Groupwise registration of richly labelled images

K O Babalola* and T F Cootes

Division of Imaging Science and Biomedical Engineering (ISBE), University of Manchester, UK

Abstract.
We describe a method of registering 3D images in which regions have been segmented and labelled. Standard registration schemes cannot be naively applied to such images and several modifications to allow their registration have been proposed such as using vector valued images with one plane for each label [1] or a *label consistency* measure [2]. However, these either lead to impractically large images when a large number of labels are being considered, or cannot be applied to groupwise registration in a straightforward manner.
The method we describe does not lead to impractically large images and can be applied to both groupwise and pairwise registration. It involves mapping each label value to a vector in a low dimensional space and applying a multi-plane registration algorithm to the resulting vector valued image. To obtain good results, the vectors representing each label should be well separated and chosen in such a way that there is minimal confusion between them. We demonstrate the method by using it to register a set of richly labelled images in a groupwise manner and use the resulting correspondences to construct a statistical shape model of a number of subcortical structures in the brain.

1 Introduction

Non-rigid registration is widely used to find correspondences between images, and between atlases and images [3, 4]. Many techniques have been proposed for registering unlabelled grey level images. Where the images have been segmented, so that each voxel contains a label indicating which structure it belongs to, one commonly has a single plane image of label values (such as an integer indicating the region type). Registering labelled images is a necessary step in the automatic construction of statistical shape models. However, standard registration algorithms should not be naively applied, as the difference between voxel values (integer labels) is no longer a meaningful metric. Instead, Tsai *et al.* [1] suggested extracting binary planes, one for each structure and encode the structures in a multi-dimensional binary vector-valued image. Thus if each voxel can be labelled with an integer in the range $[1, N]$, then N planes are created.

We propose an alternative technique. We generate a mapping $i \rightarrow \vec{v}_i$, which replaces each label index i with a vector value $\vec{v}_i$ in an m-dimensional space. The dimensionality, m, can be any positive integer, but in practise will be kept relatively small (see below). The key to the approach is to ensure that the vectors $\{\vec{v}_i\}$ are well separated in the space so that there is minimal interference between labels from different structures when they are nearby in the image. Having applied the mapping, we can smooth the images if required to encourage convergence, and then apply any multi-plane registration algorithm.

In the following we will describe the approach, showing one choice for the mapping, and use it to register a set of richly labelled images with a "groupwise" algorithm [5], which seeks to find the best registration of a group, rather than just a pair of images. Section 2 gives details of related work and section 3 gives details of the mapping process. The construction of SSMs using the modified group-wise method of Cootes *et al.* [5] is described in section 4 and quantitative analysis of the results of the registration are shown in section 5. The paper ends with a discussion.

2 Related work

Frangi *et al.* [6] propose a method of automatically constructing 3D Active Shape Models from binary images of a particular structure. This required the registration of binary images of the structure being modelled. They used a *quasi-affine* registration with nine degrees of freedom and a normalised mutual information metric to perform an initial registration to an atlas, then combined the shapes using their signed distance transforms. This method gave promising results, however, its application to the case of images containing more than one structure was not investigated.

Frangi *et al.* [2] construct statistical shape models of the ventricles of the heart from labelled images. They define two similarity measures applicable to registration of a pair of labelled images. Both vary between 0 (the worst value) and 1 (the best value). The first, *Label consistency*, measures the fraction of labels in the source shape that are correctly mapped into the target shape. The second, the κ *statistic*, is used in biomedical research to assess inter-rater agreement.

*Email: kola.babalola@manchester.ac.uk

These measures allowed the statistical shape models to be constructed, however, they are only defined for pairwise registration, and their extension to groupwise registration has not been explored.

Tsai *et al.* [1] built shape models from images with labels of multiple structures and applied the models to segmentation. However, they constructed their models by creating binary vector-valued images, with the dimensions of the vectors determined by the number of unique labels. They registered the images using an overlap based cost function and constructed the shape models using signed distance transforms of the aligned binary shapes. This approach results in memory limitations restricting the number of labels that can be considered during registration.

3 METHODS

3.1 Mapping labels to vectors

We seek to map each label value i to a unique vector, $\vec{v}_i$ in an m-dimensional space, so as to ensure that different resulting vectors are well separated. However, we also wish to arrange that if the resulting vector image is smoothed (either explicitly in order to improve registration convergence, or implicitly when interpolating values), there is a minimal chance of vector values being created which could be confused with other labels. Consider the border between two structures, labelled i and j. On one side of the border voxels are mapped to vector $\vec{v}_i$, and on the other to $\vec{v}_j$. If the vector image were smoothed, voxels around the border would have values lying on the line segment between $\vec{v}_i$ and $\vec{v}_j$, i.e. $\vec{v} = \alpha\vec{v}_i + (1-\alpha)\vec{v}_j$. We would thus like to ensure that no other labels have values near that line. Similarly at junctions between three structures, blurring produces vectors which are linear combinations of three key vectors, in a triangular region.

