Modelling Shapes with Uncertainties : Higher Order Polynomials, Variable Bandwidth Kernels and non Parametric Density Estimation.*

Maxime TARON Nikos PARAGIOS Marie-Pierre JOLLY

CERTIS,
Ecole Nationale des Ponts et Chaussées,
77455 Marne la Vallee, France
{maxime.taron ; nikos.paragios}@certis.enpc.fr

Imaging & Visualization Dept.,
Siemens Corporate Research,
08540 Princeton, USA
marie-pierre.jolly@siemens.com

Abstract

In this paper, we introduce a new technique for shape modelling in the space of implicit polynomials. Registration consists of recovering an optimal one-to-one transformation of a higher order polynomial along with uncertainties measures that are determined according to the covariance matrix of the correspondences at the zero isosurface. In the modelling phase, these measures are used to weight the importance of the training samples phase according to a variable bandwidth non-parametric density estimation process. The selection of the most appropriate kernels to represent the training set is done through the maximum likelihood criterion. Excellent results for patterns of digits, related with the registration and the modelling aspects of our approach demonstrate the potentials of our method.

1. Introduction

Domain knowledge is often available in computational vision and therefore efficient techniques are to be developed to account for it. However, modelling the geometric form of objects is a challenging task. Such a task consists of two steps, (i) registration, and (ii) statistical modelling. Prior work consists of addressing registration and modelling in a sequential fashion. Within such an approach registration errors are not accounted for and often lead to incorrect and erroneous models. We are proposing a new modelling technique to account for uncertainties in the registration process.

Shapes can be represented in many different ways. The simplest form is a collection of points. More advanced representations of shapes include B-splines as well as other form of continuous interpolation functions [22], shocks, skeletons [11] and distance transforms [1].

One can define the registration problem as follows: recover a transformation between a source and a target shape that results in meaningful correspondences between their basic elements. To this end, one (i) should select an appropriate representation for the structures of interest, (ii) define the set and the nature of plausible transformations, and (iii) determine an appropriate mathematical framework to recover the optimal registration parameters.

Registration can be either global or local. Global parametric transformations are within a restricted group, like rigid, similarity, affine, etc. The term local registration refers to a transformation with infinite degrees of freedom. Such a deformation can potentially map any finite number of points to the same number of points. However, non-rigid registration is often an underconstrained problem. Therefore in order to find a unique non-rigid transformation, we need to further constrain the problem through a regularization of the registration field. Point-based global and local registration [23] through low cost optimization techniques like the ICP [4] algorithm is the most primitive approach to shape registration. One can also refer to more advanced methods like diffeomorphic matching [5].

A different approach consists of addressing registration as a statistical estimation problem through successive steps. Within each step the uncertainty in the estimates is being computed [19] and is used to guide further steps of the overall algorithm [15]. In [18] the covariance matrix is used within an ICP algorithm to sample the correspondences so that registration is well-constrained in all directions in the parameter space. In [17] local deformation and uncertainties are simultaneously recovered for the optical flow estimation problem through a Gaussian noise assumption on the observation.

*Research work supported by the Imaging & Visualization Department of Siemens Corporate Research; Princeton USA

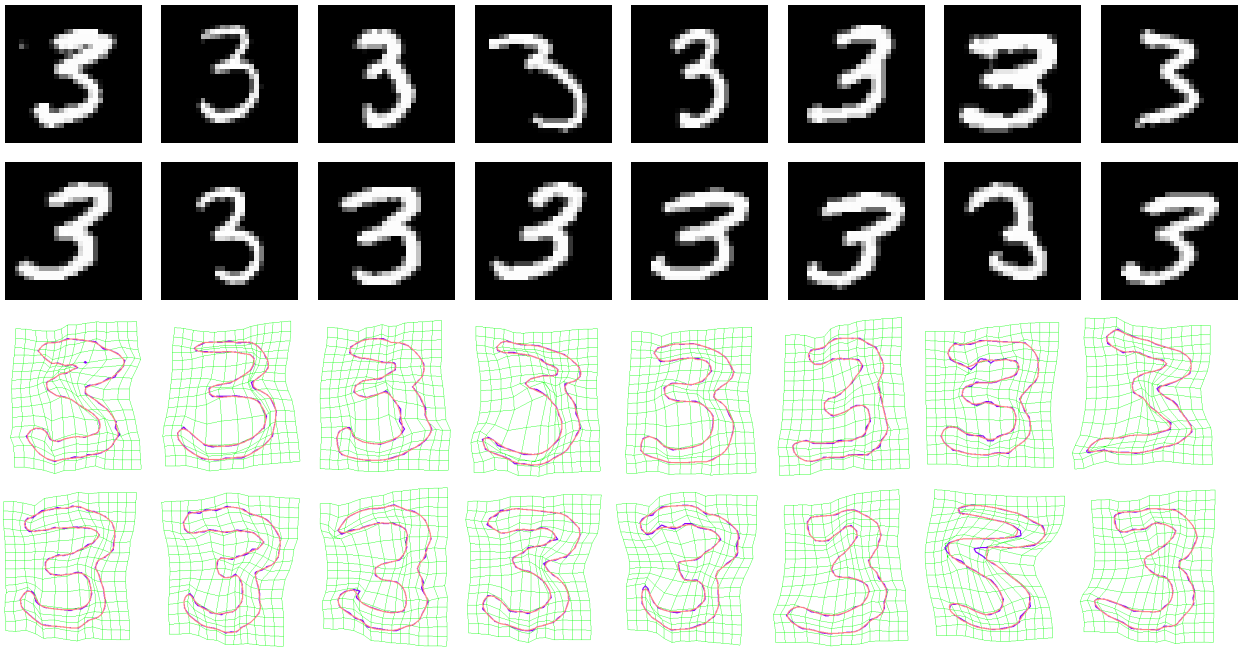

Figure 1. *Examples of registration of various samples of '3' using a* 16×14 *FFD grid.*

