

Is a Single Energy Functional Sufficient? Adaptive Energy Functionals and Automatic Initialization

Chris McIntosh and Ghassan Hamarneh

Medical Image Analysis Lab
School of Computing Science, Simon Fraser University, Canada
{cmcintos,hamarneh}@cs.sfu.ca
<http://mial.cs.sfu.ca>

Abstract. Energy functional minimization is an increasingly popular technique for image segmentation. However, it is far too commonly applied with hand-tuned parameters and initializations that have only been validated for a few images. Fixing these parameters over a set of images assumes the same parameters are ideal for each image. We highlight the effects of varying the parameters and initialization on segmentation accuracy and propose a framework for attaining improved results using image adaptive parameters and initializations. We provide an analytical definition of optimal weights for functional terms through an examination of segmentation in the context of image manifolds, where nearby images on the manifold require similar parameters and similar initializations. Our results validate that fixed parameters are insufficient in addressing the variability in real clinical data, that similar images require similar parameters, and demonstrate how these parameters correlate with the image manifold. We present significantly improved segmentations for synthetic images and a set of 470 clinical examples.

1 Introduction

Despite its importance to medical image analysis (MIA), image segmentation remains a daunting task. Though many segmentation approaches exist, see [1] for a survey, an increasing number are relying on the minimization of objective functions. Examples include several landmark papers: from the seminal paper of Snakes for 2D segmentation [2] and other explicit deformable models [3] to implicit models [4,5], graph cuts approaches [6], and numerous variants thereof. Though seemingly different, the methods share a common ground.

Each method requires four building blocks: (*i*) an objective function whose minima provide good segmentations; (*ii*) a set of parameters including weights to balance the terms of the energy function; (*iii*) a starting state and a stopping condition; and (*iv*) a method for minimization, whether it be a local or a global solver. It is in this commonality that their problems lie. Specifically, the parameter setting, initialization, and minimization phases are well known to be problematic when the fully automated segmentation of real data sets is sought.

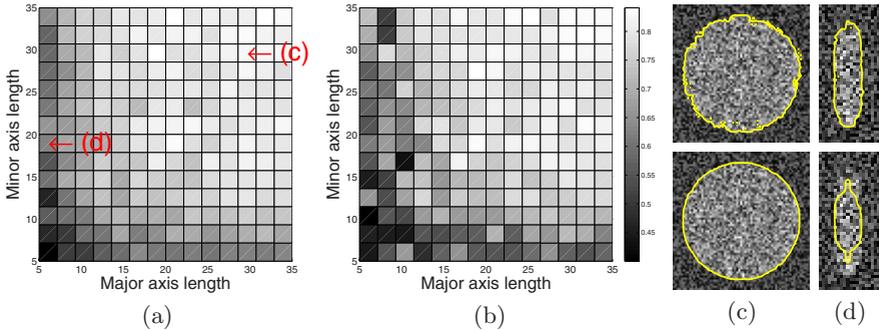


Fig. 1. Segmenting ellipses with varying lengths of major and minor axes in noisy images. (a) optimal values of w for each noisy ellipse image, *given* its known segmentation, as calculated by our algorithm (Sec. 2.3). Circles are along the diagonal. Brighter pixels imply higher values of w . (b) optimal values of w for each image, *without* knowing its segmentation, as calculated by our algorithm (Sec. 2.4). Note the similarity of (b) to (a). (a) and (b) span the intrinsic 2D space (corresponding to change in major and minor axes) of the image manifold learned from 225 ellipses. The locations of the example images in (c) and (d) are indicated with small arrows in (a). (c) Segmentation of a noisy circle image with $w = 0.347$ (optimal parameter from (d)) (top) and with its optimal parameter $w = 0.826$ (bottom). (d) segmentation of a noisy ellipse image with its optimal parameter $w = 0.347$ (top) and with $w = 0.826$ (optimal parameter from (c)) (bottom). Note the need for different weights to properly segment the different images.