Ideally we would thus like to select vector values $\{\vec{v}_i\}$ which are both well separated, and such that the linear combinations of pairs or triplets which occur at boundaries are unlikely to be mistaken for any other values.

A complete solution would involve analysing the original images to determine which structures are neighbours, and using this as a constraint in an optimisation. However, as an initial approximation we have experimented with a mapping in which the background label is mapped to zero ($\vec{v}_0 = \vec{0}$), and all other labels are mapped to points maximally separated on a unit hypersphere ($|\vec{v}_i| = 1$).

We have treated the background differently from other labels because in the general case any label could be a neighbour of the background. Without making use of neighbour information, this arrangement should lead to the least interference between different structures. However, there is a possibility of weighted sums of non-background labels representing the background. This can be averted by taking information of label neighbours into account.

3.1.1 Maximally separable points on a hypersphere

Mapping labels to 2D values is equivalent to equally spacing the labels on the radius of a unit circle. For 3D it is equivalent to equally spacing the labels on the surface of a unit sphere. For the 2D case calculating the positions are trivial. For the 3D case one could use the analogy of equally spaced charges.

However, the general case of spacing m points equally on the surface of a unit hypersphere is equivalent to the optimisation problem :

$$\text{Find } m \text{ points}$$
$$\{\mathbf{x}_i\},\ |\mathbf{x}| = 1,\ i = 1 \ldots m$$
$$\text{maximally separated} \quad (1)$$

$\mathbf{x}_i.\mathbf{x}_j = \cos\theta$ as $|\mathbf{x}_i| = |\mathbf{x}_j| = 1$ and $\mathbf{x}_i.\mathbf{x}_j = 1$ if $\mathbf{x}_i = \mathbf{x}_j$ and $\mathbf{x}_i.\mathbf{x}_j < 1$ if $\mathbf{x}_i$ is away from $\mathbf{x}_j$. We therefore seek to minimise an objective function of the form :

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathbf{x}_i.\mathbf{x}_j \qquad \text{subject to } |\mathbf{x}_i| = 1 \quad \forall\, i \tag{2}$$

Performing such an optimisation is not trivial due to the sinusiodal nature of the *dot* product. At the moment we use an iterative procedure to equally space our points. We generate the vectors of the appropriate dimensionality for the required number of points from a random multidimensional distribution and normalise them to unit length. We then iteratively find the two closest points in the set and increase their separation by a small distance normalising both vectors afterwards. This is repeated until the standard deviation of the closest distance between the points falls below a given value, or a maximum number of iterations is reached. Figure 3.1.1 shows the results of mapping 20 points to 2D and 3D using this algorithm.

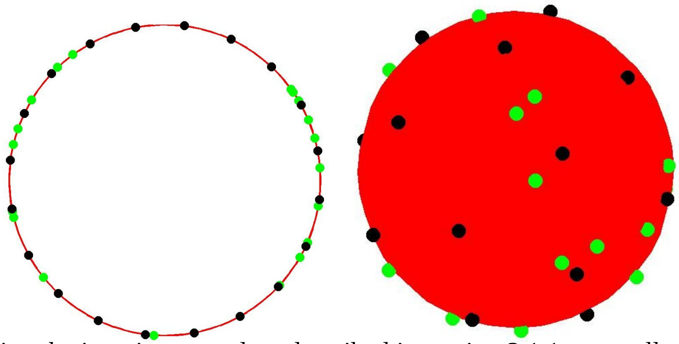


Figure 1. Results of applying the iterative procedure described in section 3.1.1 to equally spacing 20 points on a unit circle in 2D (left) and on a unit sphere in 3D (right). The green circles show the initial distribution of the random points, and the black circles show the final result.

3.2 Group-wise registration

It is often necessary to construct statistical models of the shape and shape variation of structures appearing in medical images. This requires finding correspondences across a set of images. Rather than use pair-wise techniques (such as selecting one image as a reference and independently registering each of the other images to it), there is considerable interest in *groupwise* methods, which attempt to find the optimal correspondences across the group as a whole [5, 7, 8]

We adopt a variant of the minimum-description length approach [5, 8] in which the quality of the current registration is evaluated by estimating the amount of information that would be required to encode the training images using a statistical model of shape and texture constructed using the current correspondences. This gives an objective function (a description length), which can be optimised by manipulating the correspondences.

