Computing Accurate Correspondences across Groups of Images

Timothy F. Cootes, Carole J. Twining, Vladimir S. Petrović, Kolawole O. Babalola, and Christopher J. Taylor

Abstract—Groupwise image registration algorithms seek to establish dense correspondences between sets of images. Typically they involve iteratively improving the registration between each image and an evolving mean. A variety of methods have been proposed, which differ in their choice of objective function, representation of deformation field and optimisation methods. Given the complexity of the task, the final accuracy is significantly affected by the choices made for each component. Here we present a groupwise registration algorithm which can take advantage of the statistics of both the image intensities and the range of shapes across the group to achieve accurate matching. By testing on large sets of images (in both 2D and 3D), we explore the effects of using different image representations and different statistical shape constraints. We demonstrate that careful choice of such representations can lead to significant improvements in overall performance.

Index Terms—Non-rigid registration, correspondence problem, appearance models

✦

---

1 INTRODUCTION

AUTOMATIC methods for computing dense correspondences across sets of images are now widely used to both construct statistical models and to analyse variations across a group. For instance, it is common to build a mean brain image from a group to define a standardised co-ordinate frame in which to compare different examples from the group [1]. Alternatively statistical models of shape and appearance can be built which can then be used for further image interpretation [2], [3].

A variety of different techniques for establishing correspondence have been proposed (see the review below). The majority involve iteratively updating estimates of correspondences between each image and an image or model in a reference frame, usually with the reference frame evolving towards a group average at the same time. Each technique requires three core components:

1) a representation of the deformation field
2) an objective function
3) a method of optimisation

Common choices for the deformation field are B-splines, [4], [5], piece-wise affine [6], [7] and dense (per-voxel) flow fields [8], [9].

Objective functions typically have two components, one based on image intensities (comparing the warped target image with the reference) and one based on the shape (often just a regularisation term). The intensity term is usually a simple sum of squares error term, though more robust versions have been used [6], as well as variants of mutual information [10], and voxel-wise entropy measures [5].

The shape term is usually a local measure to penalise excessive local deformation, such as a bending energy [11]. The statistical information about the shape, estimated from the group, is less often exploited [3], [6].

In this paper we describe a particular version of groupwise registration (developed from [6]), and use it to investigate the effects of choices of image features and shape regularisation on the overall accuracy of the resulting correspondences.

In particular, we show that local normalisation of the intensities can improve accuracy, and that using features other than intensity (such as smoothed gradient information) at each pixel can help achieve more robust correspondence in a coarse-to-fine framework.

We also explore using simple local shape constraints to regularise the resulting mesh, demonstrating that careful choice can improve the performance.

We apply the registration framework to two large sets of images. The first is a set of 2D images of faces from the XM2VTS database [12]. We have detailed manual annotation of feature points for this data, allowing direct comparison of the performance of the automatic system with manual correspondences. The second is a set of 270 3D MR images of the brain, for which we have voxel labels for various structures. While this does not allow explicit point by point evaluation of the correspondence accuracy, it does allow us to evaluate the system using overlap measures for each labelled structure, which is a useful indirect assessment.

The main novelty in this paper is the detailed quantitative evaluation of a group-wise registration algorithm on large 2D and 3D datasets and the demonstration of the importance of different components of the algorithm. The lessons learned about the choices of component which can significantly improve performance will be applicable to other group-wise registration methods.

In the following we review the literature on groupwise correspondence, describe our particular registration algorithm and explore some of the choices for image representation and shape regularisation. We present detailed results on the face and brain dataset, discuss the applicability of the approach and draw some conclusions.

2 BACKGROUND

Finding mappings between structures across a set of images can facilitate many image analysis tasks. One particular area of importance is in medical image interpretation, where image registration can help in tasks as diverse as anatomical atlas matching and labelling, image classification, and data fusion.

Over the past decade there has been a considerable growth in interest in the problem of groupwise registration (also called 'joint alignment'), and a range of techniques have been developed.

There are two broad approaches - those that aim to determine a dense correspondence field across a set, and those seeking to find a sparse set of points or structures present in every image.

2.1 Dense correspondence methods

The aim of these techniques is to find a set of continuous transformations of space that map each training image into a common reference frame. Almost all approaches involve iteratively updating estimates of correspondences between each image and a (possibly evolving) reference model. They generally differ in their choice of objective function, representation of deformation or optimisation method.

The earliest work on dense correspondence was that by Vetter, Jones and Poggio [13], [14] which uses a combination of model fitting and optical flow to estimate dense vector fields across sets of objects, to build 'morphable models' - statistical models of shape and appearance.

A now common approach is to use a non-rigid registration algorithm (capable of finding the deformation of one image into another) [15]–[17] to independently map each image to a chosen reference.

For instance, Guimond et al. [1] used Thirion's 'Demons' algorithm [18] to register sets of images, describing how to iteratively update the group mean. Frangi et al. used non-rigid B-spline registration [11] to correspond 3D images and build statistical shape models [10]. Duchesne and Collins [7] used a multi-resolution block matching strategy [19] to compute piecewise linear deformation fields between images, which were used to build shape and appearance models. Cootes et al. used a

composition of simple diffeomorphic deformation fields to register a group of images into a common frame [20].

The field of Computational Anatomy involves estimating diffeomorphisms to warp each example of an object into a common (average) reference frame. Typically the deformations are represented as dense displacement fields, which are the integration of dense velocity fields, estimated using a fluid-flow style algorithm. Joshi et al. [9], [21] demonstrate how to simultaneously estimate the reference shape and image in the case of large scale diffeomorphisms, where linear approximations to averages break down.

Rather than perform registration to a reference independently, information from the group can be used to help the process. Marsland et al. [22], [23] describe a groupwise algorithm which registered images using clamped-plate splines so as to minimise a description length measure of the whole set.

Baker et al. [3] also considered building an appearance model as an image coding problem. The aim is to construct a model which can encode the set of images as efficiently as possible. The model parameters are iteratively re-estimated after fitting the current model to the images, leading to an implicit correspondence defined across the data set. The model fitting takes account of the statistics of intensities and shapes across the group.

Cootes et al. [24] describe a related approach which also aims to minimise a description length measure, with the added constraint that the deformations should be diffeomorphic (smooth and invertible). They represent the deformation fields using a composition of grid-based diffeomorphic deformations.

Learned-Miller [25] performs groupwise alignment using a 'congealing' approach, in which each image is repeatedly aligned with a reference so as to minimise a per-pixel entropy measure. The entropy measure pools information from the whole group. Balci et al. [5] extended this approach to demonstrate groupwise registration using B-spline deformations.

The method we propose below (developed from [6]) takes ideas from several of the above, in that it represents deformations with a piece-wise affine mesh, uses an objective function derived from the MDL coding approach [23] and involves optimising using methods developed in non-rigid registration.

