Algorithms for accurate 3D registration of neuronal images acquired by confocal scanning laser microscopy

O. AL-KOFAHI*, A. CAN*, S. LASEK†, D. H. SZAROWSKI†,
J. N. TURNER*† & B. ROYSAM*
*Rensselaer Polytechnic Institute, Troy, NY 12180–3590, U.S.A.
†The Wadsworth Center, NY State Department of Health, Albany, NY 12201–0509, U.S.A.

Summary

This paper presents automated and accurate algorithms based on high-order transformation models for registering three-dimensional (3D) confocal images of dye-injected neurons. The algorithms improve upon prior methods in several ways, and meet the more stringent image registration needs of applications such as two-view attenuation correction recently developed by us. First, they achieve high accuracy (≈ 1.2 voxels, equivalent to 0.4 μm) by using landmarks, rather than intensity correlations, and by using a high-dimensional affine and quadratic transformation model that accounts for 3D translation, rotation, non-isotropic scaling, modest curvature of field, distortions and mechanical inconsistencies introduced by the imaging system. Second, they use a hierarchy of models and iterative algorithms to eliminate potential instabilities. Third, they incorporate robust statistical methods to achieve accurate registration in the face of inaccurate and missing landmarks. Fourth, they are fully automated, even estimating the initial registration from the extracted landmarks. Finally, they are computationally efficient, taking less than a minute on a 900-MHz Pentium III computer for registering two images roughly 70 MB in size. The registration errors represent a combination of modelling, estimation, discretization and neuron tracing errors. Accurate 3D montaging is described; the algorithms have broader applicability to images of vasculature, and other structures with distinctive point, line and surface landmarks.

Introduction

Volumetric imaging modalities such as laser-scanning confocal microscopy are now in widespread use in the life sciences.

Correspondence: Dr Badrinath Roysam. Tel.: +1 518 276 8067; fax: +1 518 276 6261; e-mail: roysam@ecse.rpi.edu

In several applications, there is a need to register a pair or several of these volumetric images as accurately as possible, and with the maximum possible level of automation.

Montage synthesis, also known as mosaicing, is a common application of image registration that produces synthetic images at high magnification as well as large spatial extent. This is useful when the structures of interest are much larger than the field of view of the microscope (Pavel & Annica, 1995; James *et al.*, 2000). In addition to the common greyscale montage, it is possible to generate artificial representations such as segmentation label montages, and ball/stick diagrams, also known as object montages (Becker *et al.*, 1996). The image registration requirements of montage synthesis are modest – adjacent images are linked by just a linear translation (Becker *et al.*, 1996; Capek & Krekule, 1999). This is especially true when a microscope system is equipped with an accurate computer-controlled motion stage.

An application with more demanding registration requirements is two-view confocal microscopy (Can *et al.*, 2003; Can *et al.*, 2000). In this application, the goal is to extend the depth of imaging utilizing two three-dimensional (3D) images recorded from different directions. The images to be registered are flipped relative to each other. The precise angle of the flip is ideally 180°. However, in practice, there is an unavoidable and unknown error (tilt) around and normal to the optical axis, related to non-parallelism of the sample mounting between the two cover slips. The directions of the tilt are unknown and difficult to control mechanically. These angular errors indicate that the spatial transformation model that is used for image registration should at least include a 3D angle (θ, ϕ, ψ), in addition to the usual translation parameters (t_x, t_y, t_z) (Hartley & Zisserman, 2000). Furthermore, accounting for the different resolution along the x, y and z-axes introduces three additional scale factors (s_x, s_y, s_z). If skew due to mechanical distortions were present, three more degrees of freedom would be needed. When a large number of images need to be registered, curvature of field and

Fig. 1. Illustration of how three trace points are chosen around a branching point for initial registration estimation. First, the angles are determined and sorted from greatest to smallest. Then, the trace points are ordered by angle. Since angles are invariant under rigid body rotation, the correspondences between the trace points from two images can be established.
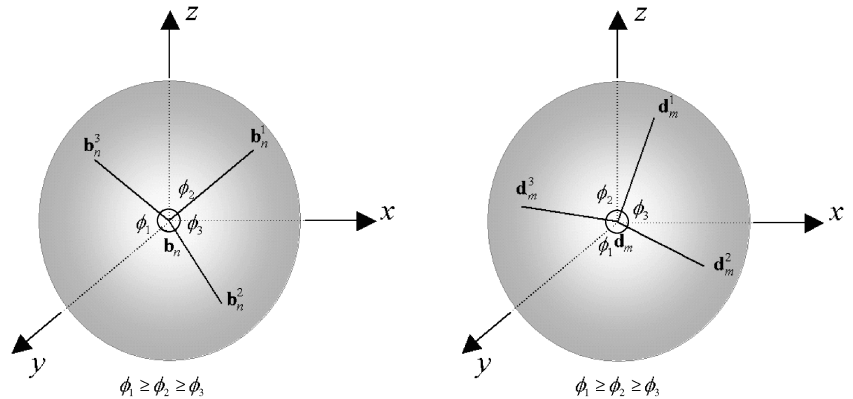
geometric distortions (Pawley, 1995; Bradbury & Bracegirdle, 1998) must be accounted for. This paper presents stable algorithms that can fit a hierarchy of transformation models ranging up to a quadratic model with 30 degrees of freedom.

A brief survey of 3D image registration methods

The literature on volumetric image registration is extensive (Brown, 1992; Lavallee, 1995; Maintz & Viergever, 1998). Broadly, two approaches are used: landmark-based methods and intensity correlation based methods. Landmark-based methods rely on matching a set of distinctive points (landmarks) in each of the images. For instance, the work of Becker *et al.* (1996) used the centroids of segmented nuclei as landmarks. Intensity-based methods proceed directly from the volumetric image data, utilizing all of the points. They compute a measure of correlation between the two images to estimate their relative spatial disposition. Finally, hybrid methods that combine ideas from landmark and intensity correlation methods are also emerging (Abu-Tarif, 2002).

