Robust point correspondence applied to two and three-dimensional image registration

GUEST, E., BERRY, E., BALDOCK, R. A., FIDRICH, M. and SMITH, M. A.

Published version

GUEST, E., BERRY, E., BALDOCK, R. A., FIDRICH, M. and SMITH, M. A. (2001). Robust point correspondence applied to two and three-dimensional image registration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23 (2), 165-179.

Robust Point Correspondence Applied to Two- and Three-Dimensional Image Registration

Elizabeth Guest, Elizabeth Berry, Richard A. Baldock, Márta Fidrich, and Mike A. Smith

Abstract—Accurate and robust correspondence calculations are very important in many medical and biological applications. Often, the correspondence calculation forms part of a rigid registration algorithm, but accurate correspondences are especially important for elastic registration algorithms and for quantifying changes over time. In this paper, a new correspondence calculation algorithm, CSM (Correspondence by Sensitivity to Movement), is described. A robust corresponding point is calculated by determining the sensitivity of a correspondence to movement of the point being matched. If the correspondence is reliable, a perturbation in the position of this point should not result in a large movement of the correspondence. A measure of reliability is also calculated. This correspondence calculation method is independent of the registration transformation and has been incorporated into both a 2D elastic registration algorithm for warping serial sections and a 3D rigid registration algorithm for registering pre and postoperative facial range scans. These applications use different methods for calculating the registration transformation and accurate rigid and elastic alignment of images has been achieved with the CSM method. It is expected that this method will be applicable to many different applications and that good results would be achieved if it were to be inserted into other methods for calculating a registration transformation from correspondences.

Index Terms—Image registration, iterative closest point, surface matching, point correspondence, image warping, image matching, serial sections, reconstruction.

1 Introduction

The ability to calculate accurate and robust correspondences is important for many applications in biomedical imaging. Examples of such applications include evaluation of aesthetic surgery to the face, growth measurement, elastic registration of acquired images to an atlas, and elastic registration of serial histological sections in order to produce a smooth three-dimensional reconstruction. Correspondence algorithms are needed for developing point distribution models [1] and are widely used in all the types of registration algorithms categorized by Lester and Arridge [2]: rigid, affine, elastic, and fluid registration.

Correspondence algorithms generally consist of two parts: a similarity measure and a cost function. The similarity measure is used to compare two points or regions in different images. The cost function analyzes the values produced by the similarity measure after a point or region in one image has been compared to many points or regions in the other image. In this paper, the term "correspondence calculation" will refer solely to the cost function.

The similarity measure can be application specific as is the case for Zhang [3], who matched megakaryocyte cells on adjacent serial sections based on the relative locations of neighboring cells. Hibbard et al. [4] have devised a method for matching and finding correspondences between median eminence microvascular nodules in the brain. However, many authors have proposed more general methods. Some match gray or color levels between images either directly [5], or using cross correlation [6], [7], mutual information [8], or gradient values [9]; others extract features such as contours [10], [11] or crest lines [12] and match properties of these features. Three-dimensional surfaces are often matched using functions based on their normals and curvatures [13], [14]. However, other types of features such as "spin images" [15] and "harmonic shape images" [16] have recently been devised. The similarity measure used for these feature images is cross correlation.

In general, correspondences are found by choosing the point with the optimum similarity value, but this can give unsatisfactory correspondences. For example, in a gray-scale image, a point in the middle of a region of texture is likely to match well to a large area, but may have as its optimum values, a point near the boundary of the region, rather than near the middle as would be expected. As a result, neighboring points will often have correspondences that are not close to each other.

Some authors attempt to overcome these problems by smoothing the calculated correspondences so that the resulting displacement vectors all point in approximately the same direction [7], or detect and discard unreliable correspondences [9]; others grow correspondences from a distinctive point [5], [17]. Methods that involve segmentation often have reliable and robust correspondence algorithms [10], [12], [18], but apart from the difficulties of reliable segmentation, these methods cannot directly calculate correspondences for all the points in the image, but

- *E. Guest is with the School of Computing, Leeds Metropolitan University, Beckett Park Campus, Leeds LS6 3QS, United Kingdom. E-mail: e.guest@lmu.ac.uk.*
- *E. Berry and M.A. Smith are with the Centre of Medical Imaging Research and Academic Unit of Medical Physics, University of Leeds, Leeds, United Kingdom. E-mail: {e.berry, m.a.smith}@leeds.ac.uk.*
- *R.A. Baldock is with the MRC Human Genetics Unit, Western General Hospital, Crewe Rd, Edinburgh EH4 2XU, Scotland. E-mail: richard@hgu.mrc.ac.uk.*
- *M. Fidrich is with the Research Group on Artificial Intelligence, József Attila University, Szeged, Hungary. E-mail: fidrich@cc.u-szeged.hu.*

Manuscript received 28 Dec. 1999; revised 7 July 2000; accepted 6 Oct. 2000. Recommended for acceptance by S. Sarkar.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 111146.

must interpolate correspondences for points not included in the segmentation.

In this paper, we propose a new method for calculating correspondences: the CSM (Correspondences by Sensitivity to Movement) algorithm. Instead of taking the point with the optimum match value, we calculate a "tentative corresponding point" by summing over the matchmap produced by matching the point of interest to several points in the other image. This summation includes a distance penalty so that more weight is given to closer matches. The basic principle of the CSM algorithm is to test the reliability of a potential correspondence by virtual movements of the point of interest to see if it always matches to the same point in the other image. If the point of interest matches well to a large area, these points will be scattered, but if the point of interest matches well to only a small region, all the tentative corresponding points will be clustered in this region (unless the distance penalty swamps the similarity values). This approach overcomes the problems of unsatisfactory correspondences outlined above and ensures that correspondences for neighboring points are close. This new method has the following advantages:

- There is no segmentation. The results of the similarity measure are not thresholded and the images do not need to be segmented into edges or regions. In the two example applications described below, points are not constrained to be distinctive (such as lie on a high contrast edge).
- Points in the middle of homogeneous regions are given "sensible" correspondences in that the correspondence will be the closest point in the corresponding homogenous region. This overcomes the problem that the best match may suggest a large displacement vector, even when two almost identical images are perfectly registered.
- A measure of reliability of the correspondence is calculated. In practice, this means that points that match well to a point or line have high reliability, whereas points that match well to a large region will have low reliability. As a result, it is not important that all match points are distinctive and, in registration applications, points that match well to a large region can be allowed to move freely over this region.
- Subresolution matching can be achieved even when matching only to subsampled points. So if a node of a (surface) mesh is matched to nodes of another mesh, the correspondence is not a mesh node, but can be any point on the surface of the second mesh.

This algorithm is independent of both the similarity measure and any method for computing a registration transformation. Therefore, we envisage that it will be widely applicable to many different problems and that it will give good results when it is inserted into many of the different published registration algorithms that require correspondences.

The performance of the CSM correspondence calculation algorithm is tested on two very different applications: 2D warping (or elastic registration) of serial histological sections of mouse embryos in order to produce a smooth reconstruction of the original object; and rigid registration of range images of human heads produced by a circular laser scanner. As the applications are so different from each other, different (novel) similarity measures had to be devised for each application. 2D matching is based on statistical information obtained from a new filter, the F-test filter [19]. The 3D similarity measure was constructed from standard 3D surface similarity measures and a new similarity measure, the relative difference in angle.

In the following sections, the correspondence calculation algorithm is described in detail and criteria which the match function must satisfy for good results are given. Then the similarity measures and other details for each application are described. Finally, we describe experiments that show that these new methods enable the aims of both applications to be achieved satisfactorily.

2 Calculating Robust Correspondences

In this section, we describe the CSM correspondence calculation algorithm in detail. We assume that a suitable similarity measure has been devised for the particular application. The criteria that the similarity measure should satisfy are given below; similarity measures for each of the two applications are described in Section 3.

There are three stages to the CSM algorithm:

1. Use an appropriate similarity measure to match a point (the matchpoint) of image A to points within a certain radius in image B. The points in image B, together with their similarity values to the matchpoint define a "matchmap." Note that if the two images are surfaces represented by a surface mesh, then a node of the surface mesh, representing image A, is matched to nodes of the surface mesh, representing image B.
2. Perform virtual movements of the matchpoint and, for each position of the matchpoint, calculate a "tentative corresponding point" on image B. Each tentative corresponding point is calculated by summing over the matchmap, while taking distances from the matchpoint to each point in the matchmap into account. Note that neither the corresponding point nor the tentative corresponding points are constrained to coincide with one of the points in the matchmap and, in general, will lie between these points.
3. Calculate a corresponding point and a measure of reliability by analyzing the distribution of the tentative corresponding points. If the tentative corresponding points are scattered along a line, the corresponding point is the closest point on the line to the matchpoint. Otherwise, the centroid of the scatter of points is used. The reliability is derived from the eigenvalues obtained by a principal axes analysis of the scatter of points.

Each of these stages will be described in detail below. In order for the correspondence calculations to work well, it is essential that the matchmap and, therefore, the similarity measure satisfies the following criteria:

- There must be good discrimination between the good and bad matches. For best results, the

similarity measure should be tailored to the type of images in the application. For example, in our 2D application, the different regions in the image can be characterized by the mean and standard deviation of their gray levels. So, a similarity measure based on these quantities is appropriate for these images (see Section 3.1.1).

