DEFORMOTION
Deforming Motion, Shape Average and the Joint Registration and Segmentation of Images*

Stefano Soatto[1] and Anthony J. Yezzi[2]

[1] Department of Computer Science, University of California, Los Angeles, Los Angeles – CA 90095
soatto@ucla.edu
[2] Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta – GA 30332
ayezzi@ece.gatech.edu

Abstract. What does it mean for a deforming object to be "moving" (see Fig. 1)? How can we separate the overall motion (a finite-dimensional group action) from the more general deformation (a diffeomorphism)? In this paper we propose a definition of motion for a deforming object and introduce a notion of "shape average" as the entity that separates the motion from the deformation. Our definition allows us to derive novel and efficient algorithms to register non-equivalent shapes using region-based methods, and to simultaneously approximate and register structures in grey-scale images. We also extend the notion of shape average to that of a "moving average" in order to track moving and deforming objects through time.

Fig. 1. *A jellyfish is* "moving while deforming." *What exactly does this mean? How can we separate its "global" motion from its "local" deformation?*

1 Introduction

Consider a sheet of paper falling. If it were a rigid object, one could describe its *motion* by providing the coordinates of one particle and the orientation of an orthonormal reference frame attached to that particle. That is, 6 numbers

* This research is supported in part by NSF grant IIS-9876145, ARO grant DAAD19-99-1-0139 and Intel grant 8029.

A. Heyden et al. (Eds.): ECCV 2002, LNCS 2352, pp. 32–47, 2002.

would be sufficient to describe the object at any instant of time. However, being a non-rigid object, in order to describe it at any instant of time one should really specify the trajectory of each individual particle on the sheet. That is, if γ_0 represents the initial collection of particles, one could provide a function f that describes how the entire set of particles evolves in time: $\gamma_t = f(\gamma_0, t)$. Indeed, if each particle can move independently, there may be no notion of "overall motion," and a more appropriate description of f is that of a *"deformation"* of the sheet. That includes as a special case a rigid motion, described collectively by a rotation matrix[1] $R(t) \in SO(3)$ and a translation vector $T(t) \in \mathbb{R}^3$, so that $\gamma_t = f(\gamma_0, t) = R(t)\gamma_0 + T(t)$ with $R(t)$ and $T(t)$ independent of the particle in γ_0. In practice, however, that is *not* how one usually describes a sheet of paper falling. Instead, one may say that the sheet is "moving" downwards along the vertical direction while "locally deforming." The jellyfish in Fig. 1 is just another example to illustrate the same issue.

But what does it even mean for a deforming object to be "moving"? From a mathematical standpoint, rigorously defining a notion of motion for deforming objects presents a challenge. In fact, if we describe the deformation f as the composition of a rigid motion $(R(t), T(t))$ and a "local deformation" function $h(\cdot, t)$, so that $\gamma_t = h(R(t)\gamma_0 + T(t), t)$, we can always find infinitely many different choices $\tilde{h}(\cdot, t), \tilde{R}(t), \tilde{T}(t)$ that give rise to the same overall deformation f: $\gamma_t = f(\gamma_0, t) = h(R(t)\gamma_0 + T(t), t) = \tilde{h}(\tilde{R}\gamma_0 + \tilde{T}(t), t)$ by simply choosing $\tilde{h}(\gamma, t) \doteq h(R\tilde{R}^T(\gamma - \tilde{T}) + T, t)$ for any rigid motion $(\tilde{R}, \tilde{T})$. Therefore, we could describe the motion of our sheet with (R, T) as well as with $(\tilde{R}, \tilde{T})$, which is arbitrary, and in the end we would have failed in defining a notion of "motion" that is unique to the event observed. So, how can we define a notion of motion for a deforming object in a mathematically sound way that reflects our intuition? For instance, in Fig. 6, how do we describe the "motion" of a jellyfish? Or in Fig. 5 the "motion" of a storm? In neuroanatomy, how can we "register" a database of images of a given structure, say the corpus callosum (Fig. 9), by "moving" them to a common reference frame?

All these questions ultimately boil down to an attempt to *separate the overall motion from the more general deformation.* Before proceeding, note that this is not always possible or even meaningful. In order to talk about the "motion" of an object, one must assume that *"something"* of the object is preserved as it deforms. For instance, it may not make sense to try to capture the "motion" of a swarm of bees, or of a collection of particles that indeed all move independently. What we want to capture mathematically is the notion of overall motion when indeed there is one that corresponds to our intuition! The key to this paper is the observation that the notion of motion and the notion of shape are very tightly coupled. Indeed, we will see that the *shape average is exactly what allows separating the motion from the deformation.*

[1] $SO(3)$ denotes the set of 3×3 orthogonal matrices with unit determinant.

1.1 Prior Related Work

The study of shape spans at least a hundred years of research in different communities from mathematical morphology to statistics, geology, neuroanatomy, paleontology, astronomy etc. In statistics, the study of "Shape Spaces" was championed by Kendall, Mardia and Carne among others [15,20,8,24,9]. Shapes are defined as the equivalence classes of N *points* in $\mathbb{R}^M$ under the similarity group[2], $\mathbb{R}^{MN}/\{SE(M) \times \mathbb{R}\}$. Although the framework clearly distinguishes the notion of "motion" (along the fibers) from the "deformation" (across fibers), the analytical tools are essentially tied to the point-wise representation. One of our goals in this paper is to extend the theory to smooth curves, surfaces and other geometric objects that do not have distinct "landmarks."

In computer vision, a wide literature exists for the problem of "matching" or "aligning" objects based on their images, and space limitations do not allow us to do justice to the many valuable contributions. We refer the reader to [32] for a recent survey. A common approach consists of matching collections of points organized in graphs or trees (e.g. [19,11]). Belongie et al. [4] propose comparing planar contours based on their "shape context." The resulting match is based on "features" rather than on image intensity directly, similar to [10].