Landmark-based methods (Arun *et al.*, 1987; Luck *et al.*, 2000; Can *et al.*, 2002) are generally fast and can scale up easily with higher-order transformation models. However, preprocessing (landmark extraction) is needed to produce the geometrical features to be used as landmarks. Although generic operators, known as interest operators, exist for extracting landmarks (Haralick & Shapiro, 1992), the best choice of landmarks is often application dependent. The incorporation of robust estimators in landmark-based registration algorithms makes them remarkably tolerant to missed and spurious landmarks (Beaton & Tukey, 1974; Holland & Welsch, 1977; Stewart, 1999; Can *et al.*, 2002). This allows them to operate when the image overlap is low.

Intensity-based methods (Belichenko & Dahlstrom, 1995; Capek & Krekule, 1999; Beck *et al.*, 2000) are in a sense more generally applicable because they use all voxels rather than a selected set of landmark points. They do not require a landmark extraction step prior to registration. However, they are susceptible to imaging artefacts, such as non-uniform illumination throughout the images, and some rectifying steps need to be made prior to registration, since the voxel intensities are used directly to solve the spatial registration problem (Oldmixon & Carlsson, 1993). They do not exploit specific knowledge of the application context to simplify the problem, or to enhance the registration speed or accuracy. Furthermore, these algorithms are naturally far more data-intensive than the landmark-based methods. Several metrics are described in the literature for measuring the quality of the registration. For example, Capek & Krekule (1999) describe sum of absolute valued differences (SAVD), normalized correlation coefficient (NCC) and mutual information function (MIF) as potential measures (Maes, 1998).

In the present work, landmark-based methods are adopted in order to achieve speed and accuracy for the applications of immediate interest, and because for many applications we have already developed tracing and segmentation algorithms that find a robust set of landmarks (Ancin *et al.*, 1996; Can *et al.*, 1999; Al-Kofahi *et al.*, 2002). One important advantage of our algorithm is that it is fully automated (Fig. 1). In most of the earlier work (Oldmixon & Carlsson, 1993; Belichenko & Dahlstrom, 1995), the user is required to specify an initial crude registration estimate.

Methods for registration of the 3D images

Overview

Our method simultaneously estimates the spatial transformation parameters as well as the landmark correspondences. First, a set of landmarks is extracted for each image. A subset of the landmarks is used to estimate a crude initial guess of the registration. This is subsequently refined by an iterative procedure that is derived by extending the well-known iterative closest points (ICP) algorithm of Besl & McKay (1992) using methods from robust statistics, with a high-dimensional image transformation model.

Extraction of Landmarks

Specimens used in our experiments (Fig. 2) were prepared as described in Can *et al.* (2003). Appropriate landmarks for the neuron images include the location of soma, dendrites and branching points of the dendritic tree. An automated 3D exploratory tracing algorithm described by Al-Kofahi *et al.* (2002) is used to trace the dendrites and axons. This algorithm first segments the soma location, and then recursively extracts the dendrites as a set of traces in 3D space. The soma location, branching points and the centrelines of these traces are stored in an output file that can be imported into widely used commercial software packages such as Neurolucida (MicroBrightfield Inc., Colchester, VT, U.S.A.). A sample neuron image is shown in Fig. 2 as a set of three orthogonal projections in the (x,y), (x,z) and (y,z) planes. The automatically extracted soma, dendrites and the branch points of the dendrites are shown in Fig. 3.

The somas are ignored, since they may not be found in all the images to be registered. In the following discussion, we will refer to the two images to be registered as I_1 and I_2, and the set of landmark points for these two images as $\mathbf{P}$ and $\mathbf{Q}$, respectively. These points include the centreline as well as the branch point locations in the image coordinate system. The branch points in the two images are also denoted $\mathbf{b}$ and $\mathbf{d}$ in the following section, where $\mathbf{b} \subset \mathbf{P}$, and $\mathbf{d} \subset \mathbf{Q}$.

Initial estimation of landmark correspondences

Most image registration algorithms require an initial manual estimate of the registration. Automated initialization is possible in our application. Specifically, we exploit the fact that the dendritic branch points are generally unique and distinctive for a given neuron. The automatic initialization procedure begins with two sets of landmarks corresponding to the dendritic branch points in the two images. Since the correspondences between the landmarks are unknown, it is necessary to test all possible landmark pairs for potential matches. For the experiments described here, the number of branch points is small, so this computation is not a burden. At this stage, since a good spatial transformation estimate is not available at the point, it is reasonable to expect false and missed matches. Accordingly, candidate matches are fed into a robust statistical estimation algorithm that estimates the initial spatial transformation. The following paragraphs describe the procedure in detail.

Suppose that $\mathbf{p}_i$ and $\mathbf{q}_i$ are observations of the same image point in I_1 and I_2, respectively. A low-order spatial transformation model can be derived by assuming that the neurons are rigid bodies that are imaged without distortions. This model includes a 3D translation, denoted $\mathbf{t}'$, and a 3D orthonormal rotation matrix $\mathbf{R}'$. This rigid body model is defined in the real world Euclidean coordinate system, whereas the landmarks are available in the image coordinate system. Accordingly, it is necessary to map all the landmark points $\mathbf{p}_i, \in \mathbf{P}$ and $\mathbf{q}_j \in \mathbf{Q}$ in the

3D image coordinate system into the real-world coordinates, denoted $\mathbf{p}'_i$ and $\mathbf{q}'_j$, as follows:

$$\mathbf{p}'_i = \mathbf{S}\mathbf{p}_i \quad \text{and} \quad \mathbf{q}'_j = \mathbf{S}\mathbf{q}_j, \tag{1a}$$

where

$$\mathbf{S} = \begin{pmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{pmatrix}. \tag{1b}$$

Here, s_x, s_y and s_z are the spatial sampling rates (usually in micrometres per voxel dimension) in x, y and z dimensions, respectively. In the real-world coordinates, corresponding points are related by the following equation:

$$\mathbf{q}'_i = \mathbf{R}'\mathbf{p}'_i + \mathbf{t}' \tag{2}$$

where $\mathbf{R}'$ is a rotation matrix, and $\mathbf{t}'$ is a translation. If the correspondences were known, estimating the rotation matrix and the translation vector would be straightforward. The following paragraphs describe a hypothesize-and-test method to estimate these for the realistic case when the correspondences are unknown.