2.2 Sparse correspondence

There is extensive work in the object recognition literature about detecting common objects or parts of objects in 2D images, giving sparse sets of corresponding points across a set. Typically this involves clustering features detected at interest points across all the images to detect possible model parts, then looking for common geometric relationships between the resulting parts. For instance, see the work of Fergus et al. [26], or for a general overview of the field [27].

Langs *et al.* [28] build sparse models by extracting many points in each training image, and using point matching algorithms to locate correspondences.

Kokkinos and Yuille [29] describe constructing both global shape models and linked parts based models using an EM style optimisation. Rather than work on raw images, they register edge and ridge primal sketches, and demonstrate good results on a range of object classes.

There is also a growing literature on estimating correspondences across groups of multiple point sets - see for instance [30].

3 METHODS

We seek to estimate a set of deformation fields which will allow us to establish dense correspondences across a set of images. The approach is an extension of our earlier work [6] in which we estimate a model in a reference frame which can explain the data as effectively as possible. That work treated the problem in the Minimum Description Length (MDL) [31] framework, which has been successfully applied to the groupwise correspondence problem for both sets of shapes [32], [33] and sets of images [23].

In the two-part coding formulation of MDL, the objective function is the length of the message required to transmit the entire training set of data, where the data is sent by encoding it according to some statistical model. The objective function can then be written as the sum of two parts thus:

$$\mathcal{L} = \mathcal{L}_{\text{model}} + \mathcal{L}_{\text{data}}, \tag{1}$$

which represents the cost of transmitting first the exact details of the model, and then the data encoded using that model. In practice, the data term tends to dominate, since it scales as the number of example images or shapes in the set, whereas the model term only scales with the number of modes included. Hence for practical purposes, it is often convenient to consider just the data term. The message length for transmitting the data encoded according to the model can be computed by using the fundamental result obtained of Shannon [34], which states that the optimum codeword length for transmitting the occurrence of some event or data value x that occurs with probability $p(x)$ is given by:

$$\mathcal{L}(x) = -\log p(x).$$

By making particular choices for the form of the model pdf $p(x)$, it is possible to generate many of the common pairwise image similarity measures within this information theoretic framework [35] (e.g., sum-of-squares, absolute difference, and both mutual information and normalised mutual information), as well as to construct inherently *groupwise* image similarity measures. In the following, we will continue with this general approach, and investigate the effects of various choices both for the form of this pdf, and the effect of including both texture and shape deformation terms in the objective function.

3.1 Basic Framework

Suppose we have a set of training images I_i, $i = 1..N$, which we wish to correspond. We will achieve this by defining a reference frame, R, and estimating a set of deformation fields $W_i(\mathbf{y})$ mapping from R to each training image[1]. The mapping from any image i to any other image j can then be achieved via the reference: $W_{i \to j}(\mathbf{y}) = W_j(W_i^{-1}(\mathbf{y}))$.

Clearly the inverse should exist, so the deformations must all be invertible.

We will assume that each deformation field is uniquely defined by the position of a set of n control points $\{\mathbf{x}_k\}$, so the deformation is a function $\mathbf{y}' = W(\mathbf{y}; \mathbf{x})$, where $\mathbf{x} = (\mathbf{x}_1^T...\mathbf{x}_n^T)^T$.

Thus the aim is to find the set of control points, $\mathbf{x}_i$, for each image, defining the correspondences on image i.

We assume that each target image can be represented as a deformed version of a model image instance in the reference frame,

$$I_s(\mathbf{y}) = I_r(W^{-1}(\mathbf{y}; \mathbf{x}); \mathbf{t}) \tag{2}$$

The image in the reference frame is given by some image function $I_r(\mathbf{y}; \mathbf{t})$ with parameters $\mathbf{t}$.

Let $\mathbf{r}$ be the vector of residual differences (one per image voxel), between the training image and the warped version of the reference image, $r_\mathbf{y} = I(\mathbf{y}) - I_s(\mathbf{y})$.

Our complete model defines

- The form of texture model $I_r(\mathbf{y}; \mathbf{t})$
- A distribution for the displacements, $p(\mathbf{x}|M)$
- A distribution for the texture parameters $p(\mathbf{t}|M)$
- A distribution for the residuals $p(\mathbf{r}|\mathbf{x}, \mathbf{t}, M)$

where M is taken to be the set of parameters defining the texture model and the various distributions (such as their means and variances etc).

Each of these components may be fixed in advance or estimated from the data itself during the optimisation process.

The overall aim of the process then becomes estimating the parameters so as to optimise the correspondences and model given the images. For a given image, this is

$$p(\mathbf{x}, \mathbf{t}|I, M) \propto p(I|\mathbf{x}, \mathbf{t}, M)p(\mathbf{x}|M)p(\mathbf{t}|M) \tag{3}$$

if we assume independent priors for shape and texture parameters. Although for many tasks shape and texture parameters are not independent (a fact exploited by the Active Appearance Model (AAM) [2], for instance), good results can be achieved with the independence assumption - this is the basis for many of the developments of the AAM by Matthews *et al.* [36], and is a fundamental assumption in almost all non-rigid registration schemes.

To find the optima, we use an iterative approach, estimating each set of parameters or model component in turn.

---

1. Here we take $W(\mathbf{y})$ to be mean spatial transformation applied to a set of points encoded in vector $\mathbf{y}$, returning a vector of transformed points.

For a given image and choice of texture parameters $\mathbf{t}$, we seek to find the best control points, $\mathbf{x}$, so wish to optimise

$$p(\mathbf{x}|I, \mathbf{t}) \propto p(I|\mathbf{x}, \mathbf{t})p(\mathbf{x}) \qquad (4)$$

where we have dropped the M term for notational brevity. For a given choice of $\mathbf{x}$ and $\mathbf{t}$, the probability of the image is given by the probability of the residual difference between target image and model image (from Eq.2), $p(I|\mathbf{x}, \mathbf{t}) = p(\mathbf{r})$, so

$$\log p(\mathbf{x}|I, \mathbf{t}) = \log p(\mathbf{r}) + \log p(\mathbf{x}) + const \qquad (5)$$

Similarly for a given image and set of control points, the texture parameters can be estimated by maximising

$$\log p(\mathbf{t}|I, \mathbf{x}) = \log p(\mathbf{r}) + \log p(\mathbf{t}) + const \qquad (6)$$

The texture model and distributions can be estimated using the current values of the correspondences.

Note that if we were only interested in the correspondences themselves, we could in theory explicitly integrate out the other parameters. However, this is likely to be extremely time consuming, given the very large number of parameters involved.

3.2 The Groupwise Registration algorithm

The aim is to estimate the positions of the control points on every image, together with the parameters, $\mathbf{t}$, defining the model image in the reference frame.

It is useful to initialise the algorithm by choosing one image as an initial reference frame, then performing pairwise registration onto every other image. As long as the chosen image is not an extreme outlier, the algorithm is fairly insensitive to the initial choice.

