

# Matching Distance Functions: A Shape-to-Area Variational Approach for Global-to-Local Registration

Nikos Paragios<sup>1</sup>, Mikael Rousson<sup>2</sup>, and Visvanathan Ramesh<sup>1</sup>

<sup>1</sup> Siemens Corporate Research,  
Imaging and Visualization Department,  
755 College Road East, Princeton, NJ 08540, USA  
{nikos,rameshv}@scr.siemens.com

<sup>2</sup> I.N.R.I.A.  
B.P. 93, 2004 Route des Lucioles,  
06902 Sophia Antipolis Cedex, France  
Mikael.Rousson@sophia.inria.fr

**Abstract.** This paper deals with the matching of geometric shapes. Our primary contribution is the use of a simple, robust, rich and efficient way to represent shapes, the level set representations according to signed distance transforms. Based on these representations we propose a variational framework for global as well as local shape registration that can be extended to deal with structures of higher dimension. The optimization criterion is invariant to rotation, translation and scale and combines efficiently a global motion model with local pixel-wise deformations. Promising results are obtained on examples showing small and large global deformations as well as arbitrary topological changes.

## 1 Introduction

The problem of registering geometric shapes [23] is a complex issue in computer vision, computer graphics and recently medical imaging. It has been studied in various forms during the last decade due to its wide application not limited to registration, recognition, retrieval, tracking, etc.

A general registration formulation can be stated as follows: given two shapes, an input  $\mathcal{D}$  and a target  $\mathcal{S}$ , and a dissimilarity measure, find the best transformation that associates to any point of  $\mathcal{D}$  a corresponding point at  $\mathcal{S}$  and minimizes the dissimilarity measure between the transformed shape  $\hat{\mathcal{D}}$  and the target  $\mathcal{S}$ . This dissimilarity can be defined either along the contour (*shape-based*) or in the entire region (*area-based*) determined by the contour.

At this point, we have to distinguish the methods that aim to provide shape recognition. In that case correspondences between the source and the target shape are considered known or can be easily recovered and the objective is to find from a given set of examples the shape that provides the lower dissimilarity measurement with the target [1, 19, 22, 26]. Towards this direction, methods that do not require correspondence and are based on the matching of some global shape characteristics have been also investigated. It is important to note that

our paper does not focus on the matching and the recognition of shapes but addresses registration with unknown correspondence.

Shape registration has been approached into a number of ways, that can be mainly categorized [14] according to three factors: (i) Nature of Transformation, (ii) Domain of Transformation, (iii) Optimization Procedure. An effort to briefly cover the existing literature on these areas will be made in the next paragraphs.

A critical component of the registration procedure is the underlying motion model (nature of transformation) that is used to map the current shape to the target. The selection of the motion model can affect drastically the performance of the registration procedure. *Rigid* transformations refer to translation and rotation and are a compromise between low complexity and fairly acceptable matching between the different structures. A step forward, the *affine* transformation, is a more complicated model that is invariant to a large variety of motions. The use of *projective* geometry has been also considered to match shapes. Finally, *curved* or *elastic* registration methods can better account for local deformations and therefore have been also considered.

As far as the domain of transformation is concerned, the distinction is easier. *Global* transformations are valid for the entire shape [1] and for any point of  $\mathcal{D}$  we can obtain the corresponding point at  $\mathcal{S}$  by applying the same transformation, while *local* transformations are applied at the pixel level [5].

Finally, the optimization procedure refers to the selection of a mathematical framework by means of finding an optimum of some functional defined on the parameter space. These functionals attempt to quantify the similarity between the two shapes and can be based either in variational [4, 9], or stochastic principles [24]. As far the optimization procedures are concerned, gradient-descent, geometric hashing, iterative closest point, etc. can be found in the literature. Finally, methods for shape registration can be categorized according to the feature space on which the matching is performed (Euclidean space, affine space, curvature space, medial axes/skeletons [19, 26], graph and shocks [21], etc.).

In this paper we propose a variational framework for shape alignment that can be extended to any arbitrary dimension. We adopt an implicit representation inspired by the level set methods [16] and refers to signed distance transforms. Furthermore, we introduce an optimization criterion that can account for global (*rigid*, *affine*) and local pixel-wise deformations. This criterion is defined in the space of signed distance transforms, and is minimized using a gradient descent method. Global and local registration parameters are recovered using different update factors that allow us to recover first the global transformation and then the local deformations. Encouraging experimental results using 2D shapes were obtained.

Distance transforms [18] have been considered for image registrations in the past [6, 8, 12, 13]. In [6] a local-registration is recovered by seeking pixel-wise the lowest cost geodesic path between the source and the target. This method is very local, does not account for any neighborhood coherency and can fail to provide a meaningful registration result unless properly initialized.

In [8,12] the Euclidean distance was considered to be the feature space. Then 2D/3D objects are to be registered using linear transformations. In [12] registration is done towards a prior shape model that refers to a collection of points obtained through uniform sampling. In [8] distance maps from edges are considered. In both cases registration is obtained through the minimization of a metric defined on the space of distance transform. Both methods focus on global transformations and cannot deal with local deformations. Furthermore, they cannot deal with important scale variations and do not make use of a feature space with less information compared with the one considered in this paper.

The remainder of this paper is organized as follows. In Section 2 we briefly introduce the selected shape representation. The registration method is presented in Section 3 while summary and discussion appear in Section 4.