Let the locations of the branching points from the two images be denoted $\mathbf{b}_n$ ($n = 0, \dots, N$) and $\mathbf{d}_m$ ($m = 0, \dots, M$), respectively. I_1 has N branching points, and I_2 has M branching points. The rigid body transformation (3D translation, and 3D rotation) requires at least three corresponding point-pairs to estimate the transformation parameters. Figure 1 illustrates that each branch point consists of three salient points – one corresponding to each branch of the trace. If an imaginary sphere is placed at the branch point, we obtain three points where the traces intersect the imaginary sphere; if a branch point has more than three branches, they are discarded. Denote these three points as $\mathbf{b}_n^k$ and $\mathbf{d}_m^k$ ($k = 1, 2, 3$), respectively (see Fig. 1). The correspondence between the trace locations from $\mathbf{b}_n$ and $\mathbf{d}_m$ can be established by using the angles formed by these trace locations. For example, the entry for $\mathbf{b}_n^1$ can be ordered such that the angle ϕ_1 is greater than the angles ϕ_2 and ϕ_3 where:

$$\phi_1 = \cos^{-1}(\mathbf{e}_n^2 \cdot \mathbf{e}_n^3); \tag{3a}$$

$$\phi_2 = \cos^{-1}(\mathbf{e}_n^1 \cdot \mathbf{e}_n^3); \tag{3b}$$

$$\phi_3 = \cos^{-1}(\mathbf{e}_n^1 \cdot \mathbf{e}_n^2); \tag{3c}$$

and $e_n^k = \dfrac{(\mathbf{b}_n^k - \mathbf{b}_n)}{\left\| (\mathbf{b}_n^k - \mathbf{b}_n) \right\|}$ is the unit vector relating $\mathbf{b}_n$ to $\mathbf{b}_n^k$.

The branch points in I_2 are processed similarly. Note that these angles are invariant under rigid-body transformations (Hartley & Zisserman, 2000). From the two branching points, $\mathbf{b}_n$, $\mathbf{d}_m$, and their corresponding trace locations $\mathbf{b}_n^k$ and $\mathbf{d}_m^k$, the rigid-body parameters are estimated as follows; first the correlation matrix is calculated:
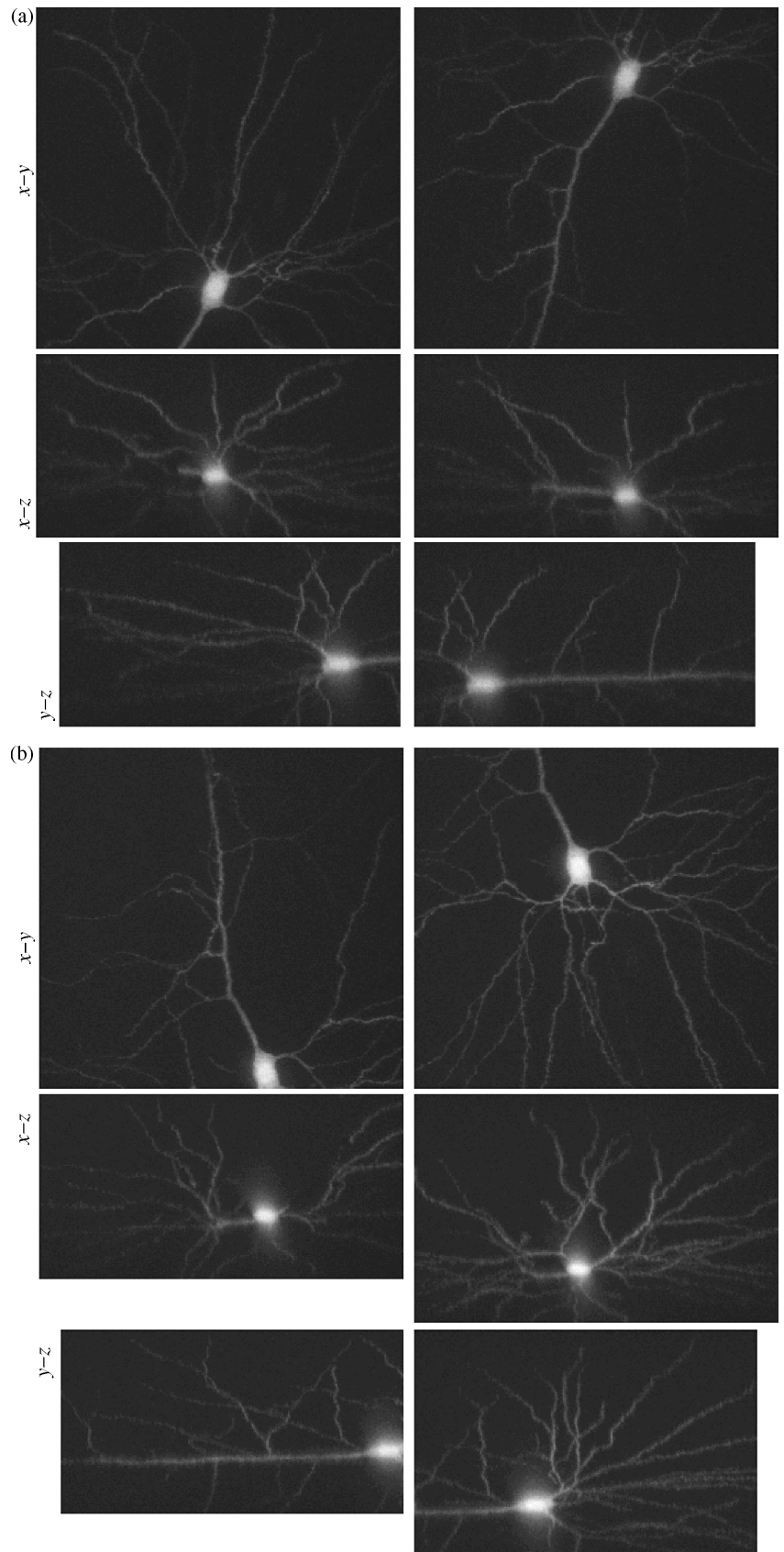
(a)

