Meaningful Correspondence, Local Contex and Shape Statistics in Deformable Templates

Byung-Woo Hong
Stefano Soatto

Computer Science Department
University of California Los Angeles, Los Angeles CA 90024, USA

UCLA CSD Technical Report ♯070007

Abstract

The paper presents a variational framework to compute first and second order statistics of an ensemble of shapes undergoing deformations. Geometrically "meaningful" correspondence between shapes is established via a kernel descriptor that characterizes local shape properties. Such a descriptor allows retaining geometric features such as high-curvature structures in the average shape, unlike conventional methods where the average shape is usually smoothed out by generic regularization terms. The obtained shape statistics are integrated into segmentation as a prior knowledge. The effectiveness of the method is demonstrated through experimental results with synthetic and real images.

1 Introduction

(a) (b) (c)

Figure 1: (a) Examples of 10 overlapped corpus callosum shapes. (b) Shape distribution with mean (black line) and deviation (dots with colors). (c) Segmentation result of corpus callosum in a MRI image using shape information.

Our goal is to compute statistics of an ensemble of shapes that capture their first-order (average) and second-order (deviation) properties. These can be used to define

and learn a prior model on the set of shapes, which in turn is key in applications such as image registration, segmentation, and shape classification. The process is illustrated pictorially in Fig. 1 and is important in medical image analysis and computational anatomy [13]. Along the way, we seek for a way to establish "meaningful" correspondence between any pair of shapes, where "meaning" entails respect of the local shape context as we will make precise shortly. We restrict our attention to planar shapes, represented as planar, closed, simple curves, not necessarily smooth, although some of the ideas naturally extend to three dimensional embedded surfaces.

There are many ways to represent shapes, and to endow them with a metric and probabilistic structure. The simplest is a (finite-dimensional) collection of points modulo the action of a finite-dimensional group. The ensuing theory has been nicely worked out (see [8] and references therein) and the inference algorithms are simple and efficient, but the power of these methods is severely limited in applications where there are no well-defined "landmark points," or there is no time to manually select them, as is the case in the analysis of massive data sets in computational anatomy.

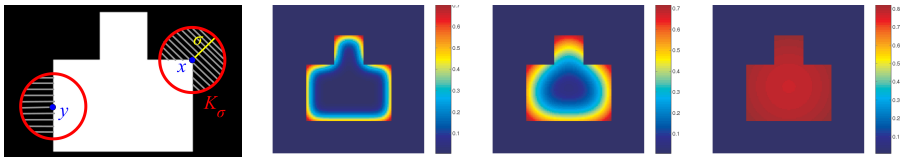
Infinite-dimensional representations of shapes as templates (grayscale or binary images representing continuous closed curves) under the action of infinite-dimensional domain deformations were pioneered by Grenander [12]. The basic premise is that each object of interest (shape) is obtained from a representative element (template) through the transitive action of the deformation group. That is, given a template, any other shape can be "reached" by applying a suitable domain deformation. The set of shapes can then be endowed with a probability structure by placing a measure on the deformation group, which allows defining a "mean," a "covariance function" and a Gibbs-type (Langevin) distribution on shapes. Alternatively, one can define a (cord) distance and define a notion of "average shape" as the element (itself a shape) that minimizes the distance from the ensemble. Such notions of "average" or "mean," however, are not well-defined in absolute terms, since the distance between any two shapes can be made arbitrarily small by a transitive action of the group. To make these notions well-posed one has to introduce regularization on the set of possible domain deformations, and the resulting statistics depend crucially on the choice of regularizer [6, 19]. In [28] the deformation is separated into a finite-dimensional group that acts at zero cost, and a diffeomorphism, whose energy is minimized. Still the resulting average is an expression of the choice of regularizer, which destroys important shape information present in the ensemble data. The problem of a meaningful shape average is bypassed in [5] by defining second-order statistics directly from pairwise distances between any two elements of the ensemble. However, first-order statistics are important in many application domains, so the problem of defining a meaningful average remains.

All these difficulties arise due to the infinite-dimensionality of the representation, and are not present in the finite-dimensional case. The assumption that any shape can be reached by a domain deformation, and the absence of local context information in the representation makes it such that any point in a template can be transformed to any other point in a target shape, and the determination of correspondence is left to generic regularization terms such as elasticity priors. Another issue that is often overlooked in the existing literature is the fact that correspondence, and hence ensemble statistics, are only meaningful when computed relative to a *scale*, so one is left with tweaking the amount of smoothness in the regularizer by guessing the amount of "noise" in the data.

In this paper, we introduce the use of local shape context in the computation of first and second-order ensemble statistics. This is, to the best of our knowledge, novel, and constitutes our contribution. We exploit existing scale-spaces of local features [15] to capture the local shape context. We exploit the resulting notion of "meaningful" correspondence [15] to define and efficiently compute ensemble statistics: We introduce a notion of "meaningful mean" and "sample covariance" of the ensemble, and exploit the linearity of the deformation field to perform principal component analysis. These statistics can be used to define a normal distribution in the set of shapes, which can be used as a prior for segmentation, registration, or classification. Our contribution is the introduction of these statistics, and simple algorithms to compute them, and to infer their values from the ensemble.

In the next subsections we will recall existing results on local features (Sect. 1.1) and meaningful correspondence (Sect. 1.3) in order to introduce our notation and terminology and make the paper self-contained. In the next two sections we introduce the first-order (Sect. 2) and second-order statistics (Sect. 3), and integrate them into segmentation as a prior knowledge (Sect. 4). Then, we will illustrate their use in experiments on real examples (Sect. 5).