- The values generated by the similarity measure should be normalized to the range [0,1].

2.1 Calculating Tentative Corresponding Points

If we match a matchpoint from image A to points within a search area in image B, we obtain a matchmap. A "tentative corresponding point" for the matchpoint can be calculated by integrating over the matchmap. One way of doing this is to attach springs from the matchpoint to each point within range in image B, where the spring constant is given by the similarity value. However, we would prefer "inverse springs" that pull more strongly for shorter distances. This can be achieved by the formula: $\mathbf{F}_i = K_i \underline{x}_i$, where $\mathbf{F}_i$ is the force exerted by spring i, $\underline{x}_i$ is the vector from the matchpoint to point i in the matchmap, and K_i is the spring constant given by

$$K_i = \frac{m_i}{1 + |\underline{x}_i|^b}, \qquad b \geq 2,$$

where m_i is the similarity value for point i and b is a constant. We have set b equal to 2.0, so that points that are far away can still influence the position of the corresponding point if they match well. Using this formula, the tentative corresponding point, $\underline{q}_i$, is given by

$$\underline{q}_i = \frac{\sum K_i \underline{x}_i}{\sum K_i}.$$

This formula is suitable for images that have been evenly subsampled. For 3D triangulated surfaces, in particular, the situation is different in that the nodes in image B may not be evenly distributed over the surface. This can cause an area where the nodes of the surface mesh are more densely packed to pull more strongly than an area where the nodes are less densely packed, irrespective of how well they match the matchpoint. To overcome this, we multiply K_i by the additional term

$$\frac{meandist_i}{n_i},$$

where $meandist_i$ is the mean distance from node i to all the nodes connected to it, and n_i is the number of nodes connected to node i. This results in the formula:

$$K_i = \frac{m_i}{1 + |\underline{x}_i|^b} \times \frac{meandist_i}{n_i}, \qquad b \geq 2.$$

2.2 Generating a Measure of Reliability and a Corresponding Point

Using the formulae given above, a single "tentative corresponding point" can be generated for each matchpoint in image A. If we (virtually) perturb each matchpoint and sum over the matchmap with the matchpoint in this new position, we will in general get a different tentative corresponding point. If this procedure is repeated for fixed displacements of the matchpoint, a scatter of tentative corresponding points will be generated. The distribution of this scatter of points provides information about the robustness of each match to image B.

In the 2D case, three scenarios are possible: the tentative corresponding points will either be clustered along a line, clustered near a point, or widely scattered in the image. In the general 3D case, there is the additional scenario: the tentative corresponding points may be scattered on a plane or a more general type of surface. In order to differentiate between these cases and to calculate a measure of reliability, we calculate the eigenvalues and eigenvectors from the moment of area matrix [20] obtained from the scatter of points.

In the 2D case, the first eigenvector gives the direction of the possible corresponding line and the second eigenvalue gives the sum of squared distances from this line. This eigenvalue is small when the corresponding points are scattered closely along the line and when they are clustered near a point. It is large only when the tentative corresponding points are widely scattered. Conversely, the first eigenvalue is large both when the tentative corresponding points are scattered along a line and when they are widely scattered. It is small only when the points are clustered around a point.

In the 3D case, there are three eigenvalues and eigenvectors. If all the eigenvalues are small, the tentative corresponding points are clustered near a point. If the second and third eigenvalues are small, the tentative corresponding points are scattered along a line; and if the third eigenvalue is small, they are scattered in a plane. When all the eigenvalues are large, the tentative corresponding points are widely scattered.

This information enables us to calculate a corresponding point. In the 2D case, the cases where the tentative corresponding points are scattered along a line or near a point can be differentiated by the ratio of the first and second eigenvalues. If this ratio is greater than a threshold, the corresponding point is set to the closest point on the line to the matchpoint; otherwise, it is given by the centroid of the scatter of points. Note that if the matchpoint corresponds to two lines, the tentative corresponding points will be widely scattered and the matchpoint will correspond to a point. Note, however, the reliability value will be low and this will allow the matchpoint to move freely during a registration calculation. If the search distance were to be decreased so that only one line is included in the matchmap, or if the distance penalty were to be increased, the matchpoint would correspond to the closest line.

In the 3D case, a matchpoint could correspond to a volume, a surface, a line (such as a blood vessel), or a point (such as a distinctive point within the image). If the images are surfaces, we may assume that the tentative corresponding points will be scattered on the surface of image B and that the third eigenvalue will be small. This is a valid assumption when the surface is not too curved and can be achieved, for smooth surfaces, by increasing the scale of the image or reducing the magnitude of the (virtual) movement of the matchpoint. Therefore, in this case, a distinction between a line and a point correspondence can be obtained by comparing the first two eigenvalues as described above for the 2D case. In the case of 3D volumes, ratios between the different eigenvalues can be used to distinguish

between the various cases. For example, if both the ratios between the first and second and between the first and third eigenvalues are above the threshold, the matchpoint will correspond to a line. If only the ratio between the first and third eigenvalues is above the threshold, the correspondence will be a plane.

In both the 2D and 3D cases, the difference between a corresponding point or region is given by the reliability value. In the 2D and 3D surface case, the reliability of a corresponding point is a function of the second eigenvalue since this value gives a measure of how closely the tentative corresponding points are clustered either along a line or near a point. In the 3D volume case, the reliability should be a function of the second eigenvalue when the matchpoint corresponds to a line, but of the third eigenvalue in all other cases since it will be small both when the correspondence is point and when it is a plane

In order to convert the appropriate eigenvalue into a more understandable and applicable indicator of the reliability of the correspondence, it was transformed to the range [0,1]. It was found, empirically, that a gaussian function provides a suitable conversion formula:

$$reliability = \exp\left\{-\frac{eigenvalue^2}{2\gamma^2}\right\}, \gamma = \frac{td^2}{a},$$

where t is the number of tentative corresponding points, d is the maximum displacement undergone by the matchpoint, and a is a variable which enables the reliability values to be tuned to the application.

The values of these parameters were chosen to fit the application. For example, if the correspondences will be used to calculate a rigid transformation, only distinctive nodes should be given a value near 1.0; the reliability of nodes in featureless regions or regions of texture should be near zero; otherwise, they will slow down the registration process. In contrast, if the images are already registered and, if the application is to warp one image onto the other, the reliability value should be greater than zero, even for unreliable correspondences. This will ensure that regions with uniform properties are included in the warping operation.

3 Applications

3.1 2D Application: Warping of 2D Serial Histological Sections

Serial sectioning of biological material is often the only way to reveal the 3D structure of an object. Serial histological sections of an object are produced by using a microtome to slice the object into thin sections (of the order of micrometers thick). Since the object is physically cut, the alignment between sections is lost and, in addition, the sections are deformed. Note also that the images may display large changes between sections, as structures appear, disappear, and change shape, or branch. The warping algorithm designed to correct small rigid registration errors and to smooth out the deformations between sections has been described in detail in [21] and [19]. Warping was achieved by modeling each section as a thin elastic plate and attaching springs to correspondences between sections, where the strength of each spring depends on the reliability of the appropriate correspondence. The finite element method was used to calculate the equilibrium position. In this paper, we will concentrate on a description of the correspondence calculation algorithm. Details of the registration calculations are given in [21] and [19].

3.1.1 2D Matching

An example of the type of images used for our 2D application is shown in Fig. 1. In these images, different regions are characterized by the mean and standard deviation of their gray levels. Therefore, for this application, matching is based on the following features for each pixel:

1. The direction of an edge segment centered on the pixel separating two regions.
2. The strength of the edge. This is 1.0 when there is no edge at the pixel and greater than 1.0 when there is an edge.
3. The means and standard deviations of the two regions separated by the edge segment.

This information is obtained for each pixel by passing a series of masks over the image. Each mask represents an edge direction and two regions, one on each side of the edge. The F-test is used to compare the standard deviations of the pixels in the two regions in the mask and the mask giving the highest response to this test gives the edge direction and the two regions. This constitutes the *F-test filter*.

Points are matched from two images, A and B, by calculating a measure of similarity between pairs of corresponding regions. The corresponding regions of the two masks are given by the orientation of the line. In the following, the regions are denoted U_{A_1}, U_{B_1} and U_{A_2}, U_{B_2}. The means of the gray values of the two regions can be compared with the T-test and the variances can be compared with the F-test. By multiplying these quantities together, we obtain a function that varies depending on the similarities between the means and variances of two regions. In order to take account of both pairs of regions, the similarity measure used to compare the means and variances is

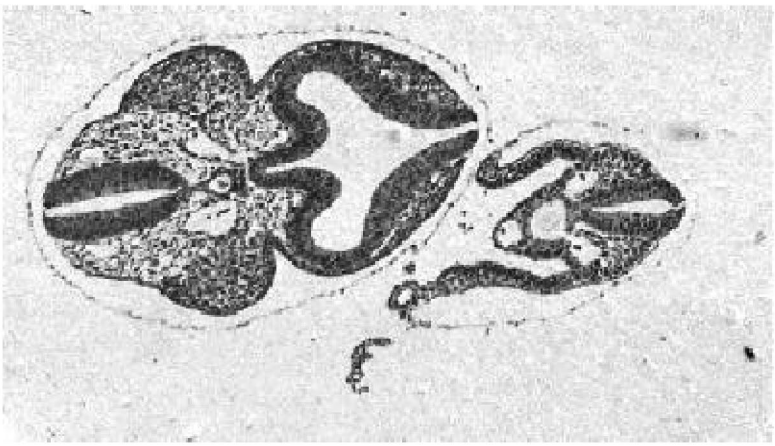
$$f = [1 + F_1^{\alpha} T_1^{\beta} + F_2^{\alpha} T_2^{\beta}]^{-1},$$

where F_i and T_i refer to the F- and T-tests applied to attributes from U_{A_i} and U_{B_i}. The values of this function are in the range [0,1], where 1 signifies that both pairs of regions have similar means and variances, and 0 signifies that these regions do not have similar means and variances. The parameters α and β were set to 1.0 and 2.0, respectively, by an experiment designed to maximize the discrimination between good and bad matches.