Correspondence is represented by the position of a set of control points on each image (piece-wise linear interpolation is used to estimate correspondence away from the control points). The general approach is as follows (though see [5] for details):

1. Select one image to be used as an initial reference
2. Initial affine registration to reference:
   (a) For each image in turn estimate (affine) movement of control points to optimise match to reference image
3. Groupwise registration:
   (a) For each image in turn
      i. Construct a shape and 'texture' model using current points in all other images
      ii. Modify the positions of the points in the target image to minimise the cost of encoding that image using the current model
   (b) Repeat until convergence

In this case 'texture' refers to vector values at each voxel.

4 EXPERIMENTS

We applied group-wise registration as described above to register 37 subjects from a dataset provided by the Centre for Morphometric Analysis, Boston. This consisted of T1 MR images and corresponding labelled images of each subject. 35 cortical and subcortical structures in the images had been labelled. We mapped label values to 3 dimensions and registered the 3 plane images of each subject. The transformations obtained from these registrations were used to establish correspondence between the surfaces of 10 subcortical structures (accumbens, amygdala, brain stem, fourth ventricle, caudate, hippocampus, lateral ventricles, pallidum, putamen and thalamus). These formed a training set from which a statistical shape model (SSM) was constructed as described in [9].

5 RESULTS

To quantitatively assess the quality of the registration a distance metric giving the registration error for each subject was obtained as follows. One subject was selected as a template and surfaces for each of the structures listed in section 4 were obtained. The surface of each structure was warped to approximate the surface of the structure in each member of the training set using the correspondences obtained during the registration. The Euclidean distance between the closest points on the original surface to each point on the warped surface was taken as the registration error of the point. If the registration was perfect these distances would be expected to be zero. Figure 2 shows a histogram of the registration error for each of the 37 subjects. A SSM is a linear model approximating a class of shapes by a main shapes and the main ways in which the shapes vary (modes of variation). The first three modes of variation of the resulting statistical shape model are shown in figure 3.

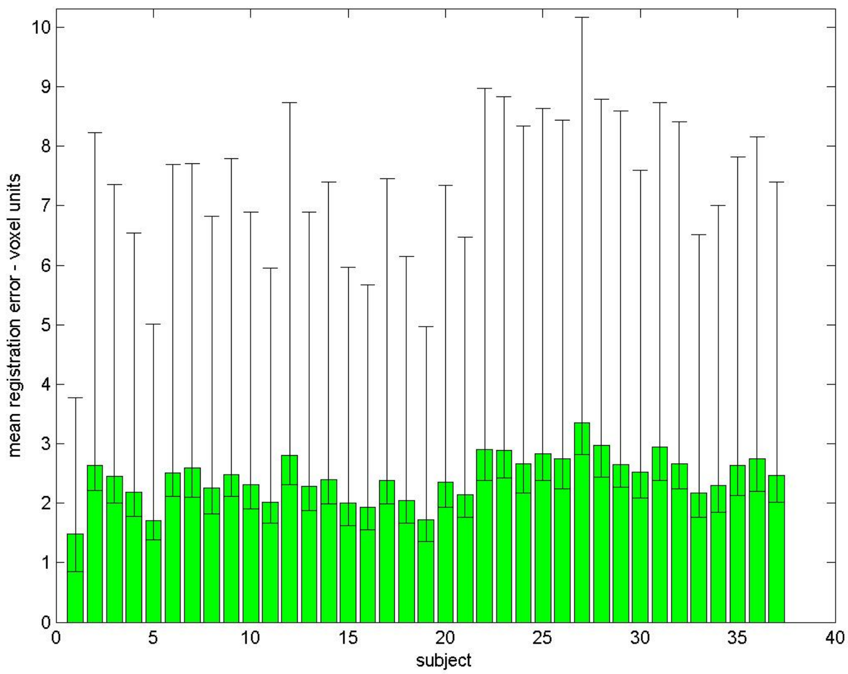


Figure 2. Mean registration error for each subject in the image set. The bars show the mean of the Euclidean distance between densely sampled points on the registered surface and the original surface. The vertical limits show the 10^{th} and 90^{th} percentiles of the point to surface error. The distances were measured in voxel units, and the voxel size being 1mm $\times$ 1.5mm $\times$ 1mm

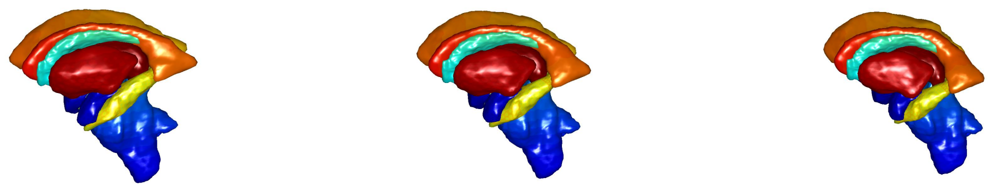
6 DISCUSSION

We have described a method of registering richly labelled data by replacing each voxel label with a vector in a low dimensional space. The values of the vectors should be chosen to ensure that those belonging to different labels are well separated, and that when smoothed there is minimal chance of confusion.

