NONPARAMETRIC SHAPE PRIORS FOR ACTIVE CONTOUR-BASED IMAGE SEGMENTATION

Junmo Kim, Müjdat Çetin, and Alan S. Willsky

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
77 Massachusetts Ave., Cambridge, MA 02139, USA

ABSTRACT

When segmenting images of low quality or with missing data, statistical prior information about the shapes of the objects to be segmented can significantly aid the segmentation process. However, defining probability densities in the space of shapes is an open and challenging problem. In this paper, we propose a nonparametric shape prior model for image segmentation problems. In particular, given example training shapes, we estimate the underlying shape distribution by extending a Parzen density estimator to the space of shapes. Such density estimates are expressed in terms of distances between shapes, and we propose two distance metrics that could be used in this framework. We then incorporate the learned shape prior distribution into a maximum *a posteriori* estimation framework for segmentation. This results in an optimization problem, which we solve using active contours. We demonstrate the effectiveness of the resulting algorithm in segmenting images that involve low-quality data and occlusions. The proposed framework is especially powerful in handling "multimodal" shape densities, involving multiple classes of objects.

1. INTRODUCTION

We consider image segmentation problems that involve limited and low-quality data. Such segmentation problems are ill-posed and require the incorporation of prior information about the objects to be segmented. When human experts segment images, they clearly make use of such prior information. For example, a radiologist can usually manually segment an organ (e.g. the prostate) in a magnetic resonance image, although the boundaries are mostly invisible to a layperson. Clearly, radiologists augment the observed data with their expertise, in other words with statistical prior information, about the shape of the organ. Existing automatic segmentation methods (either explicitly or implicitly) enforce only very simple constraints about the underlying shapes. For example, many active contour-based methods (which is the framework we also use in our work) involve a *curve length penalty*, which translates to the assumption that shorter curves are statistically more likely than longer ones. However, in many applications, more structured prior information about the shapes is available. Yet the challenge is how to construct probabilistic descriptions in the space of shapes, and then incorporate such statistical information into the segmentation process.

Early work on this problem involved landmark-based representations of shapes, and the construction of typical shapes and typical variability based on a set of training shapes via principal component analysis (PCA) [1] . The use of landmarks has the drawback that the performance of shape analysis depends on the quality of those landmarks. Recently, there has been increasing interest in using level set-based representations for shape priors [2, 3], which avoid landmarks. In [2] and [3], PCA of the signed distance functions of training data is used to capture the variability of shapes. These techniques have been applied to segmentation problems involving low SNR or occluded images successfully, especially when the shape variability is small. However, there are two major shortcomings of such techniques. First, these methods treat the signed distance functions as elements of a linear vector space, and perform operations such as averaging. Yet, the space of signed distance functions is a nonlinear manifold and is not closed under linear operations. For example, the average of signed distance functions, which is commonly used to obtain a mean shape, is not necessarily a signed distance function. Therefore, the use of linear analysis tools such as PCA gives rise to an inconsistent framework for shape modeling [3]. Second, these technique can handle only unimodal, Gaussian-like shape densities. In particular, these methods cannot deal with "multimodal" shape densities, which involve multiple classes of shapes (e.g. a shape density of handwritten digits, composed of multiple digits).

In our work, we propose a framework for constructing nonparametric shape densities from example training shapes. In particular, we assume that the training shapes are drawn from an unknown shape distribution, and we estimate the underlying shape distribution by extending a Parzen density estimator to the space of shapes. Such density estimates are expressed in terms of distances between shapes. We propose two specific distance metrics to be used for nonparametric density estimation, although other metrics could be used in our framework as well. We then formulate the segmentation problem as maximum *a posteriori* (MAP) estimation, in which we use the learned nonparametric shape density as the prior. This leads to an optimization problem for the segmenting curve, for which we develop and use an active contour-based iterative algorithm. We present experimental results of segmenting low-quality and occluded images. We also demonstrate how the proposed algorithm can successfully incorporate shape densities involving multiple object classes.

