Measuring Geodesic Distances on the Space of Bounded Diffeomorphisms

Carole Twining[1], Stephen Marsland[1] and Chris Taylor
Imaging Science and Biomedical Engineering, Stopford Building,
University of Manchester, Oxford Road, Manchester M13 9PT

Abstract

This paper considers the problem of measuring the differences between deformations. It is intended to be applied in the context of applications such as the analysis of sets of non-rigidly registered medical images, where diffeomorphic warps between pairs of images establish a dense pixel-to-pixel correspondence. We introduce a new spline with explicit boundary conditions, and show how this can be used to generate general flows of bounded diffeomorphisms. Techniques are developed to describe an arbitrary continuous diffeomorphism and to invert and interpolate between such diffeomorphisms. We show how these constructions and the geodesic distance can be used in the analysis of training sets of general diffeomorphic warps.

1 Introduction

The analysis and interpretation of medical images is a difficult, yet important task. Within one imaging modality the normal appearance of an anatomical region can vary widely across a population and over time; abnormalities associated with disease can similarly have a wide variation. Variability is further increased when combining data from different imaging modalities (e.g., CT, MRI and PET). One popular approach to dealing with this variability is to map all images into some common reference frame. While some part of this mapping may be affine, in general the complete mapping will also have some non-linear component. Generating suitable non-linear mappings is known as non-rigid registration and several methods have been developed, usually by analogy with natural processes such as fluid or elastic warping [3, 8, 12].

In this paper, we concentrate on the case where a dense (i.e., pixel-to-pixel) non-linear correspondence has been established across a set of images, by whatever method. The set of warps that define the correspondence implicitly encode information about the variability of the structures present in the images. Quantifying such anatomical variability via analysis of such sets of warps is the focus of this paper. Previous work (e.g., [10, 11]) has considered the construction of parameterised models of deformation fields. However, there are two possible problems with such approaches. Firstly, the warps they consider are not exact diffeomorphisms – the possibility exists for folding or tearing the images. Such a warp is in some sense ambiguous and is not generally invertible. Secondly, the statistical models are constructed on some space of warp parameters; the imposition of a Euclidean metric on such a space is rather ad hoc. It is well-known from the field of statistical shape modelling that arbitrary or inappropriate choice of parameters can lead to highly non-optimal models [5].

[1]Joint first authors. Email: **carole.twining@man.ac.uk, stephen.marsland@man.ac.uk**

BMVC 2002 doi:10.5244/C.16.83

Here, we consider the case where the warping functions are constrained to be smooth, whilst also not tearing or folding the image. The warp functions are then bijective, invertible and differentiable to some order. Furthermore, if the objects considered are discrete and bounded, it seems that the appropriate set of warp functions to consider will belong to the group of diffeomorphisms with some non-trivial boundary conditions. We will use a metric on the space of such bounded diffeomorphisms that is defined with reference to the fact that the space of warps is a group. Camion and Younes [4] have shown how to construct such a metric. It relies on a parameterisation of diffeomorphisms that is based on the positions of a set of control points. However in their implementation, they focus solely on the problem of inexact matching of image-based landmarks, rather than on the representation and analysis of sets of **arbitrary** diffeomorphisms. We choose to focus on the latter problem, and emphasise that the warp control points we will construct do **not** have any direct correspondence with image-based landmarks or image features.

To this end, we present a new optimisation algorithm for calculating geodesics in the space of diffeomorphisms. We show that our representation of diffeomorphisms is general, in that it can be used to approximate an arbitrary diffeomorphism to any required degree of accuracy. Furthermore, we show that within this representation, we can invert and interpolate between arbitrary warps. We then explore the relationship between this metric-based approach and conventional parameter-based models of variation.

2 Representing Arbitrary Bounded Diffeomorphisms

2.1 The Deformation Field

Suppose that we wish to construct a diffeomorphism that maps a set of control points $\{\mathbf{x}_i^{(0)} : \ i = 1 \ldots n_c\}$ to the points $\{\mathbf{x}_i^{(1)} = \mathbf{x}_i^{(0)} + \mathbf{v}_i\}$, where $\{\mathbf{v}_i\}$ are the control point displacements. One method of extending this mapping to the whole space is to interpolate the displacements. This is the approach used by Bookstein [2], who considers the thin-plate spline interpolant of the displacements $\{\mathbf{v}_i\} \rightarrow \mathbf{v}(\mathbf{x})$. The solution for the displacement field $\mathbf{v}(\mathbf{x})$ is then the minimiser of the functional Lagrangian:

$$\mathcal{L}[\mathbf{v}] = \int_{\mathbb{R}^n} d\mathbf{x} |L\mathbf{v}(\mathbf{x})|^2 + \lambda \sum_{i=1}^{n_c} |\mathbf{v}(\mathbf{x}_i^{(0)}) - \mathbf{v}_i|^2. \tag{1}$$

