Shape Representation based on Integral Kernels: Application to Image Matching and Segmentation

Byung-Woo Hong[1]   Emmanuel Prados[1]   Luminita Vese[2]
Stefano Soatto[1]

Computer Science Department[1], Mathematics Department[2]

Abstract

This paper presents a shape representation and a variational framework for the construction of diffeomorphisms that establish "meaningful" correspondences between images, in that they preserve the local geometry of singularities such as region boundaries. At the same time, the shape representation allows enforcing local shape priors in determining such region boundaries. Our representation is based on a kernel descriptor that characterizes local shape. This shape descriptor is robust to noise and forms a scale-space in which an appropriate scale can be chosen depending on the size of features of interest in the scene. In order to preserve local shape during the matching procedure, we introduce a novel constraint to traditional energy-based approach to estimate diffeomorphic deformations, and such energy is minimized in a variational framework.

1 Introduction

Enforcing prior knowledge on the shape of structures of interest is a common way to facilitate bottom-up segmentation, for instance in the detection of anatomical structures in medical images. Typically one exploits hand-segmented samples to design a density in some space where shape is represented, and such a density is used to bias the segmentation process towards shapes yielding a high posterior. This enables overcoming problems such as low contrast, occlusions, illumination variations and other unmodeled phenomena (collectively labeled as "noise") that would make purely bottom-up segmentation unsuccessful. While many researchers have been after a "universal" theory of shape, including the determination of an appropriate space, endowed with a metric and probabilistic structure, that would enable reasoning on shape in a solid analytical framework, such a theory does not exist to date. Furthermore, how "similar" two shapes are, and what portion of one shape corresponds to what portion of another, are highly dependent on the application, and a "meaningful correspondence" in one domain may be completely wrong in another [39, 26, 38, 12]. So, rather than seeking

to work in generality, we concentrate on a representation of shape and the resulting matching algorithms that enable successful image matching and segmentation in the presence of prior knowledge, encoded in a rough "template" [14].

In broad terms, shape is a geometric attribute of the image domain, for instance the locus of singularities of certain operators applied to the image (e.g. the maxima, minima, or zero-crossings of the Laplacian [22]). In some cases such locus is a set of isolated "landmark" points, and the ensuing shape spaces are simply the quotient of their positions modulo the affine group, or other finite-dimensional group [16, 19, 2, 25, 7]. This is far too restrictive for our goals, since we want to deal with data that do not exhibit obvious isolated landmarks, and we want to consider as equivalent objects that are related by more complex transformations that affine, say a deforming hand, including ones with missing fingers. At the other hand of the spectrum, deformable templates [14, 6] consider the orbit of a given geometric object under the set of all possible diffeomorphisms, which are organized into a group that reaches every point in the space (i.e. it acts transitively). This means that a human seen through her silhouette is the same as an ice-cream. In order to achieve a balance so that objects that would be reasonably considered equivalent in the domain of application of interest, we must restrict the set of allowable deformations. We choose to do so by asking such deformations to *preserve the local structure* of the template and target shape. This means that portions of a shape that have, locally, a certain geometric description (e.g. curvature), tend to match portions with similar local structure. However, we want to avoid computing curvature or other differential operator on potentially noisy data. Also, we want to avoid performing *shape analysis*, i.e. extracting the semantic (graph) structure of "parts" of a given shape, which is a computationally intensive and delicate process [41, 18, 11].[1]

To this end, we introduce a *local shape feature* to describe the local structure of a shape, using integral kernels that enjoy significant robustness to noise and naturally form a *scale-space*. From our local shape feature we design a *shape representation* that seamlessly encompasses local and global shape information. Our representation fits well with existing variational techniques for shape matching and segmentation: We introduce a new "local shape preserving" constraint on the set of matching diffeomorphisms that can be simply added to their energy functional.

We have suggested that our representation facilitates establishing "meaningful" correspondence. Naturally, "meaningful" is a tautology, since any shape matching algorithm achieves meaningful correspondence, where meaningful is defined by the chosen matching score. What we mean here is that points are matched based on the local shape structure, as measured by the local shape feature. We have no ambition of universality, we just wish to improve the results of matching and segmentation in medical images, as we illustrate with experiments on real and synthetic data. We have implemented gradient-based registration and segmentation schemes incorporating our "local shape preservation" constraints.

The strengths of our method are that (1) it yields matching that preserves local shape, without performing shape analysis to break down objects into parts, and (2) it can be used to incorporate prior knowledge for various high-level visual tasks by enforcing

---

[1]It should be mentioned that efficient and robust techniques to match point-wise representations of shape exist, e.g. [1], but they do not easily fit in a segmentation framework since they are supported on zero-measure sets.

2

local correspondence in global shape matching and segmentation.

2   Shape Representation Using Integral Kernels

In this work we restrict the analysis to closed planar regions and their boundaries, although some of the concepts carry to higher dimension. These regions can be described by a binary image χ, modulo a suitable class of (continuous and invertible) image domain transformations. It is the characterization of such domain transformations that determines the *shape space*, which we address in this section.