(b)

Fig. 2. (a) Sample views of an Alexa-stained neuron from a rat brain, imaged before (left view-1) and after translating the specimen (right view-2). Each view is $512 \times 480 \times 261$ voxels in size. The neuron in the two views is larger than the microscope field of view; the two views cover the whole neuron, maintaining high resolution. Top row, x–y projection; second and third rows, x–z and y–z projections, respectively. The sampling rates are $0.375 \, \mu\text{m}$ voxel^{-1} for the x and y dimensions and $0.5 \, \mu\text{m}$ voxel^{-1} for the z dimension. (b) Sample views of the neuron shown in (a) but after flipping. Left (view-3) is before translating the specimen and right (view-4) is after translating the specimen. The two views are different in size: left, $512 \times 480 \times 323$ voxels; right, $512 \times 480 \times 261$ voxels. The neuron in the two views is larger than the microscope field of view; the two views cover the whole neuron, maintaining high resolution. Top row, x–y projection; second and third rows, x–z and y–z projections, respectively. The sampling rates are $0.375 \, \mu\text{m}$ voxel^{-1} for the x and y dimensions and $0.5 \, \mu\text{m}$ voxel^{-1} for the z dimension.
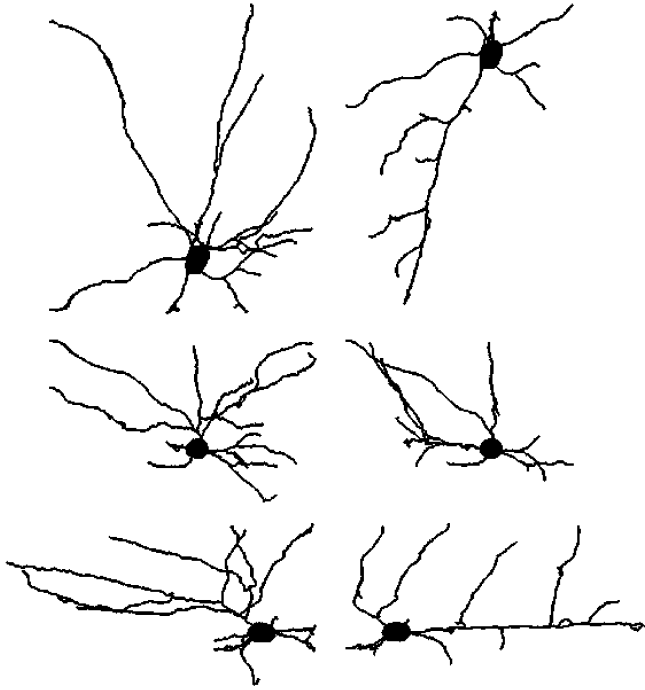
Fig. 3. Soma segmentation and exploratory tracing of the neurons shown in Fig. 2(a). Traces are ordered in the same manner as the original gray scale images in Fig. 2(a). Top row, x–y projection; second and third rows, x–z and y–z projections, respectively.

$$K_{nm} = \sum_{k=1}^{3} (\mathbf{d}_m^k - \mathbf{d}_m)(\mathbf{b}_n^k - \mathbf{b}_n)^T. \tag{4}$$

Then, a singular value decomposition (SVD) of the correlation matrix $\mathbf{K}_{nm} = \mathbf{VDU}^T$ is calculated. The rotation matrix is estimated from the decomposition as follows:

$$\hat{\mathbf{R}}'_{nm} = \mathbf{V} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det(\mathbf{VU}^T) \end{pmatrix} \mathbf{U}^T \tag{5}$$

Then, the 3D translation is computed as shown below

$$\hat{\mathbf{t}}'_{nm} = \mathbf{d}_m - \hat{\mathbf{R}}'_{nm}\,\mathbf{b}_n. \tag{6}$$

The reader is referred to Kanatani (1994) and to the survey paper by Lorusso *et al.* (1997) for further mathematical details underlying the above steps.