1.1 Definitions and Notation



(a) (b) $\sigma = 7$ (c) $\sigma = 15$ (d) $\sigma = 55$

Figure 2: (a) Schematic illustration for the calculation of shape feature. (b)-(d) Examples of feature $\mathcal{R}_\sigma$ of (a) with different scales (image size is 200×200).

In this section we introduce a representation of shape together with a notion of feature that characterizes geometrical properties of the shape up to a given scale. By (planar) *shape* we mean a closed, bounded and connected region $D \subset \mathbb{R}^2$ with finite perimeter, represented by a characteristic function

$$S(x) = S_D(x) = \begin{cases} 1 & \text{if } x \in D, \\ 0 & \text{if } x \notin D \end{cases} \tag{1}$$

defined for $x \in \Omega \subset \mathbb{R}^2$, with $D \subset \Omega$. This can be visualized as a *binary image*, where Ω is the rectangular image domain, as in Fig. 2 (a). By *feature* we mean any statistic of the data, that is any deterministic function(al) of S. In particular, we consider *multi-scale* features $\mathcal{F}$ defined by linear functionals with a family of kernels K indexed by a scale σ: Specifically, $\sigma \in \mathbb{R}^+$, $K : \mathbb{R}^2 \times \mathbb{R}^+ \to \mathbb{R}^+$; $(x, \sigma) \mapsto K_\sigma(x)$. For instance, we consider the isotropic Gaussian kernel

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

For example, a *linear feature* is a functional $\mathcal{F}_\sigma : \mathcal{L}^1(\Omega) \to \mathcal{L}^1(\mathbb{R}^2)$; $S(x)_{|_{x\in\Omega}} \mapsto K_\sigma * S(x) \doteq \mathcal{F}_\sigma(x|S)$ where the convolution is defined by

$$K_\sigma * S(x) = \int_\Omega K_\sigma(x-y)S(y)dy, \ \forall\, x \in \mathbb{R}^2. \tag{2}$$

Notice that the feature changes the domain of definition of S, although for kernels with finite support and Ω is sufficiently bigger than D this is inconsequential. Similar features have been used before in the literature by Thompson et al. [23] and Manay et al. [17], among others [11, 26, 14]. It should be noticed that in the limit where $\sigma \to 0$ they constitute a proper scale-space that can be related to the curvature of the boundary of D, ∂D, albeit it does not entail computation of derivatives (hence the name "integral invariant signatures" sometimes used for $\mathcal{F}_\sigma(x|S)$). However, the scale-space representation is only in the limit, and in general this feature destroys the boundary information present in S.

For these reasons, we prefer to work with a *non-linear feature* designed to retain boundary information:

$$\begin{aligned} \mathcal{R}_\sigma : \mathcal{L}^1(\Omega) &\to \mathcal{L}^1(\mathbb{R}^2) \\ S(x) &\mapsto \mathcal{R}_\sigma(x|S) \doteq S(x)\left(K_\sigma * (1 - S(x))\right). \end{aligned} \tag{3}$$

This feature may seem complicated at first, but it is indeed quite intuitive as illustrated in Fig. 2 (a), and the results with different scales are shown in (b)-(d). It is also more easily computed than other popular representations, for instance the (signed) distance function $d(x) = dist(x, \partial D)$, and captures the "context" of the given shape, in the sense that the value of the feature at a point is a local statistic of the shape in a neighborhood of that point. This can be thought of as extensions of ideas of [1] to infinite-dimensional representations, albeit for complexity reasons only the first order statistic (mean realtive to the kernel K) is computed, rather than the entire distribution (discrete histogram) as in [1].

1.2 Domain Deformations and Distance Between Shapes

Given two shapes, $S_1, S_2 : \Omega \to \{0, 1\}$, under the assumptions above it is possible to transform one into the other by a *warping*, that is a domain deformation $h : \Omega \to \mathbb{R}^2$ such that $h(\Omega) = \Omega$ and

$$S_1(x) = S_2(h(x)), \ \forall\, x \in \Omega. \tag{4}$$

In deformable templates, h is chosen to be an infinite-dimensional group that acts transitively on the set of shapes, so that it is possible to transform any shape (template) into any other (target), that is to achieve the equal sign in (4). The *distance* between shapes is defined as the energy of such a warping. Because h is infinite-dimensional, there are infinitely many warpings that satisfy (4), so the distance is defined as the one that minimizes such an energy in a suitably chosen class [12], for instance $d(S_1, S_2) \doteq \min_h \|h\|$ subject to (4), where h is a diffeomorphism and $\|\cdot\|$

is some chosen norm of integral form. This can be rephrased as a variational problem, for instance

$$d(S_1, S_2) = \inf_{h \in H} \int_\Omega |S_1(x) - S_2(h(x))| dx + \alpha \|h\|_H, \tag{5}$$

where H is some function space, for instance $\mathcal{L}^2(\Omega \to \Omega)$. The first term in the above sum, which we indicate with $E_{\text{data}}(S_1, S_2|h)$, describes the *data fidelity*, that is how well the warped template fits the target, whereas the second is a *regularization term* to render the problem well posed; α is a Lagrange multiplier. The regularization term, which we indicate with $E_{\text{reg}}(h)$, can be taken for example from linear elasticity

$$E_{\text{reg}}(h) = \frac{1}{2} \int_\Omega \left\{ \lambda (\text{div } h)^2 + 2\mu \sum_{i,j=1}^{2} (\epsilon_{ij}(h))^2 \right\} dx, \tag{6}$$

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

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

1.3 "Meaningful" Correspondence

A distance between shapes, as defined above, implicitly determines a *correspondence* between them: a point x in the interior of the set that defines S_1, that is $S_{S_1} = 1$, is mapped to a point $h(x)$ in the interior of the set that defines S_2, $S_{S_2} \circ h(x) = 1$; similarly for the exterior, or *background*, and the boundary of such a set.