2.1   Integral Kernels as Local Shape Features

In general, the term *feature* $\mathcal{F}_\sigma$ indicates any image statistic. In particular, *local features* are functions of the image defined on a compact subset of its domain, including "soft" versions where the "effective subset" is determined by a kernel:

$$\mathcal{F}_\sigma(\chi, x) = K_\sigma * \chi\,(x),$$

where the kernel K_σ can be, for instance, a Gaussian

$$K_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{|x|^2}{2\sigma^2}}.$$

Applied to a binary image, the value of $\mathcal{F}_\sigma$ computed at a boundary point depends on its curvature at that point [24, 33]. In fact, $\mathcal{F}_\sigma$ entails a regularized notion of curvature in a scale-space even where the boundary is not differentiable. In this sense, we call $\mathcal{F}_\sigma$ a *local shape feature* and observe that it is robust to "noise" (lack of differentiability) of the boundary. As defined in the feature function, a scale is associated with feature of interest and various levels of features can be characterized in a feature scale-space with varying scales. Note that, despite the fact that local features entail "blurring" with a kernel, the geometrical shape attributes are precisely preserved in the shape representation which will be discussed in the following section. 3.

2.2   Shape Representation

While *features* are any statistics of the image, not necessarily sufficient ones, a *representation* is a bijective function of the data and can be used to retrieve it, possibly up to a class of transformations (i.e. it contains the "same information"). Examples of direct representations of shape include parameterized curves describing the boundary, which are not intrinsic and present obvious problems. An alternative approach consists of representing shapes by maps, i.e. functions defined on $\mathbb{R}^2$ onto $\mathbb{R}$. Such representations, of which the Level Set framework [29, 23] is a prototype, are simpler to handle and better adapted to segmentation. For example, the binary image itself is a representation. Another celebrated choice is the signed distance from the contour of the shape. Its computation is based on its characterization as the unique viscosity solution ϕ of the partial differential equation (PDE) $|\nabla\phi(x)| = 1$ verifying $\phi(x) = 0$, $\forall x \in \Gamma$, $\phi(x) < 0$, for all x inside of Γ, $\phi(x) > 0$, for all x outside of Γ, where Γ is the boundary of the shape. These two representations (binary image and signed distance

function) have been used successfully for registration and segmentation [31, 32, 9, 5]. Nevertheless, these two representations suffer from the lack of characterization with respect to local geometry of the shape when dealing with *deformable shape priors* for segmentation. In order to compensate for such weakness we introduce a new representation that can handle at the same time local geometric features and global boundary information:

$$\mathcal{R}_\sigma(\chi)(x) = \chi(x)\,(K_\sigma * (1 - \chi))(x)$$
$$+ (1 - \chi)(x)\,(K_\sigma * \chi)(x).$$

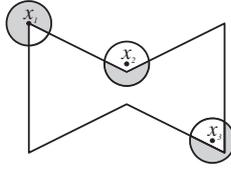Fig. 1 provides a schematic illustration of our shape representation. Our shape rep-

Figure 1: Schematic illustration for the calculation of the shape representation based on an integral (disk) kernel.
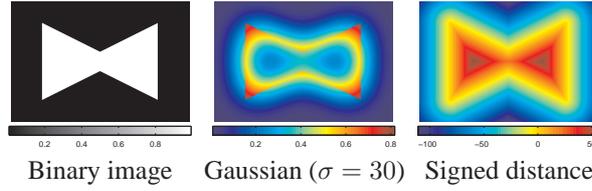

Binary image   Gaussian ($\sigma = 30$)   Signed distance

Figure 2: Comparison of the shape representation for a bow-tie shape. (image size is 300×444)
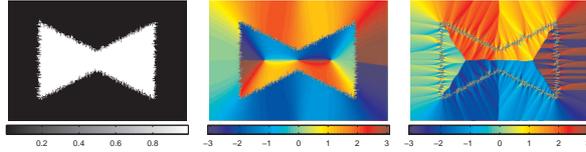

Figure 3: Comparison of shape representations for (left) a noisy bow-tie. (middle) Directions of normal vectors to the integral kernel representation using a Gaussian kernel ($\sigma = 30$). (right) Directions of normal vectors to the signed distance function. (image size is 300×444).

resentation is roughly given by the area of the intersection between the kernel and the inside or outside of the shape. For any point along the boundary (e.g. x_1) and inside the shape (e.g. x_3), the value of the shape representation is given by the area of the intersection between the kernel and the outside of the shape. For any point outside of

the shape (e.g. x_2), the value of the shape representation is given by the area of the intersection between the kernel and the inside of the shape. Fig. 2-b) shows the values of our shape representation for a bow-tie. Fig. 2-c) shows the signed distance function for comparison. Our shape representation enjoys of several desirable properties: It is **robust to "noise"**: Fig. 3 shows that our shape representation is insensitive to lack of differentiability of the contour, unlike curvature or the normal vectors to the signed distance function. This is due to the fact that regularization is implicit in our representation. Our representation **"propagates" shape information** inside and outside the boundary since its value, unlike that of a binary image, depends on the local geometry. Our shape representation is **fast and easy to compute**. It is just a convolution. The gradients of the energy can also be derived easily due to the convenient properties of the Gaussian kernel. Naturally, any kernel other than Gaussian can be used instead, and our choice is dictated by mathematical convenience. Note also that, inside the shape, our representation is the solution of the heat equation with initial data χ

$$\frac{\partial u}{\partial \sigma} = \triangle_x u, \ \ u(x, 0) = \chi(x).$$

 This establishes a connection with the work of Gorelick et al. [13] which propose representing shapes as the solution of a Poisson equation. Unlike [13] and other work based on the signed distance function, we do not solve any PDE for designing our representation.

3 Shape Matching

Establishing correspondence between shapes is an important problem in many areas of science and engineering, from neuroscience to entomology. It also plays a role in image segmentation, where we want to enforce prior knowledge on the shape of structures of interest in the image. For instance, in detecting brain structures from tomographic images prior knowledge is often crucial as the raw signal often does not exhibit distinctive enough signatures to allow detection in a purely bottom-up fashion. Unfortunately, a universal theory of shape, including the definition of an appropriate shape space with the right metric structure that captures the intuitive notion of similarity, is elusive. Here we are interested in exploiting our representation of shape to enable matching, registration and segmentation while preserving the *local* shape structure of a "template," which encodes a rough form of prior information.