The basic groupwise registration algorithm proceeds as follows:

1) Select one image to be used as an initial reference
2) Select control points on this reference image
3) Initial pairwise registration to reference:
   a) For each image in turn estimate movement of control points to optimise match to the reference image
4) Groupwise registration:
   a) Estimate a new reference frame from the mean of the current points
   b) Estimate $p(\mathbf{x}|M)$ from the position of points in all images
   c) Compute the texture model by warping all images to the reference frame
   d) Estimate the distribution of texture parameters, $p(\mathbf{t}|M)$.
   e) Estimate the distribution of residuals, $p(\mathbf{r}|M)$.
   f) For each image in turn
      i) Estimate $\mathbf{t}$ to maximise (6).
      ii) Estimate $\mathbf{x}$ to maximise (5).
   g) Repeat until happy

Typically the initial control points on the reference frame (Step 2) will be a regular mesh (such as that shown in Figure 1). Though there may be some efficiency gains in concentrating the control points around strong edges, we have found little improvement in overall accuracy compared to a sufficiently dense regular mesh.

To estimate the points in the reference frame (Step 4a), we apply Generalised Procrustes Analysis [37] to align each set of control points to a common frame, then take the mean. In cases where there is only relatively smooth deformations, the mean set of points will define a valid warp (ie the associated mesh will not self intersect). If this is not the case, the reference points can be set to the control points of the example which is closest to the mean (an approximation to the median).

For small datasets, to avoid problems with local minima, when optimising on one image we work with a model built from all images except the current one (a 'leave-one-out' approach).

In its current form the algorithm isn't guaranteed to converge in all cases - it is possible that if initialisation is poor some of the points will diverge off the image, or that the mean reference image would begin cycling though a series of different states. However, in practice we terminate after a fixed number of iterations, by which time we usually find the results have stabilised.

3.3 Piece-wise Affine Deformation Fields

An important issue for a registration algorithm is the choice of representation for the deformation field. There are three commonly used classes.

1) A dense field representing the movement of every pixel (eg fluid deformation models [14]),
2) a composition of simple warps (such as [24], [38]) or
3) fields controlled by a sparse set of control points (such as [4]). Typically the control points parameterise a set of splines (eg B-splines, thin plate splines or clamped plate splines).

The algorithm described above requires us to be able to invert the deformation efficiently, since during model construction we use $W(\mathbf{y}; \mathbf{x})$ to interpolate the training images, and when evaluating residuals we use $W^{-1}(\mathbf{y}; \mathbf{x})$ to interpolate the reference frame image. Although we have experimented with smoother interpolating splines, they cannot easily be inverted. Instead we have adopted piece-wise linear interpolation, in which the region of interest (RoI) is mosaiced by a set of triangles (in 2D) or tetrahedra (in 3D), the control points being the nodes of this mesh. For instance, see figure 1. Within each region we can use an affine approximation of the deformation field, which is easily inverted. Although the resulting piece-wise affine representation is not smooth (the derivatives are not defined at the boundaries), we have found it to produce good results. It is also simple to add constraints to prevent non-invertible mappings.
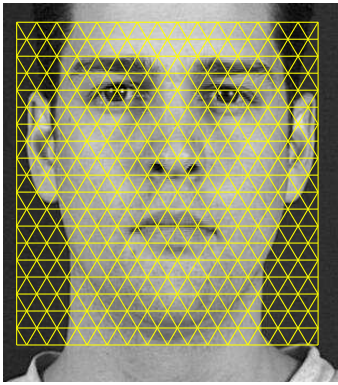
Fig. 1. Example face image and mesh of control points

3.4 Optimising the Control Point Positions

The algorithm outlined above involves performing an optimisation of control points, $\mathbf{x}$, on one image at a time. Since for a given set of points $\mathbf{x}$ we can estimate the parameters and residuals, the objective function depends only on $\mathbf{x}$ (call it $E(\mathbf{x})$).

This is a complex non-linear function. To reduce the chance of falling into local minima, we adopt a coarse to fine regime, with three distinct stages:

1) Estimate the global affine transformation
2) Optimise the positions of groups of control points
3) Optimise the positions of individual control points

3.4.1 Estimating the affine transformation

We first perform a coarse exhaustive search to estimate the translation. We then estimate affine transformation parameters using the Nelder-Mead Downhill Simplex algorithm [39].

3.4.2 Optimising groups of control points

Given an estimate of the affine transformation, we optimise a sequence of smooth transformations applied to the control points, effectively moving them in groups. This allows us to get approximately correct, smooth solutions quickly.

The most effective method we have found to do this is inspired by the registration algorithm of Lötjönen and Mäkelä [38], in which different regions of the image are selected at random and the points within the region optimised by applying a smooth local transformation.

The exact form of the local transformation is not critical during this initialisation stage, as long as it is smooth and invertible. In the 2D experiments below we use a simple variant of a grid based deformation [20]. Let $G(\mathbf{a} : \mathbf{d})$ be a transformation of space as follows

$$\mathbf{a}' = G(\mathbf{a} : \mathbf{d}) = \mathbf{a} + k(a_x)k(a_y)\mathbf{d} \tag{7}$$

where $\mathbf{a} = (a_x, a_y)^T$, $\mathbf{d}$ is the displacement of the origin (which parameterises the transformation) and $k(r)$ is a kernel defined as follows

$$k(r) = \begin{cases} 0.5(1 + \cos(\pi r)) & |r| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

This has no effect outside the region $[-1, 1][-1, 1]$. If the elements of $\mathbf{d}$ have a magnitude less than 0.4, this can be shown to be smooth and invertible (the proof involves demonstrating that the Jacobian of (7) remains positive under these conditions). An example of the effect of such a warp is shown in Figure 2.
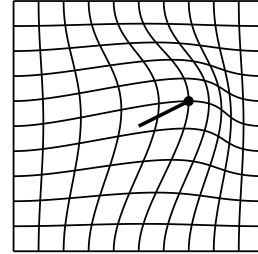

Fig. 2. Example of smooth local warp $G(\mathbf{a} : (0.4, 0.2))$ on regular grid over region [-1,1][-1,1]. The bold line shows the displacement, $\mathbf{d}$, of the origin.

If we wish to apply this deformation to a rectangular region of the current control point mesh, we compute the affine transformation $T()$ which maps the region $[-1, 1][-1, 1]$ to the desired region. We then define a deformation function

$$\mathbf{a}' = G_T(\mathbf{a} : \mathbf{d}) = T(G(T^{-1}(\mathbf{a}) : \mathbf{d})) \tag{9}$$

(which effectively maps the points to the unit frame, transforms them, then maps them back).

We can thus optimise the positions by finding the displacement $\mathbf{d}$ which optimises $E(G_T(\mathbf{x} : \mathbf{d}))$. Since $\mathbf{d}$ has only 2 parameters (or 3 in 3D), this can be quite efficient. To reduce the chance of falling into local minima, we evaluate this function on a grid of displacements and choose the best. We find this gives more robust results than using a simple downhill search.