Since the data fidelity term can be made zero by infinitely many warpings h, it alone does not establish a unique correspondence, which is therefore determined by the choice of regularization term E_{reg}. In this sense we say that the correspondence determined by a *generic* regularizer (one depending only on h), such as (6), is *meaningless*: choose a different regularizer, you get a different correspondence even for the same data. One would like a *meaningful* correspondence to map corners to corners, and straight segments to straight segments. But the regularizer is "dumb," it does not know about corners or straight edges, and so the resulting correspondence. The same considerations apply to registering grayscale templates, rather than binary ones [19].

The lack of meaning in correspondence is due to the basic premise of deformable templates, that the group is infinite-dimensional and acts transitively on the template. The problem disappears when the group is finite-dimensional, as in [8], or when the data is not in the same orbit[1] of the template [28]. Since we want to maintain the flexibility that comes from an infinite-dimensional warp h (we do not want to constrain it to be finite-dimensional, or to discretize it at the outset), we need to either make the regularizer "smart," i.e. dependent on the data S_1, S_2, or to change the data fidelity term E_{data} so that it can support a meaningful notion of correspondence.[2]

[1] Each template S defines an orbit via the group h, $S \circ h$, that has the structure of an equivalence class.

[2] Furthermore, the notion of correspondence is only meaningful if it is associated to a notion of *scale*, as argued eloquently in [17].

We opt for this second approach, where the "meaning" of correspondence is given by the notion of *local context*, captured by the feature defined in (3). For instance, if the value of the feature at a point $x \in \Omega$ of the template is "distinctive," in the sense that no other point has the same value, it can be uniquely matched on the target. For these "distinctive" points, correspondence is determined by the data fidelity term. The regularizer is only needed to fill in loci of points that have constant feature value, i.e. where the feature is uninformative. Following this rationale, we introduce a new measure of data fidelity between two shapes S_1 and S_2 as

$$E_{\text{shape}}(h|S_1, S_2) = \int_\Omega |\mathcal{R}_\sigma(x|S_1) - \mathcal{R}_\sigma(h(x)|S_2)|^2 dx \tag{7}$$

and define their *meaningful correspondence relative to the feature* $\mathcal{R}_\sigma$ with the minimizer of the ernergy

$$h^* = \arg\min_h E_{\text{shape}}(h|S_1, S_2) + E_{\text{reg}}(h),$$

where E_{reg} is a generic "dumb" regularizer such as (6). In the next section we will show how this definition can be used to compute meaningful shape statistics.

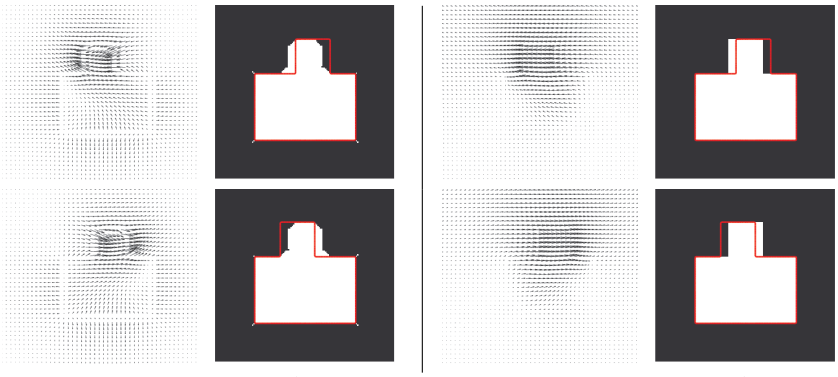
(a) Shape 1 (b) Shape 2 (c) (d)

Figure 3: (a) and (b). Examples of rectangular shapes with bumps at different locations. (c) Average shape without taking into account shape feature. (d) Average shape with taking into account shape feature.

1.4 Noise and Violations of the Model

The premise of deformable templates that each target can be reached from the template by the action of a domain deformation is challenged when the data (images or shapes) are corrupted by noise (e.g. segmentation errors for planar shapes, quantization of the pixel grid) or by unmodeled phenomena (e.g. specularities in grayscale images, occlusions), since the modeling power of the infinite-dimensional group transformation is wasted in over-fitting [12]. This issue is addressed by regularizing the warpings, again at the expense of making the final result crucially dependent on the choice of regularizer.