"Deformable Templates," pioneered by Grenander [12], do not rely on "features" or "landmarks;" rather, images are directly deformed by a (possibly infinite-dimensional) group action and compared for the best match in an "image-based" approach [34]. There, the notion of "motion" (or "alignment" or "registration") coincides with that of deformation, and there is no clear distinction between the two [5].

Another line of work uses variational methods and the solution of partial differential equations (PDEs) to model shape and to compute distances and similarity. In this framework, not only can the notion of alignment or distance be made precise [3,33,25,28], but quite sophisticated theories that encompass perceptually relevant aspects can be formalized in terms of the properties of the evolution of PDEs (e.g. [18,16]). Kimia et al. [16] describes a scale-space that corresponds to various stages of evolution of a diffusing PDE, and a "reacting" PDE that splits "salient parts" of planar contours by generating singularities. The variational framework has also proven very effective in the analysis of medical images [23,30,22]. Although most of these ideas are developed in a deterministic setting, many can be transposed to a probabilistic context [35,7,31]. Scale-space is a very active research area, and some of the key contributions as they relate to the material of this paper can be found in [14,29,17,1,2] and references therein.

The "alignment," or "registration," of curves has also been used to define a notion of "shape average" by several authors (see [21] and references therein). The shape average, or "prototype," can then be used for recognition in a nearest-neighbor classification framework, or to initialize image-based segmentation by providing a "prior." Leventon et al. [21] perform principal component analysis

[2] $SE(M)$ indicates the Euclidean group of dimension M.

in the aligned frames to regularize the segmentation of regions with low contrast in brain images.

Also related to this paper is the recent work of Paragios and Deriche, where active regions are tracked as they "move." In [27] the notion of motion is not made distinct from the general deformation, and therefore what is being tracked is a general (infinite-dimensional) deformation. Our aim is to define tracking as a trajectory on a finite-dimensional group, despite infinite-dimensional deformations. Substantially different in methods, but related in the intent, is the work on stochastic filters for contour tracking and snakes (see [6] and references therein).

Our framework is designed for objects that undergo a distinct overall "global" motion while "locally" deforming. Under these assumptions, our contribution consists of a novel definition of motion for a deforming object and a corresponding definition of shape average (Sect. 2). Our definition allows us to derive novel and efficient algorithms to register non-identical (or non-equivalent) shapes using region-based methods (Sect. 5). We use our algorithms to simultaneously approximate and register structures in images, or to simultaneously segment and calibrate images (Sect. 6). In the context of tracking, we extend our definition to a novel notion of "moving average" of shape, and use it to perform tracking for deforming objects (Sect. 4). Our definitions do not rely on a particular representation of objects (e.g. explicit vs. implicit, parametric vs. non-parametric), nor on the particular choice of group (e.g. affine, Euclidean), nor are they restricted to a particular modeling framework (e.g. deterministic, energy-based vs. probabilistic). For the implementation of our algorithms on deforming contours, we have chosen an implicit non-parametric representation in terms of level sets, and we have implemented numerical algorithms for integrating partial differential equations to converge to the steady-state of an energy-based functional. However, these choices can be easily changed without altering the nature of the contribution of this paper. Naturally, since shape and motion are computed as the solution of a nonlinear optimization problem, the algorithms we propose are only guaranteed to converge to local minima and, in general, no conclusions can be drawn on uniqueness. Indeed, it is quite simple to generate pathological examples where the setup we have proposed fails. In the experimental section we will highlight the limitations of the approach when used beyond the assumptions for which it is designed.

2 Defining Motion and Shape Average

The key idea underlying our framework is that the notion of *motion* throughout a deformation is very tightly coupled with the notion of *shape average*. In particular, if a deforming object is recognized as moving, there must be an underlying object (which will turn out to be the shape average) moving with the same motion, from which the original object can be obtained with minimal deformations. Therefore, we will model a general deformation as the composition of a group action g on a particular object, on top of which a local deformation is applied. The shape average is defined as the one that minimizes such deformations.

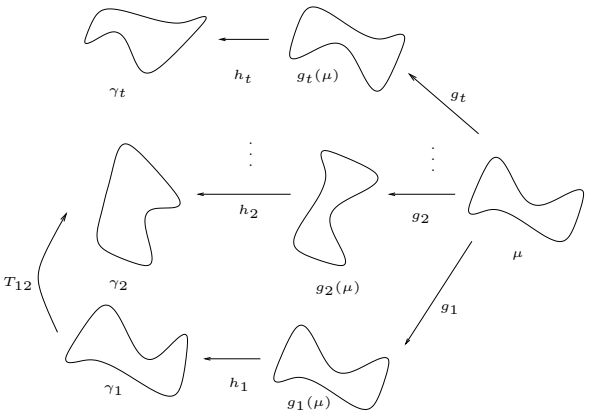

Fig. 2. *A model (commutative diagram) of a deforming contour.*

Let $\gamma_1, \gamma_2, \ldots, \gamma_n$ be n "shapes" (we will soon make the notion precise in Def. 1). Let the map between each pair of shapes be T_{ij}

$$\gamma_i = T_{ij}\gamma_j, \ i, j = 1 \ldots n. \tag{1}$$

It comprises the action of a group $g \in G$ (e.g. the Euclidean group on the plane $G = SE(2)$) and a more general transformation h that belongs to a pre-defined class $\mathcal{H}$ (for instance diffeomorphisms). The deformation h is not arbitrary, but depends upon another "shape" μ, defined in such a way that

$$\gamma_i = h_i \circ g_i(\mu), \ i = 1 \ldots n. \tag{2}$$

Therefore, in general, following the commutative diagram of Fig. 2, we have that

$$T_{ij} \doteq h_i \circ g_i \circ g_j^{-1}(\mu) \circ h_j^{-1} \tag{3}$$

so that $g = g_i g_j^{-1}$ and h is a transformation that depends on h_i, h_j and μ. Given two or more "shapes" and a cost functional $E : \mathcal{H} \to \mathbb{R}^+$ defined on the set of diffeomorphisms, the motion g_t and the shape average are defined as the minimizers of $\sum_{t=1}^{n} E(h_t)$ subject to $\gamma_t = h_t \circ g_t(\mu)$. Note that the only factors which determine the cost of h_t are the "shapes" before and after the transformation, $\mu_i \doteq g_i(\mu)$ and γ_i, so that we can write, with an abuse of notation, $E(h(\mu_i, \gamma_i)) \doteq E(\mu_i, \gamma_i)$. We are therefore ready to define our notion of motion during a deformation.