## 2 Shape Representation

A crucial component in the procedure of registering arbitrary shapes is the underlying shape representation [11] since it can significantly affect the performance of the selected registration algorithm.

The use of point-based snake models [10], deformable models/templates [2], active shapes [7], Fourier descriptors, medial axis, level set representations [3, 15] are some alternatives. Although, these representations are powerful enough to capture a certain number of local deformations, they require a large number of parameters to deal with important shape deformations, and they cannot deal with changes of topology. Also, their extension to describe structures of higher dimension than curves and surfaces is in most of the cases not trivial. Based on these considerations and given the assumption that we would like to obtain a global to local registration, we will consider the use of the level set representations that can deal with local deformations and is invariant to translation and rotation.

Level Set Representation [16] is a powerful mathematical tool that can deal with applications that share a common concern, evolving interfaces. Within these representations, contours/shapes are modeled as the zero-level set ( $\phi = 0$ ) of a function defined in a higher dimension. The most common selection for the embedding function is signed (Euclidean) distance transform.

In order to facilitate the introduction of the method, we consider the 2D case and let  $\Phi : \Omega \rightarrow \mathcal{R}^+$  be a Lipschitz function that refers to a level set representation for a given shape  $\mathcal{S}$ . This shape defines a region  $\mathcal{R}$  in the image plane  $\Omega$ . Given these definitions the following shape representation is considered:

$$\Phi_{\mathcal{S}}(x, y) = \begin{cases} 0, & (x, y) \in \mathcal{S} \\ +D((x, y), \mathcal{S}) > 0, & (x, y) \in \mathcal{R}_{\mathcal{S}} \\ -D((x, y), \mathcal{S}) < 0, & (x, y) \in [\Omega - \mathcal{R}_{\mathcal{S}}] \end{cases}$$

where  $D((x, y), \mathcal{S})$  refers to the min Euclidean distance between the grid location  $(x, y)$  and the shape  $\mathcal{S}$ . The fast marching algorithm can be used for the construction of these representations [20].

It can be easily shown that this representation  $[\Phi_S]$  is invariant to translation and rotation. Let us consider a shape  $\mathcal{D}$  that is obtained after rotating  $\mathcal{S}$  by an  $\theta$  angle and translating by a vector  $(T_x, T_y)$ . Then, if we consider a pixel  $(x, y)$  with distance  $d$  from the interface  $\mathcal{D}$ , the following relation holds:

$$d = D((x, y), \mathcal{D}) = \min_{(u,v) \in \mathcal{D}} \left\{ \sqrt{(u-x)^2 + (v-y)^2} \right\}$$

The use of the inverse transformation between  $\mathcal{D}$  and  $\mathcal{S}$  for  $(x, y)$  leads to the following equation:

$$\begin{aligned} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} &= \begin{pmatrix} x \cos(-\theta) + y \sin(-\theta) - T_x \\ -x \sin(-\theta) + y \cos(-\theta) - T_y \end{pmatrix} \\ \mathcal{S} &= \left\{ \begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} = \begin{pmatrix} u \cos(-\theta) + v \sin(-\theta) - T_x \\ -u \sin(-\theta) + v \cos(-\theta) - T_y \end{pmatrix} : (u, v) \in \mathcal{D} \right\} \end{aligned}$$

Thus, the distance between  $\begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix}$  and the interface  $\mathcal{S}$  is given by

$$\begin{aligned} D((\hat{x}, \hat{y}), \mathcal{S}) &= \min_{(\hat{u}, \hat{v}) \in \mathcal{S}} \left\{ \sqrt{(\hat{u} - \hat{x})^2 + (\hat{v} - \hat{y})^2} \right\} \\ &= \min_{(u,v) \in \mathcal{D}} \left\{ \sqrt{\begin{aligned} &((u-x) \cos(-\theta) - (v-y) \sin(-\theta))^2 \\ &+ ((u-x) \sin(-\theta) + (v-y) \cos(-\theta))^2 \end{aligned}} \right\} \\ &= \min_{(u,v) \in \mathcal{D}} \left\{ \sqrt{(u-x)^2 + (v-y)^2} \right\} = D((x, y), \mathcal{D}) \end{aligned}$$

We have shown that the selected representation is invariant to rotation and translation. On the other hand when the transformation has as a scale component  $s$  then level set shape representations are not invariant and the following relation holds between the representation of  $\mathcal{S}$  and  $\mathcal{D}$ :

$$D((x, y), \mathcal{D}) = s D((\hat{x}, \hat{y}), \mathcal{S})$$

### 3 Shape Alignment and Registration

In the registration problem the objective is to find a point-wise transformation between the current shape  $\mathcal{D}$  and the target shape  $\mathcal{S}$  that minimizes a given dissimilarity measure between the two shapes.