In our formulation of the problem, where local context features are matched, rather than naked templates, this issue is attacked on two fronts. First, the choice of feature can be made to reduce the effects of specific noise sources, for instance pixelization or irregularity of the contour [17]. Second, matching is associated with a scale σ, and



(a) (b) (c) (d)

Figure 4: Comparison of average shapes obtained from two rectangular shapes with bumps at different locations. (a) Deformation fields that deform original shapes into the average shape using the method without shape feature. (b) Average shape using the method without shape feature superimposed the original shapes (red). (c) Deformation fields that deform original shapes into the average shape using the method with shape feature. (d) Average shape using the method with shape feature superimposed the original shapes (red). (The parameters $\sigma = 7, \alpha = 0.005, \lambda = 0.5, \mu = 1$ are used and image size is 100×100).

therefore one can adaptively search for the best scale in matching pairs. Examples of feature values with varying scales are illustrated in Fig. 2. Note that σ can be thought of as a regularization parameter, but one that is adapted to the data, rather than being a generic smoothness term.

2 Shape Averaging

Given an ensemble of shapes $\{S_1, S_2, \cdots, S_n\}$, one of the simplest statistics of practical interest is their *average shape*[3] defined as yet another shape M that is on average closest to the ensemble. In other words, we look for M that minimizes

$$\sum_{i=1}^{n} d(M, S_i).$$

If we chose as distance $d(M, S) = \min_h \int_\Omega |M(x) - S(h(x))| dx$, without a regularizer, the mean would not be well-defined, because such a distance can be made zero for any shape M.[4] If we regularize the distance, the resulting average will be dependent on the choice of regularizer, hence "meaningless" in the same sense of the correspondence

[3]Note that we do not yet use the notion of *mean shape* since we have not yet defined a distribution of measures on the set of shapes.

[4]This point was also raised by [5], who bypassed it by avoiding a notion of average shape and looking instead at the matrix of pairwise distances $d_{ij} = \tilde{d}(S_i, S_j)$ with respect to a cordal distance $\tilde{d}$.

discussed above, and indeed notice that the determination of the mean also defines its correspondence to each of the given shapes S_i. Therefore, we adopt a new definition of average shape *relative to the feature* $\mathcal{R}_\sigma$ as follows.

Since we are looking for a shape M represented by a binary image, in order to guarantee its structure we write it explicitly as a Heaviside function $M = H(\phi)$ where $H(\phi) : \Omega \to \{0, 1\}$ and $\phi : \Omega \to \mathbb{R}$ is an implicit representation of M: $\phi \geq 0$ inside the average shape or on its boundary, and $\phi < 0$ outside the average shape. The average shape is then defined by the function ϕ.

In order to find the shape average, therefore, we minimize the energy

$$E = \sum_{i=1}^{n} \int_\Omega |\mathcal{R}_\sigma(x|H(\phi)) - \mathcal{R}_\sigma(h(x)|S_i)|^2 dx + \alpha E_{\text{reg}}(h), \tag{8}$$

with respect to ϕ and $h_1, \ldots, h_n$, where

$$\mathcal{R}_\sigma(x|H(\phi)) = H(\phi)(x)(K_\sigma * (1 - H(\phi))(x).$$

We do so via alternating minimization and gradient descent. The associated Euler-Lagrange equations are solved iteratively by first fixing h_i's and performing one step in the opposite direction of the gradient of E with respect to ϕ:

$$\begin{aligned}\frac{\partial \phi}{\partial t} = \delta(\phi)\Big\{ &(K_\sigma * (1 - H(\phi))) \cdot \sum_{i=1}^{n} \big[\mathcal{R}_\sigma(H(\phi); x) \\ &- \mathcal{R}_\sigma(S_i; h_i(x))\big] - K_\sigma * \big[H(\phi) \cdot \sum_{i=1}^{n} \big(\mathcal{R}_\sigma(H(\phi); x) \\ &- \mathcal{R}_\sigma(S_i; h_i(x)))\big]\Big\},\end{aligned}$$

where $t \geq 0$ is the iteration index. Then, fixing ϕ, we perform an iteration in the opposite direction of the gradient of E with respect to $h_i = (h_{i,1}, h_{i,2})$:

$$\begin{aligned}\frac{\partial h_i}{\partial t} = &(\mathcal{R}_\sigma(H(\phi); x) - \mathcal{R}_\sigma(S_i; h_i(x))) \\ &\cdot \Big\{ - (\nabla S_i \circ h_i) \cdot (K_\sigma * (1 - S_i \circ h_i)) \\ &+ (S_i \circ h_i) \cdot (K_\sigma * \nabla S_i) \circ h_i \Big\} \\ &+ \Big(\mu \triangle h + (\lambda + \mu) \nabla(\text{div } h)\Big).\end{aligned}$$

We initialize the iteration with a generic shape average (e.g. a circle) or with the average of the signed distance functions of S_i's as in [16].

3 Shape Variations

In order to analyze the variability of an ensemble $\{S_1, S_2, \cdots, S_n\}$ relative to the average shape M we compute second-order statistics by finding the principal modes