Definition 1 *Let $\gamma_1, \ldots, \gamma_n$ be smooth boundaries of closed subsets of a differentiable manifold embedded in $\mathbb{R}^N$, which we call* pre-shapes. *Let $\mathcal{H}$ be a class of diffeomorphisms acting on γ_i, and let $E : \mathcal{H} \to \mathbb{R}^+$ be a positive, real-valued functional. Consider now a group G acting on γ_i via $g(\gamma_i)$. We say that $\hat{g}_1, \ldots, \hat{g}_n$ is a* **motion** *undergone by γ_i, $i = 1 \ldots n$ if there exists a pre-shape $\hat{\mu}$ such that*

$$\boxed{\hat{g}_1, \ldots, \hat{g}_n, \hat{\mu} = \arg\min_{g_t, \mu} \sum_{i=1}^{n} E(h_i) \text{ subject to } \gamma_i = h_i \circ g_i(\mu) \quad i = 1 \ldots n} \tag{4}$$

The pre-shape $\hat{\mu}$ is called the **shape average** *relative to the group G, or G-average, and the quantity $\hat{g}_i^{-1}(\gamma_i)$ is called the* **shape** *of γ_i.*

Remark 1 (Invariance) *In the definition above, one will notice that the shape average is actually a pre-shape, and that there is an arbitrary choice of group action g_0 that, if applied to γ_i and μ, leaves the definition unchanged (the functional E is invariant with respect to g_0 because $T(g \circ g_0, h \circ g_0) = T(g, h)\ \forall\ g_0$). For the case of the Euclidean group $SE(N)$, a way to see this is to notice that the reference frame where μ is described is arbitrary. Therefore, one may choose, for instance, $\mu = h_1^{-1}(\gamma_1)$.*

Remark 2 (Symmetries) *In Def. 1 we have purposefully avoided to use the article* "the" *for the minimizing value of the group action $\hat{g}_t$. It is in fact possible that the minimum of (4) not be unique. A particular case when this occurs is* when the pre-shape γ *is (symmetric, or) invariant with respect to a particular element of the group G, or to an entire subgroup. Notice, however, that the notion of* shape average *is still well-defined even when the notion of motion is not unique. This is because any element in the symmetry group suffices to register the pre-shapes, and therefore compute the shape average (Fig. 3).*

3 Shape and Deformation of a Planar Contour

In this section we consider the implementation of the program above for a simple case: two closed planar contours, γ_1 and γ_2, where we choose as cost functional for the deformations h_1, h_2 either the set-symmetric difference Δ of their interior (the union minus the intersection), or what we call the signed distance score[3] ψ

$$\psi(\mu, \gamma) \doteq \int_{\bar{\mu}} \zeta(\gamma) d\mathbf{x} \tag{5}$$

where $\bar{\mu}$ denotes the interior of the contour μ and ζ is the signed distance function of the contour γ; $d\mathbf{x}$ is the area form on the plane. In either case, since we have an arbitrary choice of the global reference frame, we can choose $g_1 = e$, the group identity. We also call $g \doteq g_2$, so that $\mu_2 = g(\mu)$. The problem of defining the motion and shape average can then be written as

$$\hat{g}, \hat{\mu} = \arg\min_{g,\mu} \sum_{i=1}^{2} E(h_i) \text{ subject to } \gamma_1 = h_1(\mu);\ \gamma_2 = h_2 \circ g(\mu). \tag{6}$$

[3] The rationale behind this score is that one wants to make the signed distance function as positive as possible outside the contour to be matched, and as negative as possible inside. This score can be interpreted as a weighted Monge-Kantorovic functional where the mass of a curve is weighted by its distance from the boundary.

As we have anticipated, we choose either $E(h_i) = \Delta(g_i(\mu), \gamma_i)$ or $E(h_i) \doteq \psi(g_i(\mu), \gamma_i)$. Therefore, abusing the notation as anticipated before Def. 1, we can write the problem above as an unconstrained minimization

$$\boxed{\hat{g}, \hat{\mu} = \arg\min_{g,\mu} \phi(\gamma_1, \gamma_2)} \quad \text{where} \quad \boxed{\phi(\gamma_1, \gamma_2) \doteq E(\mu, \gamma_1) + E(g(\mu), \gamma_2)} \tag{7}$$

and E is either Δ or ψ. The estimate $\hat{g}$ defines the motion between γ_1 and γ_2, and the estimate $\hat{\mu}$ defines the average of the two contours.