If L is the differential operator $L = \triangle = \nabla^2$ then the first term is the approximate Willmore energy [14] (i.e., the bending energy) of the displacement fields. The second term includes a Lagrange multiplier λ to impose the **exact** constraint that the displacement field has the required value at the control points. Taking ordinary and functional derivatives of the Lagrangian, we obtain:

$$\frac{d\mathcal{L}}{d\lambda} = \sum_{i=1}^{n_c} |\mathbf{v}(\mathbf{x}_i^{(0)}) - \mathbf{v}_i|^2 = 0 \ \Rightarrow \ \mathbf{v}(\mathbf{x}_i^{(0)}) = \mathbf{v}_i, \tag{2}$$

$$\frac{1}{2}\left(\frac{\delta\mathcal{L}}{\delta\mathbf{v}(\mathbf{x})}\right) = \triangle^2\mathbf{v}(\mathbf{x}) + \lambda \sum_{i=1}^{n_c} \delta(\mathbf{x} - \mathbf{x}_i^{(0)}) \left[\mathbf{v}(\mathbf{x}) - \mathbf{v}_i\right] = 0$$

$$\Rightarrow \ \triangle^2\mathbf{v}(\mathbf{x}) \propto \sum_{i=1}^{n_c} \delta(\mathbf{x} - \mathbf{x}_i^{(0)}) \left[\mathbf{v}(\mathbf{x}) - \mathbf{v}_i\right]. \tag{3}$$

The non-affine part of the solutions of equations (2, 3) can be expanded in terms of some Green's function $G(\mathbf{x}, \mathbf{y})$ of the biharmonic equation thus:

$$\triangle^2_{(\mathbf{x})} G(\mathbf{x}, \mathbf{y}) \propto \delta(\mathbf{x} - \mathbf{y}), \ \mathbf{v}(\mathbf{x}) = \sum_{i=1}^{n_c} \boldsymbol{\alpha}_i G(\mathbf{x}, \mathbf{x}_i^{(0)}), \tag{4}$$

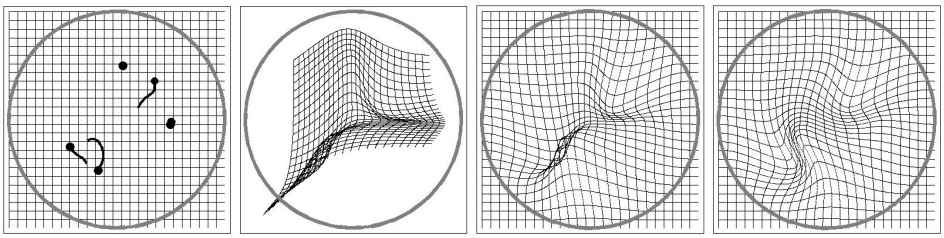

Figure 1: *From the left:* The original image, optimal control point paths and the unit circle (in gray), the thin-plate spline interpolant of the deformation field, the clamped-plate spline interpolant of the deformation field and the geodesic interpolating spline.

where the subscript indicates that derivatives are taken with respect to the first argument only. The coefficients $\{\boldsymbol{\alpha}_i\}$ are completely determined by fitting the solution to the constraints. The 2D thin-plate spline uses the Green's function:

$$\widetilde{G}(\mathbf{x},\mathbf{y}) = -|\mathbf{x}-\mathbf{y}|^2 \log|\mathbf{x}-\mathbf{y}|^2. \tag{5}$$

However, the thin-plate spline interpolant is only **asymptotically** flat, and has no specific boundary conditions. A new spline with explicit boundary conditions is the subject of the next section.

2.2 Clamped-Plate Splines

The thin-plate spline interpolant potentially affects every point in the space. In the context of an image of a discrete object or objects that may be unrelated, such global warps are unsuitable. We hence require a deformation that acts only within some specified bounded region[1], and that vanishes smoothly at the boundary of the region, that is, we specify both Dirichlet and Neumann boundary conditions. We therefore introduce a new spline interpolant, which we will call the **clamped-plate spline**. This is defined for an arbitrary number of dimensions based on the general Green's function of the biharmonic (clamped-plate) equation given by Boggio [1], which in two dimensions, takes the form:

$$G(\mathbf{x},\mathbf{y}) = |\mathbf{x}-\mathbf{y}|^2\left(\frac{1}{2}(A^2-1)-\log A\right),\ \ A(\mathbf{x},\mathbf{y}) = \frac{\sqrt{|\mathbf{x}|^2|\mathbf{y}|^2-2\mathbf{x}\cdot\mathbf{y}+1}}{|\mathbf{x}-\mathbf{y}|}. \tag{6}$$