The similarity measure must also include the difference in the directions of the edge segments centered on the two image points. This is done by means of the Gaussian

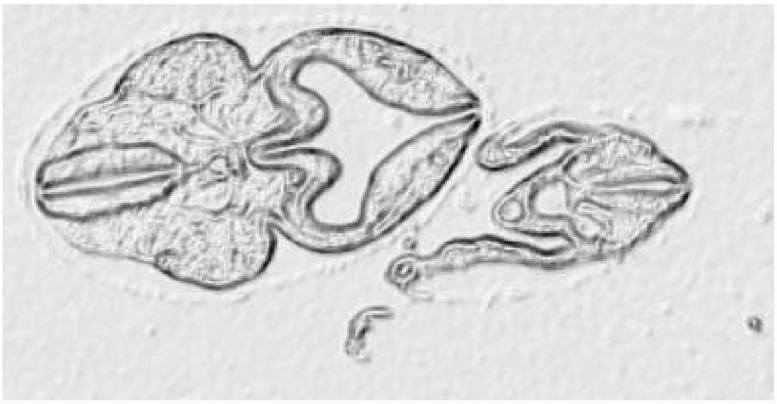
$$\exp\left(-\left[\frac{|\theta_A - \theta_B|^2}{2\omega^2}\right]\right).$$

The value of the parameter ω should be chosen so that a small angle penalty is calculated for differences in angles up to the resolution of the line direction calculation (a value of 30 degrees was found to be suitable for 5 x 5 masks). The final similarity measure is:



(a)

(b)



(c)

Fig. 1. The results of applying the F-test filter to image (a), which shows a section from the nine day mouse embryo series. (b) Shows the results of a 5 x 5 mask and (c) shows the results of a 7 x 7 mask.

$$M_{2D} = f \times \exp\left(-\left[\frac{|\theta_A - \theta_B|^2}{2\omega^2}\right]\right).$$

Matchpoints for this application were calculated by applying the F-filter to the image, thresholding the result to obtain the edges and thinning these edges using nonmaximal suppression. Finally, points from the resulting edges were chosen, starting with those with the highest values and making sure that all points were separated by a minimum distance.

3.1.2 2D Correspondence Calculation

The correspondence calculation was performed as given in Section 2. For this application, the initial correspondence for

all the matchpoints was the closest point on the corresponding line. Correspondences were allowed to slide along the lines as the warp calculation progressed (the threshold for the ratio of the second and first eigenvalues was set to 0.5). A value of 20 was found to be suitable for the parameter a. Matchpoints were perturbed by 1 and 2 pixels in the x and y directions to give 25 tentative corresponding points.

3.1.3 2D Registration

Registration in the 2D case proceeded using the finite element method. The image to be warped was triangulated in such a way as to ensure a higher resolution mesh over the object and a lower resolution over the background [19], [21]. The matchpoints (described above) were matched to adjacent images and springs were attached to their correspondences, where the spring strength depended on the reliability of the correspondence. Then, the images to be warped were modeled as thin elastic plates and the finite element method was used to calculate the equilibrium position.

This procedure was calculated iteratively for each section in the stack of serial sections where each section was warped to both its adjacent sections at the same time. More details of this registration procedure are given in [21] and [19].

3.2 3D Application: Surface Registration

The 3D surface registration application is part of a project to model the soft tissue on the face with the aim of predicting the outcome of facial surgery. For this application, accurate rigid registration and accurate correspondence calculation algorithms are essential.

3.2.1 3D Matching

The match function used for our 3D application is based on the normals and curvatures of the surfaces. Three quantities are compared:

- The curvedness [22], defined as:

$$C = \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}},$$

  where κ_1 and κ_2 and are the principal curvatures;
- The shape [22], defined by:

$$S = -\frac{2}{\pi} \arctan\left(\frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2}\right).$$

  The shape is in the range [-1,1], where -1 denotes a spherical concave surface and +1 a spherical convex surface.
- The relative angle, Θ, between the normal and the vector from the main axis of symmetry of the object to a surface point. This is a new measure that helps to localize points in regions of uniform curvature. At first, we used the vector from the centroid to the point, but as the face resembles a sphere quite well, this produces small relative angles. We have found that using a cylinder as a template works much better, giving better localization to the points.

We define a separate measure for each of these quantities and, then multiply these measures together to give the matchvalue $M_{3D} = lhg$, where l, h, and g are defined by:

- Curvedness:

$$l = min\left(\frac{1 + 1,000 \times C_1}{1 + 1,000 \times C_2}, \frac{1 + 1,000 \times C_2}{1 + 1,000 \times C_1}\right).$$

  The factor 1,000 is used to stop the values being swamped by 1.0, which has been included to prevent division by zero. Note that although the curvedness varies with scale, this match function is independent of scale.
- Shape:

$$h = \exp\{-\frac{|S_1 - S_2|^2}{2\sigma_s^2}\}.$$

- Relative angle:

$$g = \exp\{-\frac{|\Theta_1 - \Theta_2|^2}{2\sigma_\Theta^2}\}.$$

  The parameters σ_S and σ_Θ were set to $1/3$ and $\pi/18$, respectively to allow for small differences in the surfaces and small errors in the calculation of S and Θ.

3.2.2 3D Correspondence Calculation

The correspondence calculation was performed as given in Section 2. The threshold for distinguishing between corresponding points and lines was 0.5 to exclude only slightly elongated distributions of tentative corresponding points. A value of 16 was found to be suitable for the parameter a; and the matchpoints were moved 12mm in the x, y, and z directions to give 27 tentative corresponding points. The matchpoints were simply the nodes that make up the surface triangulation. In this application, we are matching smooth surfaces and because of this, points will tend to match well over a relatively large area, even for distinctive points. This tends to bring the reliabilities of the good and bad matches closer together. To improve the discrimination between reliable and unreliable correspondences and to reduce the effects of using different sized search areas for the correspondences, the function

$$h(x) = \begin{cases} 2x^2 & 0 \le x \le 1/2 \\ 1 - 2(1-x)^2 & 1/2 \le x \le 1 \end{cases}$$

was applied to the reliability value. Note that this function leaves the value 0.5 unchanged, but increases those greater than 0.5 towards 1.0, and decreases those less than 0.5 towards 0.0.

3.2.3 3D Surface Registration

A modification of the transformation calculation in the Iterative Closest Point (ICP) algorithm [23] is used to calculate the transformation. This calculation minimizes the squared distance between paired points, $\underline{p}_i$ in image A and $\underline{q}_i$ in image B

$$\min \sum |\underline{q}_i - (\mathrm{R}\underline{p}_i + \mathrm{T})|^2,$$

where R is the rotation matrix, and T the translation vector. This is done by calculating the covariance matrix [24], [25]:

$$\sum \underline{p}_i \underline{q}_i^T.$$

Now, we not only have correspondences, but also a measure of reliability for these correspondences that we would like to include in the calculation. Therefore, we solve the weighted least squares problem:

$$\min \sum w_i^2 |\underline{q}_i - (\mathrm{R}\underline{p}_i + \mathrm{T})|^2.$$

Which gives the covariance matrix:

$$\sum w_i^2 \underline{p}_i \underline{q}_i^T.$$

4 Experiments

Two sets of experiments were performed for each application. First, the matching algorithms were tested, then the performance of the registration algorithms was evaluated. Since the full validation of the 2D warping calculations is described in [21] and [19], only the main results are given here.

4.1 2D Application: Warping Serial Histological Sections

Experiments to test the 2D warping algorithm were performed on two series of sections: one series of a 7.75 day old mouse embryo and the other of a nine day old mouse embryo. All sections were digitized using a Zeiss Axioplan microscope fitted with a Xillix 1400 (Optimum Vision Ltd. Petersfield, Hampshire, UK) digital camera. The 7.75 day mouse embryo was embedded in plastic (araldite), sectioned at 2 μm, digitized at a resolution of 0.68 μm and subsampled by a factor of 3. The nine day mouse embryo was embedded in paraffin wax, sectioned at 7 μm and stained with haematoxylin and eosin. The sections were digitized at a resolution of 1.36 μm, corrected for shading and also subsampled by a factor of 3. Both series of sections were (roughly) manually registered during digitization. The matching and correspondence experiment was performed on sections from the nine day series. All experiments used the subsampled images as their starting points.