The above procedure generates an estimate of rotation and translation parameters given one pair of corresponding branch points. The accuracy of these estimates is tested by transforming all the landmarks $\mathbf{p}'_i$ and $\mathbf{q}'_j$ (centrelines and branch-points) with these estimates and calculating a cost function. For all possible branch-point correspondences, this hypothesize–test scheme is repeated and the transformation estimate that yields the minimum cost function is chosen as

the initial estimate. Mathematically, $(\hat{\mathbf{R}}', \hat{\mathbf{t}}')$ are set equal to $(\hat{\mathbf{R}}'_{nm}, \hat{\mathbf{t}}'_{nm})$ that minimize the cost function:

$$\sum_i \rho\left(\min_j \left\| \mathbf{q}'_j - \hat{\mathbf{R}}'_{nm}\mathbf{p}'_i - \hat{\mathbf{t}}'_{nm} \right\| \big/ \sigma \right), \tag{7}$$

where $\rho(.)$ is known as the robust cost function, and it is described in the next section. The argument of the cost function is the minimum Euclidean distance between a transformed point and the set of landmark points on the other image. The average minimum rigid body error using this estimator was 3.5 voxels. Notice that only the branching points are used to estimate the transformation model parameters, whereas the error is measured on the locations of the centrelines as well as the branching points. If the centrelines were also used to estimate the parameters of the rigid body transformation parameters using the robust estimator described in the next section, the registration error can be lowered to 2.5 voxels. However, the hypothesize–test scheme has enough accuracy to initialize the affine estimator described in the next section. The number of hypotheses and tests that are computed is $N \times M$, and the total complexity would therefore be $O(NM)$ times the error calculation complexity, which is $O(|\mathbf{P}|)$. Although the computation of the error measure in the test step looks more involved than the computation of the parameter estimates in the hypothesize step, Euclidean distance maps can be used to lower the error computation to $O(|\mathbf{P}|)$. Note that without using the trace locations around the branching points, the number of hypotheses and tests that are computed is $N^3 \times M^3$, and the total complexity would therefore be $O(N^3 M^3)$ times the error calculation complexity $O(|\mathbf{P}|)$, because at least three branching points are needed to solve for the rotation and translation. The resulting $\hat{\mathbf{R}}'$ and $\hat{\mathbf{t}}'$ estimates are used to initialize the algorithm for estimating a higher-order 12-parameter affine transformation model, as described in the following section.

Robust iterative closest point algorithm with affine model

The affine transformation model, consisting of a 3D translation vector $\mathbf{t}$, and an unconstrained 3×3 matrix $\mathbf{A}$ that implicitly accommodates 3D rotation, skew and non-isotropic scaling, is estimated next. Unlike the rigid body model, the affine model is capable of using points in image coordinates directly, since it has extra degrees of freedom to accommodate non-isotropic scaling. Hence, knowing the spatial sampling factors $\mathbf{S}$ from Eq. (1b), the estimated rotation and transformation parameters in image coordinates, denoted $\hat{\mathbf{R}}'$ and $\hat{\mathbf{t}}'$, can be written as follows:

$$\hat{\mathbf{R}} = \mathbf{S}^{-1}\hat{\mathbf{R}}'\mathbf{S}, \tag{8}$$

and

$$\hat{\mathbf{t}} = \mathbf{S}^{-1}\hat{\mathbf{t}}'. \tag{9}$$

Table 1. Summary of the registration algorithm. Starting from an initial transformation estimate, the point correspondences are computed by choosing the closest point as an estimate. Based on these correspondences the spatial transformation parameters are computed. These new transformation parameters are used to update the point correspondences for the next iteration.

(i)	Set the initial values: $k = 0$, $\hat{\mathbf{A}} = \hat{\mathbf{R}}$, $\hat{\mathbf{t}}$,
	Find the initial correspondences [Eq. (10)],
	Find the initial scale estimate [Eq. (11)]
(ii)	Compute the weights for each corresponding point [Eq. (15)].
(iii)	Estimate the affine transformation parameters, by solving weighted least squares [Eq. (12)].
(iv)	Estimate the scale for the next iteration,

$$\hat{\sigma}_{k+1} = \left(\frac{\sum w(u_i) \left\| \mathbf{q}_i^* - \hat{\mathbf{A}}\mathbf{p}_i - \hat{\mathbf{t}} \right\|^2}{\sum w(u_i)} \right)^{\frac{1}{2}}$$

	(The scale is updated for the first few iterations, and then fixed.)
(v)	Estimate new correspondences [Eq. (10)],
(vi)	If the robust cost function [Eq. (14)], has converged stop,
	Otherwise, set $k \leftarrow k + 1$ and go to step-ii.

A crude initial estimate for the affine parameters is to use the rotation matrix as an initial guess for the matrix $\hat{\mathbf{A}}$ (i.e. $\hat{\mathbf{A}} = \hat{\mathbf{R}}$), and the translation estimate $\hat{\mathbf{t}}$ from Eqs (5) and (6). From a given initial estimate $\hat{\mathbf{A}}$ and $\hat{\mathbf{t}}$, an estimate of the correspondences can be computed by searching the set of landmarks $\mathbf{P}$ for the closest point in $\mathbf{Q}$. Mathematically, the closest point $\mathbf{q}_i^*$ is chosen such that it minimizes the squared error given by

$$\left\| \mathbf{q}_j - \hat{\mathbf{A}}\mathbf{p}_i - \hat{\mathbf{t}} \right\|^2, \tag{10}$$

for all points $\mathbf{p}_i \in \mathbf{P}$. For computational convenience, if the correspondence of a point $\mathbf{p}_i$ is within a small neighbourhood of radius R, i.e. if $u_i = \left\| \mathbf{q}_i^* - \hat{\mathbf{A}}\mathbf{p}_i - \hat{\mathbf{t}} \right\| \leq R$, then it is declared as an inlier and $\mathbf{p}_i$ is included in a new set, $\mathbf{P}_1 \subset \mathbf{P}$; if this condition is not met, it is considered an outlier and excluded. Refer to Friedman & Finkel (1977), Blias & Levine (1995) and Rusinkiewicz & Levoy (2001) for more details on criteria for selecting corresponding pairs. Using this subset of points, an initial scale estimate is computed using robust median statistics as follows:

$$\hat{\sigma}_0 = 1.4826 \sqrt{\underset{p_i \in P_1}{median} \left\| \mathbf{q}_i^* - \mathbf{A}\mathbf{p}_i - \hat{\mathbf{t}} \right\|^2}. \tag{11}$$