The general Green's function has the boundary conditions that it is zero and has zero normal derivative on the boundary of the unit ball in $\mathbb{R}^n$. The clamped-plate spline interpolant for the deformation fields is constructed as before (note that there is no affine part).

$$\mathbf{v}(\mathbf{x}) = \sum_{i=1}^{n_c} \boldsymbol{\alpha}_i G(\mathbf{x},\mathbf{x}_i^{(0)}). \tag{7}$$

So, we now have a method for interpolating the displacement field with specified boundary conditions. However, only in the limit of small deformations are interpolants of the deformation field guaranteed to be diffeomorphic, whether we consider thin-plate splines, clamped-plate splines or the Gaussian interpolant considered by Camion and Younes [4]. An example illustrating this is shown in Figure 1, where the thin-plate spline and clamped-plate spline interpolants of the deformation field both show folding.

[1]This could be approximated to any required accuracy by using additional control points for the thin-plate spline, positioned on the required boundary. Solving for the expansion coefficients involves the inversion of a matrix whose size is given by the number of control points, hence this approximation would greatly increase the computational cost.

2.3 Geodesic Interpolating Splines

Since the deformation field interpolant is only diffeomorphic for small deformations, the usual approach to constructing large-deformation diffeomorphisms is to consider them as a sequence of small deformations [4, 9, 7, 13]. We introduce the flow-time t, where:

$$\mathbf{x}_i^{(0)} \rightarrow \mathbf{x}_i(t=0), \;\; \mathbf{x}_i^{(1)} \rightarrow \mathbf{x}_i(t=1). \tag{8}$$

The control points now trace out paths $\{\mathbf{x}_i(t)\}$, the velocities of which are constrained by a time-dependant deformation (velocity) field:

$$\frac{d\mathbf{x}_i(t)}{dt} = \mathbf{v}(t, \mathbf{x}_i(t)). \tag{9}$$

The functional Lagrangian (1) can then be generalised thus [4]:

$$\mathcal{L} = \int_0^1 dt \int_{\mathbb{R}^n} d\mathbf{x} \, |L\mathbf{v}(t, \mathbf{x}(t))|^2 + \lambda \sum_{i=1}^{n_c} \int_0^1 dt \, \left| \mathbf{v}(t, \mathbf{x}_i(t)) - \frac{d\mathbf{x}_i(t)}{dt} \right|^2. \tag{10}$$

Camion and Younes [4] use this form of the Lagrangian for the case of **inexact** landmark matching only, where the functional Lagrangian is optimised over the **whole** of the paths $\{\mathbf{x}_i(t)\}$, with the end-points of the paths being held fixed. The first term, which is the time-integrated bending energy of the velocity field, defines a metric distance for the flow in the space of diffeomorphisms. We will consider the **exact** matching case. We expand the velocity fields in terms of the Green's function as before:

$$\mathbf{v}(t, \mathbf{x}(t)) = \sum\nolimits_{i=1}^{n_c} \boldsymbol{\alpha}_i(t) G(\mathbf{x}(t), \mathbf{x}_i(t)), \tag{11}$$

where the functions $\{\boldsymbol{\alpha}_i(t)\}$ are determined by the **exact** matching constraint:

$$\mathbf{v}(t, \mathbf{x}_j(t)) = \frac{d\mathbf{x}_j(t)}{dt} = \sum\nolimits_{i=1}^{n_c} \boldsymbol{\alpha}_i(t) G(\mathbf{x}_i(t), \mathbf{x}_j(t)). \tag{12}$$

Consider an arbitrary smooth set of paths $\{\mathbf{x}_i(t)\}$ with velocities $\{\mathbf{v}_i(t) = \frac{d\mathbf{x}_i(t)}{dt}\}$:

$$\triangle_{\mathbf{x}}^2 \mathbf{v}(t, \mathbf{x}(t)) = \sum\nolimits_{i=1}^{n_c} \boldsymbol{\alpha}_i(t) \delta(\mathbf{x}(t) - \mathbf{x}_i(t)) \tag{13}$$

$$\begin{aligned} \therefore \;\; l^2\left[\{\mathbf{x}_i(t)\}\right] &= \int_0^1 dt \int_{\mathbb{R}^n} d\mathbf{x} \, |L\mathbf{v}(t, \mathbf{x}(t))|^2 = \int_0^1 dt \int_{\mathbb{R}^n} d\mathbf{x} \, \mathbf{v}(t, \mathbf{x}) \cdot \left(\triangle_{\mathbf{x}}^2 \mathbf{v}(t, \mathbf{x})\right) \\ &= \sum\nolimits_{i=1}^{n_c} \int_0^1 dt \, \boldsymbol{\alpha}_i(t) \cdot \mathbf{v}(t, \mathbf{x}_i(t)), \end{aligned} \tag{14}$$

where $l\left[\{\mathbf{x}_i(t)\}\right]$ is the metric distance in the space of diffeomorphisms along the flow path determined by the set of control point paths $\{\mathbf{x}_i(t) : i = 1, \ldots, n_c\}$. Hence, optimising the above expression for the metric distance with respect to varying the control point paths corresponds to constructing a geodesic in the space of diffeomorphisms.