Even with globally optimal graph cuts, seeds must be set and a weight that balances the regional and boundary terms in the cost function must be chosen. Classically, the underlying assumption has been made that once the function, parameters¹, initialization, and solver are in place, they can be fixed across all images. For example, consider the weight $w \in [0, 1]$ in the simplified formulation: $energy(shape, image) = w \times internal(shape) + (1 - w) \times external(shape, image)$. Segmentation is performed by finding the *shape* that minimizes the *energy* for a given *image*. As w approaches zero, the model will be attracted solely to image features. As w approaches one, the objective function will favor smooth shapes (only internal energy) with no regard to the image data. The correct w is typically chosen empirically based on the inspection of a few images in the data set. However, what if the characteristics of the images vary across the data set? For noisy images, smoothness must be emphasized, while for highly curved structures smoothness must be relaxed. Different patients have different anatomy, and may be imaged at different times with different noise. Therefore, is it realistic for a fixed set of weights to address the variability present in meaningful medical image data sets? We argue that it is not, demonstrated in Fig. 1 using a more elaborate toy example. The training data set used to provide the shape

¹ We identify two types of common parameters found in energy functionals: the weights w which balance competing terms of the functional; and the training data set used to provide the prior shape and appearance knowledge.

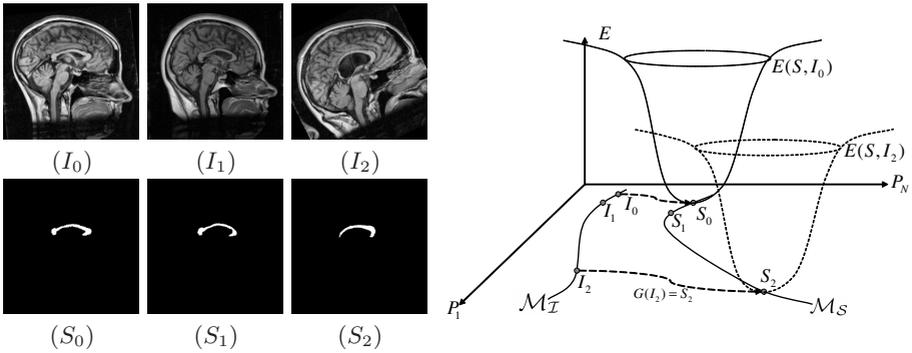


Fig. 2. Ideal energy functions $E(shape, image)$. Let graylevel image I , and shape model S , be represented as images with N pixels (P_1 to P_N). Then $\{I_1, I_2, \dots, I_N\} \in \mathbf{I}$, $\{S_1, S_2, \dots, S_N\} \in \mathbf{S}$, are sets of images and corresponding segmentations, where each element is a point in \mathbb{R}^N . Samples from \mathbf{I} and \mathbf{S} lie on manifolds \mathcal{M}_I and \mathcal{M}_S , respectively. Three (S, I) pairs (left panel), positioned on their respective manifolds (right panel). Segmenting an image is represented by a smooth mapping $G : \mathbb{R}^N \rightarrow \mathbb{R}^N$, since for each I_i there exists a corresponding segmentation S_i . Similar images, I_0 and I_1 , are nearby on \mathcal{M}_I and require similar segmentations, S_0 and S_1 . The ideal energy functional $E(S, I_0)$, is a convex functional minimized at the segmentation S_0 on \mathcal{M}_S .

and appearance priors is also typically fixed over the data set. A better approach is to limit the training data to those most similar to the correct shape. If the shape prior is, for example, a uniform distribution over a set of training shapes then restricting that set to a few likely shapes is clearly advantageous. Similarly, the initialization should also be image specific since the correct segmentation of that image is the best initialization. Though different images need different parameters similar images will likely have similar parameters.

Recent research has begun to address the idea that groups of similar images embedded in \mathbb{R}^N , where N is the number of image pixels, lie on manifolds [7]. However, there has been no work on performing segmentation in a way that respects the image manifold, i.e. method parameters reflect the data’s variability and minima of the energy functional lie on the segmentation manifold (Fig. 2).

The aforementioned problem of selecting parameters and initializing models is a serious one that remains unsolved. This work addresses the problem. The goal is to produce energy functionals with global, or at least local, minima for every $(segmentation, image)$ pair; since anything less yields a suboptimal solution for some of these pairs (Fig. 2). Our approach is to more expressly modify the energy functional with every new image to be segmented by varying its parameters.

The main contributions of this work are the following. To the best of our knowledge ours is the first work to: (1) explore the need for varying the energy functional in energy-minimizing segmentation techniques; (2) analytically derive the optimal energy functional weights, of competing energy terms, for training $(segmentation, image)$ pairs; (3) calculate optimal parameters of novel images utilizing learned image manifolds; (4) provide a general formulation that can be directly applied to any energy minimizing segmentation technique (e.g.