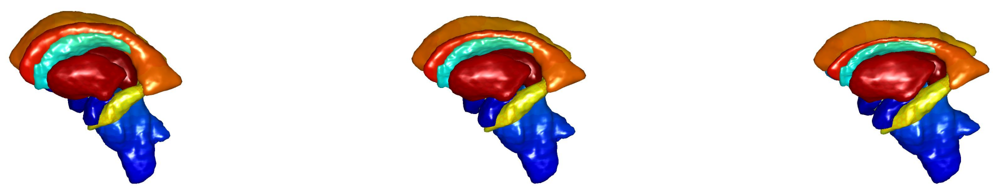
Our initial implementation simply maps points to (randomly assigned) maximally separated vectors on the unit hypersphere. When we use the technique as pre-processing for a groupwise registration algorithm, we obtained encouraging results. In future work we will explore the effect of the choice of dimension m for the vector space, and investigate methods of taking known neighbour relationships into account when assigning vectors to structure labels. We believe inclusion of the neighbour relationships will improve the accuracy of registration. We will also perform quantitative evaluation of the method and compare it with other registration methods, in particular with the pairwise method based on *label consistency* [2].

7 ACKNOWLEDGEMENTS

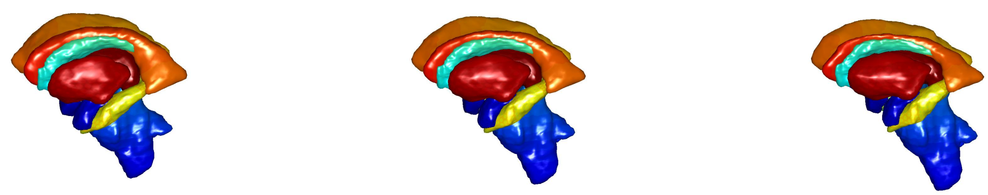
This work was funded by the EPSRC under the IBIM project, and David Kennedy provided the MR images used.



(a) +3 and -3 standard deviations from mean (centre subfigure) along mode 1



(b) +3 and -3 standard deviations from mean (centre subfigure) along mode 2



(c) +3 and -3 standard deviations from mean (centre subfigure) along mode 3

Figure 3. Mean shape and extremes of variation for the first three modes. The left subfigures are for $-2.5\sqrt{\lambda_i}$ and the right $+2.5\sqrt{\lambda_i}$. λ_i is the variance associated with mode i

References

1. A Tsai, W Wells, C Tempany, E Grimson, and A Willsky, "Mutual information in coupled multi-shape modle for medical image segmentation," *Medical Image Analysis*, vol. 8, no. 4, pp. 429–445.
2. A F Frangi, D Rueckert, J A Schnabel, and Niessen W J, "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling," *IEEE Transactions on Medical Imaging*, vol. 21, no. 9, pp. 1151–1166, 2002.
3. K K Bhatia, J V Hajnal, B K Puri, A D Edwards, and D Rueckert, "Consistent groupwise non-rigid registration for atlas construction," in *IEEE Symposium on Biomedical Imaging (ISBI) Arlington*, 2004, pp. 908–911.
4. D Rueckert, A F Frangi, and J A Schnabel, "Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 1014, 2003.
5. T F Cootes, C J Twining, V Petrovic, R Schestowitz, and C J Taylor, "Groupwise construction of appearance models using piece-wise affine deformations," in *Proceedings of 16th British Machine Vision Conference, Oxford*, 2005, pp. 879–888.
6. A F Frangi, D Rueckert, J A Schnabel, and W J Niessen, "Automatic 3D ASM construction via atlas-based landmarking and volumetric elastic registration," *LNCS (IPMI 2001 Proceedings)*, vol. 2082, pp. 78–91, 2001.
7. B Davis, P Lorenzen, and S Joshi, "Large deformation minimum mean squared error template estimation for computational anatomy," in *Proceedings of ISBI*, 2004, pp. 173–176.
8. C J Twining, T F Cootes, S Marsland, V Petrovic, R Schestowitz, and C J Taylor, "A unified information-theoretic approach to groupwise non-rigid registration and model building," in *LNCS (IPMI 2005 Proceedings)*, 2005, vol. 3565, pp. 1–14.
9. T F Cootes, C J Taylor, D H Cooper, and J Graham, "Active Shape Models - their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, Jan. 1995.