4.1.1 2D Matching, Correspondence, and Warping

The performance of the 2D matching and correspondence algorithm was tested by calculating the corresponding line for nine matchpoints, chosen to show the performance of the CSM algorithm for different kinds of regions. The nodes were matched to an adjacent image in the nine day series. The nine points and the regions they were matched to in the adjacent image are shown in Fig. 2. The chosen nodes include both external and internal boundaries of the object and areas of uniform texture. Node 2 is on an external boundary, but was chosen because its corresponding point lies in an area containing an artifact in the adjacent image.

The full warping algorithm was tested by using it to reconstruct the full 7.75 and nine day stack of sections.

4.2 2D Application: Surface Registration

In order to test our matching and registration algorithm, a Cyberware laser scanner 3030HRC (Cyberware Inc. Monterey, California) was used to scan a polystyrene bust. The bust was scanned three times: once as it is (scan "normal1"); then with a lump (made of some sticky and pliable material) attached to its face (scan "lump"); and then with the lump removed very carefully in order not to disturb the bust's position (scan "normal2"). Due to the way in which the data were obtained, we would expect the scans lump and normal2 to be better registered than scans normal1 and normal2. The reason for scanning the bust with an attached lump was to test the registration algorithm on data with significant differences.

The 3D similarity measure depends on the accuracy of the calculation of the normals and curvatures for each node on the face. According to scale-space theory [26], these quantities must be calculated using differential operators at a sufficiently large scale. That is, smoothing and differentiation must be accomplished at the same time. Trucco and Fisher [27] have shown, however, that reliable values for these quantities can be calculated when the smoothing and the curvature and normal calculations are separated. We have chosen to take the latter approach so that we could use existing software and the following two steps were carried out.

Using two programs for processing Cyberware data, "cysurf" and "decimate" (Cyberware Inc.), we first removed outliers and smoothed the data using cysurf with default parameters. Then decimate was used to convert the Cyberware data format to a triangular surface mesh and to reduce the number of nodes. In addition, the head data was cut to remove the neck and chin and the top of the head. This was to remove large areas (such as under the chin) where the range scanner could obtain no data. Note however, that regions of the lump with no data remained.

4.2.1 3D Matching

The performance of the surface match function was tested by matching various nodes on the face of scan normal1 to scan normal2. Both distinctive nodes (e.g., nodes on the nose, eyes, and mouth) and nodes in featureless areas (cheek and forehead) were chosen. Both sets of data were decimated to 1 percent of the original number of nodes to reduce the density of the meshes for display purposes.

4.2.2 3D Surface Registration

The performance of our new registration algorithm, ICP with the CSM correspondence finding algorithm (ICP-CSM), was tested against two others, which are based on the ICP registration algorithm, but have different correspondence calculation methods:

- **ICP**: The ICP algorithm with closest point matching
- **ICP-CN**: This method was included because it is similar to those used in the literature([3], [10], [11]). It consists of the ICP algorithm where the node correspondences were obtained by matching points according to their curvatures and relative angles as well as distance. All match quantities were included in the featurevector:

$$(\mathrm{x, y, z\ coordinates}; 1,000 \times k_1, \\ 1,000 \times k_2, \mathrm{relative\ angle\ in\ degrees}).$$

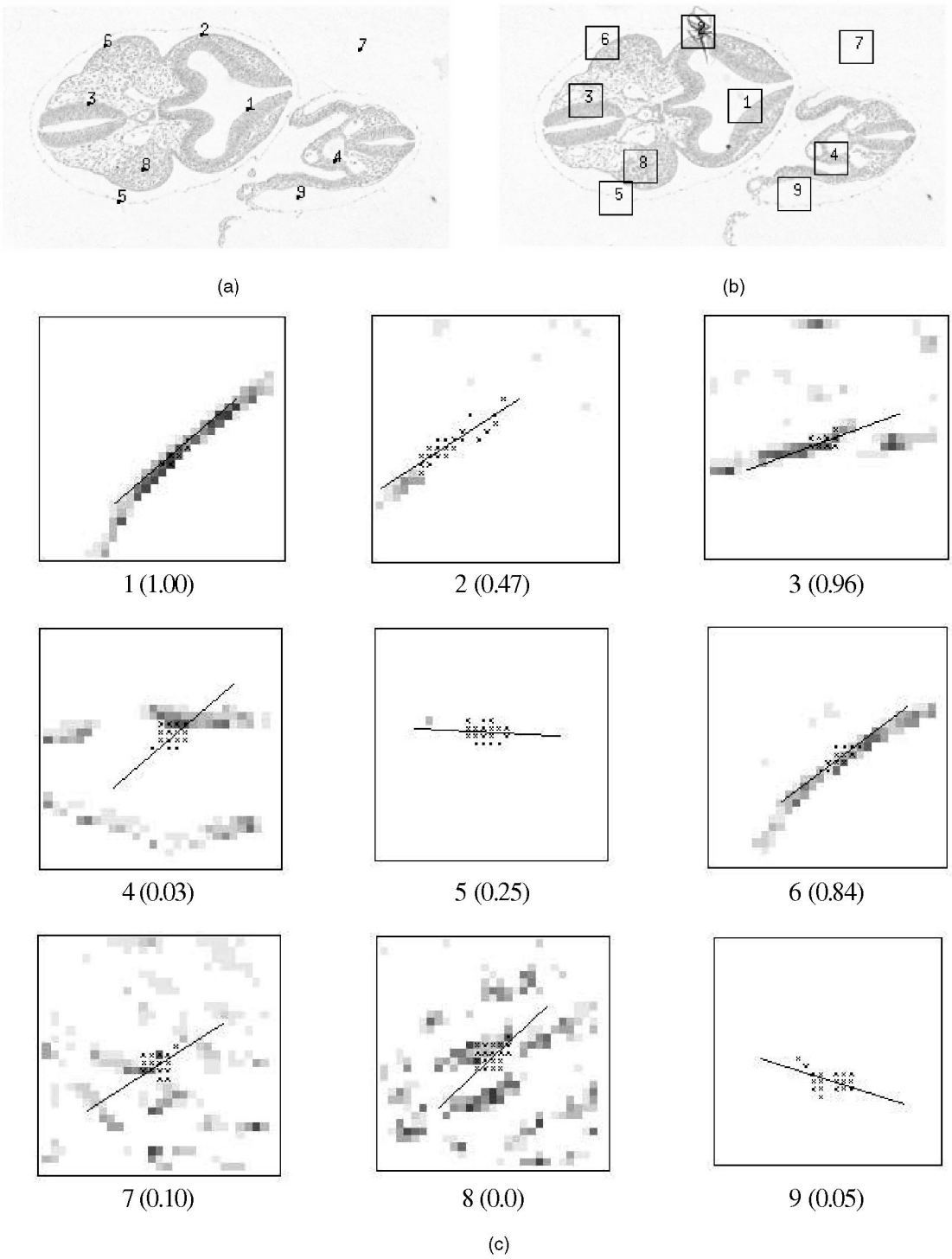


Fig. 2. The matchpoints shown in (a) were matched to the areas shown in (b) and the corresponding lines for these matchpoints are shown in (c). The black crosses denote the tentative corresponding points; the black line denotes the corresponding line. The reliability values are given in brackets.

The corresponding point was taken to be the point with the shortest distance in the featurevector. The curvatures are multiplied by 1,000 to give them more weight.

All these algorithms were tested by first rotating and translating one image of the bust (the match image) while keeping another image fixed. Two sets of tests were performed.

Test1: The transformations were applied to normal1, which was then registered to normal2.

Test2: The transformations were applied to normal2, which was then registered to lump.

For these experiments, the data was decimated to 15 percent of the original number of nodes. However, the ICP-CSM algorithm only used 5 percent of the nodes for matching. These nodes were found by decimating the data

to 5 percent, but their normals and curvatures were calculated from the 15 percent mesh. The transformations applied to the match images were as follows (the code that is used in later descriptions for each transformation is given in brackets):

Rotations.

1. Rotations of 5, 10, and 15 degrees about the z axis, the superior-inferior axis for the face data (Rz5, Rz10, Rz15).
2. A rotation of 10 degrees about the x axis (Rx10).
3. A rotation made up of rotations of 10 degrees about the x and y axes, applied in that order (Rxy10).
4. A rotation made up of rotations of 15 degrees about the x, y, and z axes, applied in that order (Rxyz15).

Translations.

1. Translations of 5.5, 11, and 16.5 mm in the x direction (Tx5, Tx11, Tx16).
2. Translations of 16.5 and 27.5 mm in both the x and y directions (Txy16, Txy27).

A benchmark image that could be used to determine the consistency of each of the registration methods was obtained by using each method to register the relevant pairs of images with no initial applied transformation. In each case, the benchmark image is the resulting transformed match image, where the match image is the image to which the test transformations were subsequently applied.

4.2.3 Evaluation of Results for 3D Surface Registration

The results of the above experiments were evaluated according to three criteria: consistency, accuracy, and speed of convergence. Consistency measures how closely each registration algorithm returns the match image to its benchmark position. We calculate this measure by comparing the transformed and registered match image with the benchmark image. The two meshes are compared by calculating the mean and standard deviation of the distances between the nodes of the transformed and registered match image and the benchmark image.