The use of level set representations changes the dynamics of the problem since these representations refer to a higher dimensional space than the original one (2D shapes). Hence, we will seek for a transformation  $A$  that creates pixel-wise intensity correspondences (level set values) between the current shape representation  $s \Phi_{\mathcal{D}}$  and the target shape  $\Phi_{\mathcal{S}}$ . Due to the fact that the selected representations  $[\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}]$  are invariant to translation and rotation we can write:

$$\begin{cases} (s, \theta, T) \\ A(x, y) = \begin{pmatrix} A_x \\ A_y \end{pmatrix} = s \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} T_x \\ T_y \end{pmatrix} \\ \forall (\mathbf{x}, \mathbf{y}) \in \Omega : s \Phi_{\mathcal{D}}(\mathbf{x}, \mathbf{y}) = \Phi_{\mathcal{S}}(\mathbf{A}(\mathbf{x}, \mathbf{y})) \end{cases}$$

Thus, registration can be viewed now as a global optimization task that involves all pixels in the image plane. Based on this assumption, several optimization criteria have been proposed to recover the unknown transformation parameters; sum of squared differences, optimization of the correlation ratio, mutual information, etc. In order to introduce and demonstrate the performance of our method, at the very beginning we will consider the simplest possible criterion, the sum of square differences.

### 3.1 Sum of Squared Differences

Initially, we assume a global rigid deformation between  $\mathcal{S}$  and  $\mathcal{D}$  that involves three parameters  $[A = (s, \theta, T)]$ ; a rotation angle  $\theta$ , and translation vector  $T = (T_x, T_y)$  and a scale factor  $s$ . Then, the optimization criterion is given by:

$$E(s, \theta, T) = \iint_{\Omega} (s\Phi_{\mathcal{D}}(x, y) - \Phi_{\mathcal{S}}(A^{\tau}(x, y)))^2 dx dy$$

However, the level set representations for the considered shapes  $\mathcal{S}$  and  $\mathcal{D}$  are not equally defined in the image plane (they depend on the positions of the initial shapes). As a consequence, we can constrain the information space and decrease the complexity of the problem by considering the area defined by two equal distance contours (inwards, outwards) from the input shapes.

$$E(s, \theta, T) = \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}(x, y), \Phi_{\mathcal{S}}(A^{\tau}(x, y))) (s \Phi_{\mathcal{D}}(x, y) - \Phi_{\mathcal{S}}(A^{\tau}(x, y)))^2 dx dy$$

where  $N_{\delta}$  is a binary function given by

$$N_{\delta}(\phi_1, \phi_2) = \begin{cases} 0, & \min(|\phi_1|, |\phi_2|) > \delta \\ 1, & \min(|\phi_1|, |\phi_2|) \leq \delta \end{cases}$$

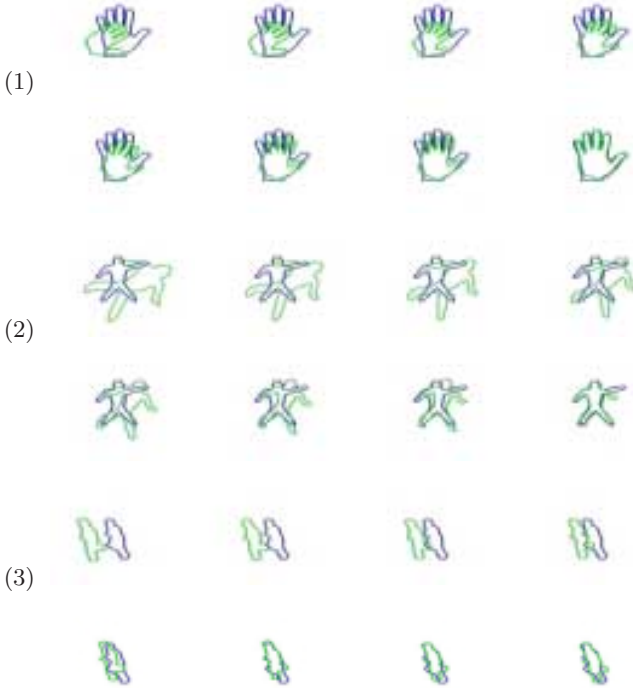
that has a simple interpretation. All pixels (isophotes), within a range of distance  $\delta$  from the actual shape are considered in the optimization process.

As a consequence, the registration is done in an augmented shape-driven (level set) space. This space is robust to very local deformations and missing data since the selected representation is obtained through a global procedure (Euclidean distance). Moreover, the proposed framework is invariant to rigid transformations and refers to multiple shape matching (isophotes) between shapes that are clones of the original ones to be registered.

Using the proposed formulation we were able to convert a geometry driven point-correspondence problem into an image-registration application where space as well feature-based (intensity) correspondences are considered.

The optimization of this criterion can be done using a gradient descent method

$$\begin{cases} \frac{d}{dt}\theta = 2 \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}})(\nabla\Phi_{\mathcal{S}} \cdot \nabla_{\theta}A^{\tau}) [(s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^{\tau}))] \\ \frac{d}{dt}s = 2 \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}})(\Phi_{\mathcal{D}} + \nabla\Phi_{\mathcal{S}} \cdot \nabla_s A^{\tau}) [(s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^{\tau}))] \\ \frac{d}{dt}T = 2 \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) \left( \nabla\Phi_{\mathcal{S}} \cdot \begin{pmatrix} \nabla_{T_x}A^{\tau} \\ \nabla_{T_y}A^{\tau} \end{pmatrix} \right) [(s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^{\tau}))] \end{cases}$$



