Zhang and J. Kwok, "Density-weighted nyström method for computing large kernel eigensystems," *Neural Computation*, vol. 21, no. 1, pp. 121–146, 2009.
- [23] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [24] D. Grosgeorge, C. Petitjean, J.-N. Dacher, and S. Ruan, "Graph cut segmentation with a statistical shape model in cardiac mri," *Computer Vision and Image Understanding*, vol. 117, no. 9, pp. 1027–1035, 2013.
- [25] E. Levina and P. Bickel, "Maximum likelihood estimation of intrinsic dimension," in *Advances in Neural Information Processing Systems 17*. MIT Press, 2005, pp. 777–784.
- [26] L. Van der Maaten, "Matlab toolbox for dimensionality reduction," 2015. [Online]. Available: <http://lvdmaaten.github.io/drtoolbox/>
- [27] S. Bagon, "Matlab wrapper for graph cuts," December 2006. [Online]. Available: <http://www.wisdom.weizmann.ac.il/~bagon>