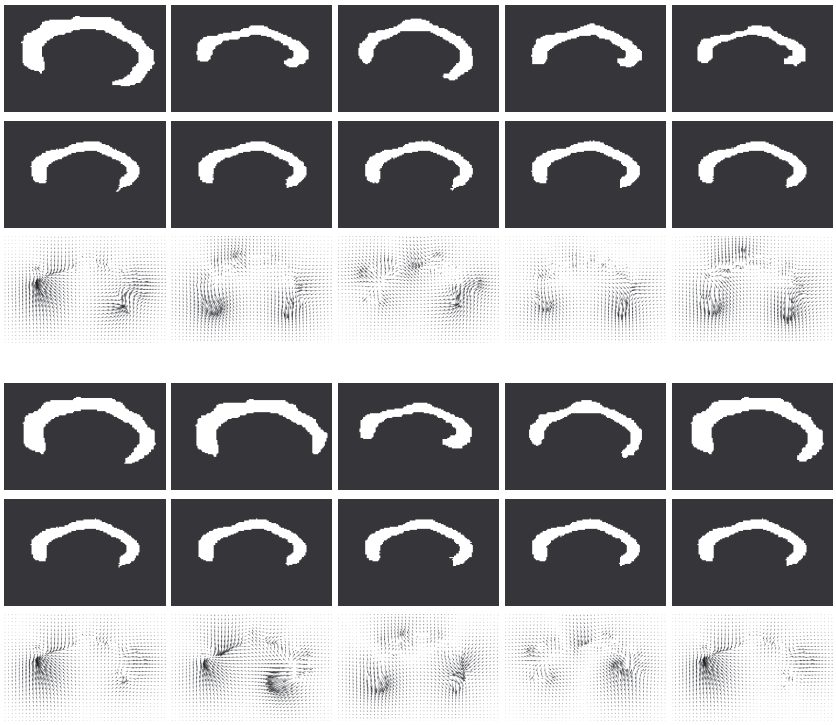


Figure 5: Average shape results obtained from 10 different corpus callosum shapes using our method. (First row) Original shapes. (Second row) Warped shapes to the average shape. (Third row) Deformation fields that represent the warping of the shapes to the average shape. (The parameters $\sigma = 7, \alpha = 0.01, \lambda = 0.5, \mu = 1$ are used and image size is 120×200).

of the sample covariance. Because the ensemble is obtained from the average via a warping, $S_i(h_i) \simeq M$, and warpings are vector fields that can be composed linearly, this can be performed by simple linear statistical analysis, that is principal component analysis (PCA).

This linear compositionality is one of the key advantages of working within the framework of deformable templates. Other shape representations, for instance signed-distance functions, are not closed under linear operations.

In order to perform PCA we first compute the average deformation field:

$$\mu = \frac{1}{n} \sum_{i=1}^{n} h_i.$$

Then, the sample covariance of the deformation fields is constructed by:

$$\Sigma = \frac{1}{n} \sum_{i=1}^{n} (h_i - \mu)(h_i - \mu)^T.$$

The principal modes of variation are the eigenvectors v_i of the covariance Σ, and a deformation h can be approximated using the first k eigenvectors v_i, $i = 1, \ldots, k$ corresponding to the largest eigenvalues λ_i by:

$$h = \mu + \sum_{i=1}^{k} \alpha_i v_i$$

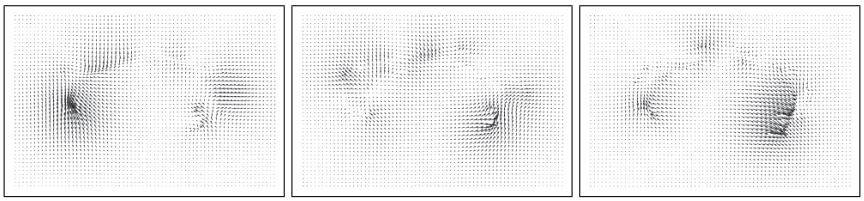
where $\alpha_i \in \mathbb{R}$ are weights for the different modes. An instance of a shape S within the ensanble can be approximated by projection onto the principal modes along with the representative average shape, that is by finding the coefficients α_i such that

$$S \simeq M(\mu + \sum_{i=1}^{k} \alpha_i v_i)$$

in the $\mathcal{L}^2$ norm. In the experimental section we will show the average shape and principal modes of variation of a number of real ensenbles of shapes.

The shape average and principal modes of variation, together with the linear structure of the deformation fields, can be used to define a Gaussian distribution in an ensenble of shapes: $\{S_i\}_{i=1,\ldots,n} \sim \mathcal{N}(\mu, \Sigma)$, with μ and Σ defined above. This can be used as a prior for segmentation as illustrated in Sect. 4.

Note that the definition of a shape average and covariance via linear statistical analysis of diffeomorphic domain deformations is not new, and in fact is one of the staples of deformable templates. What is different here is that such deformations are computed not based on the naked template, but based on the feature $\mathcal{R}_\sigma$, which includes a notion of scale and induces a correspondence based on local context.



v_1 v_2 v_3

Figure 6: Three principal modes of variation obtained from the warping in Fig. 5.

4 Segmentation using a shape prior

Analysis of medical images often requires the detection of anatomical structures. The process of "segmenting" these structures from the images using only low-level information (e.g. intensity values) is often unsatisfactory because of low contrast, noise and the intrinsic variability of the target shape, so top-down information, in the form of prior knowledge on the anatomical structure of interest learned from examples, can greatly benefit the analysis [30, 31, 9].