To ensure we cover the whole region of interest, in 2D we select overlapping rectangles in a grid, visited in a random order. In the experiments on faces below we use 3 stages, visiting the image at rectangles first in a 3x3 grid, then 7x7, then 13x13.

For 3D images we use the natural 3D extensions of the deformations and volumetric grids. In the experiments on brains we use a 3x3x3 grid before optimising individual control points.

3.4.3 Optimising individual control points

The finest stages of the search involve directly optimising the control point positions, $\mathbf{x}$. For this we use a simple gradient based search. By displacing each point independently we can efficiently estimate the first and second derivatives, g_i and h_i of the function with respect to the ith parameter. We then perform a line search along direction $\mathbf{u}$, where $u_i = -g_i/h_i$ (with a suitable check to avoid division by zero errors). Figure 4b gives an example of the final mesh.

Allowing points to move independently leads to the danger of creating a non-invertible transformation. This

happens if a triangle or tetrahedra 'flips', so that two or more such regions then overlap. To prevent this we can add a large penalty to the cost function if the signed area/volume of any element changes sign relative to the equivalent element in the reference (The signed area of triangle $\{\mathbf{0}, \mathbf{a}, \mathbf{b}\}$ is $0.5(a_x b_y - b_x a_y)$, the signed volume of tetrahedron $\{\mathbf{0}, \mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is $(\mathbf{a} \times \mathbf{b}).\mathbf{c}/6$).

Whether this is important will be data dependent - the experiments below were performed without such penalties, as we found that we did not observe problems with invertibility.

In addition the basic framework can be modified for more sophisticated regimes. For instance, it can be useful to start with a sparse set of control points, and to increase the density as the search progresses.

3.5 Image Representation

In the 'interpretation though synthesis' approach it is common to work directly with image intensities, possibly with a simple linear normalisation to deal with global lighting effects [2].

However, it is known that making more explicit use of edge information can lead to more robust matching. In the medical image registration field it is common to use mutual information measures to perform image registration, as they are highly invariant to lighting effects [40]–[42]. Kokkinos and Yuille [29] perform groupwise registration on edge and ridge primal sketches.

In the following we demonstrate the effect of using various different normalisations and gradient information on the performance of registration.

3.5.1 Local z-normalisation

In global linear normalisation, each pixel with value v is mapped to $v' = (v - \mu)/\sigma$, where μ and σ are the mean and standard deviation of pixel values across the whole image (or modelled region).

A more effective approach is to perform a more local mapping,

$$v'(\mathbf{x}) = \frac{v(\mathbf{x}) - \mu(\mathbf{x})}{\max(\sigma(\mathbf{x}), \sigma_{min})} \qquad (10)$$

where $\mu(\mathbf{x})$ and $\sigma(\mathbf{x})$ are local estimates of the mean and standard deviation obtained by smoothing the image and the square of the image, and σ_{min} is a lower bound included to avoid enhancing noise in flat regions.

This deals with both global and more localised intensity variations.

3.5.2 Gradient information

We represent gradient information by applying Sobel filters to the result of the local z-normalised image, computing the absolute values in x and y gradient, then smoothing.

We then create a three-plane image in which the first plane is the local z-normalised image, and the other two are the x and y smoothed absolute gradient images.

We have found that registering such combined images during the early stages of the algorithm leads to significant improvements in robustness, as the smoothed edge information ensures that major features are aligned approximately at the start (see results below). In practice we reduce the degree of smoothing as the algorithm proceeds, in line with a coarse-to-fine style of search.

3.6 Shape Constraints

It is possible to perform registration without any explicit constraints on the shape deformation, and in many cases surprisingly good results can be achieved (see below). This is because shape constraints only make a significant positive impact in cases where there is considerable noise or image ambiguity. Where there are sufficient features in the image texture, the shape constraints may be unnecessary.

We have explored a range of forms of shape constraints, and found that often they make the registration accuracy worse as they tend to pull examples towards a model mode, and away from the data.

The simplest effective constraint we have found is a local elastic measure:

$$\log p(\mathbf{x}) = \sum_i \log p_i(\mathbf{x}_i) \qquad (11)$$

$$\log p_i(\mathbf{x}_i) = const - 0.5|\mathbf{x}_i - (d\mathbf{x}_i + \mathbf{x}_{\mu,i})|^2/\sigma_x^2 \qquad (12)$$

where $\mathbf{x}_i$ is the position of the ith point, $\mathbf{x}_{\mu,i}$ is the average position of the points in a neighbourhood around point i (determined by the connectivity of the mesh), and $d\mathbf{x}_i$ is the offset of this point from the average for the reference (mean) shape. A global value of the position variance σ_x^2 is used for all points. If significant rotation or scaling is anticipated, the measure can be evaluated after applying a global affine transformation into a standard reference frame. This measure penalises local bending.

Although it might seem that a natural extension is to estimate separate variances for each example from the training set, or to use a full covariance matrix to model the local shape, our experience to date is that this actually reduces overall performance.

3.7 Texture Model

The simplest model of texture is just to record the mean texture in the reference frame, $\mathbf{g}_\mu$. In this case there are no other texture parameters ($\mathbf{t}$ is empty, and $p(\mathbf{t}) = 1$). Again, we find that this works surprisingly well.

To measure the match, we assume an exponential distribution (which is long tailed, leading to more robustness than a Gaussian model).

Thus

$$\log p(r) = const - |r|/\sigma_r \qquad (13)$$

where r is the difference between image pixel and equivalent reference pixel and σ_r is a distribution width parameter. Below we compare the effect of using an exponential distribution with that of a Gaussian model.

3.8 Robust Texture Estimates

Rather than estimate the reference as the mean of all the data, a more robust result can be obtained by discarding some outliers. At each stage we construct the reference shape and texture by selecting the 75% of the examples which have the smallest absolute deviation in texture from the (global) mean.

This is particularly useful for avoiding corruption by the occasional poor model match or occlusion.

3.9 Choice of Measurement frame

To evaluate a particular warp $W(\mathbf{y})$ between a reference image, I_r, and a target image, I, a natural approach is to compare each pixel in the reference frame with the corresponding pixel in the target frame, evaluating $C(I_r(\mathbf{y}), I(W(\mathbf{y})))$ where $C(a, b)$ is some cost function. This is effectively deforming the target image into the reference frame, and is often the simplest and most efficient approach to implement. However, where there is local deformation, this can lead to some pixels in the target frame not contributing as they are not sampled.

An alternative approach is to consider each pixel in a region of interest in the target frame, and compare it with the corresponding pixel in the reference frame, $C(I_r(W^{-1}(\mathbf{y}')), I(\mathbf{y}'))$. This ensures that every pixel in the target region is 'explained' by the model. The Minimum Description Length approach [22] to groupwise registration suggests that the aim should be to completely explain the target images in terms of the model - this suggests that every pixel in the RoI should be included in the cost function.