So, given such an **optimal** set of control point paths $\{\tilde{\mathbf{x}}_i(t)\}$, the optimal velocity field $\tilde{\mathbf{v}}(t, \mathbf{x}(t))$ is totally determined. We can hence determine the path of an arbitrary test point $\mathbf{x}(0)$ by a simple integration. This then defines the complete action of a diffeomorphism of the unit disc, which is parameterised by the initial and final positions of the control points. We will denote the functional form of such a geodesic interpolating spline (GIS) diffeomorphism thus:

$$\omega \equiv \omega(\{\mathbf{x}_i(0)\},\{\mathbf{x}_i(1)\}) \ \in \ \text{Diff}\left(\mathbb{D}^2\right), \tag{15}$$

where Diff $\left(\mathbb{D}^2\right)$ is the group of diffeomorphisms of the unit disc $\mathbb{D}^2$ with the boundary conditions as described above. Given a sufficiently large set of control points this representation can, in principle, be used to describe **any** element of the group of diffeomorphisms. The geodesic distance (14) can then be written as the metric function:

$$\{d : \text{Diff}\left(\mathbb{D}^2\right) \times \text{Diff}\left(\mathbb{D}^2\right) \rightarrow \mathbb{R}^+\}, \quad d\left(e, \omega\left(\{\mathbf{x}_i(0)\},\{\mathbf{x}_i(1)\}\right)\right) = l\left[\{\mathbf{x}_i(t)\}\right], \tag{16}$$

and e is the group identity element. This metric has several important properties; in particular, it is invariant under the action of a general group element thus:

$$d(g,h) \equiv d(e \circ g, e \circ h) \equiv d(f^{-1} \circ g, f^{-1} \circ h) \ \forall \ f,g,h \ \in \ \text{Diff}\left(\mathbb{D}^2\right), \tag{17}$$

where $\circ$ is the group multiplication. That is, the geodesic distance between the two warps g and h is independent of the choice of the reference warp (or image), where here we have considered the identity e or a general group element f.

3 Implementation

Camion and Younes [4] describe an algorithm for optimising the Lagrangian in equation (10) within the context of inexact matching of image-based landmarks. However, we note here that optimising with a fixed value of λ (as in [4]) is **not** the same as the analytical optimisation of the same expression, where λ is a Lagrange multiplier. Their algorithm can lead to a severe mis-match between the velocity fields and the control point paths; the paths for the control points constructed from the velocities give only approximate (and poorly-controlled) matching of the end-points, whilst the flows constructed from the exact-matching control point paths are far from optimal.

We consider the more general problem of representing an **arbitrary** diffeomorphism; we hence require exact matching of the optimal paths and flows. We have developed an algorithm to minimise the metric distance in equation (14) directly (see the Appendix for details). The Appendix also includes exact analytic results that enabled us to check our algorithm and to verify that the algorithm of Camion and Younes does not give the optimal paths or flows. We were also able to considerably improve the convergence time (by up to a factor of 30 in some cases) by using these analytic results to initialise the algorithm, rather than initialising with constant-velocity, straight line paths, as in [4].

4 Applying Diffeomorphic Warps to Images

We now consider the action of the group Diff $\left(\mathbb{D}^2\right)$ on pixelated images. We define an image to be a scalar (grayscale) or vector (colour) function I defined on a set of pixel positions. For an initial, d-dimensional unwarped image, the set of pixel positions will lie on a regular (hyper)cubic lattice of points $S_e \ \in \ \mathbb{R}^d$. A group element g acts on this to produce a warped set of pixel positions $S = g(S_e)$. Two elements $g, h \ \in \ \text{Diff}\left(\mathbb{D}^2\right)$ are equivalent with respect to their action on a pixel set S if $g(S) = h(S)$. We take the warped image I' to be that obtained by the push-forward map, so that:

$$g : I \rightarrow I', \ \text{ where } \ I'(g(S)) \equiv I(S). \tag{18}$$

Note that in the examples we will present, the warped images $I'(g(S))$ are **not** resampled, but plotted as surfaces whose facets are the **deformed** pixels. The boundary of the warped region in the images is shown by the white circle.