[4,5,8,9]) to improve its segmentation and increase its ability to generalize to larger datasets; (5) to fully automatically initialize and tune parameters optimally, demonstrated on a large ($N = 470$) data set of images; (6) to provide the automatic cropping of images prior to segmentation for a 10 to 20 times speedup.

To the best of our knowledge there are only two papers on a related topic. In [10], Gennert et al. find weights that minimize the energy functional, but do not encourage its convexity (Fig. 2), or provide a way to obtain the optimal parameters and initialization for a new image. By contrast, we use the image manifold to encourage energy functional convexity, and obtain optimal initializations and parameters for new images. In [11], Koikkalainen et al. use a nearest neighbors approach to initialize the segmentation procedure but do not define optimal parameters for the model, nor do they make use of manifold learning to calculate distances, perform interpolation, or obtain optimal parameters for the model.

2 Methods

Our approach is to use $\mathcal{M}_{\mathbf{I}}$ (Fig. 2) to automatically obtain parameters and initializations for novel images, encouraging energy functionals like those presented in Fig. 2. Hence the main concept of our paper is designing an energy functional that is more expressly a function of $\mathcal{M}_{\mathbf{I}}$. First, we formulate the segmentation problem (Sec. 2.1). Then given a set of training image and segmentation pairs we learn their respective manifolds (Sec. 2.2). With the geometry of the energy functional’s domain known, we can calculate the optimal weights, w , for each training image by minimizing a convex, quadratic objective function (Sec. 2.3). Now we assume that $w(I)$ is smooth over its domain, the space of application images, and that the mapping G , from images to segmentations, is smooth. Then for a given novel image from the same class, we identify its intrinsic coordinates on $\mathcal{M}_{\mathbf{I}}$, and assign it an optimal w and an initialization based on the optimal w ’s and segmentations of “nearby” images (Sec. 2.4).

2.1 Energy Functional and Free Parameters

The first step in our proposed framework is the identification of the form of the energy functional and the associated free parameters. Our examples use the level set framework, where the contour or surface is embedded as the zero level-set of a scalar function, $\Phi(\mathbf{x})$, defined for every point $\mathbf{x} \in \Omega$, where Ω is the domain of the image to be segmented². See [4,5,8,9] for detailed discussions, examples, and derivations. In general, level-set based energy functionals $E(\Phi, I_j, w) = \int_{\Omega} w_1 J_1(\Phi, I_j) + w_2 J_2(\Phi, I_j) + \dots + w_n J_n(\Phi, I_j) d\mathbf{x}$ for a fixed image I_j and fixed parameters w , have update equations of the form: $\Phi_t = w_1 T_1(\Phi, I_j) + \dots + w_n T_n(\Phi, I_j)$ obtained by making Φ a function of time and applying the Euler Lagrange equations, where $w = [w_1, \dots, w_n]$ are weights, and T_i is the gradient of the J_i th term. Minimizing E is then typically performed using gradient descent $\Phi(\mathbf{x}, n + 1) = \Phi(\mathbf{x}, n) + \Delta t \Phi_t(\mathbf{x}, n)$.

² In what follows we drop the dependance on \mathbf{x} for clarity.

In this work we propose functionals where the weights vary as a function of the image, yielding update equations of the form: $\Phi_t = w_1(I_j)T_1(\Phi, I_j) + \dots + w_n(I_j)T_n(\Phi, I_j)$. The method for finding $w(I_j)$ will be explained in Sec. 2.4, and is designed to encourage functions of a convex nature (Fig. 2).