While shape priors for segmentation have been studied before [21, 40, 20, 30, 9], they have concentrated on a global representation. This is not well suited for objects like the brain that exhibit a significant variability in certain structures (hence the density derived from samples would be "flat," or uninformative), while displaying significant consistency in other structures. Furthermore, we are interested in establishing *correspondence* between the shape template and the target image in the process, in a way that preserves the local "structure" of the template. This has also been studied before, but most previous studies require *shape analysis*, that is the organization of a given structure into "parts" [17, 43], a process that often relies on extracting delicate structures such as the medial axis or skeleton from the images. Correspondence is then based on discrete (graph) structures. We are interested in a simple approach that, while having no ambition to elucidate the semantic structure of shapes, would nevertheless put regions with similar local shape into correspondence.

We exploit the kernel representation derived in the previous section by integrating it

into standard energy functionals used for segmentation, to measure the energy of the diffeomorphism to match a given template. In this framework, shape similarity is measured via the shape kernel, bringing regions with similar local shape features into correspondence. Given two shapes S_1, S_2 defined in the domain Ω, matching is determined by a diffeomorphism $h : \Omega \to \mathbb{R}^2$ assigning to each point x in Ω a displacement vector $h(x) \in \mathbb{R}^2$ minimizing an error criterion between the feature of S_1 and the warped feature of S_2 under the diffeomorphism h as used in [27, 15]. The energy functional $E_{\text{shape}}(h; S_1, S_2)$ for S_1 and S_2 is defined by:

$$E_{\text{shape}}(h; S_1, S_2) = E_{\text{data}}(h; S_1, S_2) + \alpha \, E_{\text{reg}}(h),$$

where E_{data} measures the "dissimilarity" between corresponding features and E_{reg} measures the "irregularity" of h; α is a constant coefficient to be tuned as part of the design process. The well-posedness of this energy functional is discussed in [15]. The dissimilarity between S_1 and S_2 is defined in terms of their features $\mathcal{R}_{\sigma_1}(S_1)$ and $\mathcal{R}_{\sigma_2}(S_2)$:

$$E_{\text{data}}(h; S_1, S_2) = \int_\Omega |\mathcal{R}_{\sigma_1}(S_1)(x) - \mathcal{R}_{\sigma_2}(S_2)(h(x))|^2 dx.$$

This entails a choice of σ that can be performed to select the native scale of features that we want to match on the two shapes. Alternatively, comparison can be performed on the entire scale-space of σ_1 and σ_2 thus allowing hierarchical comparison of shapes across scales from coarse to fine. The feature of each shape is associated with a scale and the dissimilarity is measured up to that scale. The regularization term E_{reg} is defined by the linear elasticity [8] given by:

$$E_{\text{reg}}(h) = \frac{1}{2} \int_\Omega \left\{ \lambda(\text{div } h)^2 + 2\mu \sum_{i,j=1}^{n} (\epsilon_{ij}(h))^2 \right\} dx,$$

where $\lambda, \mu > 0$ are the Lamé coefficients of the material,

$$\text{div } h = (h_1)_{x_1} + (h_2)_{x_2}, \ \ \epsilon_{ij}(h) = \frac{1}{2}\Big((h_i)_{x_j} + (h_j)_{x_i}\Big).$$

This regularization term is designed to penalize variations of the function h and favor smoothness. The optimal correspondences given by h^* are obtained as:

$$h^* = \arg\min_h E_{\text{shape}}(h).$$

The energy minimization is performed in a variational framework using a gradient descent method. The Euler-Lagrange equation corresponding to the energy E_{shape} yields the gradient direction for h:

$$\frac{\partial h}{\partial t} = -\frac{\partial E_{\text{shape}}}{\partial h} = -\frac{\partial E_{\text{data}}}{\partial h} - \alpha \, \frac{\partial E_{\text{reg}}}{\partial h},$$

$$\frac{\partial E_{\text{data}}}{\partial h} = (\mathcal{R}_{\sigma_1}(S_1) - \mathcal{R}_{\sigma_2}(S_2)(h)) \cdot$$
$$\Big(\nabla S_2(K_{\sigma_2} * (2S_2 - 1)) + (2S_2 - 1)(K_{\sigma_2} * \nabla S_2)\Big)(h)$$

$$\frac{\partial E_{\text{reg}}}{\partial h} = -\Big(\mu \triangle h + (\lambda + \mu)\nabla(\text{div } h)\Big)$$

Given one scale (e.g. σ_2), the other scale (e.g. σ_1) is also optimized in the process of matching and the gradient can be easily derived. Shape matching results for the case of the bow-tie are shown in Fig. 4 where target shape is shown to be perfectly matched with the deformed source shape. The robustness of our representation is demonstrated in Figure 4 (second row) where the deformed bow-tie to the noisy bow-tie is matched exactly except noise using a certain scale that is above the feature of noise presented in the noisy bow-tie and below the feature of its global geometry.
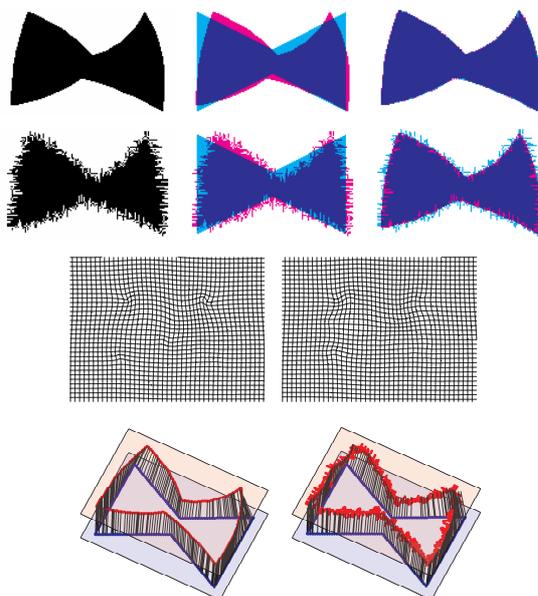