2. NONPARAMETRIC SHAPE DENSITIES

In this section, we address the problem of estimating an unknown shape probability density. Given n example training curves $C_1, \ldots, C_n$, we first align[1] them to obtain the aligned training set $\tilde{C}_1, \ldots, \tilde{C}_n$. Ideally, this operation removes the pose variability in the training data, and what remains is just shape variability. The problem then is to estimate $p_{\tilde{C}}(\tilde{C})$ from which the training samples $\tilde{C}_1, \ldots, \tilde{C}_n$ are drawn. This density is a probability density over an infinite dimensional space. We would like to leave the shape of this density unconstrained, therefore we adopt a nonparametric density estimation route. Assuming that we are given a distance metric $d_{\mathcal{C}}(\cdot, \cdot)$ in the space of curves $\mathcal{C}$, we can form a Parzen density estimator as follows:

---

[1] We use the alignment algorithm of [3], but other techniques can be used as well.

$$\hat{p}_{\tilde{C}}(\tilde{C}) \quad = \quad \frac{1}{n}\sum_{i=1}^{n}k(d_{\mathcal{C}}(\tilde{C},\tilde{C}_i),\sigma) \tag{1}$$

where $k(\cdot,\sigma)$ denotes a Gaussian kernel with kernel size σ, i.e. $k(x,\sigma)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{x^2}{2\sigma^2})$.

Conceptually, the nonparametric density estimate in (1) can be used with a variety of distance metrics, yet the key issue is what kind of metrics make sense. This is related to the question of *what it means for two shapes to be similar*. In the following sections, we consider two specific metrics, namely the template metric[2] and the Euclidean (or L_2) distance between signed distance functions. The template metric is given by the area of the set-symmetric difference between interior regions of two shapes, and can be expressed as a norm of difference between two binary maps representing the shapes. On the other hand, the L_2 metric we use is a norm of the difference between two signed distance functions. From a practical standpoint, the key difference between these two metrics is that the template metric puts equal weight on pixels, whereas the L_2 distance between signed distance functions puts variable weight.

2.1 Template Metric

We now consider the Parzen density estimate in (1) with a specific metric, namely the template metric $d_T(\tilde{C},\tilde{C}_i) = \text{Area}(R_{\text{inside }\tilde{c}}\triangle R_{\text{inside }\tilde{c}_i})$ [5], where $\triangle$ denotes set symmetric difference. The density estimate with the template metric is given by

$$\hat{p}_{\tilde{C}}(\tilde{C}) = \frac{1}{n}\sum_{i=1}^{n}k(d_T(\tilde{C},\tilde{C}_i),\sigma). \tag{2}$$

In [6], we also consider a variant of this metric, namely its square-root, as an alternative distance measure.

2.2 Euclidean Distance between Signed Distance Functions

We now consider representing each shape $\tilde{C}_i$, $i \in \{1,...,n\}$, by its corresponding signed distance function $\tilde{\phi}_i$, and constructing a shape density estimate $\hat{p}_{\tilde{C}}(\tilde{C}) = \hat{p}_{\tilde{\phi}}(\tilde{\phi})$.[3] We observe that the space of signed distance functions $\mathcal{D}$ is a subset of an infinite dimensional Hilbert space $\mathcal{L} \triangleq \{\phi|\phi:\Omega\to\mathcal{R}\}$, where Ω denotes the image domain. We can define an inner product and an induced L_2 distance in $\mathcal{L}$ as follows:

$$\langle\phi_1,\phi_2\rangle = \int_{\Omega}\phi_1(x)\phi_2(x)dx, \tag{3}$$

$$d_{L_2}(\phi_1,\phi_2) = \sqrt{\langle\phi_1-\phi_2,\phi_1-\phi_2\rangle}. \tag{4}$$

Since the space $\mathcal{D}$ is embedded in a Hilbert space, a natural metric $d(\phi_1,\phi_2)$ for this space would be a minimum geodesic distance, i.e. the distance of the shortest path from ϕ_1 to ϕ_2 lying in $\mathcal{D}$. If one could compute such minimum geodesic distances $d_{\text{geodesic}}(\cdot,\cdot)$, the corresponding Parzen density estimate based on aligned training samples $\{\tilde{\phi}_i\}$ would be

$$\hat{p}_{\tilde{\phi}}(\tilde{\phi}) = \frac{1}{n}\sum_{i=1}^{n}k(d_{\text{geodesic}}(\tilde{\phi},\tilde{\phi}_i),\sigma). \tag{5}$$

---

[2]We note that the work in [4] (which came to our attention while preparing this manuscript) has commonalities with our template metric-based approach.

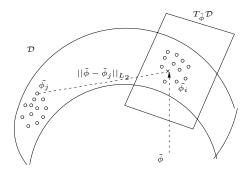[3]ϕ and $\tilde{\phi}$ denote the signed distance functions before and after alignment, respectively.

Figure 1: Notional illustration of the space of signed distance functions $\mathcal{D}$, in a scenario where example shapes form two clusters, corresponding to two classes of shapes.

However, computing a geodesic distance in an infinite dimensional manifold is a challenging problem. There is some previous work on computing geodesic distances in the space of curves [7, 8], but there is little work when the shape is represented by signed distance functions.