2.2 Learning the Image Manifold

Manifold learning methods are a special class of nonlinear dimensionality reduction techniques that enable the calculation of geodesic distances between data points. Given a set of expert-segmented images we apply manifold learning to learn the image manifold, $\mathcal{M}_{\mathbf{I}}$, and the segmentation manifold, $\mathcal{M}_{\mathbf{S}}$. The learned manifolds are then used to estimate geodesic distances between pairs of images or segmentations. Let $d^{\mathcal{M}_{\mathbf{I}}}(I_a, I_b)$ be the learned geodesic distance function for $\mathcal{M}_{\mathbf{I}}$, and $d^{\mathcal{M}_{\mathbf{S}}}(\Phi^a, \Phi^b)$ for $\mathcal{M}_{\mathbf{S}}$ (See Fig. 2 for definitions). Let $\mathcal{N}_{\mathbf{I}}$ be a neighborhood of points on $\mathcal{M}_{\mathbf{I}}$, and $\mathcal{N}_{\mathbf{S}}$ a neighborhood of points on $\mathcal{M}_{\mathbf{S}}$. Then define $F_{\mathcal{N}}(a, b)$ as a geodesic weighting function for a given neighborhood, \mathcal{N} , normalized to sum to one. It stands to reason that similar images will have similar segmentations, which implies a smooth mapping between manifolds. In this case $d^{\mathcal{M}_{\mathbf{I}}}(I_a, I_b) \propto M(d^{\mathcal{M}_{\mathbf{S}}}(\Phi^a, \Phi^b))$ where $M : \mathbb{R} \rightarrow \mathbb{R}$ is monotonically increasing.

2.3 Estimation of Optimal Weights for Training Data

For each $\{I_j, S_s\} \in \{\mathbf{I}, \mathbf{S}\}$ (Fig. 2), the task is to find the optimal values for the free weights $w(I_j)$. This section explores the notion of ‘optimal’. Essentially, after this phase of the algorithm we will have $|\mathbf{I}|$ samples of $w(I)$, from which we can interpolate to find values at new points (images).

One potential, but computationally intractable, approach for finding the optimal parameter is to try all possible parameter combinations against all possible (or likely) initializations and run the segmentation method then select the parameters with the least error. A better approach for PDE-based methods is to find the parameter that minimizes the magnitude of the time derivative of the shape model, Φ_t , at the solution³ Φ^s . Doing so encourages the correct solution, Φ^s , to be a stopping point of the simulation (i.e. $\Phi_t^s=0$). Since Φ_t is itself a scalar field and $w_i(I_j)$ a scalar function (see below), we measure its magnitude as $|\Phi_t^s|^2$.

Better still is to minimize the time derivative of the energy given the current pair and maximize it for all other possible segmentations (in a direction toward the optimal solution), thereby encouraging Φ^s to be the minimum of a convex function (Fig. 2). For a given shape Φ^i , a point in \mathbb{R}^N , $(\Phi^s - \Phi^i)$, taken as a vector⁴, represents the direction towards Φ^s . Since Φ_t^i is the vector in \mathbb{R}^N dictating in what direction, and by what amount, the solution will change at the point Φ^i , a normalized dot-product will measure how much in the right direction Φ_t^i points. For computational feasibility, we settle for nearby segmentations $\mathcal{N}_{\mathbf{S}}$; a reasonable step given our method for initializing is an interpolation over $\mathcal{N}_{\mathbf{S}}$.

³ The Euler Lagrange constraints ensure such a point will be a stationary point of E .

⁴ Using $(\Phi^i - \Phi^s)$ would encourage Φ^s to be a maximum.

So for an energy function with a form like those in Sec. 2.1, we need to find the parameters $w(I_j)$ that minimize

$$\arg \min \left(|\Phi_t^s|^2 - \sum_{i \in \mathcal{N}_S} F_{\mathcal{N}_S}(\Phi^s, \Phi^i) \frac{\Phi_t^i \cdot (\Phi^s - \Phi^i)}{|\Phi^s - \Phi^i|} \right) \quad (1)$$

where $F_{\mathcal{N}_S}$ is used to give more weight to nearby segmentations as a function of their geodesic proximity. The neighborhood term is negative and $|\Phi_t^i|$ is omitted from the normalized dot-product to reward large steps in the correct direction, and to allow fast solutions for w . Weights w are embedded in Φ_t .