To compare the accuracy of the registration algorithms, the volume of the set difference of the solids contained within each pair of surface meshes was calculated. The solids are calculated by calculating a series of solid slices for each mesh, making sure that the slices are taken at the same z coordinate. If there is no closed loop, the loop is closed by a straight line. Having obtained the solids, the set difference is calculated slice by slice. Note that discretization errors are attenuated for each slice by multiplying the surface node coordinate values by 100. This enables a resolution of less than 0.08 cubic cm to be achieved for the volumes calculated here.

The final measure for comparing the different registration algorithms is the speed of convergence. This is simply the number of iterations needed to achieve convergence, which is defined to be when the mean square displacement of the mesh nodes during an iteration is less than 0.05 mm.

5 Results

5.1 2D Application: Warping Serial Histological Sections

5.1.1 2D Correspondence Calculation

The results of matching the points shown in Fig. 2a to an adjacent image are shown in Fig. 2c. As required, the good (darker gray levels) and bad matches are well differentiated. Points on the outside boundaries (points 1 and 6) correspond to the appropriate line in the adjacent image and this is also true for point 2 despite the very limited information resulting from the artifact in the adjacent image. The line for node 3, which is on an internal boundary, is not perfect because of noise to one side of the boundary, but it is still good. In the cases where there should be a verdict of "no match" (nodes 7, 8, and 9), the line direction is arbitrary, but the tentative corresponding points are widely scattered showing that the correspondence is unlikely to be correct (and this is reflected in the reliability value). This is also the case where there is a choice of corresponding line (point 4), but in this case, we expect that as point 4 gets closer to one of the lines, the corresponding line would match the appropriate line.

5.1.2 2D Warping

The results of reconstructing the 7.75 and nine day mouse embryos are shown in Fig. 3. Following warping, the structures have become better differentiated and the boundaries have become smooth. Warping has significantly improved the visual quality of the arbitrary section through the stack and, indeed, after warping the quality of an arbitrary section is similar to that of a section taken in the original cutting direction.

5.2 3D Application: Surface Registration

5.2.1 3D Matching

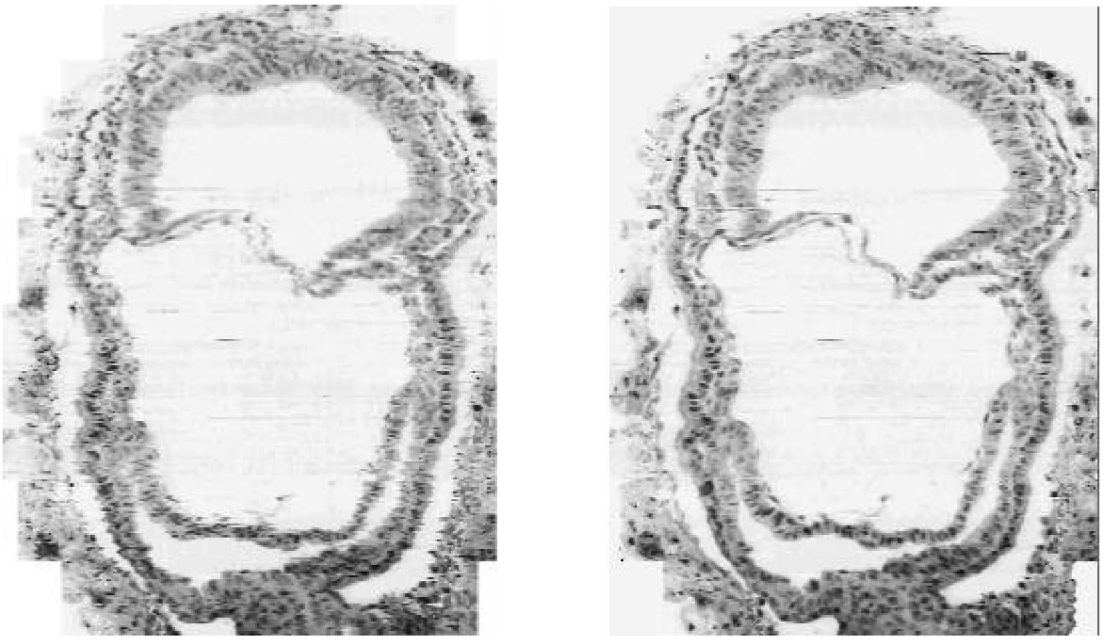
The results of matching nodes from scan normal1 to scan normal2 are shown in Fig. 4. We see that there is good differentiation between good and bad matches. The effects of the relative angle component of the match function can be seen in the results for node 8, which is located in the middle of the forehead. Without this component, the wide line of good matches extends to cover most of the forehead.

5.2.2 3D Registration

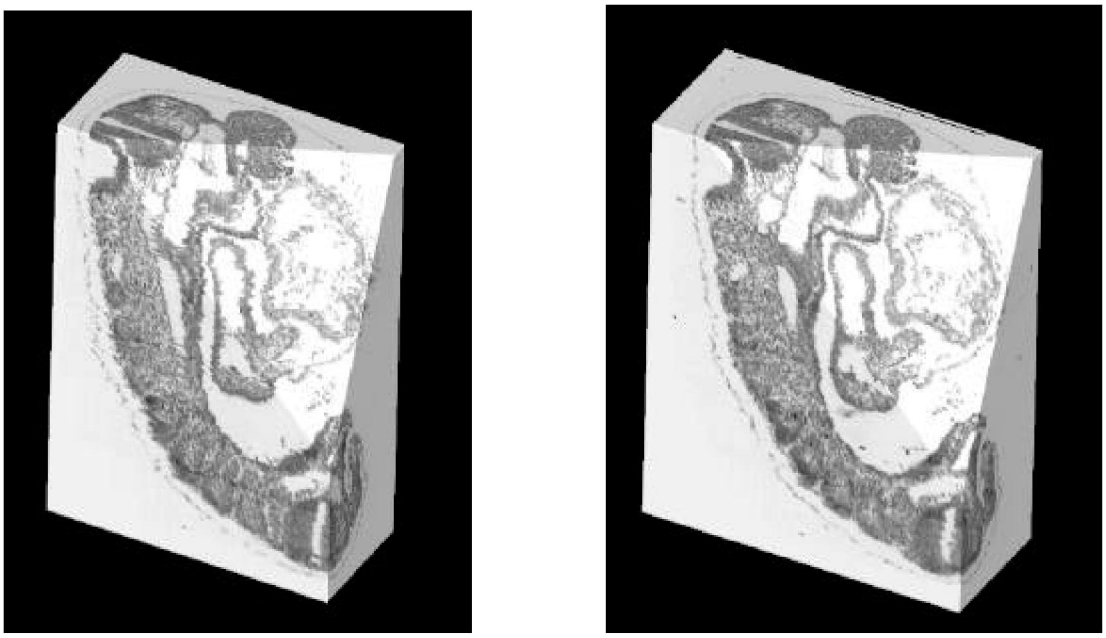
Fig. 5 shows the mean distances between the registered image calculated after the various transformations were applied and the benchmark image for tests 1 and 2. In both cases, the ICP-CSM algorithm gives the most consistent results for all 11 transformations tested. This algorithm also converges in the fewest number of iterations (Fig. 6).

The results for the ICP method are sparse because this method failed to register the surfaces when one surface was rotated about the z axis. This is because, due to the nature of the surfaces, rotating about the z axis does not move one image away from the other, but only superimposes the two sets of data at a different position. Since the ICP algorithm only uses distance to locate corresponding points, a local minimum can be found using these new neighboring points.

Fig. 7 shows the set differences in volume (in cm^3) for the registered surfaces. In all cases, the results are shown for



(a)



(b)

Fig. 3. The results of applying the 2D warping algorithm to (a) images of the 7.75 and (b) nine day mouse embryos before (left) and after (right) warping. In (b), the top surface of each block shows a section taken in the original cutting direction; the other surfaces show orthogonal sections and a section cut in an arbitrary direction.

registration with no applied transformation. As expected, scans lump and normal2 had better initial alignment than scans normal1 and normal2. The mean set difference in volume over all 11 transformations was smaller for the ICP-CSM method than for ICP-CN or ICP for both Test1 and Test2. For example, for ICP-CSM vs ICP-CN, the mean set difference between methods was smaller by 2.6 cm^3 for Test1 and by 2.7 cm^3 for Test2. This difference between methods was significant ($p<0.05$, paired t-test) for all comparisons.