The modelling aspect consists of (i) selecting the nature of the density function, and (ii) recovering the parameters of such a function so it approximates the registered data. Parametric linear models like Gaussian densities are often employed through either an EM algorithm or a singular value decomposition. One can claim that such models refer to efficient compact approximations when the selected model fits the data. Non-parametric approaches of fixed bandwidth kernels like Parzen windows [9, 7] are a more efficient technique to approximate data that do not obey any particular rule. However, important attention is to be paid on the selection of their bandwidth. Also, kernel selection schemes are available to reduce the computational complexity.

In this paper we propose a novel technique to shape modelling that exploits registration uncertainties. To this end shapes are represented in an implicit fashion and are registered using a free form deformation (FFD) model according to a topology-preservation algorithm. We show how we can compute uncertainty measures by estimating a covariance matrix at the zero iso-surface. These measures are then used within a variable bandwidth kernel-based density function to estimate a probability density function of the class of objects. We use an iterative algorithm to reduce the number of kernels and speed up the computation time. Given a new unknown example, registration is first applied and uncertainties are estimated. We then propose a measure do explicitly encode the estimates and their uncertainties and evaluate the probability of the subject under consideration being part of the family of the model.

The remainder of the paper is organized in the following fashion. In section 2 we briefly present shape registration in the space of implicit polynomials while the estimation of uncertainties is described in section 3. The objective of building compact non-parametric densities to describe shapes is addressed in section 4. Experimental results are presented in section 5. Conclusions and future directions are provided in section 6.

2 Registration through Implicit Polynomials

In the present framework, a shape $\mathcal{S}$ is represented in an implicit fashion using the Euclidean distance transform $\mathcal{D}$ [1]. In the 2D case, we consider the function defined on the image domain Ω and $\mathcal{R}_{\mathcal{S}}$ is the region enclosed by $\mathcal{S}$:

$$\phi_{\mathcal{S}}(x,y) = \begin{cases} 0, & (x,y) \in \mathcal{S} \\ +\mathcal{D}((x,y),\mathcal{S}), & (x,y) \in \mathcal{R}_{\mathcal{S}} \\ -\mathcal{D}((x,y),\mathcal{S}), & (x,y) \in \bar{\mathcal{R}_{\mathcal{S}}} \end{cases}$$

Such a space is invariant to translation and rotation and can also be modified to account for scale variations. In the most general case an apparent relation between the distance function of the source and the target is not present.

Now consider a smooth diffeomorphism defined on the image domain Ω and with the vector of parameters $\Theta \in \mathbb{R}^n$:

$$\mathcal{L}(\Theta, .) : \Omega \rightarrow \Omega$$

Standard point-based curve registration consists of applying $\mathcal{L}$ to the source shape $\mathcal{S}$ and minimizing the curve inte-

gral along $\mathcal{S}$ such that some metric error between the transformed source and the target is minimal:

$$E_0(\mathcal{L}(\Theta)) = \oint_{\mathcal{S}} \rho(\phi_{\mathcal{T}}(\mathcal{L}(\Theta, \mathbf{x}))d\mathbf{s}$$

where ρ is a robust estimator. One can extend registration within a band of information along numerous image isosurfaces:

$$E_\alpha(\mathcal{L}(\Theta)) = \iint_\Omega \chi_\alpha(\phi_{\mathcal{S}}(\mathbf{x}))\rho\left(\phi_{\mathcal{S}}(\mathbf{x}) - \phi_{\mathcal{T}}(\mathcal{L}(\Theta, \mathbf{x}))\right) d\mathbf{x}$$

where we introduce the indicator function:

$$\chi_\alpha(x) = \begin{cases} 1/(2\alpha) & \text{if } x \in [-\alpha, \alpha] \\ 0 & \text{else} \end{cases}$$

Within such a process the selection of the parameter α is crucial since to some extent it refers to the scale of the shapes to be registered. On the other hand, it is natural when converging to the optimal solution that α tends to 0. Therefore, we assume a finite number of decreasing set of radii $\{\alpha_0 > ... > \alpha_t > ... > \alpha_n \approx 0\}$ that is equivalent to a scale-space decomposition of the process. If Θ is too large, there is a high risk of converging to a local minimum. So, we progressively increase the complexity of the transformation and therefore the size of Θ as α_k decreases.