The RoI is defined by the points on the borders of the mesh. During early stages of the optimisation we perform comparisons in the reference frame and allow the border points to move freely, thus locating the region of interest in the target frame. In the later stages we perform comparisons in the target frame, but fix the border points in place, ensuring the RoI, and therefore the number of pixels/voxels sampled, remains constant. Without this constraint there would be a tendency to reduce the area covered by the mesh to a single point as this would reduce the description length to zero.

In the following we demonstrate that the latter approach, measuring the differences in the target frame, leads to improved overall registration performance.

4 EXPERIMENTS

We demonstrate the approach to group-wise registration on two data sets. We apply it to 2D images of faces, for which we have accurate manual annotation, allowing evaluation of the accuracy of the resulting correspondences. We also demonstrate it on 3D MR images of the brain. For this data we have labelled voxel data, which allows evaluation by computing overlap measures, which are a surrogate for measures of correspondence accuracy.

4.1 2D Data: Face Images

We have performed a set of experiments to evaluate the effect of using the different choices of method above. In each we compare the result of registration with manual annotations which define a ground truth.

We use the first image of every person in Session 1 of the XM2VTS database [12], 293 images in total. Note that 121 of these images (41%) have either facial hair or glasses occluding some facial features. Each image was manually cropped to the facial region (see Figure 3).
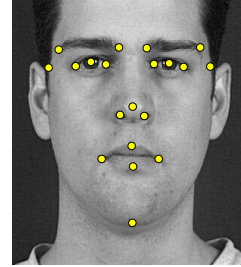

Fig. 3. Face image with 20 manually placed landmarks used for evaluation

Figure 4 gives an example of the results for one image in the set. Registering all 293 images takes approximately 20 minutes on a dual-core laptop. In this case the robust reference was used, see Figure 6b.

a) Target image   b) Final mesh

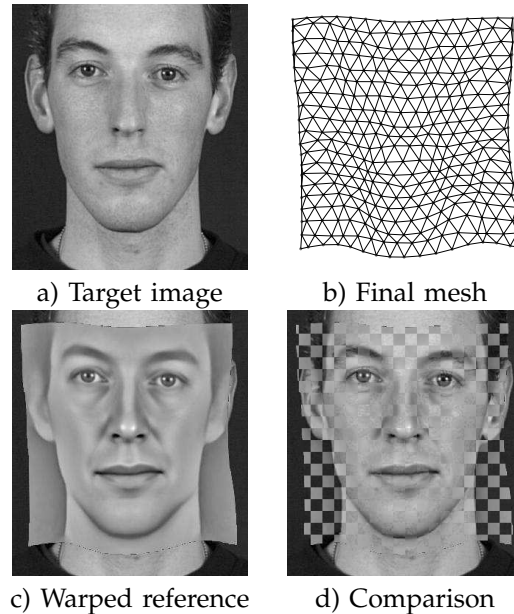c) Warped reference   d) Comparison

Fig. 4. Example final result for a face. (a) Target image. (b) Final mesh defining warp from reference to target (c) Result of warping reference image with mesh (d) Comparison: alternative squares from (a) and (c)

To evaluate the performance we used the publicly available manual ground truth points [2]. For each image we select the 20 feature points shown in Figure 3. We use an evaluation metric similar to that described by

2. http://personalpages.manchester.ac.uk/staff/timothy.f.cootes

Cristinacce *et al.* [43] for measuring the accuracy of facial feature search.

After applying the registration, we use the resulting warp fields to map all these ground truth points into the reference frame, where we compute their mean positions, giving an estimate of the 'true' position of the feature point on the reference face. Thus if $\mathbf{p}_{ij}$ is the position of landmark j on image i, we compute the mean reference point positions as

$$\bar{\mathbf{p}}_j = \frac{1}{N} \sum_{i=1}^{N} W_i^{-1}(\mathbf{p}_{ij}) \tag{14}$$

We then use the warp fields to project these 'average' manual points back to each image, and compare them with the actual point positions. The error measure for a single image, i, is then

$$d_i = \frac{1}{20 s_{eyes}} \sum_{j=1}^{20} |\mathbf{p}_{ij} - W_i(\bar{\mathbf{p}}_j)| \tag{15}$$

where s_{eyes} is the separation between the manual eye centre points (the inter-ocular distance). For examples of the value of this measure for different results, see Figure 5.

The tables below give the median and the 90th percentile for such point errors over the set of 293 images.

We apply the group-wise algorithm described above to the data, varying

- the image normalisation and representation,
- the shape constraints,
- whether the errors are measured in target or reference frame and
- the way in which the mean is estimated

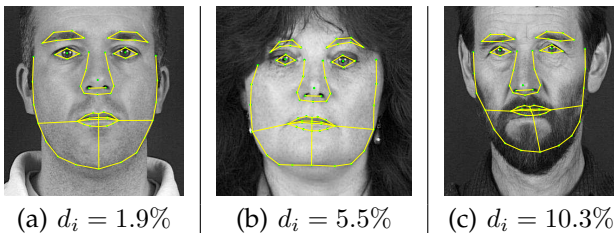| (a) $d_i = 1.9\%$ | (b) $d_i = 5.5\%$ | (c) $d_i = 10.3\%$ |

Fig. 5. Results of projecting mean annotation onto different examples, demonstrating difference errors.

Figure 5 shows the results of projecting the mean of the manual points (Eq.14) onto each of three images using the estimated deformation fields from the best performing test. Figure 5a shows one of the best results, 5c shows one of the worst results of the 293 images. Qualitatively the results are good. Since the method is essentially a local optimisation, good initialisation is required. For this data it was sufficient to estimate the translation, scale and orientation for each image using a coarse exhaustive search. Table 1 summarises the results of evaluating the accuracy of the registration.

The results demonstrate that

Image Features	Reference Frame		Target Frame	
	No shape	elastic	No shape	elastic
Linear Norm.	7.6 (13.2)	5.1 (10.4)	6.3 (12.1)	4.8 (10.3)
Histo. Eq.	7.4 (12.6)	4.8 (9.5)	6.3 (11.1)	4.6 (9.3)
Local z-norm	5.1 (9.7)	3.8 (7.5)	4.7 (8.9)	3.7 (7.0)
z-norm+grad	4.8 (7.4)	3.5 (5.5)	4.2 (6.8)	**3.5 (5.3)**

TABLE 1

Median and (90%ile) Point-Point registration errors (%) for different choices of normalisation, shape constraint and sampling frame. Std.Error≈ 0.1

1) Histogram equalisation gives better results than linear normalisation
2) Local z-normalisation gives better results than histogram equalisation
3) The gradient normalisation gives better results than local z-normalisation
4) Including local elastic shape constraints is beneficial
5) Measuring errors in the target frame gives better results than measuring them in the reference frame.

Figure 6a shows the evolution of the model mean for the best results, demonstrating that the registration gets more accurate as the algorithm progresses. Since a significant number of the images include people with glasses, a shadow of the average glasses appears on the final mean.

In order to investigate sensitivity to initialisation, we registered the set using 10 different images as the first template. The results were almost unchanged, with a standard deviation of 0.06 on the median errors.