This criterion is remarkable in that it is robust up to 50% outliers. The constant 1.4826 ensures consistency with the normal distribution. Using the initial scale estimate, $\hat{\sigma}_0$, and the initial estimate for $\hat{\mathbf{A}}$ and $\hat{\mathbf{t}}$, the overall algorithm for simultaneous correspondence and parameter estimation is given in Table 1. Following the work of Can *et al.* (2002), the robust approach to estimating the affine parameters is based on the well-known M-estimator (see Hampel *et al.*, 1986; Stewart, 1999) from robust statistics. Specifically, the M-estimate of the unknown parameters can be computed by minimizing a robust cost function. Mathematically, $(\hat{\mathbf{A}}, \hat{\mathbf{t}})$ are set equal to $(\mathbf{A}, \mathbf{t})$ that minimize the cost function

$$\sum_i \rho \left(\left\| \mathbf{q}_i^* - \mathbf{A}\mathbf{p}_i - \mathbf{t} \right\| \middle/ \hat{\sigma} \right). \tag{12}$$

$\hat{\sigma}$ is described in step-4 in Table 1. Here we used the Beaton–Tukey bi-weight function (Beaton & Tukey, 1974)

$$\rho(u_i) = \begin{cases} \frac{a^2}{6} \left[1 - \left(1 - \left(\frac{u_i}{a} \right)^2 \right)^3 \right] & |u_i| \leq a \\ \frac{a^2}{6} & |u_i| > a \end{cases} \tag{13}$$

where $u_i = \left\| \mathbf{q}_i^* - \hat{\mathbf{A}}\mathbf{p}_i - \hat{\mathbf{t}} \right\| \middle/ \hat{\sigma}$ is a 'scale normalized residual'. (Typically, $a \approx 4.0$; Holland & Welsch, 1977). The above minimization is performed using the iteratively re-weighted least-squares (IRLS) algorithm (see Holland & Welsch, 1977). This is an iterative algorithm that works by minimizing a cost function of the following form at each iteration:

$$\sum_i w(u_i) \left\| \mathbf{q}_i^* - \hat{\mathbf{A}}\mathbf{p}_i - \hat{\mathbf{t}} \right\|^2, \tag{14}$$

where, $w(.)$ is a weight function that is given by

$$w(u_i) = \left(\frac{d}{du_i} \rho(u_i) \right) \middle/ u_i.$$

For the Beaton–Tukey bi-weight function, the weights reduce to the following expression:

$$w(u_i) = \begin{cases} \left(1 - \left(\frac{u_i}{a} \right)^2 \right)^2 & |u_i| \leq a \\ 0 & |u_i| > a \end{cases} \tag{15}$$

The value of u_i for each match is calculated using the estimate of $\hat{A}$, $\hat{t}$ from the previous iteration. Typically, this iterative algorithm converges in less than 10 iterations, and the scale in step-5 in Table 1 is updated only for the first few iterations. The radius is updated such that the point correspondences that yield residuals with zero weights are excluded from the error computations, $R = a \cdot \sigma$.

The above method can be continued to estimate higher-order transformation models, such as a quadratic model with 30 degrees of freedom. The quadratic model has enough degrees of freedom to accommodate for curvature of field and geometric distortions, if present. When this was done, the quadratic model was found to reduce the registration error by 15% over the affine. Now, this decrease in the error measure may be due to an over fitting of the data in some cases. A more precise model for a given application can be derived based on the optics of the microscope, supplemented by generalized cross-validation techniques. Different reduced quadratic models for different applications (such as flipped vs. translated set-ups) can be the right choice. As a general registration model we prefer to use the affine model to avoid over-fitting, and maintain stability. The errors reported in the next section for the affine model were satisfactory, i.e. 1.2 voxels. This includes an unavoidable error of about 0.7 voxels attributable to spatial descretization (Can *et al.*, 2002).

Experimental results

The example illustrates registration and montaging of four images of a 4% Alexa 594 (Molecular Probes, Eugene, OR, U.S.A.) dye-injected neuron from a rat brain, acquired from a symmetrically prepared specimen (mounted between a pair of cover slips) using the two-view imaging method (Can *et al.*, 2003). The images involve different types of movement, translation and flip. The images in Fig. 2(a) are two views that involve translation before flipping the specimen, and the images in Fig. 2(b) are acquired in the same manner but after flipping. The 3D views are shown as a set of three orthogonal projections. Each of the four views is of size $512 \times 480 \times 323$ voxels (rows × columns × slices), except for view-3, which has 261 slices, with a sampling rate of 0.375 μm voxel^{-1} in the lateral plane, and 0.5 μm in the axial direction. The goals of using this example are to augment the effective axial depth of imaging, and to increase the effective field of view in the *x–y* direction, while maintaining a high spatial resolution.

Figure 3 shows the results of automated soma segmentation and tracing of the neuron images shown in Fig. 2, using the algorithm of Al-Kofahi *et al.* (2002).

Figures 4 and 5 show the result of registering the first pair of views (before flipping). The accuracy of the registration is measured by the mean squared distance between the two corresponding sets of trace points after they have been spatially transformed. The two sets of traces are colour coded, red, green and blue, for the first view, second view and the overlapping traces, respectively. These traces are then superimposed over a montage of the two views as shown in Fig. 4.

Figure 6 shows the joint mosaic of the four views, where the goals for the experiment are clearly met. The views before and after flipping are corrected for depth attenuation using the two-view attenuation correction method; the reader is referred to the companion article, Can *et al.* (2003), for details of the two-view attenuation correction.