A decrease in the set difference in volume after registration for Test2, was seen only for the ICP-CSM algorithm. Since, in all cases, the set difference in volume is greater than the volume of the lump (7.8 ± 0.1 cm^3), this decrease cannot be explained by an increase in the misregistration. Therefore, we conclude that since the set

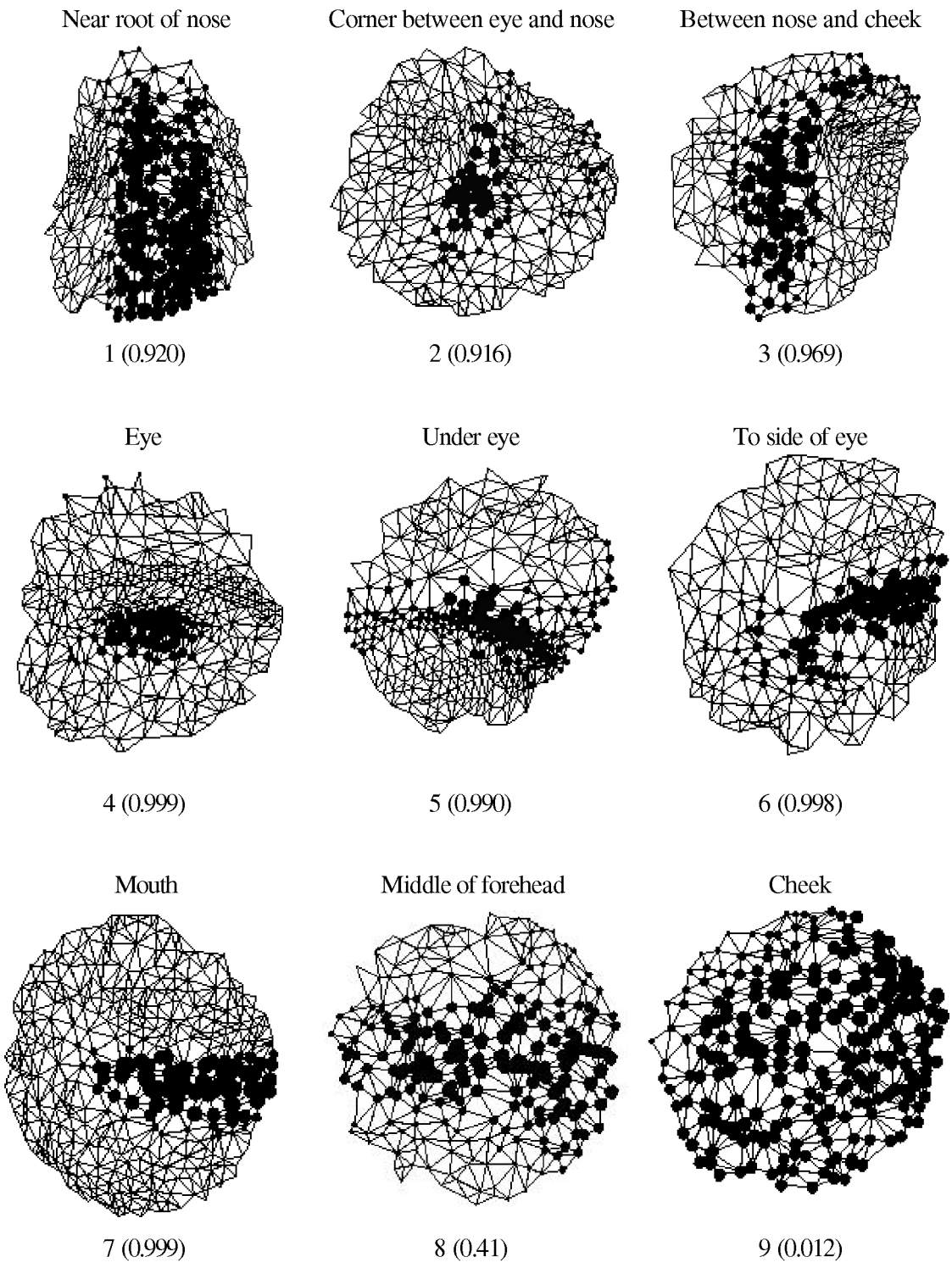
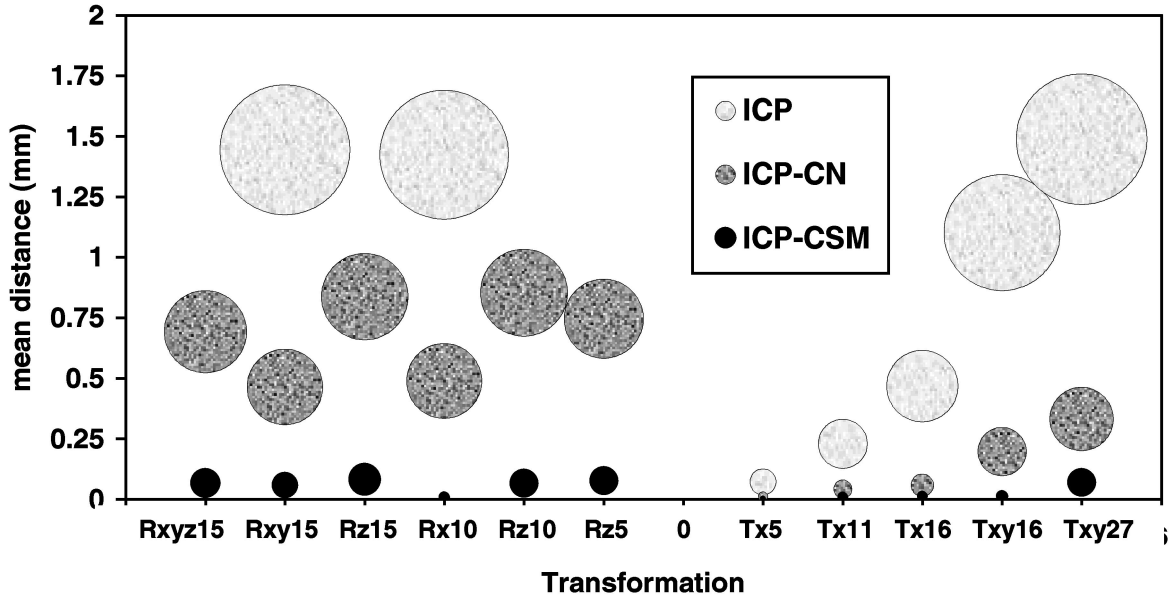


Fig. 4. The results of matching some points from scan normal1 to scan normal2. Good matches are given by large discs; bad matches by smaller (or no) discs. The reliability value is given in brackets.

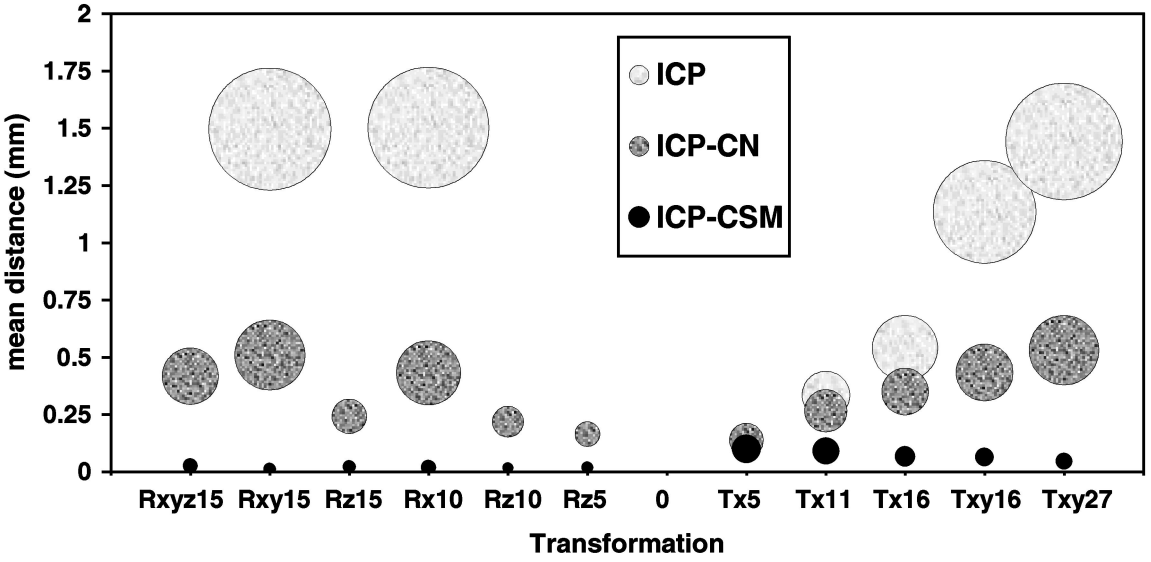
difference in volume was smallest for the ICP-CSM algorithm, this algorithm gives the most accurate results.

6 Discussion

The experiments on surfaces that compare the three methods of calculating correspondences in the context of ICP rigid registration enable us to draw some conclusions about the behavior of the CSM correspondence calculation algorithm. Figs. 5 and 6 suggest that the calculation of robust correspondences using this algorithm reduces the number of local minima in the registration space. Fig. 5 shows that registration by this method is less likely to get stuck in a local minimum and Fig. 6 shows that the gradient descent is steeper since fewer iterations are required for convergence.