4.1.1 Using a robust reference

To reduce the effects of occlusions and outliers, we repeat the above experiment using a robust estimate of the mean (building from the best 75%). Figure 6b shows the evolution of the robust mean - the glasses, being outliers, are removed from the mean. Table 2 compares this with using the global mean (note that the evaluation is performed over all images, not just the best 75%). This shows that in practice using a robust mean for this data makes little overall difference to performance.

	Reference Frame		Target Frame	
	No shape	elastic	No shape	elastic
Mean	4.8 (7.4)	3.5 (5.5)	4.2 (6.8)	3.5 (5.3)
Robust Mean	4.7 (7.4)	3.5 (5.5)	4.3 (6.9)	**3.4 (5.3)**

TABLE 2

Median and (90%ile) Point-Point registration errors when using mean and robust mean (using z-norm+grad features)

4.1.2 PDF of Intensity Residual

We compared the use of an exponential PDF for the texture errors with a Gaussian. The former tends to be

a) Standard Mean   b) Robust Mean

Fig. 6. Evolution of the mean of 293 face images.

more robust to outliers. Table 3 shows that the use of the exponential distribution generally gives slightly better results, but that the effect is not large enough to be significant.

PDF Form	Reference Frame		Target Frame	
	No shape	elastic	No shape	elastic
Exponential	4.8 (7.4)	3.5 (5.5)	4.2 (6.8)	**3.5 (5.3)**
Gaussian	4.8 (7.6)	3.6 (5.9)	4.5 (7.0)	**3.5 (5.3)**

TABLE 3
Median and (90%ile) Point-Point registration errors when
using different PDFs for texture residuals

4.1.3 Appearance Models

Statistical shape and appearance (shape+texture) models [2] have been constructed from the data using the best registration results. Figure 7 shows the first four modes

of appearance variation. Note that the inclusion of faces with glasses leads to a shadowing effect.
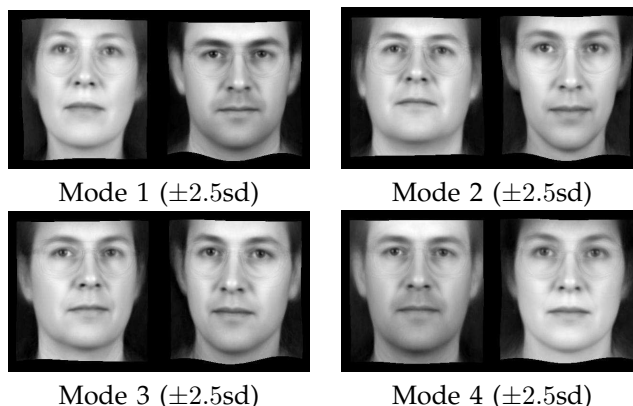

Mode 1 ($\pm$2.5sd)   Mode 2 ($\pm$2.5sd)

Mode 3 ($\pm$2.5sd)   Mode 4 ($\pm$2.5sd)

Fig. 7. First four modes of an appearance model built from 293 faces

4.1.4 Effect of choice of mesh density

The results above were generated with meshes of size 16 x 20 nodes. To investigate the effect of the choice of mesh density we repeated the registration with different numbers of control points (see Figure 8). In each case we used elastic shape constraints and measured errors in the target frame.

Figure 9 shows the results. This demonstrates that coarse meshes cannot capture the details sufficiently accurately. Beyond a certain size, no further improvement in performance is observed.
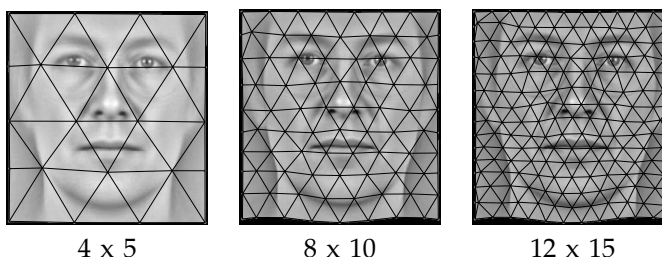

4 x 5   8 x 10   12 x 15

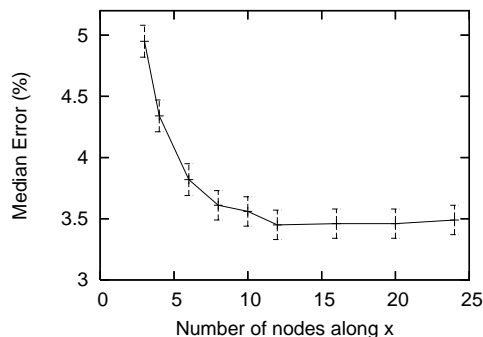Fig. 8. Examples of meshes and resulting mean

Fig. 9. Median point to point error (%) as a function of mesh density

4.2 3D Data: MR Images of the Brain

We have applied a 3D version of the algorithm to a set of 270 MR images of the brain, for which we have manual annotations. The images were T1-weighted, and had been acquired at different times with different scanners. The imaged cohorts include both control subjects as well as subjects with Alzheimer's Disease, Schizophrenia, Attention Deficit Hyperactivity Disorder (ADHD), and prenatal drug exposure. Their ages ranged from 4 years to 83 years. The segmentation of subcortical structures was performed at the Centre for Morphometric Analysis, Boston. The protocol is semi-automated and has been extensively described (for example see [44], [45]).

To evaluate the groupwise registration, we applied the algorithm to the complete set of 270 images. This resulted in an estimate of the mean shape and the mean intensity features in the reference frame defined by the mean shape, together with piece-wise linear deformation fields for every image into this reference. Figure 10 shows the evolution of the mean of the model. The registration took 23 hours on a 20 node (2 GHz) computing cluster. For each of several labelled structures we used the deformation fields to map each label image into the reference and compare it with the average.

More precisely, for a given structure, let $L_i(\mathbf{x})$ be a boolean label image ($L_i(\mathbf{x}) = 1$ if $\mathbf{x}$ is inside the structure in image i, $L_i(\mathbf{x}) = 0$ otherwise). Let $W_i(\mathbf{x})$ be the warp from reference frame to image i. We compute a mean probability image in the reference frame,

$$P(\mathbf{x}) = \frac{1}{270} \sum_{i=1}^{270} L_i(W_i(\mathbf{x})) \tag{16}$$

To evaluate the quality of the result on a particular image we compute the Dice overlap [46] between $P(\mathbf{x})$ (thresholded at 0.5), and the normalised label image $L_i(W_i(\mathbf{x}))$ (thresholded at 0.5).

Table 4 shows the mean Dice overlap for several structures when measuring errors in either the reference or target frame, with and without elastic shape constraints. In each case local z-normalisation of the intensity was applied.

The value of the position variance, σ_x^2 was chosen after experiments on a small subset (10%) of the data. Automatic selection of a suitable value, which affects the weighting between shape and texture components, is difficult and the subject of further research.

The table demonstrates that the best results are achieved when using the elastic shape constraints, and when measuring errors in the target frame.