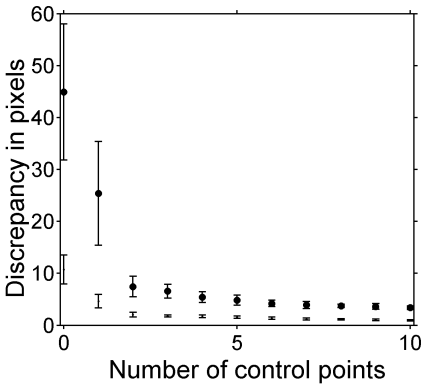

Figure 2: The maximum (•) and mean (·) pixel discrepancies between a warp and its approximant as a function of n_c.

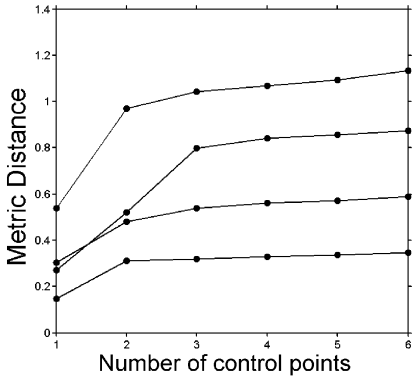

Figure 3: The geodesic distance $d(e,\omega)$ as a function of n_c for 4 random warps.

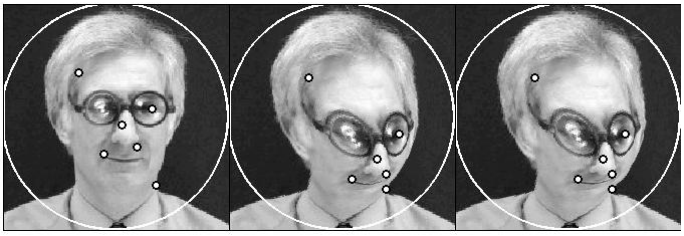

Figure 4: *Left:* Original image, *Centre:* Cauchy warp, *Right:* Approximant with control points shown. Image size 166×166. Maximum discrepancy 3.2 pixels, mean 0.77 pixels.

4.1 Approximating an Arbitrary Bounded Diffeomorphism

Suppose that we have an arbitrary diffeomorphism g that acts on some dense pixel set S_e. We can iteratively construct a GIS approximant $\omega(S_e) \approx g(S_e)$ to this warp. Taking as the initial approximant $\omega = e,\ n_c = 0$:

- Find the pixel position $\mathbf{y} \in S_e$ where the difference $|\omega(\mathbf{y}) - g(\mathbf{y})|$ is maximal
- Add this point to the existing set of control points of ω, with endpoints $\mathbf{y}$ and $g(\mathbf{y})$
- Update the warp ω
- Iterate until convergence

The algorithm was tested by applying it to a set of random exact diffeomorphisms. The test diffeomorphisms were generated using a generalisation of the 1D techniques described in [6]. The cumulative distribution function of a wrapped Cauchy function was used to create an **exact** diffeomorphism of the unit disc with the required boundary conditions (see Figure 4 for an example). The parameters of the Cauchy warps were chosen at random, and a concatenation of several such warps used to generate each test warp. The test pixel set S_e was of size 166×166 over the square of side length 2. The discrepancies between the two warped pixel sets were then compared over the area of the inscribed unit circle (shown in Figure 2). As can be seen from the figure, the approximant quickly converges, giving a mean discrepancy of size 0.9 ± 0.1 pixels ($(5.4 \pm 0.6) \times 10^{-3}$ in units of the circle radius) when 10 control points are included. As is shown in Figure 3, the calculated geodesic distance $d(e,\omega)$ also converges rapidly.

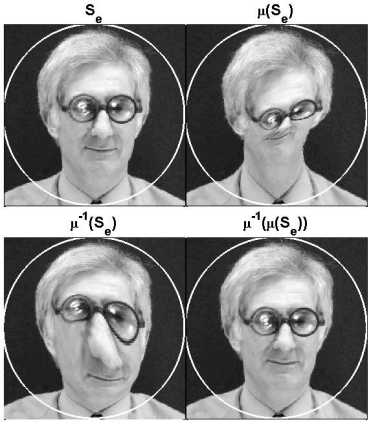

Figure 5: An example warp and its inverse, with a maximum discrepancy of 0.6 pixels.