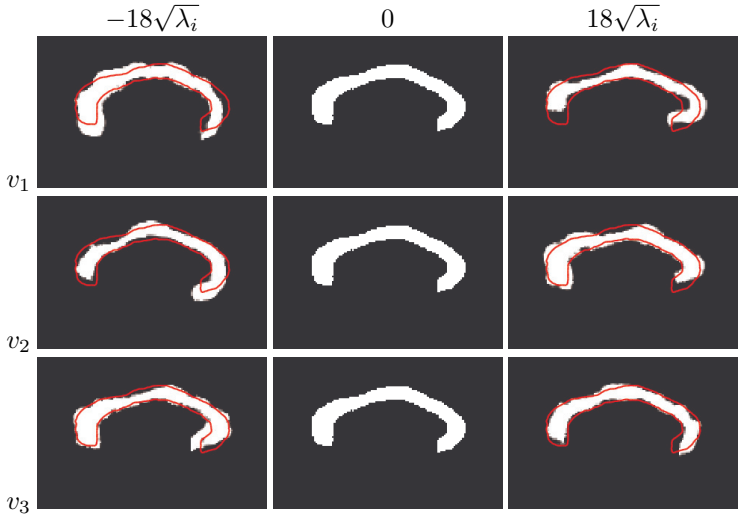


Figure 7: Shape statistics from the shapes shown in Fig. 5. Three principal modes v_i, $i = 1, 2, 3$ shown in Fig. 6 are used with weights $\alpha_i \in \{-18\sqrt{\lambda_i}, 0, 18\sqrt{\lambda_i}\}$. $\alpha = 0$ means the average.

Incorporating prior knowledge in segmentation is an important area of work in medical imaging with many significant contributions. In this section we explore the use of our representation and shape statistics in segmentation using the level set framework [21] as commonly done in segmentation [2, 22, 3, 29, 27]. We evolve the boundary of the shape via an embedding function ϕ, which is a convenient mean to compute our feature, and also allows easy computation of a measure of dissimilarity to the target shape [25].

The basic underlying model we use is [20] and the numerical scheme is [3]. Our prior is essentially a Gaussian density learned using PCA from a set of training shapes, incorporated into level-set segmentation, following the line of work pioneered by [29, 16, 7, 18]. Our goal here is to illustrate the role of the shape context, hence we keep the statistical model as simple as possible. Finite-dimensional models for registration have also been used in [24, 4].

To illustrate our approach, assume that we have a training set of shapes $S_1, ..., S_n$, and compute their average shape M as in Sect. 2. As a byproduct, we get the deformations $h_1, ..., h_n$ that warp each S_i into the average shape M. Following the derivation in Sect. 3 we construct a deformable template $M(\mu + \sum_{i=1}^{k} \alpha_i v_i)$, allowing the scalar weights α_i to vary, as illustrated in Figures 7 and 10. Now we wish to detect a shape defined by the unknown binary function $H(\phi)$ in a given image I, such that $H(\phi)$ is also close to $M(\mu + \sum_{i=1}^{k} \alpha_i v_i)$. We consider the energy function

$$\inf_{\phi, \alpha_i} E(\phi, \alpha_i) = E_{\text{seg}}(\phi) + \beta E_{\text{prior}}(\phi, \alpha_i),$$

where for the first term we use [3] as given by:

$$E_{\text{seg}}(\phi) = \gamma \int_\Omega |\nabla H(\phi)| dx$$
$$+ \int_\Omega |I(x) - c^+|^2 H(\phi) dx + \int_\Omega |I(x) - c^-|^2 H(-\phi) dx,$$

where γ is a constant and

$$E_{\text{prior}}(\phi, \alpha_i) = \int_\Omega \Big| H(\phi) - M(\mu + \sum_{i=1}^{k} \alpha_i v_i) \Big|^2 dx,$$

where α_i is optimized within the range $[-\eta\sqrt{\lambda_i}, \eta\sqrt{\lambda_i}]$ for a constant η as in [10] where they used $\eta = 3$.

We minimize $E(\phi, \alpha_i)$ by an alternating procedure using the Euler-Lagrange equations and time-dependent regularization $t > 0$:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \Big[\gamma \text{div}\Big(\frac{\nabla \phi}{|\nabla \phi|}\Big) - |I - c^+|^2 + |I - c^-|^2$$
$$-2\beta \Big(H(\phi) - M(\mu + \sum_{i=1}^{k} \alpha_i v_i) \Big) \Big],$$

and for $i = 1, ..., k$

$$\frac{\partial \alpha_i}{\partial t} = 2\beta \int_\Omega (H(\phi) - M(\mu + \sum_{i=1}^{k} \alpha_i v_i))(\nabla M \cdot v_i) dx,$$

where ∇M is evaluated at $\mu + \sum_{i=1}^{k} \alpha_i v_i$.

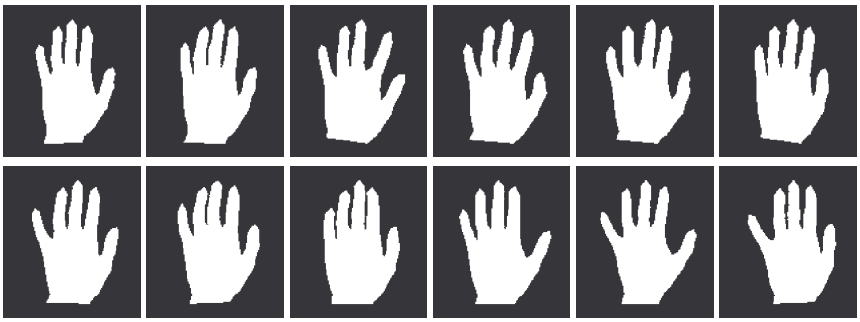


Figure 8: A set of 12 different hand shapes.