Table 5 shows the effect of different features and normalisation methods. In this case the "z-norm+grad" is a two plane image in which the first plane is a locally z-normalised version of the original, and the second is a smoothed absolute gradient image.

In this case errors are measured in the target frame, and elastic shape constraints used. It demonstrates that the z-normalisation significantly improves performance
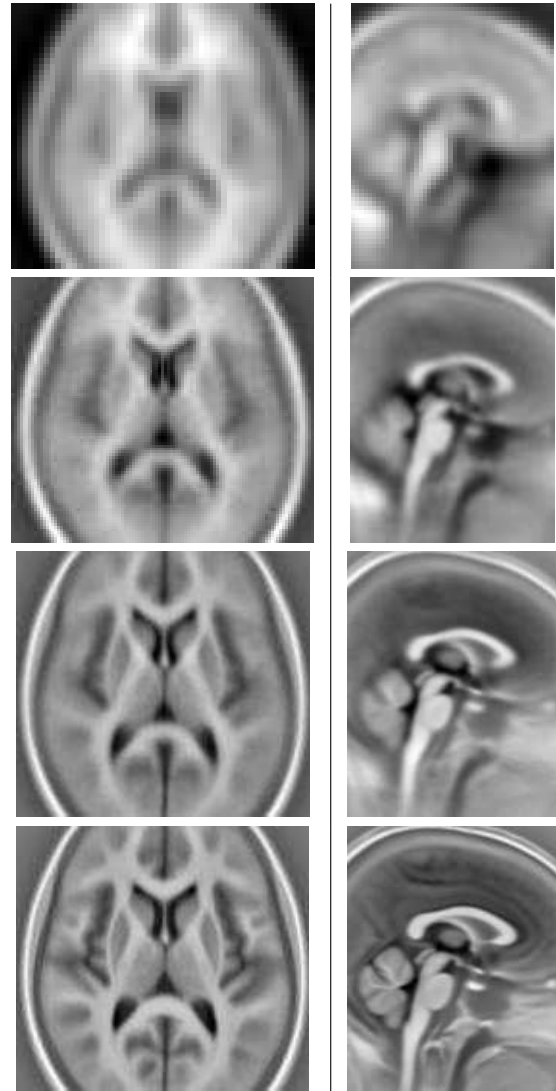

Fig. 10. Evolution of the mean of 270 MR images of the brain. Local intensity normalisation was used to pre-filter the images.

over the simple linear normalisation. This is because the data set includes images from different sites and scanners, leading to significant differences in intensities for similar structures.

Again we find that adding the gradient information improves the overall performance.

5 DISCUSSION AND CONCLUSIONS

We have demonstrated that accurate correspondences can be obtained across large groups of images. There are many choices which can be made for the features, normalisation and statistical modelling, which can significantly affect the overall performance of the system.

Extensive experiments suggest that effective results can be obtained by using local intensity normalisation and edge based features, using a simple locally elastic shape constraint and by measuring the errors in the target frame rather than the reference frame.

Structure	Reference Frame		Target Frame	
	No shape	elastic	No shape	elastic
Accumbens	68.6%	71.9%	68.4%	**72.4%**
Amigdala	70.2%	73.9%	71.5%	**74.2%**
Brain Stem	89.8%	90.2%	89.9%	**90.3%**
Caudate	81.1%	80.5%	80.7%	**81.3%**
Hippocampus	77.7%	79.3%	78.5%	**79.7%**
Ventricle	84.2%	83.2%	**84.8%**	83.8%
Pallidum	79.2%	81.4%	79.8%	**81.5%**
Putamen	87.8%	88.9%	88.1%	**89.0%**
Thalamus	87.0%	87.7%	87.0%	**88.1%**
Mean over all	80.0%	81.3%	80.3%	**81.6%**

TABLE 4
Mean of Dice overlap with average in the reference
frame for different brain structures (270 images).
Standard error ≈ 0.6

Structure	Linear	z-norm	z-norm+grad
Accumbens	14.7%	72.4%	**73.3%**
Amigdala	52.8%	74.2%	**75.0%**
Brain Stem	86.5%	90.3%	**91.0%**
Caudate	66.0%	81.3%	**82.5%**
Hippocampus	58.0%	79.7%	**80.7%**
Ventricle	76.5%	83.8%	**85.1%**
Pallidum	46.8%	81.5%	**82.3%**
Putamen	57.4%	89.0%	**89.8%**
Thalamus	81.1%	88.1%	**88.7%**
Mean	58.9%	81.6%	**82.5%**

TABLE 5
Mean of Dice overlap with average in the reference
frame for different brain structures (270 images), when
varying image features and normalisation used.
Standard error ≈ 0.6

The algorithm described above is essentially a local optimisation, requiring a 'good enough' initialisation. In the experiments initialisation involved estimating the affine correspondence between one image and the rest by an exhaustive search across translations, followed by downhill simplex optimisation of the orientation, scaling and shear parameters. This proved sufficient for all the cases we explored. Clearly, if the position of the object of interest is less constrained in the image, more care must be taken at this stage.

During the non-linear stages of the registration process there is a danger that the algorithm will fall into local minima. This can in part be mitigated by semi-exhaustive search during the early stages (see Section 3.4.2), but also relies on a good choice of objective function. One reason that the smoothed gradient image features improve results is because they seem to reduce the number of times the algorithm falls into local minima (such as the eyes matching to eyebrows or vice-versa in the case of faces).

In common with many non-rigid registration schemes, the approach assumes a one-to-one correspondence between all images. There are often cases where structures appear in only a subset of images (for instance if some of the faces had open mouths). With the current approach such structures either get blurred out or cause local

distortions of the deformation field - dealing with them properly is the subject of ongoing research. In the case of faces the method can deal with modest out of plane head rotation, but too much will cause features to become occluded and the method will fail. Also exaggerated facial expressions can cause problems, partly because initialisation becomes less reliable.

In future work we will explore more sophisticated shape and texture models, and other initialisation regimes. A key unresolved problem is that of choosing the optimal weighting between the shape components and the texture components of the model. In the above we estimated values based on results of a small subset. This would not be satisfactory for a fully automatic system.

However, the correspondences obtained from the relatively simple methods described above are accurate enough to allow the automatic construction of detailed models of shape and appearance which can be applied to many different problems.