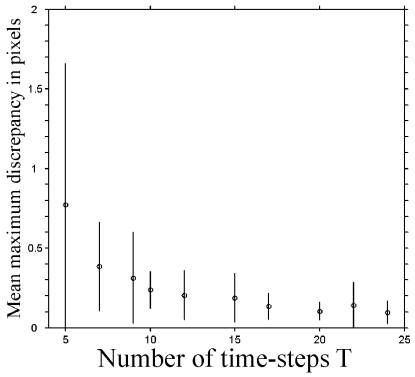

Figure 6: The mean maximum discrepancy in pixels as a function of the number of time-steps. Test image size 166×166 pixels.

4.2 Inverting an Arbitrary Bounded Diffeomorphism

Consider an arbitrary GIS warp $\mu = \omega(\{\mathbf{x}_i(0)\}, \{\mathbf{x}_i(1)\})$. The exact inverse of this warp is given by:

$$\mu^{-1} = \omega(\{\mathbf{x}_i(1)\}, \{\mathbf{x}_i(0)\}), \tag{19}$$

that is, we just reverse the initial and final positions of the control points. The numerical accuracy of our algorithm with respect to this result was tested by taking a random set of GIS warps $\{\mu_a : a = 1 \ldots 20\}$ with 4 control points. The initial positions of the control points were held fixed, and the final positions chosen at random. For each warp, the inverse was calculated as above, and the accuracy of the inversion calculated by comparing $\mu_a^{-1}(\mu_a(S_e))$ and S_e (see Figure 5 for an example). The algorithm was run for a range of values of T, the number of time-steps, where a new random set of 20 warps was generated for each value of T. As can be seen from Figure 6, the mean maximum discrepancy is well below half a pixel for values of T of 10 and above.

4.3 Interpolating Between Arbitrary Bounded Diffeomorphisms

Suppose that we have some training set of warps $\{g_a : a = 1 \ldots n\}$. As we have shown in section 4.1, such a set can be represented by GIS warps to any required degree of accuracy, provided the number of control points n_c is large enough. So, we now consider the equivalent set of GIS warps $\{\mu_a\}$, where the control points have been chosen so that the initial positions are the same across the set. That is:

$$g_a \approx \mu_a = \omega(\{\mathbf{p}_i\}, \{\mathbf{y}_i^a\}), \;\; a = 1 \ldots n, \;\; i = 1 \ldots n_c, \tag{20}$$

where $\{\mathbf{p}_i\}$ are fixed reference points. The information about the distribution of the set of warps in the space of diffeomorphisms is encoded by the set of geodesic distances between all pairs of warps:

$$d(\mu_a, \mu_b) \equiv d(e, \mu_a \circ \mu_b^{-1}) \equiv d(e, \mu_b \circ \mu_a^{-1}). \tag{21}$$

To calculate these distances, we need to be able to construct warps of the form $\mu_b \circ \mu_a^{-1}$, which is the warp which interpolates between μ_a and μ_b. We will consider an approximant to this warp thus:

$$\nu_{ab} = \omega(\{\mathbf{y}_i^a\}, \{\mathbf{y}_i^b\}). \tag{22}$$

It is an approximant in the sense that it is equivalent to the exact warp $\mu_b \circ \mu_a^{-1}$ when we consider its action on the control point positions. Since the control points tend to be the set of points for which the displacement generated by the warp is largest we would expect

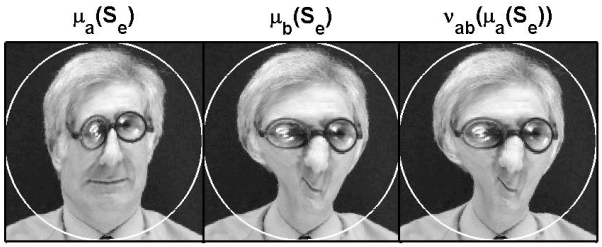

Figure 7: Warps $\mu_a(S_e)$ and $\mu_b(S_e)$, and their interpolant ν_{ab} applied to $\mu_a(S_e)$ (see text for details). The maximum discrepancy between the 2nd and 3rd images is 1.5 pixels.

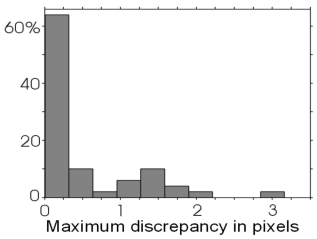

Figure 8: The distribution of maximum discrepancies in pixels across a test set.