Instead, we now consider the Parzen density estimate with the L_2 distance in $\mathcal{L}$:

$$\hat{p}_{\tilde{\phi}}(\tilde{\phi}) = \frac{1}{n}\sum_{i=1}^{n}k(d_{L_2}(\tilde{\phi},\tilde{\phi}_i),\sigma). \tag{6}$$

Now we discuss how the density estimate in (6) might be a good approximation to the one with geodesic distance in (5). If the example shapes are of small variation then the part of the manifold supporting the example shapes can be approximately flat provided that the manifold does not have too much curvature. This is why methods involving linear analysis tools such as PCA of signed distance functions [2, 3] work reasonably well in the case of small shape variation. Due to the same phenomenon, the L_2 distance used in the Parzen density estimate can be close to the geodesic distance. Now consider the case where the example shapes belong to multiple classes and form multiple clusters as illustrated in Fig. 1. In this case, the part of the manifold supporting the samples is no longer flat, and PCA would no longer be a valid approach. In contrast, the density estimate with L_2 distance can still be a good approximation of (5) for the following reasons. For two shapes that are in the same cluster (e.g. $\tilde{\phi}$ and $\tilde{\phi}_i$ in Fig. 1), hence are of small variation, the L_2 norm will be a good approximation of the geodesic distance. On the other hand, for shapes that are in different clusters (e.g. $\tilde{\phi}$ and $\tilde{\phi}_j$ in Fig. 1), there will be an error in approximation of the geodesic distance, but the overall error in density estimation will be small as long as the kernel size σ is small compared to the distance $d_{L_2}(\tilde{\phi},\tilde{\phi}_j)$.

3. SHAPE-BASED SEGMENTATION

Now we combine the nonparametric shape prior with a data term, and formulate the segmentation problem as MAP estimation. This leads to the following energy functional to be minimized for segmentation[4]:

$$E(C) = -\log p(\text{data}|C) - \log p_C(C). \tag{7}$$

The data term we use here is based on [9], and the second term is based on the nonparametric shape priors introduced in Section 2. We minimize this functional iteratively by gradient descent, for which we need the gradient flow for the curve C or the corresponding signed distance function ϕ.

---

[4]In the rest of the paper, we drop the hat for simplicity in density estimate $\hat{p}$.

1. Evolve C without the shape prior for time $t \in [0, t_0]$
2. For the curve $C|_{t=t_0}$, compute the pose $\mathbf{p}|_{t=t_0}$ by aligning $C|_{t=t_0}$ with respect to $\{\tilde{C}_i\}$
3. Iterate until convergence:
   (a) fix $\mathbf{p}$ and
      i. compute $\tilde{C} = T[\mathbf{p}]C$
      ii. compute $\frac{\partial \tilde{C}}{\partial t}$ for the shape prior $\log p_{\tilde{C}}(\tilde{C})$
      iii. compute $\frac{\partial C}{\partial t}$ from $\frac{\partial \tilde{C}}{\partial t}$ by $\frac{\partial C}{\partial t} = T^{-1}[\mathbf{p}]\frac{\partial \tilde{C}}{\partial t}$
   (b) update C by both the data and the shape force
   (c) fix C and update $\mathbf{p}$ using an alignment technique

Algorithm 1: Iterative algorithm for updating the pose $\mathbf{p}$ and the curve C.

The gradient flow for the data term is computed as is done in [9]. So we only focus on the gradient flow $\frac{\partial C}{\partial t}$ or $\frac{\partial \phi}{\partial t}$ for the shape term $\log p_C(C)$, where t denotes an artificial iteration time.

Note that any given curve C is not necessarily aligned with the training data, whereas the shape densities we proposed in Section 2 are based on aligned curves $\tilde{C}$ and $\{\tilde{C}_i\}$. As a result, in order to compute $\frac{\partial C}{\partial t}$, we first compute $\frac{\partial \tilde{C}}{\partial t}$, and then compute $\frac{\partial C}{\partial t}$ from $\frac{\partial \tilde{C}}{\partial t}$. To this end, let $\tilde{C} = T[\mathbf{p}]C$, where $T[\mathbf{p}]$ is a similarity transformation with the pose parameter $\mathbf{p}$ capturing translation, rotation, and scale. We update C and $\mathbf{p}$ iteratively as described in Algorithm 1. We discuss Step 3-(a)-ii below for the two distance metrics we consider.