If one thinks of contours and their interior, represented by a characteristic function χ, as a binary image, then the cost functional above is just a particular case of a more general cost functional where each term is obtained by integrating a function inside and a function outside the contours

$$\boxed{\phi = \sum_{i=1}^{2} \int_{\bar{\mu}_{in}} f_{in}(\mathbf{x}, \gamma_i) d\mathbf{x} + \int_{\bar{\mu}_{out}} f_{out}(\mathbf{x}, \gamma_i) d\mathbf{x}} \tag{8}$$

where the bar in $\bar{\mu}$ indicates that the integral is computed on a *region* inside or outside μ and we have emphasized the fact that the function f depends upon the contour γ_i. For instance, for the case of the set symmetric difference we have $f_{in} = (\chi_\gamma - 1)$ and $f_{out} = \chi_\gamma$. To solve the problem, therefore, we need to minimize the following functional

$$\int_{\bar{\mu}_{in}} f_{in}(\mathbf{x}, \gamma_1) + f_{in}(g(\mathbf{x}), \gamma_2)|J_g| d\mathbf{x} + \int_{\bar{\mu}_{out}} f_{out}(\mathbf{x}, \gamma_1) + f_{out}(g(\mathbf{x}), \gamma_2)|J_g| d\mathbf{x} \tag{9}$$

where $|J_g|$ is the determinant of the Jacobian of the group action g. This makes it easy to compute the component of the first variation of ϕ along the normal direction to the contour μ, so that we can impose

$$\nabla_\mu \phi \cdot N = 0 \tag{10}$$

to derive the first-order necessary condition. If we choose $G = SE(2)$, an isometry, it can be easily shown that

$$\boxed{\nabla_\mu \phi = f_{in}(\mathbf{x}, \gamma_1) - f_{out}(\mathbf{x}, \gamma_1) + f_{in}(g(\mathbf{x}), \gamma_2) - f_{out}(g(\mathbf{x}), \gamma_2)} \tag{11}$$

3.1 Representation of Motions and Their Variation

For the specific case of matrix Lie groups (e.g. $G = SE(2)$), there exist twist coordinates ξ that can be represented as a skew-symmetric matrix $\widehat{\xi}$ so that[4]

$$\boxed{g = e^{\widehat{\xi}}} \quad \text{and} \quad \boxed{\frac{\partial g}{\partial \xi_i} = \frac{\partial \widehat{\xi}}{\partial \xi_i} g} \tag{12}$$

[4] The "widehat" notation $\widehat{\ }$, which indicates a lifting to the Lie algebra, should not be confused with the "hat" $\hat{\ }$, which indicates an estimated quantity.

where the matrix $\frac{\partial \widehat{\xi}}{\partial \xi_i}$ is composed of zeros and ones and the matrix exponential can be computed in closed form using Rodrigues' formula.

To compute the variation of the functional ϕ with respect to the group action g, we notice that the first two terms in ϕ do not contribute since they are independent of g. The second two terms are of the generic form $A(g) \doteq \int_{g(\bar{\mu})} f(\mathbf{x}) d\mathbf{x}$. Therefore, we consider the variation of A with respect to the components of the twist ξ_i, $\frac{\partial A}{\partial \xi_i}$, which we will eventually use to compute the gradient with respect to the natural connection $\nabla_G \phi = \widehat{\left(\frac{\partial \phi}{\partial \xi}\right)} g$. We first rewrite $A(g)$ using the change of measure $\int_{g(\bar{\mu})} f(\mathbf{x}) d\mathbf{x} = \int_{\bar{\mu}} f \circ g(\mathbf{x}) |J_g| d\mathbf{x}$ which leads to $\frac{\partial A(g)}{\partial \xi_i} = \int_{\bar{\mu}} \frac{\partial}{\partial \xi_i} (f \circ g(\mathbf{x})) |J_g| d\mathbf{x} + \int_{\bar{\mu}} (f \circ g(\mathbf{x})) \frac{\partial}{\partial \xi_i} |J_g| d\mathbf{x}$ and note that the Euclidean group is an isometry and therefore the determinant of the Jacobian is one and the second integral is zero. The last equation can be re-written, using Green's theorem, as $\int_{g(\mu)} \left\langle f(\mathbf{x}) \frac{\partial g}{\partial \xi_i} \circ g^{-1}(\mathbf{x}), N \right\rangle ds = \int_{\mu} \left\langle f \circ g(\mathbf{x}) \frac{\partial g}{\partial \xi_i}, g_* N \right\rangle ds$ where g_* indicates the push-forward. Notice that g is an isometry and therefore it does not affect the arc length; we then have

$$\frac{\partial A(g)}{\partial \xi_i} = \int_{\mu} f(g(\mathbf{x})) \left\langle \frac{\partial \widehat{\xi}}{\partial \xi_i} g, g_* N \right\rangle ds \tag{13}$$

After collecting all the partial derivatives into an operator $\frac{\partial \phi}{\partial \xi}$, we can write the evolution of the group action.

3.2 Evolution

The algorithm for evolving the contour and the group action consists of a two-step process where an initial estimate of the contour $\hat{\mu} = \gamma_1$ is provided, along with an initial estimate of the motion $\hat{g} = e$. The contour and motion are then updated in an alternating minimization where motion is updated according to

$$\boxed{\frac{d\hat{g}}{dt} = \widehat{\left(\frac{\partial \phi}{\partial \xi}\right)} \hat{g}} \tag{14}$$

Notice that this is valid not just for $SE(2)$, but for any (finite-dimensional) matrix Lie group, although there may not be a closed-form solution for the exponential map like in the case of $SE(3)$ and its subgroups. In practice, the group evolution (14) can be implemented in local (exponential) coordinates by evolving ξ defined by $g = e^{\widehat{\xi}}$ via $\frac{d\xi}{dt} = \frac{\partial \phi}{\partial \xi}$. In the level set framework, the derivative of the cost function ϕ with respect to the coordinates of the group action ξ_i can be computed as the collection of two terms, one for f_{in}, one for f_{out} where $\frac{\partial \phi}{\partial \xi_i} = \int_{g(\gamma_{1,2})} \left\langle \frac{\partial g(\mathbf{x})}{\partial \xi_i}, f_{\{in,out\}}(g(\mathbf{x}), \gamma_{1,2}) J(g_* T) \right\rangle ds$. As we have anticipated in Eq. (11), the contour $\hat{\mu}$ evolves according to

$$\boxed{\frac{d\hat{\mu}}{dt} = (f_{in}(\mathbf{x}, \gamma_1) - f_{out}(\mathbf{x}, \gamma_1) + f_{in}(g(\mathbf{x}), \gamma_2) - f_{out}(g(\mathbf{x}), \gamma_2)) N} \tag{15}$$

As we have already pointed out, the derivation can be readily extended to surfaces in space and to multiple objects, as we show in Sect. 6.