Generally each parameter can vary spatially, or with time. For many energy functionals this can cause Eq. 1 to itself become a PDE. Consequently, we limit this work to scalar parameters that are fixed in time, causing Eq. 1 to become a convex, quadratic optimization problem in \mathbb{R}^n ; shown as follows. Letting $C_i = F_{\mathcal{N}_S}(\Phi^s, \Phi^i) |\Phi^s - \Phi^i|^{-1}$, and expanding the norm and the dot-product in Eq. 1 to expressly sum over the image domain gives: $\arg \min \left(\sum_{\mathbf{x} \in \Omega} \left[\left(\sum_{p=1}^n w_p T_p(\Phi^s) \right)^2 - \sum_{i \in \mathcal{N}_S} \left(C_i \sum_{p=1}^n w_p T_p(\Phi^i) \right) \right] \right)$. Define two $1 \times n$ vectors: \mathbf{T} with entries $T_p = \sum_{\mathbf{x} \in \Omega} T_p(\Phi^s)$; and V where $V_p = \sum_{\mathbf{x} \in \Omega} \sum_{i \in \mathcal{N}_S} (C_i T_p(\Phi^i))$. Collecting like terms and simplifying yields a quadratic equation in standard form: $\arg \min (w^T \mathbf{T}^T \mathbf{T} w - V w)$, where $\mathbf{T}^T \mathbf{T}$ is by definition positive semi-definite; ensuring the optimization problem is convex.

2.4 Optimal Parameters and Initialization for a Novel Image

Our approach assumes the optimal parameters for a novel image will be “similar” to the optimal parameters of “similar” images; we assume that $w(I)$ is a smooth function. Therefore, for a given novel image we identify its coordinates on the learned manifold, and then assign it an optimal w and an initialization by smoothly interpolating \mathcal{N}_I using a normalized Gaussian kernel on $F_{\mathcal{N}_I}$. For the priors, we limit the training data to \mathcal{N}_I , \mathcal{N}_S , since we are more confident that the correct shape and appearance information lies in those neighborhoods.

3 Results

We tested our method on a synthetic data set of 225 noisy ellipse images, where only the major and minor axes lengths of the ellipses are varied; a 2-manifold. To learn the manifolds we used a MATLAB implementation of k-ISOMAP [12] from <http://isomap.stanford.edu/>. As this paper is about the application of manifolds to segmentation, issues related to learning the manifold will not be addressed in this work. Sample images are shown in Fig. 1 (c) and (d), while (a) shows how the smoothness weight w varies gradually as a function of the manifold, with more eccentric ellipses having the least smoothness requirement. For this example we set $k = 7$ in k-ISOMAP, and reduce the image space to a 2-manifold. Adopting the Chan-Vespe energy functional of [5] and using leave-one-out validation, we

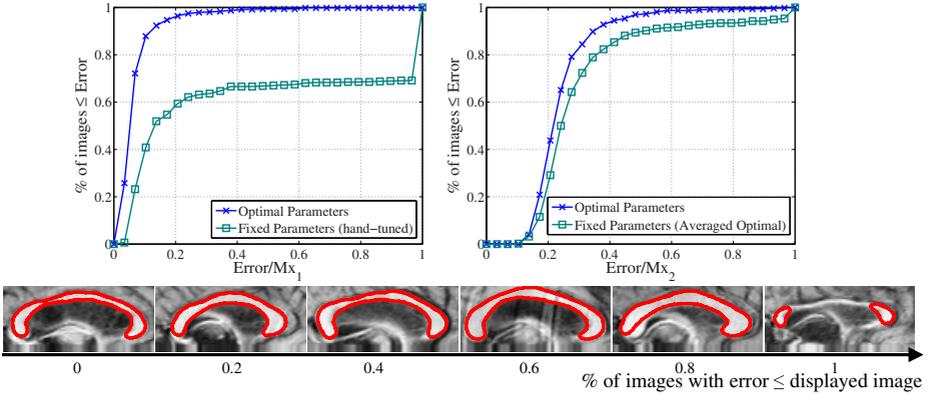


Fig. 3. CC segmentation results. (Top) Error plots where Mx_1 and Mx_2 are the maximum values of the optimal parameter error result for experiment (i) and (ii), respectively. (Bottom) Automatic, optimal parameter segmentations demonstrating the full range of error.

compare between two scenarios: (i) fixed initialization with the average ellipse and fixed parameters obtained by averaging the set of optimal parameters; and (ii) adaptively chosen optimal parameters and initializations as functions of the coordinates of the novel image on \mathcal{M}_I . We found a 47.1% reduction in average error⁵ using optimal parameters over fixed (0.190 vs 0.358). This means that 50% of the error was due to erroneous parameters and/or initializations.