Let Θ_{t-1} be the parameters defining the transformation $\mathcal{L}_{t-1} = \mathcal{L}(\Theta_{t-1}, .)$ for which the energy was minimum at scale $t-1$. Also let $\mathcal{S}^{t-1} = \mathcal{L}_{t-1} \circ \mathcal{S}$. The registration between shapes is then equivalent to iteratively minimizing :

$$E_{\alpha_t}(\mathcal{L}(\Theta)) = \iint_\Omega \chi_{\alpha_t}(\phi_{\mathcal{S}}(\mathbf{x}))\rho(\phi_{\mathcal{S}^{t-1}}(\mathcal{L}_{t-1}(\mathbf{x})) - \phi_{\mathcal{T}}(\mathcal{L}(\Theta, \mathbf{x})))d\mathbf{x}$$

where a correction process is applied when refining scales through the modification of the distance transform that describes the source shape $\phi_{\mathcal{S}^{t-1}}()$. Within such a formulation the integration domain is always related to the initial source shape and does not depend on the number of iteration or the parameter α_t. Moreover when using the Euclidean distance and α_t tends to 0, $E_{\alpha_t}(\mathcal{L}(\Theta))$ is equivalent to the point based registration ($E_{\alpha_\infty}(\mathcal{L}(\Theta)) = E_0(\mathcal{L}(\Theta))$).

Such an objective function can be used to address global registration as well as local deformations. We use an affine transformation (with six degrees of freedom) to represent the global transformation and a free form deformation as in [8] to address the local deformations. Cubic B-spline based free form deformations are an efficient way to model locally smooth transformations on images [16]. Deformations of shapes (and their implicit representation $\phi_{\mathcal{S}}$) are recovered by evolving a square control lattice $\mathbf{P}$ that is overlaid on the initial distance transform structure. Let us consider the control lattice points $\{\mathbf{P}_{m,n}\}$ defining the initial regular grid. The displacement of any of control point will induce a local and $\mathcal{C}^2$ field of deformation:

$$\mathcal{L}(\Theta, \mathbf{x}) = \sum_{k=-1}^{2} \sum_{l=-1}^{2} B_k(u) B_l(v)(\mathbf{P}_{i+k,j+l} + \delta\mathbf{P}_{i+k,j+l})$$

where $\mathbf{x} = (u, v)$ and B_k is the k^{th} basis function of the cubic B-spline. This local transformation is a compromise between global and local registration and its parameters consist of the displacement of the control points ($\Theta = \{\delta\mathbf{P}_{m,n}\}$). In [8], such a framework is introduced using implicit functions defined on the complete domain Ω.

To recover a smooth transformation and avoid folding, we adopt a regularization term motivated by the thin plate energy functional [21] to control the spatial variations of the displacement:

$$E_{\text{smooth}}(\mathcal{L}(\Theta)) = \iint_\Omega \left(|\mathcal{L}_{xx}|^2 + 2\,|\mathcal{L}_{xy}|^2 + |\mathcal{L}_{yy}|^2\right) d\Omega$$

that can be further simplified in the case of the cubic B-spline to the quadratic form $[E_{\text{smooth}}(\mathcal{L}(\Theta)) = \Theta^T \mathrm{C} \Theta]$ with C a symmetric matrix.

The objective function $[E_{\alpha_\infty}(\mathcal{L}(\Theta)) + wE_{\text{smooth}}(\mathcal{L}(\Theta))]$ is optimized using a standard gradient descent method leading to exceptional results. The method was tested for 2000 digits of the number '3' from the MNIST database and we qualitatively judged the registration results to be good in 98.2% or the cases. Figure 1 shows some registration examples. The top two rows show the original image examples. Each example was globally aligned to the model using an affine transformation. Then, the FFD grid associated with the model was deformed to align to the example. The bottom three rows of Figure 1 show the deformed model, the affine transformed example, and the deformation grid. What we observe is that the two contours coincide very well, which shows that the registration results are excellent. Some examples of cases where the method has failed are shown in Figure 2. In the left example, the model did not deform enough. In the middle example, the model is perfectly aligned with the example, but the deformation grid contains a few irregularities. Finally, in the right example, a loop appears in the deformed model. These cases are very rare though, only 1.8% of the 2000 cases of '3's.

However, one can claim that the local deformation field is not sufficient to characterize the registration between two shapes. Often data is corrupted by noise while at the same time outliers exist in the training set. Therefore recovering measurements of the quality of the registration is an eminent condition for accurate shape modelling.

3. Estimation of Registration Uncertainties

Several attempts to build statistical models on noisy set of data in order to infer the properties of a certain model have been proposed in the literature. In [10], various techniques were reported to extract feature points in images along with uncertainties due to the inherent noise. In [15], an iterative estimation method was proposed to handle uncertainty estimates of rigid motion on sets of matched

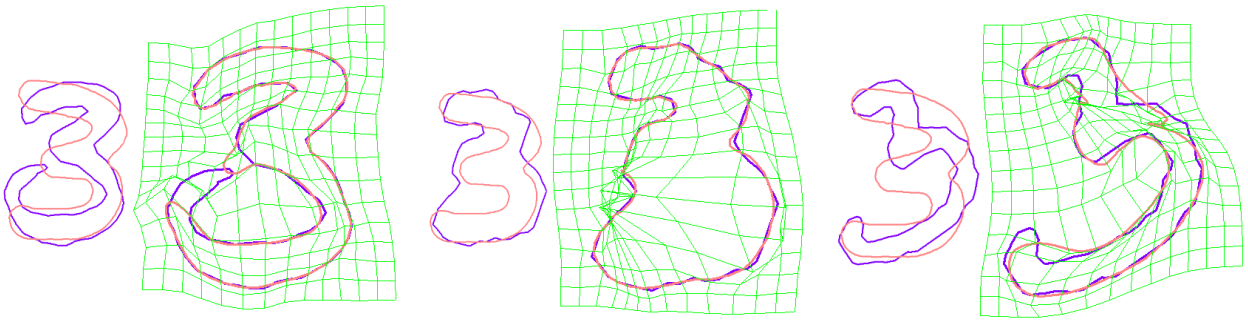

Figure 2. *Examples of erroneous registration results.*

points. In [18] an iterative technique to determine uncertainties within the ICP [4] registration algorithm was proposed. In a quite different context, [17] introduced uncertainties within the estimation of dense optical flow, that can be seen as a form of registration between images.