5 Experiments

In this section, we present experimental results to illustrate the performance of our model in matching, in the computation of shape statistics, and in their exploitation

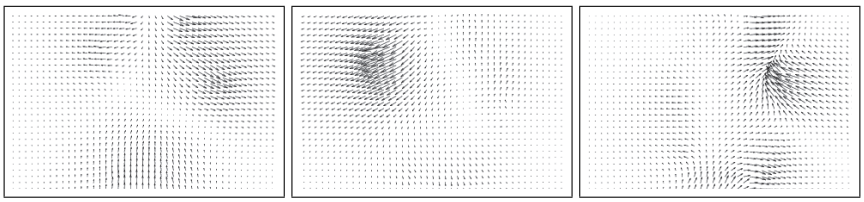


Figure 9: Three principal modes of variation obtained from the warping in Fig. 8.

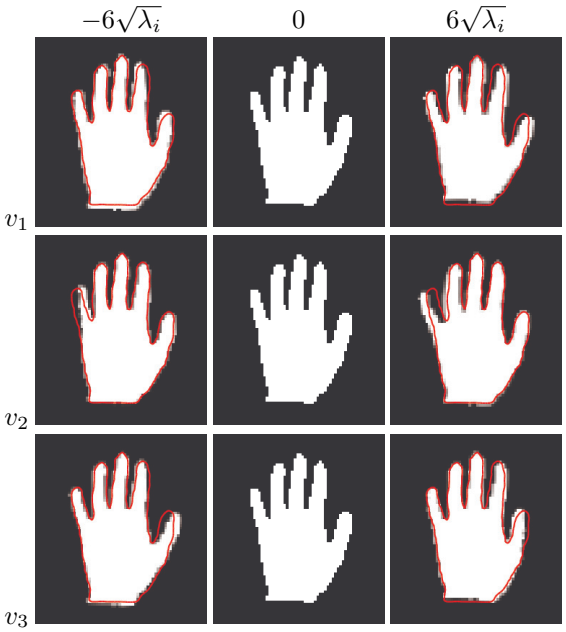


Figure 10: Shape statistics from the shapes shown in Fig. 8. Three principal modes v_i, $i = 1, 2, 3$ shown in Fig. 9 are used with weights $\alpha_i \in \{-6\sqrt{\lambda_i}, 0, 6\sqrt{\lambda_i}\}$. $\alpha = 0$ means the average.

to perform segmentation. We begin with simple but demonstrative examples using synthetic shapes that have distinctive geometric features, then move on to real images.

In Fig. 4 we compare our proposed method with a conventional scheme without shape context to compute the average shape of two rectangular objects with "bumps" at different locations. The average shape obtained by conventional warping is affected by the choice of regularizer, and has therefore smooth corners; our average shape, instead, preserves significant geometric features that are present in the original shapes. This is due to the meaningful correspondence enforced by matching context features, instead of naked templates.

We now apply the proposed method to compute statistics of an ensemble of real data of corpi callosi from 10 different individuals shown in the first row of Fig. 5. We

first compute the average shape shown at the middle column in Fig. 7, together with the warped shapes in the second row in Fig. 5 and the deformation fields between each shape and the average are presented in the third row of Fig. 5. We apply PCA to the obtained deformation fields and the first three principal modes are shown in Fig. 6. In Fig. 7 we illustrate the shape variability captured by the second-order statistics relative to the average (middle column) via the variations (first and last columns) based on varying weights $\alpha_i \in \{-18\sqrt{\lambda_i}, 18\sqrt{\lambda_i}\}$ of the first three principal modes v_i, $i = 1, 2, 3$.

We repeat these experiments on 12 different hand shapes shown in Fig. 8. The average shape is computed together with the domain deformations. The three principal modes of variation v_1, v_2, v_3 are shown in Fig. 9. In Fig. 10, we illustrate the variability from the average shape by changing the weights α_i for v_i, with $\alpha_i \in \{-6\sqrt{\lambda_i}, 6\sqrt{\lambda_i}\}$.

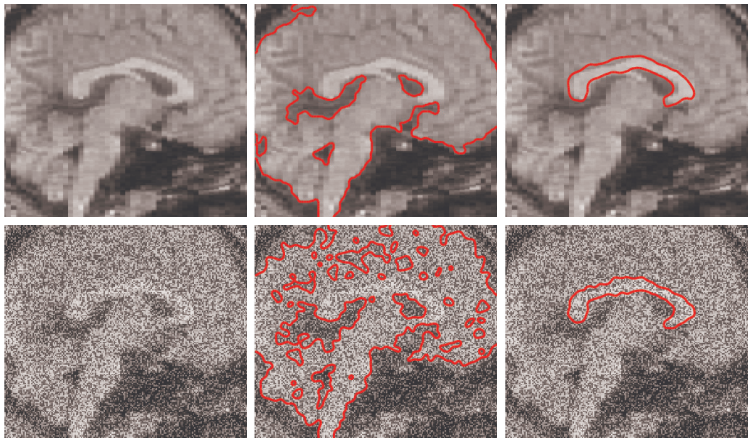


Figure 11: Segmentation results on a brain MRI image (top row) and its noisy version (bottom row). (Left column) Source images. (Middle column) Segmentations without shape prior. (Right column) Segmentations with shape prior. The average shape shown in Fig. 7 and the three principal modes shown in Fig. 6 are used for the shape prior model. (The parameters $\gamma = 10^5$, $\beta = 4500$ (top), $\beta = 5000$ (bottom) are used and image size is 200×200).