When we use the template metric-based shape prior $\hat{p}_{\tilde{C}}(\tilde{C})$ of (2) in (7), we obtain the following gradient flow [6]:

$$\frac{\partial \tilde{C}}{\partial t} = \frac{1}{p_{\tilde{C}}(\tilde{C})}\frac{1}{n}\frac{1}{\sigma^2}\sum_i k(d_T(\tilde{C}, \tilde{C}_i), \sigma)d_T(\tilde{C}, \tilde{C}_i)(1 - 2H(\phi_i))\vec{N} \quad (8)$$

where $\vec{N}$ is the outward normal direction with respect to the curve, and $H(\cdot)$ is the Heaviside function, i.e. $H(\phi) = 1$ if $\phi \geq 0$ and $H(\phi) = 0$ if $\phi < 0$. Note that the gradient flow is composed of a weighted average of several directions, where the i-th direction is an optimal (gradient) direction that decreases the distance between the i-th training shape and the evolving shape. We implement (8) using level set methods.

Next, we consider the Euclidean distance-based shape prior $p_{\tilde{\phi}}(\tilde{\phi})$ of (6). Let us first consider evolving $\tilde{\phi}$ in the space $\mathcal{L}$ along the gradient of $\log p_{\tilde{\phi}}(\tilde{\phi})$ w.r.t. $\tilde{\phi}$. The resulting gradient flow would be [6]:

$$\frac{\partial \tilde{\phi}}{\partial t} = \frac{1}{p_{\tilde{\phi}}(\tilde{\phi})}\frac{1}{\sigma^2}\frac{1}{n}\sum_i k(d_{L_2}(\tilde{\phi}, \tilde{\phi}_i), \sigma)(\tilde{\phi}_i - \tilde{\phi}). \quad (9)$$

This flow involves a weighted average of $\{\tilde{\phi}_i - \tilde{\phi}\}_{i=1}^n$, where $\tilde{\phi}_i - \tilde{\phi}$ is the direction toward the i-th training shape $\tilde{\phi}_i$. Note that the weight $k(d_{L_2}(\tilde{\phi}, \tilde{\phi}_i), \sigma)$ for the velocity component $\tilde{\phi}_i - \tilde{\phi}$ increases as $\tilde{\phi}$ gets closer to $\tilde{\phi}_i$.

One issue with the evolution in (9) is that the evolving level set function does not necessarily remain on the manifold of signed distance functions $\mathcal{D}$. The evolution of the zero level set (i.e. the curve) in such a case can be less stable than the case where the evolving level set function remains a signed distance function [10]. Hence, it is desirable to constrain the evolving level set function to stay on the manifold.

Based on these observations, we modify the evolution equation (9). The goal here is to extract relevant information for shape evolution from the velocity field in (9) and to construct a new velocity field such that the resulting trajectory is on $\mathcal{D}$. First we observe that the only components of
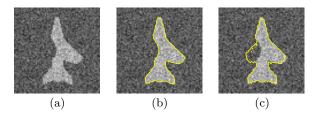

Figure 2: Segmentation of an occluded aircraft image. (a) Test image. (b) Result without a shape prior. (c) Result with the nonparametric shape prior.

the velocity field in (9) that directly impact the shape evolution are those defined at the points on the boundary $\tilde{C}$, i.e. where the $\tilde{\phi} = 0$. So we do not modify those components. The key property we exploit is the following: if we start with a signed distance function and evolve it with a velocity field that is constant along the direction normal to the corresponding curve $\tilde{C}$, the evolving level set function remains a signed distance function [11]. So we construct a modified velocity field by copying the velocities on the boundary from (9), and then extending these values in the direction normal to the boundary.

4. EXPERIMENTAL RESULTS

Now we present experimental results demonstrating our segmentation method based on nonparametric shape priors. We first show results for segmenting occluded objects. Here the shape prior involves a single class of object shapes. Next, we present experimental results on a more challenging problem, where the prior involves multiple classes of object shapes. In particular, we consider the problem of segmenting handwritten digits (with low quality or missing data), where the prior density involves all digits, and the algorithm does not know which digit class the test data belong to. For the sake of brevity we present results with only the L_2 metric; the results with the template metric are similar.

4.1 Segmentation of Occluded Objects

In this section, we demonstrate our shape-based segmentation algorithm on synthetic aircraft images. As example shapes of this class, we have a set of 11 binary images (not shown here), whose boundaries provide the training curves $C_1, \ldots, C_n$ ($n = 11$). We present segmentation results on the image of an aircraft whose shape is not included in the training set. Fig. 2(a) shows the noisy aircraft test image with an occluded left-wing. Fig. 2(b) shows the segmentation of this image without a shape prior (the shape prior is replaced by a curve length penalty term). In Fig. 2(c) we show the result of our proposed technique, where the shape prior information allows us to recover the wing of the aircraft. We also note that our algorithm does not have access to any information about which parts of the image contain occlusions.