In the present case curves are considered using implicit representation, therefore uncertainty does not lie in the relative position of points but of an isosurface and therefore, the problem can be seen as equivalent to the "'aperture problem"' in optical flow estimation. Inspired by the work in [2, 18] we aim to recover uncertainties on the vector Θ while being able to use only the zero iso-surface of $\phi_{\mathcal{S}}$, defining the shape itself. To this end, we use a discrete formulation of the energy $E_0 = E_{\alpha_\infty}$, by summing along points regularly spaced on the source contour :

$$E_0(\Theta) = \sum_{i=1}^{K} \rho(\phi_{\mathcal{T}}(\mathcal{L}(\Theta, \mathbf{x}_i)) = \sum_{i=1}^{K} \rho(\phi_{\mathcal{T}}(\mathbf{x}'_i))$$

Let us consider $\mathbf{q}_i$ to be the closest point on the target contour from $\mathbf{x}'_i$. Since $\phi_{\mathcal{T}}$ is assumed to be a Euclidean distance transform, it satisfies the condition $\|\nabla\phi_{\mathcal{T}}(\mathbf{x}'_i)\| = 1$. Therefore one can express the values of $\phi_{\mathcal{T}}(\mathbf{x}'_i)$:

$$\phi_{\mathcal{T}}(\mathbf{x}'_i) = \|\mathbf{x}'_i - \mathbf{q}_i\| = (\mathbf{x}'_i - \mathbf{q}_i)\nabla\phi_{\mathcal{T}}(\mathbf{x}'_i)$$

Then, one has a first order approximation of $\phi_{\mathcal{T}}(\mathbf{x})$ in the neighborhood of $\mathbf{x}'_i$, in the form:

$$\begin{aligned}\phi_{\mathcal{T}}(\mathbf{x}'_i + \delta\mathbf{x}'_i) &= \phi_{\mathcal{T}}(\mathbf{x}'_i) + \delta\mathbf{x}'_i\nabla\phi_{\mathcal{T}}(\mathbf{x}'_i) \\ &= (\mathbf{x}'_i + \delta\mathbf{x}'_i - \mathbf{q}_i)\nabla\phi_{\mathcal{T}}(\mathbf{x}'_i)\end{aligned}$$

that reflects the condition that a point to curve distance is adopted rather than a point to point. Under the assumption that $E_0(\mathcal{L}(\Theta)) = \circ(1)$ we can neglect the second order term in the development of $\phi_{\mathcal{T}}$ and therefore write the following second order approximation of E_0 in quadratic form:

$$E(\mathcal{L}(\Theta)) = \sum \left[(\mathcal{L}(\Theta, \mathbf{x}_i) - \mathbf{q}_i)\nabla\phi_{\mathcal{T}}(\mathbf{x}'_i)\right]^2$$

A free form deformation is a linear transformation with respect to the parameters $\Theta = \delta\mathbf{P}_{i,j}$. Therefore one can rewrite this transformation over the image domain in a rather compact form:

$$\begin{aligned}\mathcal{L}(\Theta, \mathbf{x}) &= \mathbf{x} + \sum_{k=-1}^{2}\sum_{l=-1}^{2} B_k(u)B_l(v)\delta\mathbf{P}_{i+k,j+l} \\ &= \mathbf{x} + \mathcal{X}(\mathbf{x})\Theta.\end{aligned}$$

where $\mathcal{X}(\mathbf{x})$ is a matrix of dimensionality $2 \times N$ with N being the size of Θ. One now can substitute this term in the objective function:

$$E(\Theta) = (\hat{\mathcal{X}}\Theta - \mathbf{y})^T(\hat{\mathcal{X}}\Theta - \mathbf{y})$$

with

$$\hat{\mathcal{X}} = \begin{pmatrix} \eta_1^T\mathcal{X}(\mathbf{x}_1) \\ \vdots \\ \eta_K^T\mathcal{X}(\mathbf{x}_K) \end{pmatrix} \text{ and } \mathbf{y} = \begin{pmatrix} \eta_1^T(\mathbf{q}_1 - \mathbf{x}_1) \\ \vdots \\ \eta_K^T(\mathbf{q}_K - \mathbf{x}_K) \end{pmatrix}$$

and $[\eta_i = \nabla\phi_{\mathcal{T}}(\mathbf{x}'_i)]$. We assume that $\mathbf{y}$ is the only random variable. Such assumption is equivalent with saying that errors in the point positions are only quantified along the normal direction. This accounts for the fact that the point set is treated as samples extracted from a continuous manifold. One can take the derivative of the objective function in order to recover a linear relation between Θ and $\mathbf{y}$:

$$\hat{\mathcal{X}}^T\hat{\mathcal{X}}\Theta = \hat{\mathcal{X}}^T\mathbf{y}$$

Last, assume that the components of $\mathbf{y}$ are independent and identically distributed. In that case, the covariance matrix of $\mathbf{y}$ has the form $\sigma^2\mathbf{I}$ of magnitude σ^2 with $\mathbf{I}$ being the identity. In the most general case one can claim that the matrix $\hat{\mathcal{X}}^T\hat{\mathcal{X}}$ is not invertible due to the fact that the registration problem is underconstrained. Additional constraints are to be introduced towards the estimation of the covariance matrix of Θ through the use of an arbitrarily small positive parameter γ :

$$E(\Theta) = (\hat{\mathcal{X}}\Theta - \mathbf{y})^T(\hat{\mathcal{X}}\Theta - \mathbf{y}) + \gamma\,\Theta^T\Theta$$

Then the covariance matrix of the parameter estimate is :

$$\Sigma_\Theta = \sigma^2(\hat{\mathcal{X}}^T\hat{\mathcal{X}} + \gamma\mathbf{I})^{-1}$$

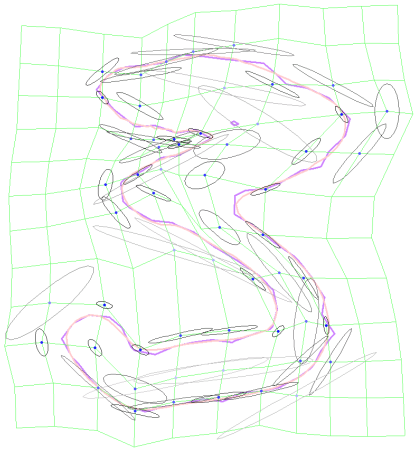

Figure 3. *Projection of the covariance matrix Σ_Θ on the grid points.*

Some examples of such estimates are shown in Figure 3 where 2×2 projections of the $N \times N$ uncertainty matrices are drawn on the grid points.

Modelling the registered examples according to some density function is the next step. To this end, two critical issues are to be addressed: the form of the PDF as well as the procedure to determine the corresponding parameters. In the most general case deformations of shapes that refer to objects of particular interest cannot be modeled with simple parametric models like Gaussians. Therefore within our approach we propose a non-parametric form of the PDF.

4 Variable Bandwidth Density Estimation [§]

Let $\{\mathbf{x}_i\}_{i=1}^M$ denote a random sample with common density function f. The fixed bandwidth kernel density estimator consists of:

$$\hat{f}(\mathbf{x}) = \frac{1}{M}\sum_{i=1}^{M} \mathrm{K}_{\mathbf{H}}\left(\mathbf{x}-\mathbf{x}_i\right) = \frac{1}{M}\sum_{i=1}^{M} \frac{1}{\|\mathbf{H}\|^{1/2}} \mathrm{K}\left(\mathbf{H}^{-1/2}(\mathbf{x}-\mathbf{x}_i)\right)$$

where $\mathbf{H}$ is a symmetric definite positive - often called a bandwidth matrix - that controls the width of the kernel around each sample point $\mathbf{x}_i$. The fixed bandwidth approach often produces undersmoothing in areas with sparse observations and oversmoothing in the opposite case. Usefulness of varying bandwidths is widely acknowledged to estimate long-tailed or multi-modal density functions with kernel methods.

[§]The author gratefully acknowledge Dr. Jean-Yves Audibert for fruitful discussion regarding the probabilistic models.

In the literature, kernel density estimation methods that rely on such varying bandwidths are generally referred to as adaptive kernel density estimation methods [20]. Two useful state of-the-art variable bandwidth kernels are the *sample point estimator* and the *balloon estimator*.

The *sample point estimator* refers to a covariance matrix depending on the repartition of the points constituting the sample :

$$\hat{f}_S(\mathbf{x}) = \frac{1}{M}\sum_{i=1}^{M} \frac{1}{\|\mathbf{H}(\mathbf{x}_i)\|^{1/2}} \mathrm{K}\left(\mathbf{H}(\mathbf{x}_i)^{-1/2}(\mathbf{x}-\mathbf{x}_i)\right)$$

where a common selection of $\mathbf{H}$ is $\mathbf{H}(\mathbf{x}_i) = h(\mathbf{x}_i)\cdot\mathbf{I}$ with $h(\mathbf{x}_i)$ being the distance of point $\mathbf{x}_i$ from the k^{th} nearest point. One can consider various alternatives to determine the bandwidth function. Such estimator may be directly used with the uncertainties calculated in section 3 and $\mathbf{H}(\mathbf{x}_i) = \mu\Sigma_{\Theta_i}$ as proposed in [3, 6].

The *balloon estimator* adapts to the point of estimation depending on the shape of the sampled data according to:

$$\hat{f}_B(\mathbf{x}) = \frac{1}{M}\sum_{i=1}^{M} \frac{1}{\|\mathbf{H}(\mathbf{x})\|^{1/2}} \mathrm{K}\left(\mathbf{H}(\mathbf{x})^{-1/2}(\mathbf{x}-\mathbf{x}_i)\right)$$

where $\mathbf{H}(\mathbf{x})$ may be chosen with the same model as for the *sample point estimator*. Such function may be seen as the average of a density associated with the estimation point $\mathbf{x}$ on all the sample points $\mathbf{x}_i$. One should point out that such a process could lead to estimates on $\hat{f}(\mathbf{x})$ that do not refer to density function because it might be discontinuous and its integral is infinity.