Figure 4: Shape matching results for the deformed bow-tie. [First row] (left) Deformed bow-tie (target); (middle) deformed bow-tie superimposed to the original bow-tie (source); (right) overlap of the warped source and the target after the matching. [Second row] (left) Deformed "noisy" bow-tie (target); (middle) deformed noisy bow-tie superimposed on the original bow-tie (source); (right) overlap of the warped source and the target after the matching. [Third row] Displacement vector field representing the diffeomorphism for the matching between (left) deformed bow-tie (right) deformed noisy bow-tie and the original bow-tie. [Fourth row] Dense correspondences along the boundary obtained by the matching between shapes (left) deformed (right) deformed with noise. In the experiments, $\sigma_1 = \sigma_2 = 7$ for the deformed shape and $\sigma_1 = \sigma_2 = 14$ for the noisy deformed shape, and $\alpha = 0.005, \mu = 1, \lambda = 0$ are used and image size is 192×256.

In order to illustrate what we mean by "meaningful" correspondence, in Figs. 5-6 we show the correspondence for shapes that have distinct "parts." Ideally we want our algorithm to put them into correspondence, but without conducting a semantic analysis
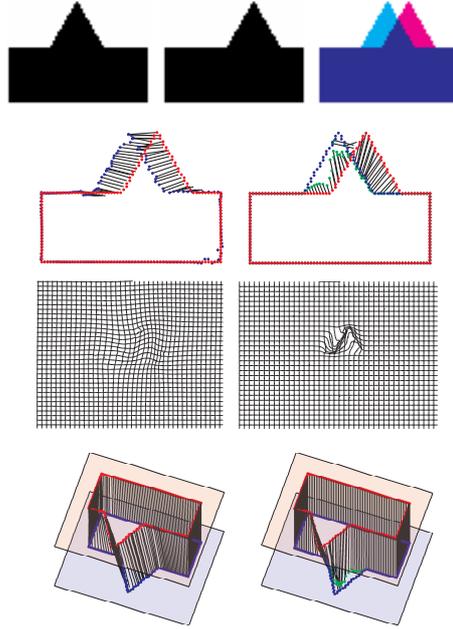
7

Figure 5: Matching results of rectangles with moving spikes. [First row] (left) Rectangle with a spike on the middle; (middle) rectangle with a spike moved to the right; (right) overlap of the two shapes. [Second row] Correspondences along the boundary based on (left) the Gaussian kernel representation; (right) the binary representation. [Third row] Displacement vector field formed by the optimal diffeomorphism based on (left) the Gaussian kernel representation; (right) the binary representation. [Fourth row] Dense correspondences along the boundary obtained by the optimal diffeomorphism based on (left) the Gaussian kernel representation and (right) the binary representation. In the experiments, $\sigma_1 = \sigma_2 = 7, \alpha = 0.0005, \mu = 1, \lambda = 0$ are used and image size is 128×160.

of each shape. For instance, in the case of a rectangle with a spike, even though there is no "right or wrong" way to move the spike (the spike could disappear from the left figure and reappear slightly displaced on the right; or, the spike could simply translate to the right between the two figures), we would like the tip of the spike in one figure to correspond to the tip of the spike in the other. For the case of a human figure, we would like the tip of various limbs to correspond in the two figures. Our representation of shape enables this behavior, since it encodes a local scale-space of curvature (the local shape feature) and attempts to match regions with local features to each other. This can be seen qualitatively in Figs. 5-6.

Figure 6: Matching results of human body with different hand and leg positions. Three scales $\sigma = 12, 10, 7$ are used in a coarse-to-fine manner and $\alpha = 0.0005, \mu = 1, \lambda = 0$. Image size is 128×160. The shape representation used enables different "parts" to be put in correspondence, without performing explicit shape analysis.

4 Non-rigid Template-based Segmentation

In this section we are interested in exploiting the shape representation devised in previous sections to perform bottom-up segmentation under the guidance of a "template." Such a template can be a model shape that is the result of aggregate training data (e.g. a "shape average") or simply a shape obtained through expert knowledge, e.g. by hand. In this sense we think of the template as a shape prior, since it will eventually bias the segmentation towards the given template. Note that this approach is not in contrast with other segmentation techniques based on shape priors, where the entire density is learned, for instance [20, 10]: We encode prior knowledge in a simple template, which makes sense only if the prior is unimodal, and refer the reader to more complex algorithms when it is not.[2]