3.3 Distance between Shapes

The definition of motion $\hat{g}$ and shape average $\hat{\mu}$ as a minimizer of (6) suggests defining the distance[5] between two shapes as the "energy" necessary to deform one into the other via the average shape:

$$d(\gamma_i, \gamma_j) \doteq E(\gamma_i, T(\hat{g}, \hat{h})\gamma_j). \tag{16}$$

For instance, for the case of the set-symmetric difference of two contours, we have

$$d_{\Delta}(\gamma_1, \gamma_2) \doteq \int \chi_{\hat{\mu}}\chi_{\gamma_1} + \chi_{\hat{g}(\hat{\mu})}\chi_{\gamma_2} d\mathbf{x} \tag{17}$$

and for the signed distance transform we have

$$d_{\psi}(\gamma_1, \gamma_2) \doteq \int_{\hat{\tilde{\mu}}} \zeta(\gamma_1) d\mathbf{x} + \int_{\hat{g}(\hat{\tilde{\mu}})} \zeta(\gamma_2) d\mathbf{x}. \tag{18}$$

In either case, a gradient flow algorithm based on Eq. (14) and (15), when it converges to a global minimum, returns an average shape and a set of group elements g_i which minimize the sums of the distances between the contours γ_i and any other common contour modulo the chosen group.

4 Moving Average and Tracking

The discussion above assumes that an unsorted collection of shapes is available, where the deformation between any two shapes is "small" (modulo G), so that the whole collection can be described by a single average shape. Consider however the situation where an object is evolving in time, for instance Fig. 5. While the deformation between adjacent time instants could be captured by a group action and a small deformation, as time goes by the object may change so drastically that talking about a global time average may not make sense.

One way to approach this issue is by defining a notion of *"moving average"*, similarly to what is done in time series analysis. In classical linear time series, however, the uncertainty is modeled via additive noise. In our case, the uncertainty is an infinite-dimensional deformation h that acts on the measured contour. So the model becomes

$$\begin{cases} \mu(t+1) = g(t)\mu(t) \\ \gamma(t) = h(\mu(t)) \end{cases} \tag{19}$$

[5] Here we use the term distance informally, since we do not require that it satisfies the triangular inequality. The term pseudo-distance would be more appropriate.

where $\mu(t)$ represents the moving average of order $k = 1$. A similar model can be used to define moving averages of higher-order $k > 1$. The procedure described in Sect. 3, initialized with $\mu(0) = \gamma_1$, provides an estimate of the moving average of order 1, as well as the *tracking* of the trajectory $g(t)$ in the group G, which in (19) is represented as the model parameter. Note that the procedure in Sect. 3 simultaneously estimates the state $\mu(t)$ and identifies the parameters $g(t)$ of the model (19). It does so, however, without imposing restrictions on the evolution of $g(t)$. If one wants to impose additional constraints on the motion parameters, one can augment the state of the model to include the parameters g, for instance $g(t+1) = e^{\widehat{\xi}(t)} g(t)$. This, however, is beyond the scope of this paper. In Fig. 5 we show the results of tracking a storm with a moving average of order one.

5 Simultaneous Approximation and Registration of Non-equivalent Shapes

So far we have assumed that the given shapes are obtained by moving and deforming a common underlying "template" (the average shape). Even though the given shapes are not *equivalent* (i.e. there is no group action g that maps one exactly onto the other), g is found as the one that minimizes the cost of the deviation from such an equivalence. In the algorithm proposed in Eq. (14)-(15), however, there is no explicit requirement that the deformation between the given shapes be small. Therefore, the procedure outlined can be seen as an algorithm to register shapes that are not equivalent under the group action. A *registration* is a group element $\hat{g}$ that minimizes the cost functional (4).

To illustrate this fact, consider the two considerably different shapes shown in Fig. 7, γ_1, γ_2. The simultaneous estimation of their average μ, for instance relative to the affine group, and of the affine motions that best matches the shape average onto the original ones, g_1, g_2, provides a registration that maps γ_1 onto γ_2 and vice-versa: $g = g_2 g_1^{-1}$.

If instead of considering the images in Fig. 7 as binary images that represent the contours, we consider them as gray-scale images, then the procedure outlined, for the case where the score is computed using the set-symmetric difference, provides a way to simultaneously jointly segment the two images and register them. This idea is illustrated in Fig. 9 for true gray-scale (magnetic resonance) images of brain sections.

6 Experiments

Fig. 3 illustrates the difference between the motion and shape average computed under the Euclidean group, and the affine group. The three examples show the two given shapes γ_i, the mean shape registered to the original shapes, $g_i(\mu)$ and the mean shape μ. Notice that affine registration allows to simultaneously capture the square and the rectangle, whereas the Euclidean average cannot be registered to either one, and is therefore only an approximation.