(a)



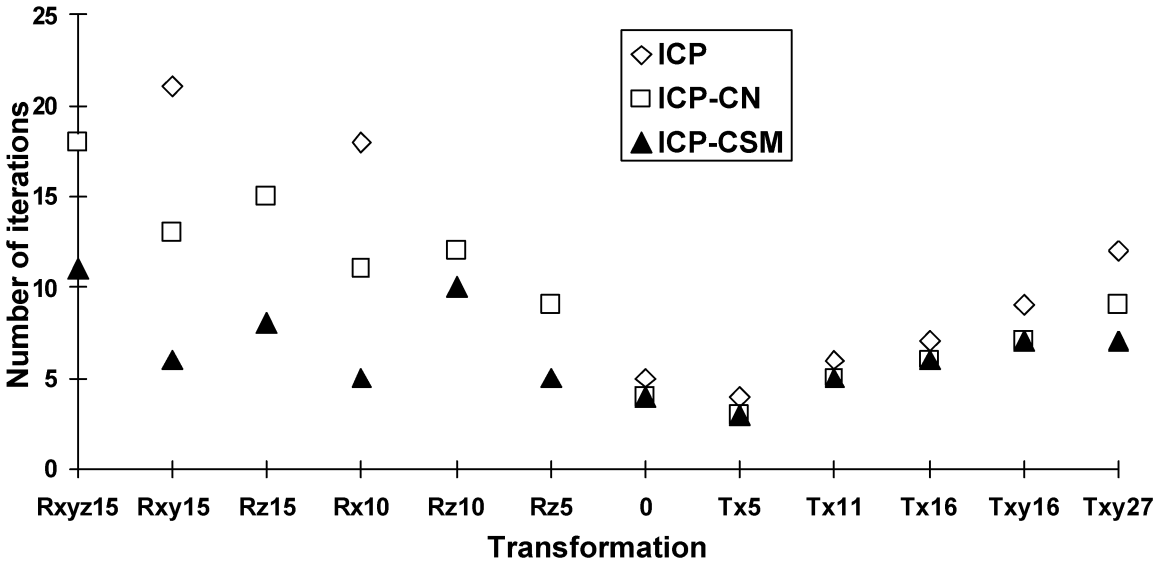
(b)

Fig. 5. The mean distance between the registered image calculated after the various transformations were applied and the benchmark image for (a) Test1 (scans normal1 and normal2) and (b) Test2 (scans normal2 and lump). The standard deviation is given by the area of the shaded disks.

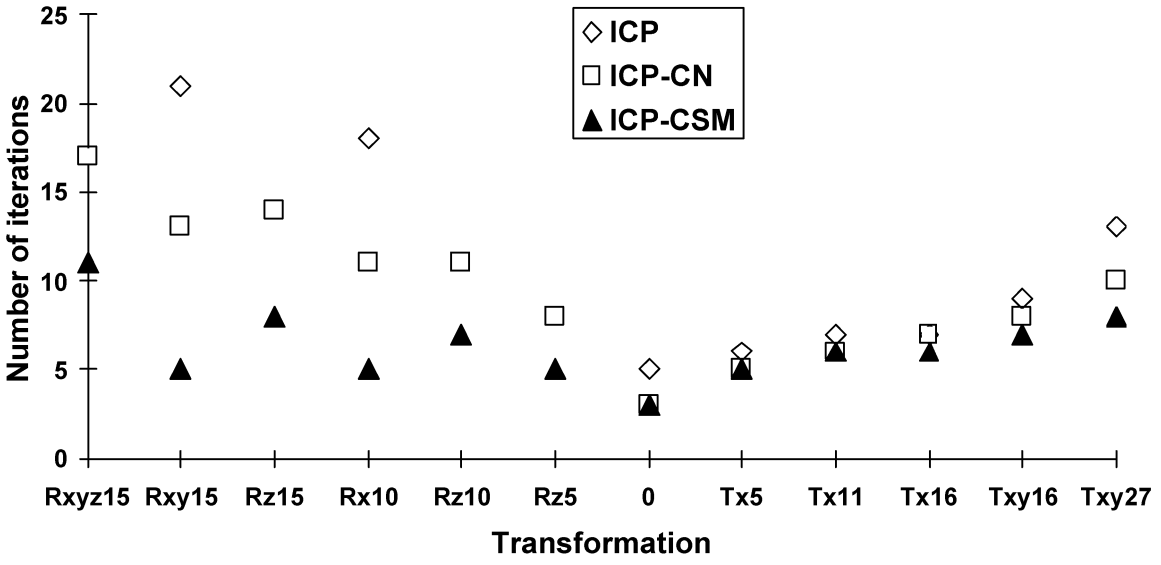
We see in Fig. 5 that the ICP method fails for quite small rotations, whereas the other two methods are more robust in this respect. The failure modes of the ICP-CSM method have not been fully tested. We expect that it will show a similar behavior to ICP in that the images have to be roughly registered prior to using this technique, but there is some evidence that ICP-CSM is robust as it has successfully registered two nonoverlapping images where the difference in orientation was 90 degrees. However, it fails if there are no distinctive features on the surface. This happens, for example, when scans of the forehead are registered without including parts of the eye sockets or nose.

ICP and ICP-CN perform particularly badly for rotations about the z axis (from top to bottom of the head). This is because correspondences in these algorithms are weighted towards closer nodes. When the bust is rotated in this way, most nodes correspond to nodes that are very close. Although the matching in the ICP-CN method may cause distinctive points to correspond to the appropriate points in the other image, the contribution from these vectors is swamped because all correspondences have equal weight

The ICP-CSM algorithm is the only algorithm to improve on the initial alignment in Test2. This is because CSM calculates correspondences that can be any point on the surface. In the other two methods, the corresponding point



(a)



(b)

Fig. 6. The number of iterations required to register (a) scans normal1 and normal2 (Test1) and (b) scans normal2 and lump (Test2) for each applied transformation.

must be a mesh node. The CSM algorithm is comparatively slow due to its computational complexity, but when it is incorporated into a registration algorithm, the resulting registration is more accurate than that for other correspondence calculation methods. Therefore, it is most suited to applications where accuracy is important.

We have demonstrated the use of the CSM algorithm both for 2D images and 3D surfaces. The fact that the method works well for the 2D warping application suggests that it will also work well for 3D warping. We also expect that the CSM correspondence calculation algorithm can be applied to many applications, including 3D volumes.

7 Conclusion

From the above experiments, we conclude that our new correspondence calculation algorithm, CSM (both 2D and 3D versions) enables accurate alignment of both 2D and 3D images. In 2D, we have successfully applied the CSM algorithm to the problem of warping serial histological

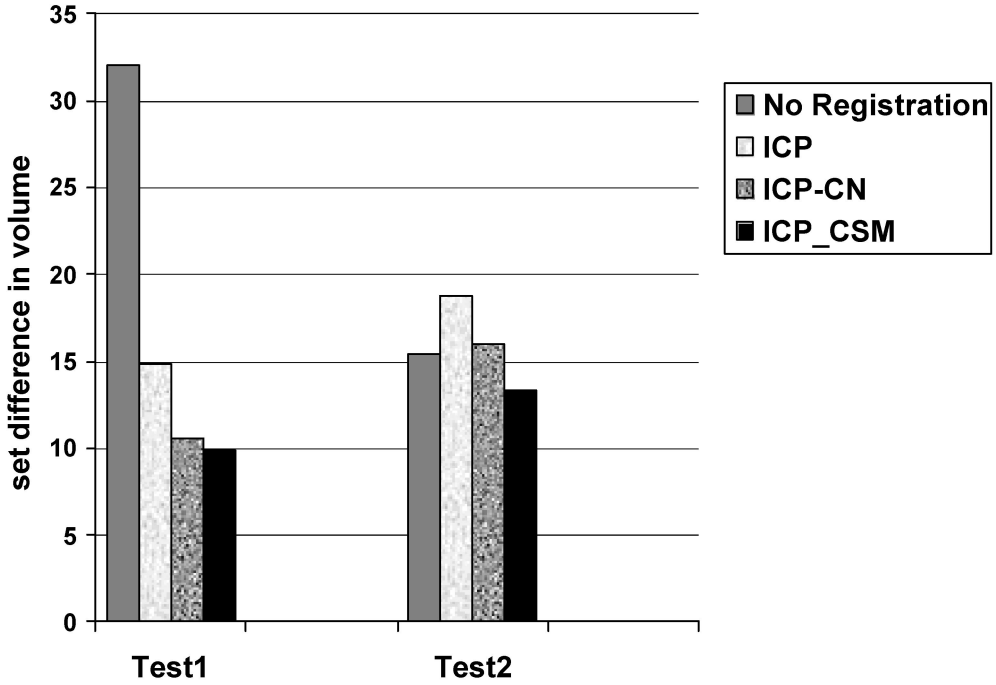


Fig. 7. The set difference in volume (in cm^3) following registration of the surfaces for zero applied transformation.

sections in order to produce a smooth reconstruction of the original object. With the CSM correspondence algorithm, sections cut perpendicular to the original cutting direction appear of similar quality to the original sections.

In 3D, we have shown that the CSM algorithm enables the ICP registration method to calculate accurate and consistent registration transformations. Compared with the two other methods of calculating correspondences, the ICP-CSM algorithm gives the most consistent results, converges in the fewest number of iterations, and results in the smallest difference in volume between registered surfaces.

The CSM algorithm is heavily dependent on the underlying similarity measure and new similarity measures had to be devised for each application. A completely new filter, the F-test filter, was devised to aid matching in the 2D application; the 3D match function combined standard similarity measures along with a new similarity measure, the relative difference in angle, in a new way. We expect that our novel match functions can be applied to other applications, but for some applications (such as 3D volumes), the developer will have to design a suitable similarity measure.