Our goal is not different from other schemes to include shape priors in segmentation: we want to improve the performance over a purely bottom-up approach. However, in addition to introducing bias in the energy, we also want to establish correspondence between the template and the target shape, so that the ensuing warping is "meaningful" in the sense that we discussed in the introduction. To this end we introduce "local shape preserving diffeomorphisms" and show that they yield improved correspondence and segmentation due to the "shape meaningful" restriction of the allowable space of the template leading to a more compact solution space of the segmentation. From the computational standpoint we work within the level set framework using the Mumford-Shah functional as proposed in [28, 3, 40]. For the segmentation, we use the Chan-Vese energy model [3]:

$$E_{\mathrm{cv}}(\phi; I) = \int_{\Omega} |I(x) - c_1|^2 H(\phi(x)) dx$$
$$+ \int_{\Omega} |I(x) - c_2|^2 (1 - H(\phi(x))) dx + \mu \int_{\Omega} \delta(\phi(x)) |\nabla \phi(x)| dx.$$

where I is image to segment, H is a Heaviside function, δ is a delta function and c_1, c_2 are constant. In order to impose prior shape information, the energy E_{cv} is complemented by a shape energy E_{shape} as follows:

$$E_{\mathrm{seg}}(\phi; I, T) = E_{\mathrm{cv}}(\phi; I) + \beta\, E_{\mathrm{shape}}(h, H(\phi), T).$$

---

[2]Although if a density in the space of shapes is multi-modal, why not further break down the domain and assign an "object" to each mode?

9

where T is the shape prior (template) and E_{shape} is our shape matching energy. The combination of segmentation energy with shape energy has been used with various function spaces in which the transformation h stays [36, 4, 10, 5, 34, 35, 31].

4.1   Feature Preserving Diffeomorphism

In this section we discuss the regularization imposed on the diffeomorphisms. Note that for practical purposes we first perform global registration by restricting h to be affine [37] before activating the full diffeomorphism to deal with deformations. The shape energy E_{shape} incorporated into the segmentation energy penalizes local deviations of the target shape from the template. However, unless properly restricted, a general diffeomorphism does not preserve the local shape of the template, and leads instead to a perfect matching to the target regardless of its shape (provided that it is in the same diffeomorphic equivalence class). Our goal, on the other hand, is to preserve correspondence of local shape during the entire evolution of the diffeomorphism from the template to the target shape. In the shape energy, restriction of the shape space under the diffeomorphism is usually addressed by generic regularizers, such as linear elasticity, that does not preserve local shape. Therefore, we propose an additional regularization term designed to preserve local shape during global deformation. Let us denote by $\overline{T}(x) = T(h(x))$ the deformed template.

$$E_{\text{diff}}(h; T) = \int_\Omega \left[\, \mathcal{R}_\sigma(\overline{T}(.))(x) - \mathcal{R}_\sigma(T(.))(h(x)) \,\right]^2 dx.$$

$$= \int_\Omega \left[\, \mathcal{R}_\sigma(T \circ h(.))(x) - \mathcal{R}_\sigma(T(.))(h(x)) \,\right]^2 dx$$

This energy penalizes a large deformation of the template in terms of the local shape feature. Its gradient is given by:

$$\frac{\partial E_{\text{diff}}}{\partial h} = \left[\, \mathcal{R}_{\sigma_2}(T \circ h(.))(x) - \mathcal{R}_{\sigma_2}(T(.))(h(x)) \,\right]$$
$$\left(\, \nabla T(h(x))\{K_{\sigma_2} * (1 - 2T(h(x)))\} \,\right)$$
$$+ \nabla T(h(x))\{K_{\sigma_2} * \left[\, \mathcal{R}_{\sigma_2}(T \circ h(.))(x) - \mathcal{R}_{\sigma_2}(T(.))(h(x)) \,\right]$$
$$(1 - 2T(h(x)))\}$$

The CV global segmentation model, as it was proposed originally, has the advantages of detecting interior contours, and other contours away from the main front can appear. Moreover, the shape and the topology during the evolution rapidly changes, and at a faster speed than in the classical snakes or gradient based local models. However, in our application of segmentation with shape prior, we would like to keep some of the advantages of the CV model (speed, robustness, simplicity), but we would also like to preserve, as much as possible, the topology of the initial contour (the template) and to stay close to the shape prior during the time evolution.

The segmentation is obtained by minimizing the energy E_{seg} which consists of E_{cv}, E_{shape}, and E_{diff}. The Euler-Lagrange equations yield the gradient direction and the following PDE:

$$\frac{\partial \phi}{\partial t} = -\frac{\partial E_{\text{seg}}}{\partial \phi} = -\left(\frac{\partial E_{\text{cv}}}{\partial \phi} + \beta \, \frac{\partial E_{\text{shape}}}{\partial \phi} + \gamma \, \frac{\partial E_{\text{diff}}}{\partial \phi} \right)$$

$$\frac{\partial E_{\mathrm{cv}}}{\partial \phi} = \delta(\phi)\Big[(I - c_1)^2 - (I - c_2)^2 - \mu \operatorname{div}\Big(\frac{\nabla \phi}{|\nabla \phi|}\Big)\Big]$$

$$\frac{\partial E_{\mathrm{shape}}}{\partial \phi} = \delta(\phi)\Big[\Big(\mathcal{R}_{\sigma_1}(H(\phi)) - \mathcal{R}_{\sigma_2}(T)(h)\Big)\cdot$$
$$\Big(K_{\sigma_1} * (1 - 2H(\phi))\Big)$$
$$+ K_{\sigma_1} * \Big((\mathcal{R}_{\sigma_1}(H(\phi)) - \mathcal{R}_{\sigma_2}(I)(h))(1 - 2H(\phi))\Big)\Big]$$

Note that $\partial E_{\mathrm{diff}}/\partial \phi = 0$ which means the energy E_{diff} only affects the shape matching process.