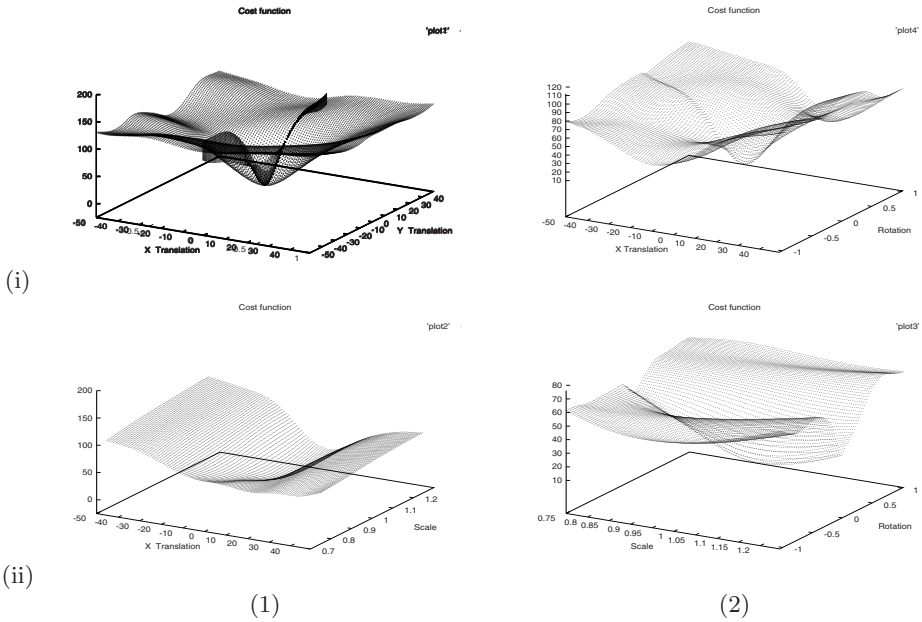
**Fig. 1.** (1) *Rigid Hand Registration (Synthetic Data)*:  $\{s = 1.27, \theta = 71.38^\circ, T_x = -19.65, T_y = 21.32\}$ . (2) *Rigid Body Registration (Synthetic Data)*:  $\{s = 0.63, \theta = 59.94^\circ, T_x = -13.97, T_y = -14.09\}$  (3) *Rigid Body Registration (Real Data)*.

The performance of the proposed module is shown in [FIG. (1)]. The selected representation is powerful and can provide encouraging results using a quite simple optimization criterion like the sum of squared differences. However, a validation of the method is required.

The characteristics of the cost function is a good indicator regarding the stability of the problem. Non-convex optimization criteria (like the one consider in this paper) suffer from the initial conditions. In our approach, we have considered a very strong feature space, the signed distance transforms and therefore one would expect that the performance of the method will be satisfactory.

In order to perform a study on this performance, we can constrain the unknown parameter space in two dimensions. We have considered the examples shown in [FIG. (1.1)]. Then, from the 4-dimensional parameter space we have studied the following cases:

- Scale and rotation are known, translation is to be recovered [FIG. (2.(i.1))],
- Translation in  $x$  and scale are known, translation in  $y$  and rotation are to be recovered [FIG. (2.(i.2))],
- Translation in  $x$  and rotation are known, translation in  $y$  and scale are to be recovered [FIG. (2.(ii.1))],



**Fig. 2.** Empirical evaluation of the cost function: (i.1) Unknown translation  $[x, y]$ , (i.2) Unknown translation  $[x]$  and rotation, (ii.1) Unknown translation  $[x]$  and scale, (ii.2) Unknown scale and rotation.

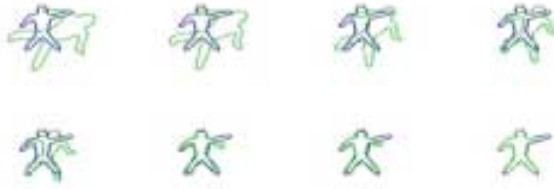
– Translation in  $x$  and translation in  $y$  are known, rotation and scale are to be recovered [FIG. (2.(ii.2))].

Then, we have quantized the search space using uniform sampling (100 samples) for all unknown parameters in each case. Translation transformations in  $(x, y)$  were in the range of  $[-50, 50] \times [-50, 50]$ , scale in  $[0.75, 1.25]$  and rotation in  $[-\frac{\pi}{3}, \frac{\pi}{3}]$ . Then, one can estimate the cost function in the space of two unknown parameters, by considering all possible combinations derived from the sampling strategy (the other two parameters are fixed) [FIG. (2)]. The resulting functional as shown in [FIG. (2)] has some nice properties; it is smooth and exhibits a single global minimum.

One can claim that the cost function [FIG. (2)] has a convex form for all cases with two unknown variables. This convexity cannot be guaranteed when the registration problem is considered in its full dimensionality, the four variables of rigid transformations. However, the form of this function in a reduced variable space is a good indicator for a well-behaved optimization criterion with smooth properties.

### 3.2 Global-to-Local Registration

The framework described in the previous section is limited to global rigid transformation and cannot account for non-rigid objects (shapes) and local deforma-



**Fig. 3.** *Global-to-Local Rigid Body Registration.*  $\{s = 0.63, \theta = 59.94^\circ, T_x = -13.97, T_y = -14.09\}$

tions. In order to overcome this limitation we will incorporate to the method the notion of local deformations by assuming that the observed shape is a rigid transformation  $A$  of the target combined with some local deformations  $(u, v)$ . Under this assumption, we can write:

$$\begin{cases} (s, \theta, T) \\ (u(x, y), v(x, y)) : (x, y) \in \Omega \\ \forall (\mathbf{x}, \mathbf{y}) \in \Omega : \mathbf{s}\Phi_{\mathcal{D}}(\mathbf{x}, \mathbf{y}) = \Phi_{\mathcal{S}}(\mathbf{A}^\tau + (\mathbf{u}, \mathbf{v})) \end{cases}$$

Now, we can decompose the previous hypothesis into two terms and define the following registration criterion:

$$\begin{aligned} E(s, \theta, T, (u, v)) &= \alpha \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^\tau))^2 \\ &+ (1 - \alpha) \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^\tau + (u, v)))^2 \end{aligned}$$

that has a simple interpretation: registration errors according to the rigid transformation are corrected using the local deformation field. The performance of this additional term is shown in [fig. (3)]. For demonstration purposes, we consider the same input as the one that was used to validate the performance of the criterion that accounts only for a global rigid transformation.