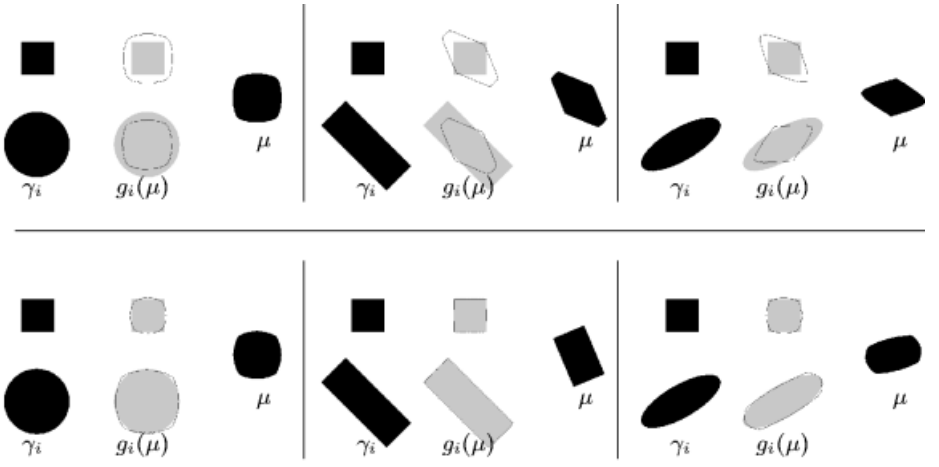

Fig. 3. Euclidean (top) vs. affine (bottom) registration and average. *For each pair of objects γ_1, γ_2, the registration $g_1(\mu), g_2(\mu)$ relative to the Euclidean motion and affine motion is shown, together with the Euclidean average and affine average μ. Note that the affine average can simultaneously "explain" a square and a rectangle, whereas the Euclidean average cannot.*

Fig. 4 compares the effect of choosing the signed distance score (left) and the set-symmetric difference (right) in the computation of the motion and average shape. The first choice results in an average that captures the common features of the original shapes, whereas the second captures more of the features in each one. Depending on the application, one may prefer one or the other.

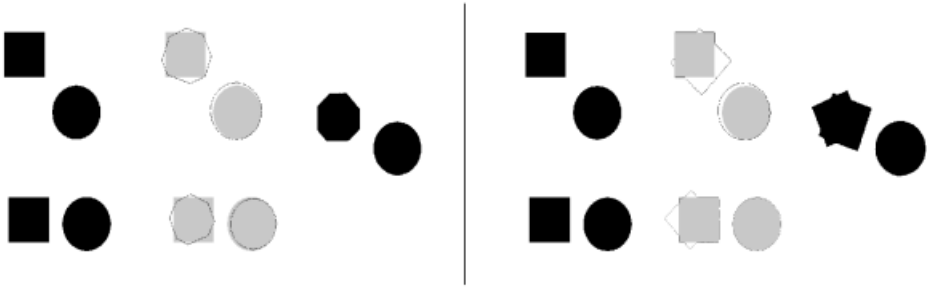

Fig. 4. Signed distance score (left) vs. set-symmetric difference (right). *Original contours (γ_1 on the top, γ_2 on the bottom), registered shape $g_i(\mu)$ and shape average μ. Note that the original objects are not connected, but are composed by a circle and a square. The choice of pseudo-distance between contours influences the resulting average. The signed distance captures more of the features that are common to the two shapes, whereas the symmetric difference captures the features of both.*

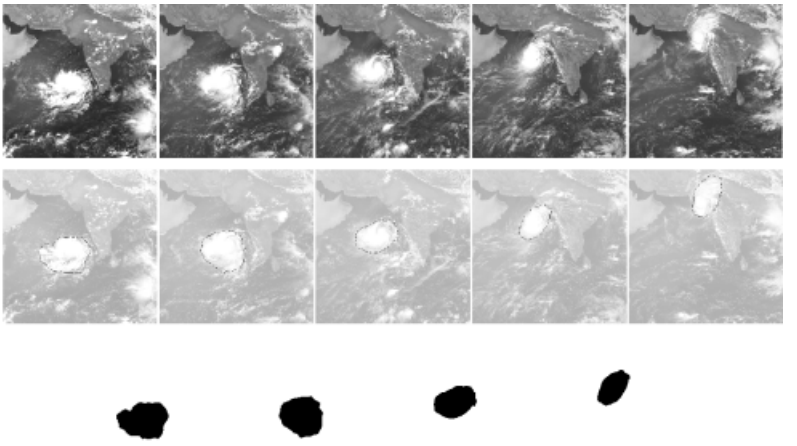

Fig. 5. Storm *(first row) a collection of images from EUMETSAT ©2001, (second row) affine motion of the storm based on two adjacent time instances, superimposed to the original images, (bottom) moving average of order 1.*

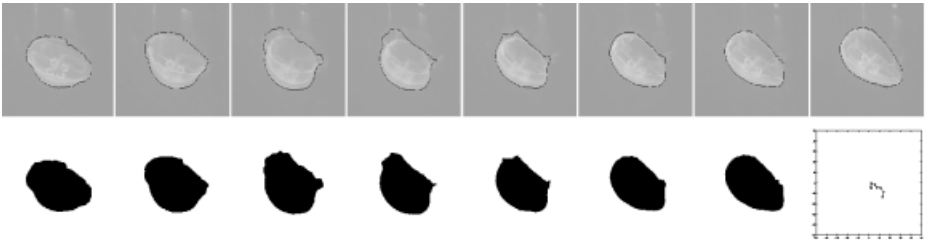

Fig. 6. Jellyfish. *Affine registration (top), moving average and affine motion (bottom) for the jellyfish in Fig. 1. (Bottom right) trajectory of the jellyfish (affine component of the group).*

Fig. 5 shows the results of tracking a storm. The affine moving average is computed, and the resulting affine motion is displayed. The same is done for the jellyfish in Fig. 6.