The final energy $E_{seg}(\phi, h) = E_{\mathrm{cv}} + \beta E_{\mathrm{shape}} + \gamma E_{\mathrm{diff}}$ has to be minimized with respect to the unknowns ϕ and h. We perform this using the so-called alternating minimization. First we write down the coupled system of PDE's:

$$\frac{\partial \phi}{\partial t} = -\frac{\partial E_{\mathrm{seg}}(\phi, h)}{\partial \phi} \tag{1}$$

$$\frac{\partial h}{\partial t} = -\frac{\partial E_{\mathrm{seg}}(\phi, h)}{\partial h}. \tag{2}$$

We start at $t = 0$ with ϕ^0 and h^0. If ϕ^n and h^n have been previously estimated, for $n \geq 0$, then we update ϕ^{n+1} and h^{n+1} in the following way: we compute ϕ^{n+1} using (1) with $h = h^n$ and we compute h^{n+1} using (2) with $\phi = \phi^n$, and we repeat. It is possible to show that the energy $E_{seg}(\phi(\cdot, t), h(\cdot, t))$ decreases in time, if ϕ and h satisfy (1)-(2).

In Figure 7, the segmentation results are presented using E_{cv} (third column), $E_{\mathrm{cv}} +$
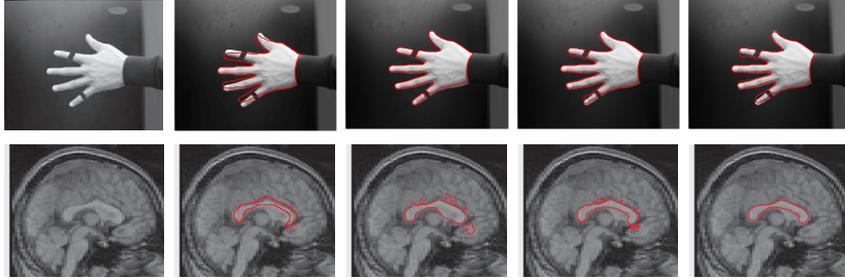

Figure 7: Segmentation results of a partially occluded hand [Top row] and corpus callosum [Bottom row]. (first column) Image with an object to be segmented. (second) Shape prior superimposed on the image. (third) Segmentation E_{cv} without using a shape prior. (fourth) Segmentation with a shape prior introducing additional E_{shape} energy. (fifth) Segmentation with a shape prior introducing both E_{shape} and E_{diff} energy.

βE_{shape} (fourth column) and $E_{\mathrm{cv}} + \beta E_{\mathrm{shape}} + \gamma E_{\mathrm{diff}}$ (fifth column) where the results

with all three terms show significant improvement. This is due to the further restriction of the allowable shape space of the template in the minimization of the segmentation, which helps to avoid local attraction basins following the geodesic of the allowable template shape space. In order to appreciate the difference between the behavior of the different models, we have used the same parameters and initial conditions in all experiments. Due to the additional constraint to the diffeomorphism by E_{diff}, the speed of segmentation is improved and convergence to local minima less frequent.

5 Conclusions and Discussion

We have presented a local shape descriptor and a global shape representation that facilitates registration while matching points with similar local shape. The same representation is used to facilitate segmentation by biasing the evolution of a region towards a template, including explicit correspondence between the two. In the matching process, we have introduced a shape preserving diffeomorphism which restricts the allowable shape space of the template so that the minimization of the segmentation can be improved.

One issue that we have not discussed at length is the choice of scales in the kernels. In our energy functional, the dissimilarity of shapes is measured by the difference between their corresponding shape feature values obtained by integral kernels of which scale is linear to the size of the shape. This requires an optimization procedure for the scale of the integral kernel. However, this can be avoided by employing a statistical measure such as cross correlation as a dissimilarity measure between the feature values as used in [15]. The technique we have introduced can be useful to improve segmentation in medical imaging, but also potentially for warping in computer graphics and for tracking and motion estimation. A statistical measure of local variations of shape also can be obtained by diffeomorphisms from the mean shape [42].

A Energy Minimization

A.1 Shape Energy E_{shape}: minimization with respect to h

We choose a test function $\tilde{h}$ taking real values in the same space with h and $\epsilon > 0$. Let us denote $F(\epsilon) = E_{\text{shape}}(h + \epsilon\tilde{h})$, $F_1(\epsilon) = E_{\text{data}}(h + \epsilon\tilde{h})$, and $F_2(\epsilon) = E_{\text{reg}}(h + \epsilon\tilde{h})$. Then, we will impose $F'(0) = 0$ for any test function $\tilde{h}$. We begin by computing $F_1'(\epsilon)$: (we use a common scale σ)

$$F_1'(\epsilon) = \int_\Omega -\Big(\mathcal{R}_\sigma(S_1) - \mathcal{R}_\sigma(S_2)(h + \epsilon\tilde{h})\Big)\frac{\partial}{\partial\epsilon}\mathcal{R}_\sigma(S_2)(h + \epsilon\tilde{h})dx.$$

$\frac{\partial}{\partial\epsilon}\mathcal{R}_\sigma(S_2)(h + \epsilon\tilde{h}) = \Big[\nabla S_2(\cdot) \cdot (K_\sigma * (1 - S_2))(\cdot)$
$- S_2(\cdot) \cdot (K_\sigma * \nabla S_2)(\cdot) - \nabla S_2(\cdot) \cdot (K_\sigma * S_2))(\cdot)$
$+ S_2(\cdot) \cdot (K_\sigma * \nabla S_2))(\cdot)\Big](h + \epsilon\tilde{h}).$
By setting $\epsilon = 0$, we obtain:
$F_1'(0) = \int_\Omega -\Big(\mathcal{R}_\sigma(S_1) - \mathcal{R}_\sigma(S_2)(h)\Big) \cdot$

$$\left[\nabla S_2 \cdot (K_\sigma * (1 - S_2)) - S_2 \cdot (K_\sigma * \nabla S_2) \right.$$
$$\left. - \nabla S_2(\cdot) \cdot (K_\sigma * S_2)) + S_2 \cdot (K_\sigma * \nabla S_2)) \right](h)dx.$$

For the computation of $F_2'(0)$, refer to [15].