Finally, in Fig. 11 we illustrate the use of our prior for segmentation using the average shape and the three principal modes as a shape prior. The corpus callosum in the original data does not exhibit enough contrast to allow successful segmentation using only low-level (intensity) data. However, using a shape prior allows successful segmentation and adaptation to the target shape.

6 Conclusions and Discussion

We have exploited a local feature that captures shape context to establish meaningful correspondence between shapes in an ensemble, and used that to compute shape statistics of the first (average) and second order (variance). These are computed using

infinite-dimensional deformation fields and an integral kernel that is robust to noise and deviation from the ideal deformable template model. Also, we have computed the statistics by performing principal component analsyis on the set of deformation fields from each shape to the average. Satisfactory experimental results on real and synthetic data have been shown to validate the proposed method when compared with analysis of the naked templates. We also have introduced a segmentation algorithm using a shape prior that is modeled by the proposed shape statistics via meaningful correspondences and shown promising segmentation results.

References

[1] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *PAMI*, 24, 2002. 4

[2] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *ICCV*, pages 694–699, 1995. 11

[3] T. Chan and L. Vese. Active contours without edges. *IEEE Trans. Image Processing*, 10(2):266–277, 2001. 11, 12

[4] T. Chan and W. Zhu. Level set based shape prior segmentation. In *Proc. CVPR'05*, pages 1164–01170, 2005. 11

[5] G. Charpiat, O. Faugeras, and R. Keriven. Image statistics based on diffeomorphic matching. In *ICCV*, 2005. 2, 7

[6] G. Christensen, R. Rabbitt, and M. Miller. Deformable template using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1437–1447, 1996. 2

[7] D. Cremers, T. Kohlberger, and C. Schnörr. Shape statistics in kernel space for variational image segmentation. *Pattern Recognition*, 36(9):1929–1943, 2003. 11

[8] I. Dryden and K. Mardia. *Statistical Shape Analysis*. John Wiley & Son, 1998. 2, 5

[9] P. Felzenszwalb. Representation and detection of deformable shapes. *IEEE Trans. on PAMI*, 27, 2005. 10

[10] T. Fletcher, C. Lu, S. M. Pizer, and S. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE TMI*, 23(8):995–1005, 2004. 12

[11] 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. 4

[12] U. Grenander. *General Pattern Theory*. Oxford University Press, 1993. 2, 4, 6

[13] U. Grenander and M. Miller. Computational anatomy: an emerging discipline. *Quart. Appl. Math.*, 56(4), 1998. 2

[14] H. Guo, A. Rangarajan, S. Joshi, and L. Younes. Non-rigid registration of shapes via diffeomorphic point matching. In *ISBI'04*, 2004. 4

[15] B.-W. Hong, E. Prados, S. Soatto, and L. Vese. Shape representation based on integral kernels: Application to image matching and segmentation. In *Proc. CVPR'06*, 2006. 3

[16] M. Leventon, E. Grimson, and O. Faugeras. Statistical Shape Influence in Geodesic Active Contours. In *In Proc. CVPR*, pages 316–323, 2000. 8, 11

[17] S. Manay, D. Cremers, B.-W. Hong, A. Yezzi, and S. Soatto. Shape matching via integral invariants'. *PAMI*, 28(10):1602–1617, 2006. 4, 5, 6

[18] R. Mikaël, P. Nikos, and D. Rachid. Active shape models from a level set perspective. *TR 4984 INRIA*, 2003. 11

[19] M. Miller and L. Younes. Group actions, homeomorphisms, and matching : A general framework. *IJCV*, 41, 2001. 2, 5

[20] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. on Pure and Applied Math.*, 42:577–684, 1989. 11

[21] S. Osher and J. Sethian. Fronts propagation with curvature dependent speed: Algorithms based on hamilton jacobi formulations. *J. of Comp. Phys.*, 1988. 11

[22] N. Paragios and R. Deriche. Geodesic active regions and level set methods for supervised texture segmentation. *IJCV*, 46(3):223–247, 2002. 11

[23] A. Pitiot, H. Delingette, and P. Thompson. Learning object correspondences with the observed transport shape measure. In *Prof. IPMI 2003*, 2003. 4

[24] T. Riklin-Raviv, N. Kiryati, and N. Sochen. Unlevel-sets: Geometry and prior-based segmentation. In *ECCV*, 2004. 11

[25] M. Rousson and N. Paragios. Shape priors for level set representations. In *Proc. of ECCV'02*, pages 78–92, 2002. 11

[26] T. Sebastian, P. Klein, and B. Kimia. On aligning curves. *IEEE Trans. on PAMI*, 25, 2003. 4

[27] K. Siddiqi, Y. Berube, A. Tannenbaum, and S. Zucker. Area and length minimizing flows for shape segmentation. *IEEE Trans. on IP*, 7, 1998. 11

[28] S. Soatto and A. Yezzi. DEFORMOTION, deforming motion, shape average and the joint registration and segmentation of images. In *Proc. ECCV'02*, pages 32–47. 2, 5

[29] A. Tsai, A. Yezzi, et al. Model-based curve evolution technique for image segmentation. In *Proc. CVPR'01*, 2001. 11

[30] B. Vemuri, J. Ye, Y. Chen, and C. Leonard. Image registration via level-set motion: applications to atlas-based segmentation. *MIA*, 7:1–20, 2003. 10

[31] T. Yu, J. Luo, A. Singhal, and N. Ahuja. Shape regularized active contour based on dynamic programming for anatomical structure segmentation. In *SPIE Med Im*, 2005. 10