4.2 Segmentation of Multiple Classes of Objects

We now consider the case where the shape prior density involves multiple classes of shapes. This is a scenario which cannot be readily handled by most existing shape-based segmentation techniques. In particular, we consider the problem of segmenting handwritten digits, where there are 10 classes of digits, i.e. 0, 1, . . . , 9. We use a training set of 100 sample images with 10 segmented images of each digit.

Let us consider the low-SNR test images (not included in the training set) in Fig. 3(a). Segmentations without a shape prior (and with a curve length penalty instead) are shown in Fig. 3(b). The results of our shape-based segmentation
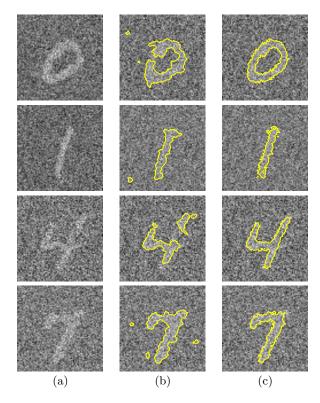
Figure 3: Segmentation of low-SNR digit images. (a) Test images. (b) Results without a shape prior. (c) Results with the proposed shape prior.

method are shown in Fig. 3(c), which appear to be quite accurate despite the low-quality data. Finally, we consider an example involving test images with missing data, as shown in Fig. 4(a). Our shape-based segmentation results shown in Fig. 4(c) provide accurate segmentations despite the data limitations.

5. CONCLUSION

We have considered the problem of estimating shape prior densities from example shapes and proposed a shape-based segmentation method. In particular, we have developed a framework for estimating shape priors from training shapes nonparametrically. Based on such nonparametric shape priors, we have formulated the shape-based segmentation problem as a MAP estimation problem. Evaluation of the nonparametric shape prior for a candidate curve for segmentation is given in terms of distances between the candidate curve and the training curves. We considered the template metric and the L_2 distance between signed distance functions, but other metrics can also be used for nonparametric shape priors. We have derived curve evolution equations based on the nonparametric shape priors. We have presented experimental results of segmenting partially-occluded images. We have considered the case in which the training shapes form multiple clusters, and demonstrated that our nonparametric shape priors model such shape distributions successfully without requiring prior knowledge on the number of clusters.

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.

[2] M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours,"
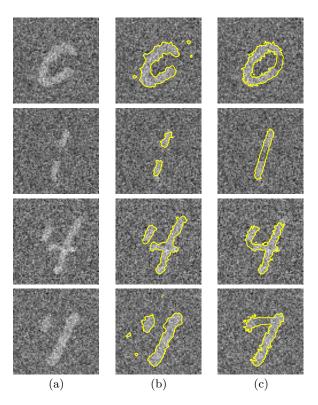
Figure 4: Segmentation of digit images with missing data. (a) Test images. (b) Results without a shape prior. (c) Results with the proposed shape prior.

in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2000, pp. 316–323.

[3] A. Tsai, A. Yezzi, Jr., W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. L. Grimson, and A. S. Willsky, "A shape-based apprach to the segmentation of medical imagery using level sets," *IEEE Trans. on Medical Imaging*, vol. 22, no. 2, pp. 137–154, Feb. 2003.

[4] D. Cremers, S. Osher, and S. Soatto, "Kernel density estimation and intrinsic alignment for knowledge-driven segmentation: Teaching level sets to walk," in *Pattern Recognition*, Aug. 2004, vol. 3175 of *Lecture Notes in Computer Science*, pp. 36–44.

[5] D. Mumford, "Mathematical theories of shape: Do they model perception?," in *SPIE Proceedings Vol. 1570 Geometric Methods in Computer Vision*, 1991, pp. 2–10.

[6] J. Kim, *Nonparametric Statistical Methods for Image Segmentation and Shape Analysis*, Ph.D. thesis, Massachusetts Institute of Technology, Feb. 2005.

[7] P. W. Minchor and D. Mumford, "Riemannian geometries on spaces of plane curves," *J. Eur. Math. Soc. (JEMS)*, to appear.

[8] E. Klassen, A. Srivastava, W. Mio, and S. H. Joshi, "Analysis of planar shapes using geodesic paths on shape spaces," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 3, pp. 372–383, Mar. 2004.

[9] T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. on Image Processing*, vol. 10, no. 2, pp. 266–277, February 2001.

[10] S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*, Springer, 2003.

[11] H.-K. Zhao, T. Chan, B. Merriman, and S. Osher, "A variational level set approach to multiphase motion," *Journal of Comp. Phys.*, vol. 127, pp. 179–195, 1996.