A major advantage of the CSM algorithm is the robust and consistent pairing of points without any prior segmentation. The use of tentative corresponding points means that points in featureless parts of the image may also contribute to the match, which is particularly important in elastic registration. In addition, this algorithm is particularly suited to matching subsampled surface meshes as it allows mesh nodes to correspond to any point on the surface, without interpolation.

Acknowledgments

Resources for the 2D part of this work were provided by the MRC Human Genetics Unit, funding was provided by the SERC, the University of Edinburgh, and (indirectly) by the Dutch Government. The 3D part of this research was performed in the Brite Euram project PISA (NR. BRPR CT97 0378). Partners in this project are: Materialise NV, Belgium (coordinator); Philips Medical Systems, Nederland BV, EasyVision Modules—Advanced Development, the Netherlands; DePuy International Ltd., United Kingdom; Ceka NV, Belgium; Katholieke Universiteit Leuven, Laboratory for Medical Imaging Research and Division of Biomechanics and Engineering Design, Belgium; University of Leeds, CoMIR, School of Medicine, United Kingdom; OBL, Paris, France. The Cyberware Scanner (Cyberware Inc.) is supported by the Wellcome Trust (047808/Z/96/Z). Part of the work was undertaken with support from the Engineering and Physical Sciences Research Council (GR/L55889).

References

[1] A. Brett and C. Taylor, "A Framework for Automated Landmark Generation for Automated 3D Statistical Model Construction," *Proc. Int'l Conf. Information Processing in Medical Imaging '99,* pp. 376-381, 1999.

[2] H. Lester and S.R. Arridge, "A Survey of Hierarchical Non-Linear Medical Image Registration," *Pattern Recognition,* vol. 32, pp. 129-149, 1999.

[3] Y. Zhang, "3D Image Analysis System and MegaKaryocyte Quantitation," *Cytometry,* vol. 12, pp. 308-315, 1991.

[4] L. Hibbard, T. Arnicar-Sulze, B. Dovey-Hartman, and R. Page, "Three-Dimensional Reconstruction of Median Eminence Microvascular Modules," *Computers in Biology and Medicine,* vol. 16, no. 6, pp. 411-421, 1986.

[5] A.E. Johnson and S. Kang, "Registration and Integration of Textured 3D Data," *Proc. Int'l Conf. Recent Advances in 3D Digital Imaging and Modeling,* pp. 234-241, 1997.

[6] R. Bajcsy and S. Kovacic, "Multiresolution Elastic Matching," *Computer Vision, Graphics and Image Processing,* vol. 46, pp. 1-21, 1989.
[7] J. Woodward, "Change Detection in MRI Brain Scan Data," MSc thesis, Dept. Artificial Intelligence, Univ. of Edinburgh, 1990.
[8] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality Image Registration by Maximization of Mutual Information," *IEEE Trans. Medical Imaging,* vol. 16, no. 2, pp. 187-198, 1997.
[9] A. Goshtasby, D. Turner, and L. Ackerman, "Matching of Tomographic Slices for Interpolation," *IEEE Trans. Medical Imaging,* vol. 11, no. 4, pp. 507-516, 1992.
[10] F. Verbeek, "Three-Dimensional Reconstruction of Biological Objects from Serial Sections Including Deformation Correction," PhD thesis, Technische Universiteit Delft, 1995.
[11] W. Ibrahim and F. Cohen, "Registering Coronal Histological 2D Sections of a Rat Brain with Coronal Sections of a 3D Brain Atlas Using Geometric Curve Invariants and B-Spline Representation," *IEEE Trans. Medical Imaging,* vol. 17, no. 6, pp. 957-966, 1999.
[12] G. Subsol, J.-P. Thirion, and N. Ayache, "A Scheme for Automatically Building Three-Dimensional Morphometric Anatomical Atlases: Application to a Skull Atlas," *Medical Image Analysis,* vol. 2, no. 1, pp. 37-60, 1998.
[13] J. Feldmar and N. Ayache, "Rigid, Affine, and Locally Affine Registration of Free-Form Surfaces," *Int'l J. Computer Vision,* vol. 18, no. 2, pp. 99-119, 1996.
[14] R. Pito, "A Registration Aid," *Proc. Int'l Conf. Recent Advances in 3D Digital Imaging and Modeling,* pp. 85-92, 1997.
[15] A.E. Johnson and M. Hebert, "Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 21, no. 5, pp. 433-449, May 1999.
[16] D. Zhang and M. Hebert, "Experimental Analysis of Harmonic Shape Images," *Proc. Second Int'l Conf. Recent Advances in 3D Digital Imaging and Modeling,* pp. 209-218, 1999.
[17] R. Bolles and R. Cain, "Recognising and Locating Partially Visible Objects: The Local-Feature-Focus Method," *Int'l J. Robotics Research* vol. 1, pp. 57-82, 1982.
[18] H. Chui, J. Rambo, J. Duncan, R. Schultz, and A. Rangarajan, "Registration of Cortical Anatomical Structures via Robust 3D Point Matching," *Proc. Int'l Conf. Information Processing in Medical Imaging '99,* pp. 168-181, 1999.
[19] E. Guest, "Automatic Reconstruction From Serial Sections," PhD thesis, Dept. Artificial Intelligence, Univ. of Edinburgh, 1994.
[20] D. Rees, *Mechanics of Solids and Structures.* McGraw-Hill, 1990.
[21] E. Guest and R. Baldock, "Correcting Distortions from Serial Sections," *Bioimaging,* vol. 3, pp. 154-167, 1996.
[22] J. Koenderink, *Solid Shape.* MIT Press, 1990.
[23] P. Besl and N. McKay, "A Method for Registration of 3D shapes," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 14, no. 2, pp. 239-256, Feb. 1992.
[24] K.S. Arun, T.S. Huang, and S.D. Blostein, "Least-Squares Fitting of Two 3D Point Sets," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 9, no. 5, pp. 698-700, May 1987.
[25] S. Umeyama, "Least-Squares Estimation of Transformation Parameters Between Two Point Patterns," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 13, no. 4, pp. 376-380, 1991.
[26] B. ter Haar Romeny, L. Florack, J. Koenderink, and M.A. Viergever, "Scale-Space: Its Natural Operators and Differential Invariants," *Lecture Notes in Computer Science 511: Information Processsing in Medical Imaging,* pp. 239-254, 1991.
[27] E. Trucco and R.B. Fisher, "Experiments in Curvature-Based Segmentation of Range Data," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 17, no. 2, pp. 177-182, 1995.

Elizabeth Guest received the BSc degree in mathematics from the University of Edinburgh in 1989. She received the PhD degree in artificial intelligence from the University of Edinburgh in 1995. After a brief sojourn in Africa doing linguistic research, she worked for two years at the University of Leeds on a project aimed at modeling the soft tissue in the face in order to predict the outcome of surgery to the underlying bone. Currently, she is a lecturer in computing at Leeds Metropolitan University. Her research interests include segmentation of biomedical images, 3D image analysis, registration, image warping, segmentation, 3D image reconstruction, and calculation of point correspondences.

Elizabeth Berry received the BSc degree in physics from the University of Hull in 1982. She received the degree PhD in applied optics from Imperial College London in 1986. She is currently a senior lecturer in medical imaging at the University of Leeds, in the Centre of Medical Imaging Research and Academic Unit of Medical Physics. Her research interests cover a spectrum from basic research in medical image production, analysis, and understanding of health technology assessment in imaging.

Richard A. Baldock received the BSc degree in mathematics and physics from the University of Bristol in 1975. He received the PhD degree in theoretical physics from the Australia National University in 1980. He is currently a senior research scientist at the MRC Human Genetics Unit in Edinburgh, where he is working on 3D reconstruction, a digital atlas of mouse embryo development, and a spatio-temporal database of gene-expression and other spatial data in the mouse. His research interests include knowledge-based systems, segmentation of biomedical images, 3D reconstruction and visualization, and databases.

Márta Fidrich received the BSc and MSc degrees in computer science from the József Attila University of Szeged, Hungary in 1989 and 1991, respectively. She received the PhD degree in computer science from the University of Paris in 1997. After completing her PhD, she spent a year as a research fellow in the Department of Statistics at Leeds University. She is currently a research associate in the Research Group in Artificial Intelligence at József Attila University. Her research interests include feature detection, image segmentation and registration, facial reconstruction and modeling, and graphics.

Mike A. Smith received the BSc degree in physics from Bristol University, the MSc degree in medical physics from Aberdeen University, and the PhD degree from the Faculty of Medicine at Edinburgh University. He is currently a professor and head of medical physics and Dean of Research in the Faculty of Medicine, Dentistry, Psychology, and Health at the University of Leeds. He is a fellow of both the Institute of Physics and the Institute of Physicists and Engineers in Medicine and was president of the British Institute of Radiology in its centenary year. His research is concerned with the development of noninvasice investigative techniques in medicine. These include methods for the measurement of bone and body composition since the mid-1970s, magnetic resonance imaging since the early 1980s, and medical image analysis since the early 1990s. More recently, his research has broadened to encompass health technology assessment and evidence-based diagnostics.