Fig. 7 and 8 are meant to challenge the assumptions underlying our method. The pairs of objects chosen, in fact, are not simply local deformations of one another. Therefore, the notion of shape average is not meaningful *per se* in this context, but serves to compute the change of (affine) pose between the two shapes (Fig. 7). Nevertheless, it is interesting to observe how the shape average allows registering even apparently disparate shapes. Fig. 8 shows a representative example from an extensive set of experiments. In some cases, the shape average contains disconnected components, in some other it includes small parts that are shared by the original dataset, whereas in others it removes parts that are not consistent among the initial shapes (e.g. the tails). Notice that our framework is not meant to capture such a wide range of variations. In particular, it does

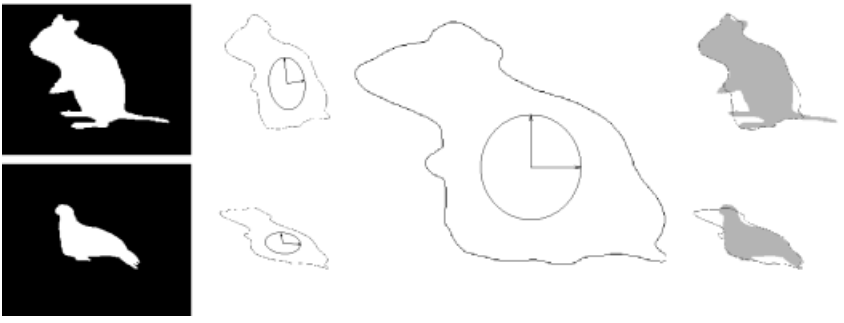

Fig. 7. Registering non-equivalent shapes. *Left to right: two binary images representing two different shapes; affine registration; corresponding affine shape; approximation of the original shapes using the registration of the shape average based on the set-symmetric difference. Results for the signed distance score are shown in Fig. 8. This example is shown to highlight the limitations of our method.*

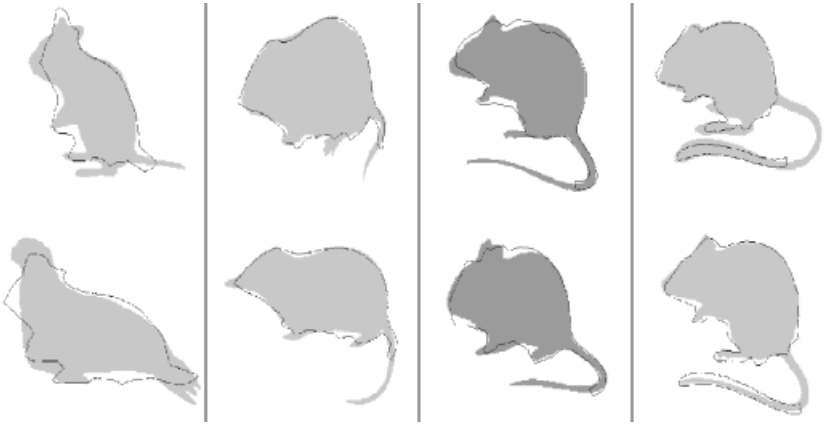

Fig. 8. Biological shapes. *For the signed distance score, we show the original shape with the affine shape average registered and superimposed. It is interesting to notice that different "parts" are captured in the average only if they are consistent in the two shapes being matched and, in some cases, the average shape is disconnected.*

not possess a notion of "parts" and it is neither hierarchical nor compositional. In the context of non-equivalent shapes (shapes for which there is no group action mapping one exactly onto the other), the *average shape serves purely as a support to define and compute motion in a collection of images of a given deforming shape.*

Fig. 9 shows the results of simultaneously segmenting and computing the average motion and registration for 4 images from a database of magnetic resonance images of the corpus callosum.

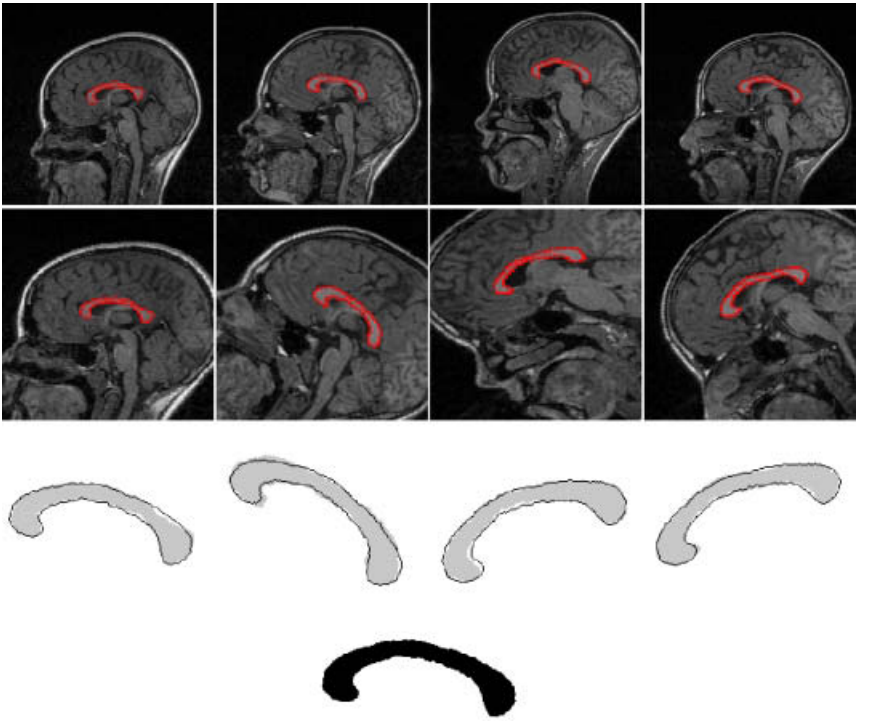

Fig. 9. Corpus Callosum *(top row) a collection of (MR) images from different patients (courtesy of J. Dutta), further translated, rotated and distorted to emphasize their misalignment (second row). Aligned contour (second row, dark gray, superimposed to original regions) and shape average (bottom) corresponding to the affine group.*

Finally, Fig. 10 shows an application of the same technique to simultaneously register and average two 3D surfaces. In particular, two 3D models in different

Fig. 10. 3D Averaging and registration *(left) two images of 3D models in different poses (center) registered average (right) affine average. Note that the original 3D surfaces are not equivalent. The technique presented allows "stitching" and registering different 3D models in a natural way.*

poses are shown. Our algorithm can be used to register the surfaces and average them, thus providing a natural framework to integrate surface and volume data.