A natural registration assumption refers to the smoothness of the transformation field. Local deformations cannot be independent in a pixel level and therefore this condition has to be also considered during the recovery of the corresponding field. This can be done by introducing a constraint (in terms of a penalty function) that accounts for smoothness on the field  $(u, v)$ :

$$\begin{aligned} E(s, \theta, T, (u, v)) &= \alpha \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^\tau))^2 \\ &+ (1 - \alpha)\beta \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^\tau + (u, v)))^2 \\ &+ (1 - \alpha)(1 - \beta) \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) (u_x^2 + u_y^2 + v_x^2 + v_y^2) \end{aligned}$$

Let us now try to interpret the above functional. The first term aims at finding pixel-wise intensity (level set representation) correspondences using a global



motion model (rigid transformation). The second term, aims at correcting the correspondences in a pixel level using a local deformation model on top of the existing global model, while the third term constrains the deformation field to be locally smooth.

The minimization of this functional is done using a gradient descent:

$$\left\{ \begin{aligned} \frac{d}{dt}s &= 2 \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) (\Phi_{\mathcal{D}} + \nabla \Phi_{\mathcal{S}} \cdot \nabla_s A^{\tau}) \\ &\quad [\alpha (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^{\tau})) + \beta(1 - \alpha) (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^{\tau} + (u, v)))] \\ \frac{d}{dt}T &= 2 \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) \left( \nabla \Phi_{\mathcal{S}} \cdot \begin{pmatrix} \nabla_{T_x} A^{\tau} \\ \nabla_{T_y} A^{\tau} \end{pmatrix} \right) \\ &\quad [\alpha (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^{\tau})) + \beta(1 - \alpha) (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^{\tau} + (u, v)))] \\ \frac{d}{dt}\theta &= 2 \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) (\nabla \Phi_{\mathcal{S}} \cdot \nabla_{\theta} A^{\tau}) \\ &\quad [\alpha (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^{\tau})) + \beta(1 - \alpha) (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^{\tau} + (u, v)))] \\ \frac{d}{dt}u &= 2(1 - \alpha)(1 - \beta) (u_{xx} + u_{yy}) + \\ &\quad 2\beta(1 - \alpha) \nabla_x \Phi_{\mathcal{S}}(A^{\tau} + (u, v)) (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^{\tau} + (u, v))) \\ \frac{d}{dt}v &= 2(1 - \alpha)(1 - \beta) (v_{xx} + v_{yy}) + \\ &\quad 2\beta(1 - \alpha) \nabla_y \Phi_{\mathcal{S}}(A^{\tau} + (u, v)) (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^{\tau} + (u, v))) \end{aligned} \right.$$

The performance of the complete system is shown in [fig. (4)].

The selection of the  $\alpha$  parameter is a crucial component of the algorithm since it controls global-to-local factor. This is clearly shown in the motion equations where the construction of the local deformation field is delayed (actually it is done very slowly) until the parameters of the global rigid model are properly estimated.

The same property can be obtained through a two stage approach that involves the estimation of the global model first, and then the local estimates of the deformation field. However, such selection involves the use of time measures to control the gradient descent method and cannot be done automatically. One can claim that a similar argument is also valid for the approach presented in this paper. We can argue that the selection of  $\alpha$  parameter does affect the registration parameters (global and local estimates) but not the final registration result since by modifying  $\alpha$ , we theoretically change the estimates of the unknown parameters (global and local deformation measures) but not their joint estimates. This is clearly shown in [fig. (4)] where three different cases have been considered independently:

- A global rigid transformation [fig. (4.1)],
- A global-to-local rigid transformation with local deformations [fig. (4.2)],
- A local (pixel-wise registration) model [fig. (4.3)].

Based on the obtained results, we can claim that the registration performance of this representation for the three different type of transformations is comparable/similar. However, in order to overcome this limitation one can consider

modifying the objective function to penalize the formation of a rich local deformation field. The use of the local deformation field magnitude is a common selection to perform this task leading to the following example:

$$\begin{aligned}
 E(s, \theta, T, (u, v)) &= \alpha \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^T))^2 \\
 &+ (1 - \alpha)\beta \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) (s\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A^T + (u, v)))^2 \\
 &+ (1 - \alpha)(1 - \beta) \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) [(u_x^2 + u_y^2 + v_x^2 + v_y^2) + \gamma(u^2 + v^2)]
 \end{aligned}$$

However, it is important to note that the introduction of local deformations cannot guarantee the proper handling of the shape to be registered. For example several points or parts of the source shape may be mapped to the same point or part of the target even in the case where scale variations are not present [fig. (4)]. Thus, the transformed shape may be quite different or even an open structure compared with the original one. In order to deal with this issue, we can consider the use membership functions for each pixel that enforces a "one-to-one" pixel-wise correspondence between the current and the target shape or