Let us consider $\{\mathbf{x}_i\}_{i=1}^M$ a multi-variate set of measurements where each sample $\mathbf{x}_i$ exhibits uncertainties in the form of a covariance matrix Σ_i. Our objective can be stated as follows: estimate the probability of a new measurement $\mathbf{x}$ that is associated with covariance matrix Σ.

Let $\mathbf{X}$ be the random variable associated with the training set and assume a density function f. f may be estimated with $\hat{f}$ using the *sample point estimator*. Therefore $\hat{f}$ may be expressed in the form $\hat{f} = \sum \hat{f}_i$ where $\hat{f}_i$ are densities associated with a single kernel $\{\mathbf{x}_i, \mathbf{H}(\mathbf{x}_i)\}$. Let $\mathbf{Y}$ be a random variable for the new sample with estimated density $\hat{g}$.

Then one can claim that in order to estimate the probability of the new sample, one should first determine for all possible $\mathbf{u} \in \mathbb{R}^N$ their *distance* $\hat{f}(\mathbf{u})$ from the existing kernels of the training set $\mathbf{X}$ and weight them according to their fit with the estimated density function of $\mathbf{Y}$:

$$p_{\mathbf{Y}} = \int \hat{f}(\mathbf{u})\hat{g}(\mathbf{u})d\mathbf{u} = \int \left[\sum_{i=1}^{M} f_i(t)\right] g(t)dt = \sum_{i=1}^{M}\left[\int f_i(t)g(t)dt\right]$$

In the case of gaussian kernels for g (centered at $\mathbf{x}$) and the f_i (centered at $\mathbf{x}_i$)the following expression is recovered:

$$\hat{f}_G(\mathbf{x}) = \frac{1}{M(2\pi)^{d/2}} \sum_{i=1}^{M} \frac{1}{\|\Sigma + \Sigma_i\|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}_i))^\top (\Sigma + \Sigma_i)^{-1}(\mathbf{x} - \mathbf{x}_i)\right)$$

Such an expression has a simple mathematical interpretation: Consider two points $\{\mathbf{x}_1, \mathbf{x}_2\}$ with associated uncertainty $\{\Sigma_1, \Sigma_2\}$. Assuming that these are the parameters (mean and variance) of two independent random variables with normal distribution

$$\{\mathbf{X}_1 \sim N(\mathbf{x}_1, \Sigma_1), \mathbf{X}_2 \sim N(\mathbf{x}_2, \Sigma_2)\}$$

Then the random variable $\mathbf{Z} = \mathbf{X}_1 - \mathbf{X}_2$ follows a distribution $N(\mathbf{x}_1 - \mathbf{x}_2, \Sigma_1 + \Sigma_2)$ and the density at $\mathbf{Z} = 0$ is given by

$$p(\mathbf{X}_1 = \mathbf{X}_2) = \frac{1}{(2\pi)^{d/2}\|\Sigma_1 + \Sigma_2\|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}_1 - \mathbf{x}_2))^T (\Sigma_1 + \Sigma_2)^{-1}(\mathbf{x}_1 - \mathbf{x}_2)\right)$$

The present concept could be relaxed to address the case of non-gaussians kernels according to a *hybrid* estimator that is considered in the present paper :

$$\hat{f}_H(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^{M} \mathcal{K}(\mathbf{x}, \Sigma, \mathbf{x}_i, \Sigma_i)$$
$$= \frac{1}{M} \sum_{i=1}^{M} \frac{1}{\|\mathbf{H}(\Sigma, \Sigma_i)\|^{1/2}} \mathrm{K}(\mathbf{H}(\Sigma, \Sigma_i)^{-1/2}(\mathbf{x} - \mathbf{x}_i)$$

Such a density estimator takes into account the uncertainty estimates both on the sample points themselves as well as on the estimation of point $\mathbf{x}$ as introduced in [14]. The outcome of this estimator may be seen as the average of the probabilities that the estimation measurement is equal to the sample measurement, calculated over all sample measurements. Consequently, the density estimation decreases more slowly in directions of large uncertainties.

This measure can now be used to assess the probability for a new sample of being part of the modeled class in an approach that accounts for the non-parametric form of the observed density. The problem however is that this technique is very time consuming since the computation is linear in the number of samples in the training set. Therefore, there is an eminent need on decreasing the cardinality of the set of retained kernels.

The goal is to select a subset of kernels to maximize the likelihood of the training set. Consider a set $\mathcal{Z}_K = \{X_1, X_2, \ldots, X_K\}$ of kernels extracted from the training set. These have associated mean and uncertainties $\{X_j = \{\mathbf{x}_i \Sigma_i\}\}_{i=1}^{K}$. The log likelihood of the entire training set according to this model is:

$$C_{\mathrm{K}} = \sum_{i=1}^{M} \log\left(\frac{1}{K} \sum_{(\mathbf{x}_j, \Sigma_j) \in \mathcal{Z}_{\mathrm{K}}} \mathcal{K}(\mathbf{x}_j, \Sigma_j, \mathbf{x}_i, \Sigma_i)\right)$$