A.2 Feature preserving Energy E_{diff}: minimization with respect to h

In the same way as the previous derivation,

$$F(\epsilon)' = \int_\Omega G \cdot \left[\frac{\partial}{\partial \epsilon} \mathcal{R}_\sigma (T \circ (h(\cdot) + \epsilon \tilde{h}(\cdot)))(x) \right.$$
$$\left. - \frac{\partial}{\partial \epsilon} \mathcal{R}_\sigma (T)(h(\cdot) + \epsilon \tilde{h}(\cdot)))(x) \right] dx.$$
$$G = \mathcal{R}_\sigma (T \circ (h(\cdot) + \epsilon \tilde{h}(\cdot)))(x) - \mathcal{R}_\sigma (T)(h(\cdot) + \epsilon \tilde{h}(\cdot)))(x).$$

$$\frac{\partial}{\partial \epsilon} \mathcal{R}_\sigma (T \circ (h(\cdot) + \epsilon \tilde{h}(\cdot)))(x) = \nabla T(h(\cdot) + \epsilon \tilde{h}(\cdot)))(x) \cdot$$
$$\left\{ K_\sigma * (1 - T(h(\cdot) + \epsilon \tilde{h}(\cdot)))) \right\}(x)\tilde{h}(x) + T(h(\cdot) + \epsilon \tilde{h}(\cdot)))(x) \cdot$$
$$\left\{ K_\sigma * (-\nabla T(h(\cdot) + \epsilon \tilde{h}(\cdot))\tilde{h}(\cdot))(x)) \right\}(x) - \nabla T(h(\cdot) + \epsilon \tilde{h}(\cdot)))(x) \cdot$$
$$\left\{ K_\sigma * (T(h(\cdot) + \epsilon \tilde{h}(\cdot)))) \right\}(x)\tilde{h}(x) + (1 - T(h(\cdot) + \epsilon \tilde{h}(\cdot)))(x)) \cdot$$
$$\left\{ K_\sigma * (\nabla T(h(\cdot) + \epsilon \tilde{h}(\cdot))\tilde{h}(\cdot))(x)) \right\}(x)$$

$$\int_\Omega G(x)T(h(\cdot) + \epsilon \tilde{h}(\cdot))(x) \left[K_\sigma * (-\nabla T(h(\cdot) + \epsilon \tilde{h}(\cdot))\tilde{h}(\cdot))(x)) \right] dx$$
$$= \int_\Omega G(x)T(h(\cdot) + \epsilon \tilde{h}(\cdot))(x) \int K_\sigma(x - y) \cdot$$
$$(-\nabla T(h(\cdot) + \epsilon \tilde{h}(\cdot))(y)\tilde{h}(\cdot))(y))dydx$$
$$= \int -\nabla T(h(\cdot) + \epsilon \tilde{h}(\cdot))(y)\tilde{h}(\cdot))(y) \int K_\sigma(y - x) \cdot$$
$$(G(x)T(h(\cdot) + \epsilon \tilde{h}(\cdot))(x))dxdy$$
$$= \int -\nabla T(h(\cdot) + \epsilon \tilde{h}(\cdot))(y)\tilde{h}(\cdot))(y) \left[K_\sigma(y) * \right.$$
$$\left. (G(y)T(h(\cdot) + \epsilon \tilde{h}(\cdot))(y)) \right] dy$$

The last step holds due to the symmetricity of the kernel K_σ. For the derivation of the other terms, it is similar to the steps above.

References

[1] S. Belongie, J. Malik, and J. Puzicha. Matching shapes. In *Proc. of the IEEE Intl. Conf. on Computer Vision*, 2001. 2

[2] T. K. Carne. The geometry of shape spaces. *Proc. of the London Math. Soc. (3) 61*, 3(61):407–432, 1990. 2

[3] T. Chan and L. Vese. Active contours without edges. *IEEE Trans. Image Processing*, 10(2):266–277, 2001. 9

[4] T. Chan and W. Zhu. Level set based shape prior segmentation. Technical Report 03-66, Computational Applied Mathematics, UCLA, Los Angeles, 2003. 10

[5] T. Chan and W. Zhu. Level set based shape prior segmentation. In *Proc. CVPR'05*, pages 20–26, 2005. 4, 10

[6] G. Christensen, R. Rabbitt, and M. Miller. Deformable templates using large-deformation kinematics. 5(10):1435–1447, 1996. 2

[7] H. Chui and A. Rangarajan. A new algorithm for non-rigid point matching. In *Proc. of the IEEE Intl. Conf. on Comp. Vis. and Patt. Recog.*, pages 44–51, 2000. 2

[8] P. Ciarlet. *Mathematical Elasticity*, volume 1. North Holland, 1988. 6

[9] D. Cremers, T. Kohlberger, and C. Schnörr. Shape statistics in kernel space for variational image segmentation. *Pattern Recognition*, 36(9):1929–1943, Sept. 2003. Special Issue on Kernel and Subspace Methods in Computer Vision. 4, 5