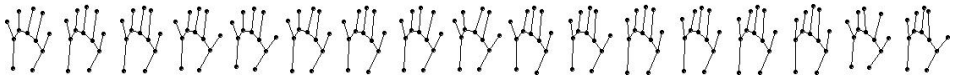

Figure 9: The training set of hand shapes.

that ν_{ab} should be a reasonable approximation to the exact warp. This was tested by generating a set of warps $\{\mu_a\}$ as in section 4.1. For each pair of warps, the approximant ν_{ab} was calculated. The accuracy of the approximation was calculated by comparing $\nu_{ab}(\mu_a(S_e))$ and $\mu_b(S_e)$ since:

$$\left(\mu_b \circ \mu_a^{-1}\right)\left(\mu_a(S_e)\right) \equiv \mu_b(S_e). \tag{23}$$

An example is shown in Figure 7. The results for the maximum discrepancy per warp pair across a test set with 50 independent pairs of warps is shown in Figure 8. Note that in only one case is the maximum discrepancy over 3 pixels, and in 80% of cases the maximum discrepancy is less than one pixel.

5 Building Models

In this section, we consider the case where the warps are generated from real data. We take as our training set a Statistical Shape Model (SSM) [5] built from 17 shape examples, each with 12 landmark points. The shapes were extracted from images of a real hand in motion (Figure 9). The landmark points defining the shapes become the control points for our warps; hence the GIS warps so constructed map the landmark points of the shapes

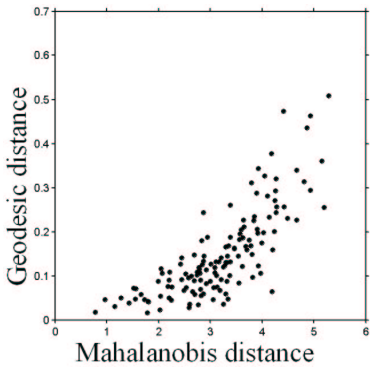

Figure 10: Geodesic vs. Mahalanobis distance for all pairs.

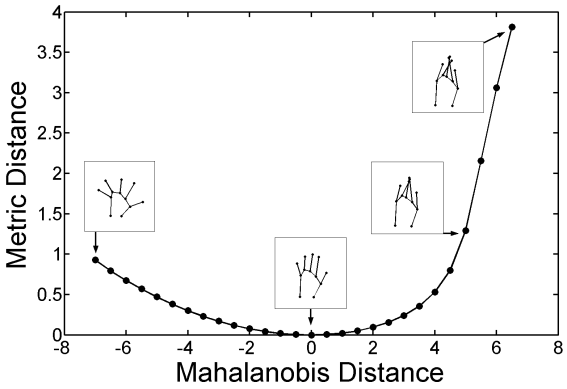

Figure 11: Geodesic vs. Mahalanobis distance for variation of the first mode of the SSM.

between the training examples (note that the lines in Figure 9 are for the purposes of illustration only). We then compare the Mahalanobis distances between pairs of shapes with the geodesic distance of the GIS warp that interpolates between them (see Figure 10). The relationship is approximately monotonic. In Figure 11, we study the effect of varying the first linear mode of the SSM. This corresponds to opening and closing of the fingers. Note that the geodesic distance penalises those variations that cause a change in connectivity (i.e., crossing of the fingers), whilst the linear SSM (which exactly corresponds to a linear model built on the space of warp parameters), does not.

6 Conclusions

We have presented a novel interpolating spline, the clamped-plate spline, which (unlike the thin-plate spline), generates warps that are strictly limited in extent. By composing sequences of such warps, we can generate exact diffeomorphisms, which are defined in terms of a small number of control points. We show that this representation of diffeomorphisms is **general**, and can be used to approximate an arbitrary diffeomorphism to any required degree of accuracy. Furthermore, we show how this representation includes a definition of a metric distance on the space of diffeomorphisms, allowing us to compute a geodesic distance between warps that respects the properties of the diffeomorphism group. This geodesic distance will enable us to develop meaningful statistics of deformations. Finally, we have compared using metric distance and parameter-based approaches to modelling the space of warps. We have shown that the metric correctly penalises deformations that alter the topology, whereas a linear model on the parameter space does not.

The techniques are likely to prove widely applicable, particularly in medical image analysis where object shape and shape change is frequently studied. The mathematical basis of these techniques exists for an arbitrary number of dimensions, hence can in principle be extended higher numbers of dimensions (e.g., 3D voxellated images).

Acknowledgements

The authors wish to thank Tim Cootes for many interesting discussions. This research was supported by the MIAS IRC project, EPSRC grant number GR/N14248/01.