### 3.3 Supervised Registration and Joint Learning

We now consider the registration problem in more complex scenario where the target is not a shape, but a shape model with local degrees of variability. We assume the existence of this model that refers to a distance function  $[\Phi_{\mathcal{S}}(x, y)]$  associated with some variability measurements  $[\sigma_{\mathcal{S}}(x, y)]$  :

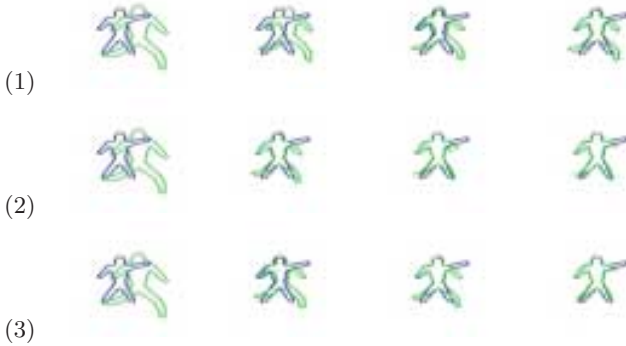
$$\Phi_{\mathcal{S}}(x, y) = \begin{cases} 0, & (x, y) \in \mathcal{S} \\ +D((x, y), \mathcal{S}) > 0, & (x, y) \in \mathcal{R}_{\mathcal{S}} \\ -D((x, y), \mathcal{S}) < 0, & (x, y) \in [\Omega - \mathcal{R}_{\mathcal{S}}] \end{cases}$$

Then, for a given pixel location  $(x, y)$  and a given value  $\phi$  the conditional probability of having this value at the location in  $\mathcal{S}$  is given by:

$$p_{\mathcal{S}(x,y)}(\phi) = \frac{1}{\sqrt{2\pi}\sigma_{\mathcal{S}}(x, y)} e^{-\frac{(\phi - \Phi_{\mathcal{S}}(x, y))^2}{2\sigma_{\mathcal{S}}^2(x, y)}}$$

The construction of this model can be done using variational framework that assumes the existence of a training set of shapes that are registered with respect to an arbitrary element of this set [17].

Given this model, a more challenging and promising task is to register a given shape  $\mathcal{D}$  by maximizing the maximum likelihood density between the model and the input shape. If we assume that the conditional densities of the model are independent across pixels, then the optimization criterion is equivalent with the



**Fig. 4.** (1) *Global*  $\{s = 0.79, \theta = 1.34^\circ, T_x = -16.34, T_y = -15.76\}$ , (2) *Global-to-Local*  $\{s = 0.81, \theta = 2.07^\circ, T_x = -14.97, T_y = -15.43\}$ , (3) *Local Registration with regularization constraints.*

minimization of:

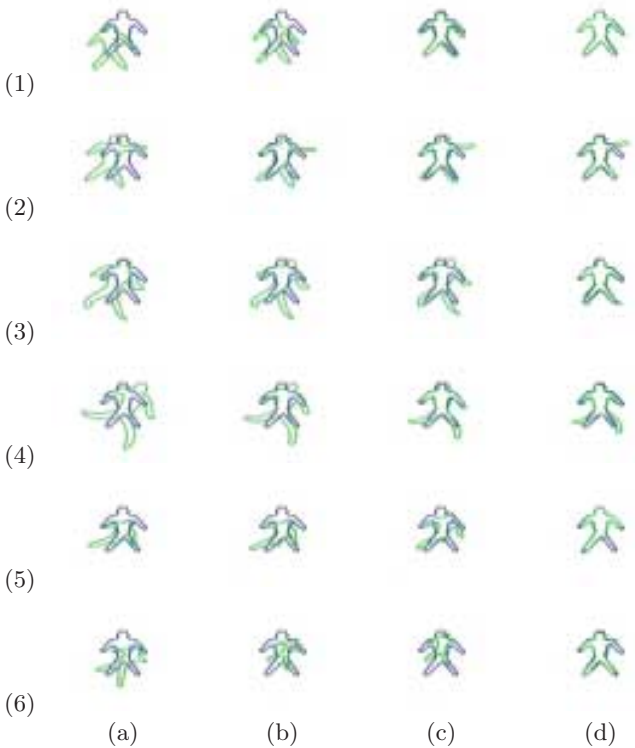
$$\begin{aligned}
 E(s, \theta, T, (u, v)) &= \alpha \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) \left( \log(\sigma_{\mathcal{S}}(A)) + \frac{(\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A))^2}{2\sigma_{\mathcal{S}}^2(A)} \right) \\
 &+ (1 - \alpha)\beta \iint_{\Omega} N_{\delta}(\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) \left( \log(\sigma_{\mathcal{S}}(A + (u, v))) + \frac{(\Phi_{\mathcal{D}} - \Phi_{\mathcal{S}}(A + (u, v)))^2}{2\sigma_{\mathcal{S}}^2(A + (u, v))} \right) \\
 &+ (1 - \alpha)(1 - \beta) \iint_{\Omega} N_{\delta}(s\Phi_{\mathcal{D}}, \Phi_{\mathcal{S}}) [u_x^2 + u_y^2 + v_x^2 + v_y^2 + \gamma(u^2 + v^2)]
 \end{aligned}$$

One can easily interpret this function. Shape components are considered according to their variability (denominator factor) in the registration process. Therefore, the global registration component will be recovered using the rigid parts of the shape model, while local deformations will be estimated accordingly in a less significant manner due to the contribution of the variability estimates.