Where $\{Y_j = \{\mathbf{x}_j, \Sigma_j\}\}_{j=1}^{M}$ denote the kernels of the training set. We use an efficient sub-optimal iterative algorithm to update the set $\mathcal{Z}_K$. A new kernel $Y = \{\mathbf{x}, \Sigma\}$ is extracted from the training set as the one maximizing the quantity C_{K+1} associated with $\mathcal{Z}_{K+1} = \mathcal{Z}_K \bigcup Y$. One kernel may be chosen several times in order to preserve a decreasing order of C_K when adding new kernels. Consequently the selected kernels X_i in $\mathcal{Z}_{\mathrm{K}}$ are also associated with a weight factor w_i. Once such a selection has been completed, the hybrid estimator is evaluated over $\mathcal{Z}_{\mathrm{K}}$:

$$\hat{f}_H(\mathbf{x}, \Sigma) = \sum_{(\mathbf{x}_i, \Sigma_i, w_i) \in \mathcal{Z}_{\mathrm{K}}} w_i . \mathcal{K}(\mathbf{x}, \Sigma, \mathbf{x}_i, \Sigma_i)$$

5 Validation

The proposed method is intended to provide efficient models for family of shapes with important variation. Handwritten digits exhibit a very large variation among individual examples. Based on this observation, we have learned the shape of three different digits, namely, 3, 4, and 9. We used 2000 examples of each digit from MNIST digit database [12] to build the model. We then used the kernel selection algorithm to retain 50 kernels.

To verify that our method can encode the shape properties of the class of objects of interest, we ran a cross validation test, where each of the 3 models was registered to all 6000 digits. We then computed the hybrid estimator for the probability of the digit belonging to the class of the model. Figures 4 to 6 show the results. Figure 4 represents the matching of 3's and 4's. The X-axis is the likelihood that an example belongs to the class of '3' ($-log(\hat{f}(x, \Sigma))$) and the Y-axis is the likelihood that an example belongs to the class '4'. It can be seen that the two classes are very well separated. To demonstrate the separation, we used a simple support vector machine classifier [13] to linearly separate the two classes in the space of likelihood measured. The linear boundary in also shown in Figure 4. The correct classification rate was 99.17%. Figure 5 illustrates the separation between classes 3 and 9, the correct classification rate was 98.73%. Finally, Figure 6 illustrates the separation between classes 4 and 9, the correct classification rate was 94.83%. Table 1 shows the overall confusion matrix.

The results are consistent with what was expected. The lowest classification rate was obtained when comparing the 4's and the 9's. These digits are indeed very similar when handwritten by Americans, as can be seen from Figure 7. We can also see that 3 and 9 look more alike than 3 and

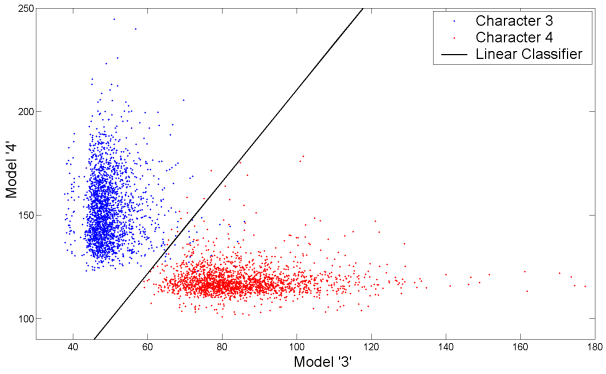

Figure 4. *Distribution of the digits 3 and 4 in the space of likelihoods of belonging to the classes '3' and '4'.*

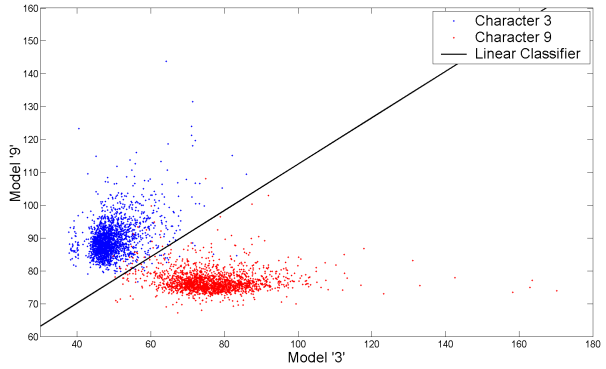

Figure 5. *Distribution of the digits 3 and 9 in the space of likelihoods of belonging to the classes '3' and '9'.*

4. It is important to note that the proposed method is not intended for such an application. However, given this validation we claim that such a model can capture samples of increasing complexity. Also, the use of deformations along with uncertainties provide efficient density estimators.

6 Conclusion

We have introduced an original framework to estimate uncertainty in the process of registration of shapes. We take advantage of this additional knowledge to build an efficient probabilistic descriptor of a certain class of shapes.

Future directions exists in the registration as well as the modeling aspect of our approach.

First, in the registration process, uncertainties could be propagated through scale when updating the transformation. We shall also notice that the uncertainties calculated on a certain FFD-grid could be extended to any finer grid and therefore qualify the density probability of any image transformation without the limitation of the choice of parameters.