It is important to note that only the points that were estimated to correspond in Eq. (14) were included in this calculation. This approach ensures that only the traces in the overlapping volume covered by the two views are included in the error calculation. Furthermore, it ensures that tracing errors, especially errors that result in missed or artifactual traces, do not contribute to the error measure, since they do not contribute to the registration. This measure was found to be 1.2 voxels for the example shown in Fig. 4, based on 2100 corresponding points. After taking the voxel dimensions into account, this translates to 0.4 μm. Note that this error includes the discretization error of the centreline locations, centreline-tracing errors, modelling error and estimation errors.

Discussion

We have presented an effective, robust and accurate algorithm for fully automated feature-based registration of 3D images using landmarks determined by preprocessing algorithms.

An important aspect of the proposed algorithm is that it can be used to estimate high-order models such as quadratic models. The low-order models are used to estimate the correspondences that are used to initialize the higher-order models. This hierarchical approach overcomes an important disadvantage of high-order models by simplifying the simultaneous correspondence and transformation estimation problem. Despite their ability to accommodate complex modelling assumptions such as object non-rigidity, they always require a good initial estimate. The hierarchical approach allows the method to be automated, and yet be able to utilize the high-order transformation models. In this regard, the proposed methods represent a significant advance over prior work in this area.

Most of the earlier work has been based on greyscale intensity-based registration. Such approaches are susceptible to imaging artifacts, such as non-uniform illumination throughout the image, and some rectifying steps need to be made prior to registration, since the voxel intensities are used directly to solve the spatial registration problem. The landmarks produced by our tracing algorithm (Al-Kofahi *et al.*, 2002) are less affected by imaging system artifacts, and are more invariant between images. Furthermore, landmarks such as traces are much fewer in number, significantly reducing
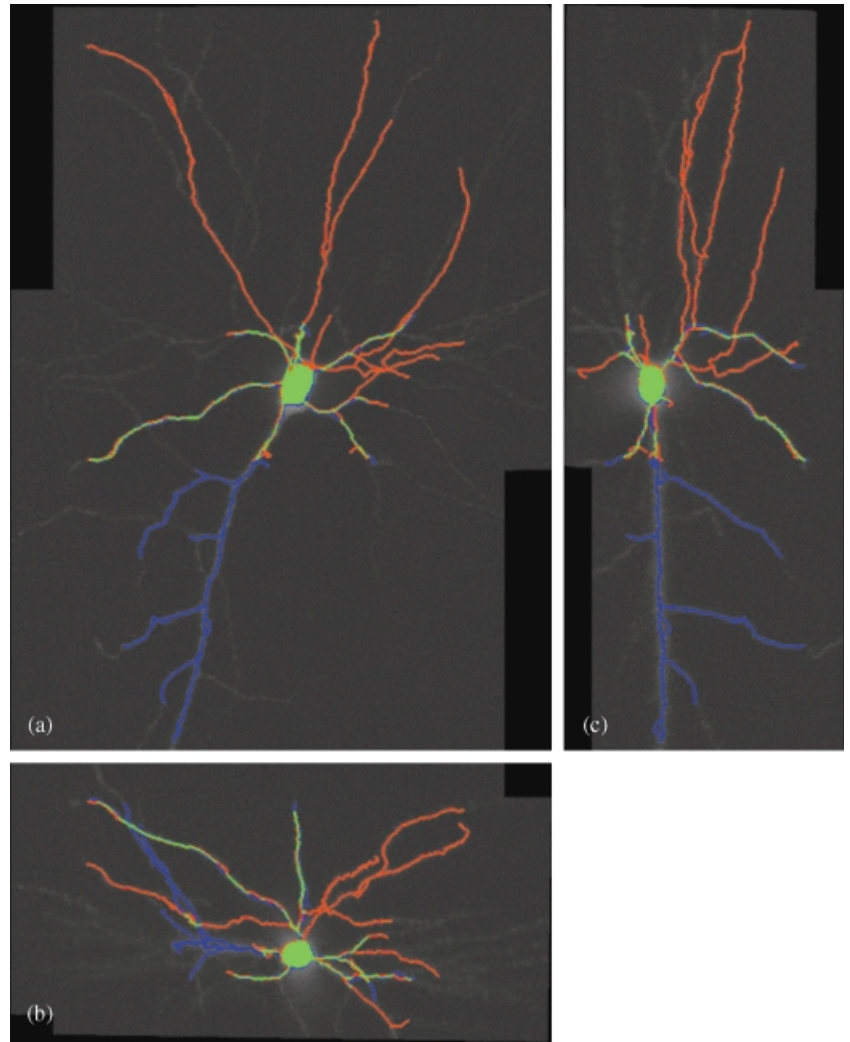
Fig. 4. Result of registering the two images presented in Fig. 2(a), using automatically generated traces shown in Fig. 3 as spatial landmarks. The 3D result is displayed as a set of three projections in panels (a), (b) and (c) as x–y, x–z and y–z, respectively. The traces are overlaid on the image. Traces from view-1 are shown in red, and traces from view-2 in blue. The fully overlapped traces appear in green. The average registration error computed from traces is 1.2 voxels (0.4 μm).

the computation complexity. The algorithm converges in the first 10 iterations, registering two images approximately 70 MB in size within less than a minute on a standard Dell 900 MHz Pentium III computer. The exploratory tracing algorithm is very fast and it takes around 10 s to detect all the centrelines and the soma.

The landmarks from the tracing algorithm add to the errors in the data sets to be registered, and hence robust estimation methods are used in our algorithm to accommodate for those and other sources of errors.

Our algorithm is fully automated. The spatial affine transformation parameters are evaluated automatically from the landmarks, in this case the traces. All that is necessary is at least one good branch point that occurs in both views to initialize registration. For the complete algorithm to register a pair of images, traces common to both images must exist. The soma segmentation need not be successful, since there might be cases where it is not even present in an image.