[10] D. Cremers, S. Osher, and S. Soatto. Kernel density estimation and intrinsic alignment for knowledge-driven segmentation: Teaching level sets to walk. In *Proc. of Pattern Recognition*, volume 3175, pages 36–4, 2004. 9, 10

[11] M. Fischler and R. Elschlager. The representation and matching of pictorial structures. *IEEE Transactions on Computers*, 22(1):67–92, 1973. 2

[12] P. Giblin. *Graphs, Surfaces and Homology*. Chapman and Hall, 1977. 1

[13] L. Gorelick, M. Galun, E. Sharon, R. Basri, and A. Brandt. Shape representation and classification using the poisson equation. In *In Proc. of CVPR'04*, pages 61–67, 2004. 5

[14] U. Grenander. *General Pattern Theory*. Oxford University Press, 1993. 2

[15] G. Hermosillo, C. Chefd'hotel, and O. Faugeras. Variational methods for multimodal image matching. *The International Journal of Computer Vision*, 50(3):329–343, Nov. 2002. 6, 12, 13

[16] D. G. Kendall. Shape manifolds, procrustean metrics and complex projective spaces. *Bull. London Math. Soc.*, 16, 1984. 2

[17] B. Kimia, A. R. Tannenbaum, and S. W. Zucker. Shapes, schoks and deformations i: The components of two-dimensional shape and the reaction-diffusion space. *IJCV*, 15:189–224, 1995. 5

[18] M. Lades, C. Borbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. Wurtz, and W. Konen. Distortion invariant object rcognition in the dynamic link architecture. *IEEE Trans. on Computers*, 42(3):300–311, 1993. 2

[19] H. Le and D. G. Kendall. The riemannian structure of euclidean shape spaces: a novel environment for statistics. *The Annals of Statistics*, 21(3):1225–1271, 1993. 2

[20] M. Leventon, E. Grimson, and O. Faugeras. Statistical Shape Influence in Geodesic Active Contours. In *In Proc. CVPR*, pages 316–323, 2000. 5, 9

[21] M. Leventon, E. Grimson, O. Faugeras, S. Wells, and R. Kikinis. Knowledge-based segmentation of medical images. In *Geometric Level Set Methods in IVG*. 2003. 5

[22] T. Lindeberg. Scale space for discrete signals. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(3):234–254, 1990. 2

[23] R. Malladi, J. Sethian, and B. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Trans. on PAMI*, 17:158–175, 1995. 3

[24] S. Manay, B. Hong, A. Yezzi, and S. Soatto. Integral invariant signatures. In *ECCV'04*, volume LNCS 2034, pages 87–99, 2004. 3

[25] K. V. Mardia and I. L. Dryden. Shape distributions for landmark data. *Adv. appl. prob.*, 21(4):742–755, 1989. 2

[26] G. Matheron. *Random Sets and Integral Geometry*. Wiley, 1975. 1

[27] M. Miller and L. Younes. Group action, diffeomorphism and matching: a general framework. In *Proc. of SCTV*, 1999. 6

[28] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42:577–684, 1989. 9

[29] S. Osher and J. Sethian. Fronts propagating with curvature dependent speed: algorithms based on the Hamilton–Jacobi formulation. *J. of Comp. Physics*, 79:12–49, 1988. 3

[30] N. Paragios and R. Deriche. Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Trans. PAMI*, 22:266–280, Mar. 2000. 5

[31] N. Paragios, M. Rousson, and V. Ramesh. Matching distance functions: A shape-to-area variational approach for global-to-local registration. In *Proc. ECCV'02*, volume 2, pages 775–789, May 2002. 4, 10

[32] N. Paragios, M. Rousson, and V. Ramesh. Non-rigid registration using distance functions. *Computer Vision and Image Understanding*, 89:142–165, 2003. 4

[33] A. Pitiot, H. Delingette, and P. Thompson. Learning object correspondences with the observed transport shape measure. In *Prof. IPMI 2003*, 2003. 3

[34] T. Riklin-Raviv, N. Kiryati, and N. Sochen. Unlevel-sets: Geometry and prior-based segmentation. In *ECCV'04*, volume LNCS 2034, pages 50–61, 2004. 10

[35] T. Riklin-Raviv, N. Kiryati, and N. Sochen. Prior-based segmentation by projective registration and level sets. In *ICCV05*, pages I: 204–211, 2005. 10

[36] M. Rousson and N. Paragios. Shape priors for level set representations. In *Proc. of ECCV'02*, pages 78–92, 2002. 10

[37] J. Shi and C. Tomasi. Good features to track. In *Proceedings of CVPR'94*, pages 593–600, 1994. 10

[38] R. Thom. *Structural Stability and Morphogenesis*. Benjamin; Reading, 1975. 1

[39] D. W. Thompson. *On Growth and Form*. Dover, 1917. 1

[40] A. Tsai, A. Yezzi, and A. Willsky. Curve evolution implementation of the mumford-shah functional for image segmentation, denoising, interpolation, and magnification. *IEEE Trans. on IP*, 10:1169–1186, 2001. 5, 9

[41] R. C. Veltkamp and M. Hagedoorn. State of the art in shape matching. Technical Report UU-CS-1999-27, University of Utrecht, 1999. 2

[42] A. Yezzi and S. Soatto. Deformotion: deforming motion, shape average and the joint segmentation and registration of images. *Intl. J. of Comp. Vis.*, 53(2):153–167, 2003. 12

[43] S.-C. Zhu and A. Yuille. Forms: a flexible object recognition and modeling system. *IJCV*, 20:187–212, 1996. 5