## 4 Conclusions

In this paper, we have proposed a novel simple framework based on variational principles for global to local shape registration. The proposed framework makes use of a powerful mathematical tool, the level set representations that is integrated with variational (sum of squared differences) and stochastic principles (Supervised Registration and Joint Learning) resulting in a scale, translation and rotation invariant paradigm for shape registration. The extension of the method to deal with structures of higher dimension is a straight-forward step that is currently under investigation, as well as to deal with open structures.

The efficiency of the proposed formulation is demonstrated using a very simple optimization criterion, the sum of squared differences. Therefore, we can presume that the selected shape representation is very powerful, and has strong discrimination power. Very encouraging [fig. (5)] experimental results were obtained using all variations of the proposed framework.



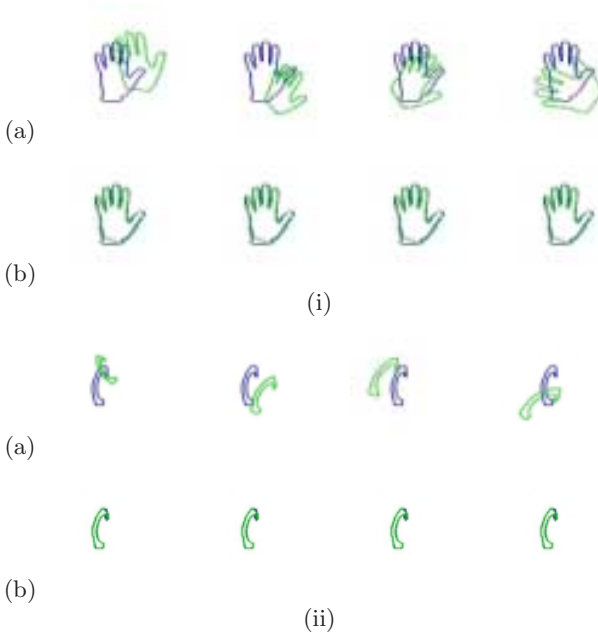
**Fig. 5.** *Global Registration for non-rigid objects: (1,2,3,4,5,6) Different examples. (a) Input Shapes, (b,c) Intermediate Results, (d) Registration Result.*

Conventional methods were considered to implement the obtained PDEs in this paper. Therefore, small time steps are required to guarantee stability and convergence. The estimation of the rigid transformation can be done in real time, due to the small number of involved parameters once an appropriate numerical approximation methods are used [25]. The current implementation takes up to a couple of seconds modulo the initial position of shapes. Recovering the complete local deformations field is very expensive.

In order to validate our approach, we have considered four different examples [fig. (6,7)] to validate the performance of the method. Cases with severe local deformations and topological changes resulting to missing and occluded parts have been also investigated. Towards this end, two fingers were progressively removed from the hand input shape [fig. (7)].

In order to perform reliable tests, the input shape has been globally and locally deformed. We have considered a four dimensional random variable  $(s, \theta, T_x, T_y)$  defined in the following space:

$$(\theta, s, T_x, T_y) \in \left( \left[-\frac{\pi}{3}, \frac{\pi}{3}\right], [0.8, 1.2], [-30, 30], [-30, 30] \right)$$

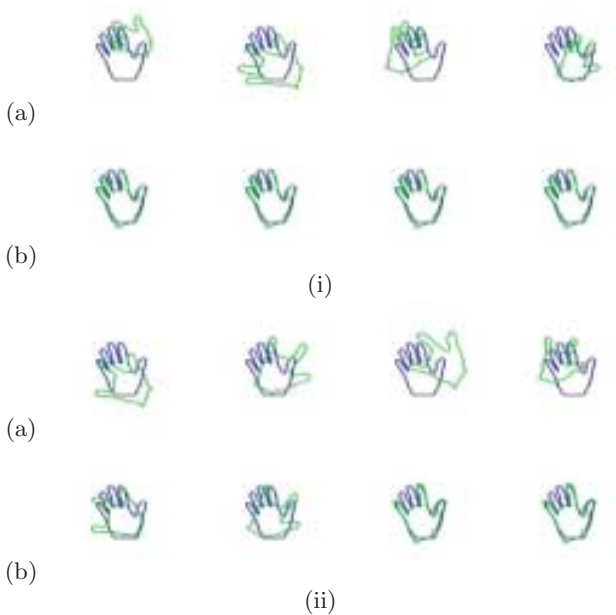


**Fig. 6.** Empirical Evaluation (a) Initial Condition, (b) Registration Result. Performance (Registration Ratio): (i) 100 %, (ii) 100 %.

For all examples we have ran a 100 trials using a random generation process for the parameters of the rigid transformation. The obtained results are shown in [fig. (6,7)]. Each column corresponds to a random trial. The first row (a), refers to the initial condition and the second row (b) to the final registration result. The registration performance of the algorithm is also shown.

As far the future directions of the proposed approach are concerned, several issues remain open. The acceleration of the method to perform real time registration is a natural step to be done. Moreover, the integration of image/intensity features into the shape registration algorithm can further improve the performance of the proposed framework. An hybrid approach that makes use primarily of the shape information and secondly of the image features can be a valuable element to the segmentation of medical structures where the global shape variability is not discriminant. Also, the investigation of more complex motion structures and global transformations is step to be done. Shape recognition is also an interesting application. One can consider (after alignment) the use of these representations as feature space to the recognition process.