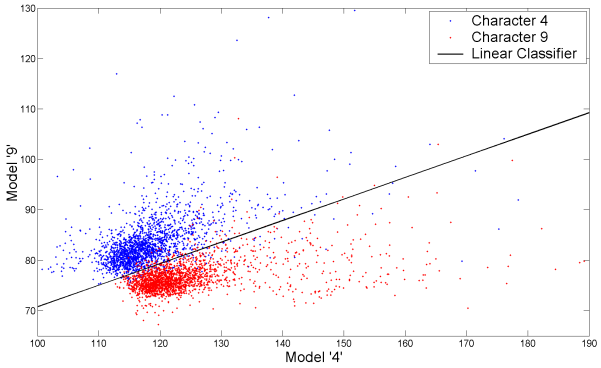

Figure 6. *Distribution of the digits 4 and 9 in the space of likelihoods of belonging to the classes '4' and '9'.*

	'3'	'4'	'9'
'3'	0.9845	0.0065	0.0090
'4'	0.0045	0.9385	0.0570
'9'	0.0145	0.0425	0.943

Table 1. *Confusion matrix between the three classes of digits 3, 4, and 9.*

Another path will be the exploration of the kernel used to make a Parzen-Window like density estimation into more advanced kernel-based learning methods such as kernel-PCA. The issue of defining the right Mercer kernel has however to be addressed.

Last but not least, this evaluation of densities using uncertainty has to be exported to the more general problem of image registration with prior knowledge. Consider an original image used as a model with the region of interest manually delineated. Then, registration can be performed

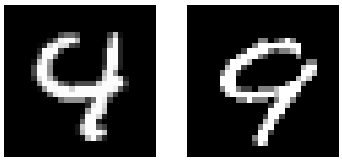

Figure 7. *Digits 4 and 9 can be very similar.*

with a shape term that directly handle the parameters of the transformation $\mathcal{L}$. Eventually, a calculation of uncertainties qualifying the present image registration may enhance the confidence for this term when using the *hybrid* estimator.

References

[1] G. Borgefors. Distance Transformations in Digital Images. *Computer Vision, Graphics, and Image Processing*, 34:344–371, 1986.

[2] A. Can, C. Stewart, B. Roysam, and H. Tanenbaum. A Feature-Based, Robust, Hierarchical Algorithm for Registering Pairs of Images of the Curved Human Retina. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:347–364, 2002.

[3] H. Chen and P. Meer. Robust computer vision through kernel density estimation. In *ECCV (1)*, pages 236–250, 2002.

[4] Y. Chen and G. Medioni. Object Modeling by Registration of Multiple Range Images. *Image and Vision Computing*, 10:145–155, 1992.

[5] H. Chui and A. Rangarajan. A New Algorithm for Non-Rigid Point Matching. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages II: 44–51, 2000.

[6] D. Comaniciu. Nonparametric information fusion for motion estimation. In *CVPR*, volume 1, pages 59–66, 2003.

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

[8] X. Huang, N. Paragios, and D. Metaxas. Establishing Local Correspondences towards Compact Representations of Anatomical Structures. In *Medical Imaging Copmuting and Computer-Assisted Intervention*, 2003.

[9] P. Huber. *Robust Statistics*. John Wiley & Sons, 1981.

[10] K. Kanatani. Uncertainty modeling and model selection for geometric inference. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(10):1307–1319, 2004.

[11] B. Kimia, A. Tannenbaum, and S. Zucker. Shocks and Deformations i: The Components of two-dimensional Shape and the reaction-diffusion space. *International Journal of Computer Vision*, 15:189–224, 1995.

[12] Y. LeCun. The MNIST database of handwritten digits. http://yann.lecun.com/exdb/mnist/.

[13] J. Ma, Y. Zhao, and S. Ahalt. OSU SVM classifier matlab toolbox. http://www.ece.osu.edu/∼maj/osu_svm/.

[14] A. Mittal and N. Paragios. Motion-based background substraction using adaptive kernel density estimation. In *Computer Vision and Pattern Recognition*, volume 2, pages 302–309, 2004.

[15] X. Pennec and J.-P. Thirion. A framework for uncertainty and validation of 3-d registration methods based on points and frames. *Int. J. Comput. Vision*, 25(3):203–229, 1997.

[16] T. Sederberg and S. Parry. Free-form deformation of solid geometric models. *Proceedings SIGGRAPH '86*, 20:151–160, 1986.

[17] E. Simoncelli. Bayesian multi-scale differential optical flow. In B Jähne, H Haussecker, and P Geissler, editors, *Handbook of Computer Vision and Applications*, volume 2, chapter 14, pages 397–422. Academic Press, San Diego, April 1999.

[18] C. Stewart, C.-L. Tsai, and B. Roysam. The dual bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE Trans. Med. Img.*, 22:1379–1394, 2003.

[19] A. Stoddart, S. Lemke, A. Hilton, and T. Renn. Estimating Pose Uncertainty for Surface Registration. *Image and Vision Computing*, 16:111–120, 1998.

[20] M. Wand and M. Jones. *Kernel Smoothing*. Chapman & Hall, 1995.

[21] Z. Xie and G. Farin. Deformation with hierarchical b-splines. *Mathematical Methods for Curves and Surfaces: Oslo 2000*, pages 545–554, 2001.

[22] L. Younes. Computable elastic distances between shapes. *SIAM Journal of Applied Mathematics*, 58:565–586, 1998.

[23] Z. Zhang. Iterative Point Matching for Registration of Free-Form Curves and Surfaces. *International Journal of Computer Vision*, 13:119–152, 1994.