For medical data we used a set of 470 256×256 affine registered mid-sagittal MRI, with corresponding expert-segmented corpora callosa (CC) (Fig. 2). We demonstrate that existing works can be extended to real clinical applications by defining the energy functional as a weighted combination of terms from [4,5,13] with an update equation in the form of:

$$\Phi_t = \begin{cases} w_1(I)g(c + \kappa) |\nabla\Phi| + w_2(I)\nabla\Phi \cdot \nabla g + w_3(I) \left(H(\Phi) |I - c_1|^2 \dots \right. \\ \left. \dots + (1 - H(\Phi)) |I - c_2|^2 \right) + w_4(I)\nabla_\Phi E_{shape}^{nonlinear} \end{cases} \quad (2)$$

$H(\Phi)$ is a Heaviside function, and c_1, c_2 are average intensities as defined in [5]. We set $c = -2.0$, since w_1 will scale accordingly regardless of its value, and use g as defined in [4]. For k-ISOMAP we set $k = 10$, and reduce the image space to a 5-manifold; chosen as the elbow of the scree plot. To demonstrate our method’s performance with different energy functionals two leave-one-out validation experiments were performed: (i) optimal parameters vs hand-tuned parameters, no outside intensity term ($H(\Phi) |I - c_1|^2$), and no auto-cropping (Fig. 3 left); (ii) optimal parameters vs averaged optimal parameters, the complete function and auto-cropping enabled (Fig. 3 right). With auto-cropping enabled segmentation took approximately 25 seconds per image, as opposed to 3 minutes, on a 2.4GHz

⁵ Error is measured as $\varepsilon = Area(Auto \cup GT - Auto \cap GT) / Area(GT)$, where $Auto$ and GT are the binary automatic segmentation and the ground truth, respectively.

AMD Opteron. As before significant reductions in average error are obtained: (0.17 vs 0.51) in experiment (i), and (0.16 vs 0.22) in (ii).

4 Conclusions

On the practical side, we motivated and outlined a new fully-automatic technique for MIA, which addresses the parameter-setting, and initialization problems that typically plague deformable model-based MIA approaches. Our approach requires training, mimicking the reliance of humans on example images to learn, and incorporating into the learning new images as they are segmented. Our results clearly demonstrate our method's ability to segment a large number of unseen images, which we foresee carrying over to other datasets and objective functions. On the theoretical side, we explored the consequences of using fixed parameters for energy-minimizing MIA techniques, defined the notion of optimal weights that favor convex energy functionals, and demonstrated a relation between image manifolds and energy-based segmentation.

Acknowledgements. CM was supported by the CIHR. The CC MRI data was provided by the MS/MRI Research Group at the University of British Columbia.

References

1. Pham, D., Xu, C., Prince, J.: Current methods in medical image segmentation. *Annu. Rev. Biomed. Eng.* 2, 315–337 (2000)
2. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. *IJCV* 1(4), 321–331 (1987)
3. Lobregt, S., Viergever, M.: A discrete dynamic contour model. *IEEE TMI* 14, 12–24 (1995)
4. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. *IJCV* 22(1), 61–79 (1997)
5. Chan, T., Vese, L.: Active contours without edges. *IEEE TIP*, 266–277 (2001)
6. Boykov, Y., Kolmogorov, V.: Computing geodesics and minimal surfaces via graph cuts. In: *ICCV* (2003)
7. Seung, H.S., Lee, D.: The manifold ways of perception. *Science* 290, 2268–2269 (2000)
8. Leventon, M., Grimson, E., Faugeras, O.: Statistical shape influence in geodesic active contours. In: *CVPR* (2000)
9. Pluempitiwiriyaewej, C., Moura, J., Wu, Y.J.L., Ho, C.: Stacs: new active contour scheme for cardiac mr image segmentation. *IEEE TMI* 24, 593–603 (2005)
10. Gennert, M., Yuille, A.: Determining the optimal weights in multiple objective function optimization. *Computer Vision* 5, 87–89 (1998)
11. Koikkalainen, J., Ljtjnen, J.: Model library for deformable model-based segmentation of 3-d brain mr-images. In: Dohi, T., Kikinis, R. (eds.) *MICCAI 2002*. LNCS, vol. 2488, pp. 540–547. Springer, Heidelberg (2002)
12. Tenenbaum, J.B., de Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. *Science* 290, 2319–2323 (2000)
13. Dambreville, S., Rathi, Y., Tannenbaum, A.: Shape-based approach to robust image segmentation using kernel pca. *CVPR* 1, 977–984 (2006)