Appendix - Implementation

The continuum variables $\mathbf{x}_i(t), \mathbf{v}_i(t) = \frac{d\mathbf{x}_i}{dt}$ are approximated by piecewise linear variables thus:

$$\begin{aligned} \mathbf{x}_i(t) &\rightarrow \mathbf{q}_i(\tau),\ \tau = 1 \ldots T+1,\ \ i = 1 \ldots n_c,\ \text{ where } t = \frac{\tau - 1}{T} \\ \mathbf{v}_i(t) &\rightarrow \mathbf{D}\mathbf{q}_i(\tau) = T\left(\mathbf{q}_i(\tau+1) - \mathbf{q}_i(\tau)\right),\ \tau = 1 \ldots T. \end{aligned} \tag{24}$$

The coefficients $\boldsymbol{\alpha}_i(\tau)$ are then calculated by inverting the equation:

$$\mathbf{D}\mathbf{q}_i(\tau) = \sum_{j=1}^{n_c} \boldsymbol{\alpha}_j(\tau) G(\mathbf{q}_j(\tau), \mathbf{q}_i(\tau)), \tag{25}$$

which enables us to calculate the metric distance:

$$l^2\left[\{\mathbf{q}_i\}\right] = \sum_{i=1}^{n_c} \sum_{\tau=1}^{T} \boldsymbol{\alpha}_i(\tau) \cdot \mathbf{D}\mathbf{q}_i(\tau). \tag{26}$$

The optimal control point paths are then found by optimising the metric distance with respect to the variable parts of the control point paths $\{\mathbf{q}_i(\tau), \tau = 2 \ldots T\}$, where we use standard MATLAB optimisation routines.

The only remaining issue is the choice of initial values for the paths $\{\mathbf{q}_i(\tau)\}$. For the case of one control point, we have obtained an exact analytic solution of equation (14) for the optimisation

of the metric distance. Using polar coordinates, and taking the polar angle θ as the dependent variable rather than the time t, the general solution has the form:

$$r(\theta) = k\cos(\theta-\psi) - \sqrt{k^2\cos^2(\theta-\psi)-1} \quad (27)$$

$$t(\theta) = A\tanh^{-1}\left[\left(\sqrt{k^2-1}\cot(\theta-\psi)\right)^{\pm 1}\right] + B, \quad (28)$$

where the parameters k, ψ, A and B are determined by fitting to the end-points. The general form of $r(\theta)$ describes the arc of a circle of varying centre and radius that intersects the unit circle at right angles[2]. We have found that initialising the algorithm with such paths for the case of multiple control points gives a considerable improvement in the convergence time over the case of initialising with constant-velocity straight-line paths. It also gives us an independent check of the validity of our algorithm when applied to the case of one control point.

References

[1] T. Boggio. Sulle funzioni di green d'ordine m. *Rendiconti - Circolo Matematico di Palermo*, 20:97–135, 1905.

[2] F. L. Bookstein. Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[3] M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. In *Proceedings of Visualization in Biomedical Computing (VBC)*, pages 267–276, 1996.

[4] V. Camion and L. Younes. Geodesic interpolating splines. In M. Figueiredo, J. Zerubia, and A. K. Jain, editors, *Proceedings of EMMCVPR'01*, volume 2134 of *Lecture Notes in Computer Science*, pages 513–527. Springer, 2001.

[5] Rh. Davies, T. F. Cootes, and C. J. Taylor. A minimum description length approach to statistical shape modelling. In *Proceedings of IPMI'01*, volume 2082 of *Lecture Notes in Computer Science*, pages 50–63. Springer, 2001.

[6] Rh. Davies, T. F. Cootes, C. J. Twining, and C. J. Taylor. An information theoretic approach to statistical shape modelling. In *Proceedings of the British Machine Vision Conference (BMVC'01)*, volume 1, pages 3–12. BMVA Press, 2001.

[7] P. Dupuis, U. Grenander, and M. I. Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*, 56(3):587–600, 1998.

[8] J. Gee, M. Reivich, and R. Bajcsy. Elastically deforming 3D atlas to match anatomical brain images. *Journal of Computer Assisted Tomography*, 17(2):225–236, 1993.

[9] S. C. Joshi and M. M. Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Transactions on Image Processing*, 9(8):1357–1370, 2000.

[10] L. LeBriquer and J. Gee. Design of a statistical model of brain shape. In *Proceedings of IPMI'97*, volume 1230 of *Lecture Notes in Computer Science*, pages 477–482. Springer, 1997.

[11] D. Rueckert, A. F. Frangi, and J. A. Schnabel. Automatic construction of 3D statistical deformation models using non-rigid registration. In *Proceedings of MICCAI'01*, volume 2208 of *Lecture notes in Computer Science*, pages 77–84. Springer, 2001.

[12] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. Non-rigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.

[13] A. Trouvé. Diffeomorphisms groups and pattern matching in image analysis. *International Journal of Computer Vision*, 28(3):213–221, 1998.

[14] T. J. Willmore. A survey on Willmore immersions. In *Geometry and Topology of Submanifolds, IV*, pages 11–16. World Scientific Publishers, 1992.

[2] It is interesting to note that this corresponds to the Poincaré model of the hyperbolic disc.