We wish to thank S. Belongie and B. Kimia for test data and suggestions.

References

1. L. Alvarez and J. M. Morel. Morphological approach to multiscale analysis: From principles to equations. In *In Geometric–Driven Diffusion in Computer Vision*, 1994.
2. L. Alvarez, J. Weickert, and J. Sanchez. A scale-space approach to nonlocal optical flow calculations. In *In ScaleSpace '99*, pages 235–246, 1999.
3. R. Azencott, F. Coldefy, and L. Younes. A distance for elastic matching in object recognition. *Proc. 13th Intl. Conf. on Patt. Recog*, 1:687–691, 1996.
4. S. Belongie, J. Malik, and J. Puzicha. Matching shapes. In *Proc. of the IEEE Intl. Conf. on Computer Vision*, 2001.
5. D. Bereziat, I. Herlin, and L. Younes. Motion detection in meteorological images sequences: Two methods and their comparison. In *Proc. of the SPIE*, 1997.
6. A. Blake and M. Isard. *Active contours*. Springer Verlag, 1998.
7. C. Bregler and S. Omohundro. Surface Learning with Applications to Lip Reading. In J. D. Cowan, G. Tesauro and J. Alspector (eds), *Advances in Neural Information Processing Systems* (6), San Francisco, CA: Morgan Kaufmann Publishers, 1994.
8. T. K. Carne. The geometry of shape spaces. *Proc. of the London Math. Soc. (3) 61*, 3(61):407–432, 1990.
9. T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active Shape Models - Their Training and Application. *Computer Vision and Image Understanding*, 61(1), 1995, 38-59.
10. 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.
11. M. Fischler and R. Elschlager. The representation and matching of pictorial structures. *IEEE Transactions on Computers*, 22(1):67–92, 1973.
12. U. Grenander. *General Pattern Theory*. Oxford University Press, 1993.
13. U. Grenander and M. I. Miller. Representation of knowledge in complex systems. *J. Roy. Statist. Soc. Ser. B*, 56:549–603, 1994.
14. P.T. Jackway and R. Deriche. Scale-space properties of the multiscale morphological dilationerosion. *IEEE Trans. PAMI*, 18(1):38–51, 1996.
15. D. G. Kendall. Shape manifolds, procrustean metrics and complex projective spaces. *Bull. London Math. Soc.*, 16, 1984.
16. B. Kimia, A. Tannenbaum, and S. Zucker. Shapes, shocks, and deformations i: the components of two-dimensional shape and the reaction-diffusion space. *IJCV*, 15:189–224, 1995.
17. R. Kimmel. Intrinsic scale space for images on surfaces: The geodesic curvature flow. In *Lecture Notes In Computer Science: First International Conference on Scale-Space Theory in Computer Vision*, 1997.
18. R. Kimmel, N. Kiryati, and A. M. Bruckstein. Multivalued distance maps for motion planning on surfaces with moving obstacles. *IEEE TAC*, 14(3):427–435, 1998.

19. M. Lades, C. Borbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. Wurtz, and W. Konen. Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans. on Computers*, 42(3):300–311, 1993.
20. 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.
21. M. Leventon, E. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours, 2000.
22. R. Malladi, R. Kimmel, D. Adalsteinsson, V. Caselles G. Sapiro, and J. A. Sethian. A geometric approach to segmentation and analysis of 3d medical images. In *Proc. Mathematical Methods in Biomedical Image Analysis Workshop*, pages 21–22, 1996.
23. R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE PAMI*, 17(2):158–175, 1995.
24. K. V. Mardia and I. L. Dryden. Shape distributions for landmark data. *Adv. appl. prob.*, 21(4):742–755, 1989.
25. M. I. Miller and L. Younes. Group actions, diffeomorphisms and matching: a general framework. In *Proc. of SCTV*, 1999.
26. S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi equations. *J. of Comp. Physics*, 79:12–49, 1988.
27. N. Paragios and R. Deriche. Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE PAMI*, 22(3):266–280, 2000.
28. C. Samson, L. Blanc-Feraud, G. Aubert, and J. Zerubia. A level set model for image classification. In *in International Conference on Scale-Space Theories in Computer Vision*, pages 306–317, 1999.
29. B. ter Haar Romeny, L. Florack, J. Koenderink, and M. Viergever (Eds.). Scale-space theory in computer vision. In *LLNCS*, Vol 1252. Springer Verlag, 1997.
30. P. Thompson and A. W. Toga. A surface-based technique for warping three-dimensional images of the brain. *IEEE Trans. Med. Imaging*, 15(4):402–417, 1996.
31. K. Toyama and A. Blake. Probabilistic tracking in a Metric Space. In *Proc. of the Eur. Conf. on Comp. Vision*, 2001.
32. R. C. Veltkamp and M. Hagedoorn. State of the art in shape matching. Technical Report UU-CS-1999-27, University of Utrecht, 1999.
33. L. Younes. Computable elastic distances between shapes. *SIAM J. of Appl. Math.*, 1998.
34. A. Yuille. Deformable templates for face recognition. *J. of Cognitive Neurosci.*, 3(1):59–70, 1991.
35. S. Zhu, T. Lee, and A. Yuille. Region competition: Unifying snakes, region growing, energy /bayes/mdl for multi-band image segmentation. In *ICCV*, pages 416–423, 1995.