ACKNOWLEDGEMENT

REFERENCES

[1] A. Guimond, J. Meunier, and J.-P. Thirion, "Automatic computation of average brain models," in *Proc. MICCAI*, 1998, pp. 631–640.
[2] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, 2001.
[3] S. Baker, I. Matthews, and J.Schneider, "Automatic construction of active appearance models as an image coding problem," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 10, pp. 1380–84, 2004.
[4] D. Rueckert, A. Frangi, and J. Schnabel, "Automatic construction of 3D statistical deformation models using non-rigid registration," *IEEE Trans. Medical Imaging*, vol. 22, no. 8, pp. 1014–25, 2003.
[5] S. Balci, P. Golland, M. Shenton, and W. Wells, "Free-form B-spline deformation model for groupwise registration," in *Proc. MICCAI Statistical Registration Workshop*, 2007, pp. 23–30.
[6] T. Cootes, C. Twining, V.Petrović, R.Schestowitz, and C. Taylor, "Groupwise construction of appearance models using piece-wise affine deformations," in *16th British Machine Vision Conference*, vol. 2, 2005, pp. 879–888.
[7] S. Duchesne, J. Pruessner, and D. Collins, "Appearance-based segmentation of medial temporal lobe structures," *NeuroImage*, vol. 17, pp. 515–531, 2002.
[8] M. J. Jones and T. Poggio, "Multidimensional morphable models : A framework for representing and matching object classes," *International Journal of Computer Vision*, vol. 2, no. 29, pp. 107–131, 1998.
[9] S.Joshi, B.Davis, M.Jomier, and G.Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage*, vol. 23, pp. S151–S160, 2004.
[10] A. Frangi, D. Rueckert, J. Schnabel, and W. Niessen, "Automatic 3D ASM construction via atlas-based landmarking and volumetric elastic registration," in 17th *Conference on Information Processing in Medical Imaging*, 2001, pp. 78–91.
[11] 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 Trans. Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[12] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, "XM2VTSdb: The extended m2vts database," in *Proc. 2nd Conf. on Audio and Video-based Biometric Personal Verification*. Springer Verlag, 1999, pp. 72–77.

[13] T. Vetter, M. Jones, and T. Poggio, "A bootstrapping algorithm for learning linear models of object classes," in *Computer Vision and Pattern Recognition Conference 1997*, 1997, pp. 40–46.

[14] M. J. Jones and T. Poggio, "Multidimensional morphable models," in 6th *International Conference on Computer Vision*, 1998, pp. 683–688.

[15] J. V. Hajnal, D. L. Hill, and D. J. Hawkes, *Medical Image Registration*. CRC Press, 2001.

[16] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.

[17] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[18] J. P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, 1998.

[19] D.L.Collins, C. Holmes, T. Peters, and A. Evans, "Automatic 3D model-based neuroanatomical segmentation," *Human Brain Mapping*, vol. 3, pp. 190–208, 1995.

[20] T. Cootes, C. Twining, K. Babalola, and C. Taylor, "Diffeomorphic statistical shape models," *Image and Vision Computing*, vol. 26, pp. 326–332, 2008.

[21] P. Lorenzen, B. Davis, and S. Joshi, "Unbiased atlas formation via large deformations metric mapping," in *Proc. MICCAI*, 2005, pp. 411–418.

[22] S. Marsland, C. Twining, and C. Taylor, "Groupwise non-rigid registration using polyharmonic clamped-plate splines," in *MIC-CAI*, ser. Lecture Notes in Computer Science, 2003.

[23] S. Marsland, C. J. Twining, and C. J. Taylor, "A minimum description length objective function for groupwise non-rigid image registration," *Image and Vision Computing*, vol. 26, no. 3, pp. 333–346, 2008.

[24] T. Cootes, S. Marsland, C. Twining, K. Smith, and C. Taylor, "Groupwise diffeomorphic non-rigid registration for automatic model building," in 8th *European Conference on Computer Vision*, vol. 4. Springer, 2004, pp. 316–327.

[25] E. Learned-Miller, "Data driven image models through continuous joint alignment," *Trans. Pattern Analysis and Machine Intelligence*, vol. 28, pp. 236–250, 2006.

[26] R. Fergus, P. Perona, and A. Zisserman, "Weakly supervised scale-invariant learning of models for visual recognition," *International Journal of Computer Vision*, vol. 71, pp. 273–303, 2007.

[27] J. Ponce, M. Hebert, C. Schmid, and A. Zisserman, Eds., *Towards Category-Level Object Recognition*. Springer-Verlag, 2006.

[28] G.Langs, R.Donner, Peloschek, and H.Bischof, "Robust autonomous model learning from 2D and 3D data sets," in *MICCAI*, vol. 1, 2007, pp. 968–976.

[29] I. Kokkinos and A.Yuille, "Unsupervised learning of object deformation models," in *International Conference on Computer Vision*, 2007.

[30] F.Wang, B.V.Vemuri, A.Rangarajan, and S.J.Eisenschenk, "Simultaneous nonrigid registration of multiple point sets and atlas construction," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30, no. 11, pp. 2011–2022, 2008.

[31] J. Rissanen, *Stochastic Complexity in Statistical Inquiry*. World Scientific Press, Singapore, 1989, 1989.

[32] R. Davies, C.Twining, T. Cootes, and C. Taylor, "A minimum description length approach to statistical shape modelling," *IEEE Trans. on Medical Imaging*, vol. 21, pp. 525–537, 2002.

[33] R. Davies, C.Twining, T. Cootes, J. Waterton, and C. Taylor, "3D statistical shape models using direct optimisation of description length," in *European Conference on Computer Vision*, vol. 3. Springer, 2002, pp. 3–20.

[34] C. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, 1948.

[35] C.J.Twining, S.Marsland, and C.J.Taylor, "A unified information-theoretic approach to the correspondence problem in image registration," in *International Conference on Pattern Recognition*, vol. 3, 2004, pp. 704–709.

[36] I. Matthews and S. Baker, "Active appearance models revisited," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 135 – 164, November 2004.

[37] I. Dryden and K. V. Mardia, *The Statistical Analysis of Shape*. Wiley, London, 1998.

[38] J. Lötjönen and T. Mäkelä, "Elastic matching using a deformation sphere," in *Proc. MICCAI*, 2001, pp. 541–548.

[39] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C (2nd Edition)*. Cambridge University Press, 1992.

[40] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. on Medical Imaging*, vol. 16, no. 2, pp. 187–198, April 1997.

[41] P. Viola and W. W. III, "Alignment by maximization of mutual information," *International Journal of Computer Vision*, vol. 24, no. 2, pp. 137–154, 1997.

[42] T. Gaens, F. Maes, D. Vandermeulen, and P. Suetens, "Non-rigid multimodal image registration using mutual information," in *Proc. MICCAI*, 1998, pp. 1099–1106.

[43] D.Cristinacce and T.F.Cootes, "Automatic feature localisation with constrained local models," *Pattern Recognition*, vol. 41, no. 10, pp. 3054–3067, 2008.

[44] P. Filipek, C. Richelme, D. Kennedy, and V. Caviness, "The young adult human brain: An MRI-based morphometric analysis," *Cerebral Cortex*, vol. 4, pp. 344–360, 1994.

[45] M. Nishida, N. Makris, D. N. Kennedy, M. Vangel, B. Fischl, K. S. Krishnamoorthy, V. S. Caviness, and P. E. Grant, "Detailed semiautomated MRI based morphometry of the neonatal brain: Preliminary results," *Neuroimage*, vol. 32, pp. 1041–1049, 2006.

[46] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, pp. 297–302, 1945.