The proposed framework can be considered static since it is based on already extracted shapes. In many cases, solving the segmentation problem cannot be done trivially and therefore shape registration becomes more complicated. To overcome this constraint, a joint optimization approach has to be investigated where both, the problems of segmentation and registration are treated simultaneously.



**Fig. 7.** *Empirical Evaluation (a) Initial Condition, (b) Registration Result. Performance (Registration Ratio): (i) 100 %, (ii) 77 %.*

**Acknowledgments.** The authors would like to thank C. Xu for fruitful discussions, the anonymous reviewers for their suggestions and Alok Gupta, the director of the IM Department at Siemens Corporate Research for encouraging the exploration of the topic and providing financial support to Mikael Rousson. We are also thankful to T. Sebastian and B. Kimia from Brown University for providing their shape database [19] and Professor Yunmei Chen, Department of Mathematics, Dr. Richard W. Briggs, Department of Radiology and Brain Institute, University of Florida and his group for providing the medical data.

## References

1. S. Belongie, J. Malik, and J. Puzicha. Matching Shapes. In *IEEE ICCV*, pages 456–461, Vancouver, Canada, 2001.
2. A. Blake and M. Isard. *Active Contours*. Springer-Verlag Press, 1997.
3. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *IEEE ICCV*, pages 694–699, Boston, USA, 1995.
4. C. Chefd’Hotel, G. Hermosillo, and O. Faugeras. A Variational Approach to Multi-Modal Image Matching. In *IEEE Workshop on Variational and Level Set Methods*, pages 21–28, 2001.
5. H. Chui and A. Rangarajan. A New Algorithm for Non-Rigid Point Matching. In *IEEE CVPR*, pages II: 44–51, Hilton Island, USA, 2000.

6. I. Cohen and I. Herlin. Curve Matching Using Geodesic Paths. In *IEEE CVPR*, pages 741–746, Santa Barbara, USA, 1998.
7. T. Cootes, C. Taylor, D. Cooper, and J. Graham. Active Shape Models - their traing and applications. *CVGIP: Image Understanding*, 61, 1995.
8. A. Fitzgibbon. Robust Registration of 2D and 3D Point Sets. volume 2, pages 411–420, 2001.
9. S. Joshi and M. Miller. Ladmark Matching via Large Deformation Diffeomorphism. *IEEE TIP*, 9:1357–1370, 2000.
10. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. In *IEEE ICCV*, pages 261–268, 1987.
11. B. Kimia, A. Tannenbaum, and S. Zucker. Shocks and Deformations i: The Components of two-dimensional Shape and the reaction-diffusion space. *IJCV*, 15:189–224, 1995.
12. D. Kozinska, O. Tretyak, J. Nissanov, and C. Ozturk. Multidimensional Alignment Using the Euclidean Distance Transform. *Graphical Models and Image Processing*, 6:373–385, 1997.
13. S. Lavalley and R. Szilinski. Recovery of the Position and Orientation of free-form Objects from Image Contours using 3D Distance Maps. *IEEE PAMI*, 17:378–390, 1995.
14. J. Maintz and M. Viergever. A Survey for Medical Image Registration. *Medical Image Analysis*, 2:1–36, 1998.
15. R. Malladi and J. Sethian. A unified framework for shape segmentation representation, and recognition. Technical Report LBL-36069 UC-405, Lawrence Berkeley Laboratory, Berkeley, 1994.
16. S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed : algorithms based on the hamilton-jacobi formulation. *Journal of Computational Physics*, 79:12–49, 1988.
17. N. Paragios and M. Rousson. Shape Priors for Level Set Representations. Copenhagen, Denmark, 2002.
18. A. Rosenfeld and J. Pfaltz. Distance Functions on Digital Pictures. *Pattern Recognition*, 1:33–61, 1968.
19. T. Sebastian, P. Klein, and B. Kimia. Recognition of Shapes by Editing Shock Graphs. In *IEEE ICCV*, pages 755–762, Vancouver, Canada, 2001.
20. J. Sethian. *Level Set Methods*. Cambridge University Press, 1996.
21. K. Siddiqi, A. Shokoufandeh, S. Dickinson, and S. Zucker. Shocks Graphs and Shape Matching. *IJCV*, 35:13–32, 1999.
22. K. Siddiqi, A. Shokoufandeh, S. Dikinson, and S. Zucker. Shock Graphs and Shape Matching. In *IEEE ICCV*, pages 222–229, Bombay, India, 1998.
23. R. Veltkamp and M. Hagedoorn. State-of-the-art in Shape Matching. Technical Report UU-CS-1999-27, Utrecht University, Sept. 1999.
24. P. Viola and W. Wells. Aligment by Maximization of Mutual Information. In *IEEE ICCV*, pages 16–23, Boston, USA, 1995.
25. J. Weickert, B. M. t. Haar Romeny, and M. Viergener. Efficient and Reliable Scheme for Non-Linear Diffusion and Filtering. *IEEE TIP*, 7:398–410, 1998.
26. S. Zhu and A. Yuille. FORMS: A flexible object recognition and modeling system. *IJCV*, 20:187–212, 1996.