The example presented shows our algorithm was able to detect the correct transformation for different geometric relationships between the two 3D images. For other applications, the proposed algorithm design strategies presented here could be adapted and extended to achieve similar capabilities. All that is necessary is a set of reliable landmarks. For instance, one could utilize surfaces of objects instead of the traces used here. Simply, it is possible to achieve full automation and high registration accuracy using hierarchical landmark-based registration, when the landmarks are chosen suitably for a given application, or a class of applications where large-scale automated image registration is necessary. We expect to publish examples of such adaptations in the near future.
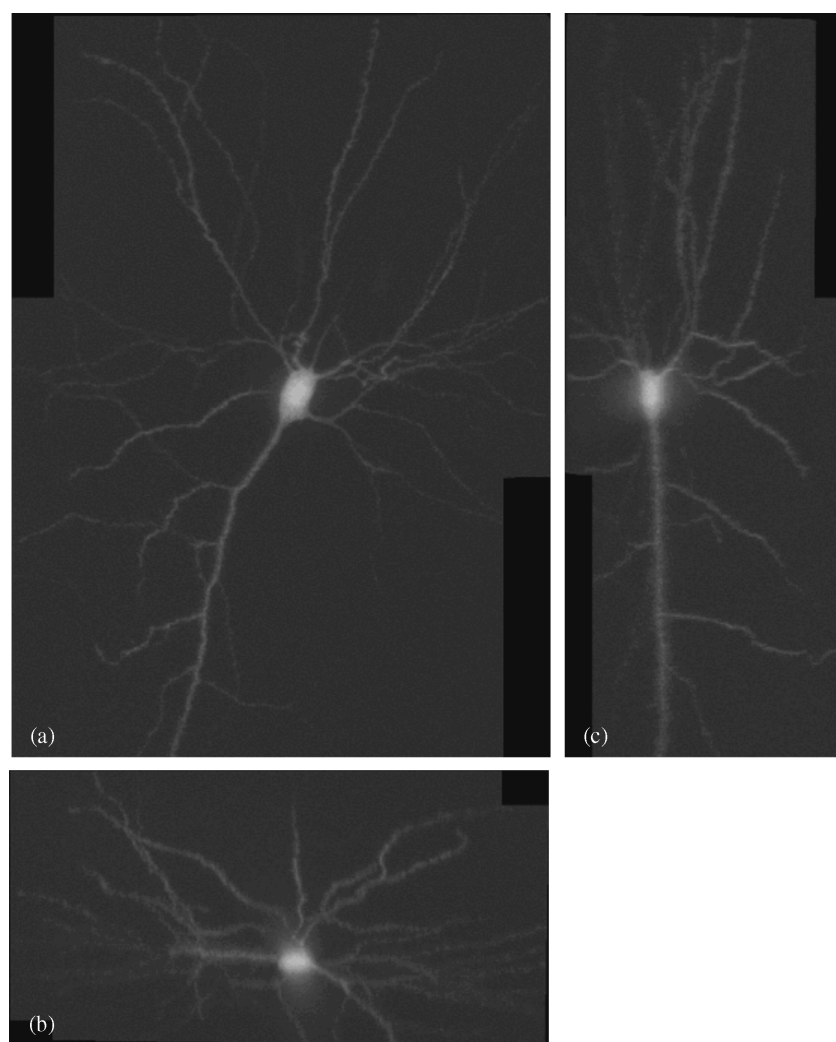
Acknowledgements

Fig. 5. Mosaic image constructed from the two views in Fig. 2(a). The 3D result is displayed as a set of three projections in panels (a), (b) and (c) as x–y, x–z and y–z, respectively. The result of the montaging is a high-resolution image for the whole neuron with increased effective imaging depth.

References

Abu-Tarif, A. (2002) *Volumetric registration based on intensity and geometry features. PhD thesis*, Rensselaer Polytechnic Institute, Troy, NY.

Al-Kofahi, K., Lasek, S., Szarowski, D., Pace, C., Nagy, G., Turner, J.N. & Roysam, B. (2002) Rapid automated three-dimensional tracing of neurons from confocal image stacks. *IEEE Trans. Information Technol. Biomedicine*, **6**, 171–187.

Ancin, H., Roysam, B., Dufresne, T.E., Chestnut, M.E., Ridder, G.M., Szarowski, D.H. & Turner, J.N. (1996) Advances in Automated 3D Image Analysis of Cell Populations Imaged by Confocal Microscopy. *Cytometry*, **25**, 221–234.

Arun, K.S., Huang, T.S. & Blostein, S.D. (1987) Least-square fitting of two 3-D point sets. *IEEE Trans. Pattern Anal Machine Intelligence*, **9**, 698–700.

Beaton, A.E. & Tukey, J.W. (1974) The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics*, **16**, 147–185.

Beck, J.C., Murray, J.A., Dennis, A.O., Willows, B. & Cooper, M.S. (2000) Computer-assisted visualizations of neural networks: expanding the field of view using seamless confocal montaging. *J. Neurosci. Meth.* **98**, 55–163.

Becker, D.E., Ancin, H., Szarowski, D.H., Turner, J.N. & Roysam, B. (1996) Automated 3-D montage synthesis from laser-scanning confocal images: application to quantitative tissue-level cytological analysis. *Cytometry*, **25**, 235–245.

Belichenko, P.V. & Dahlstrom, A. (1995) Confocal laser scanning microscopy and 3-D reconstructions of neuronal structures in human brain cortex. *J. Neuroimage*, **2**, 201–207.

Besl, P.J. & McKay, N.D. (1992) A method for registration of 3-D shapes. *IEEE Trans. PAMI*, **14**, 239–256.

Blais, G. & Levine, M.D. (1995) Registering multi-view range data to create 3D computer objects. *Trans. PAMI*, **17**, 820–824.
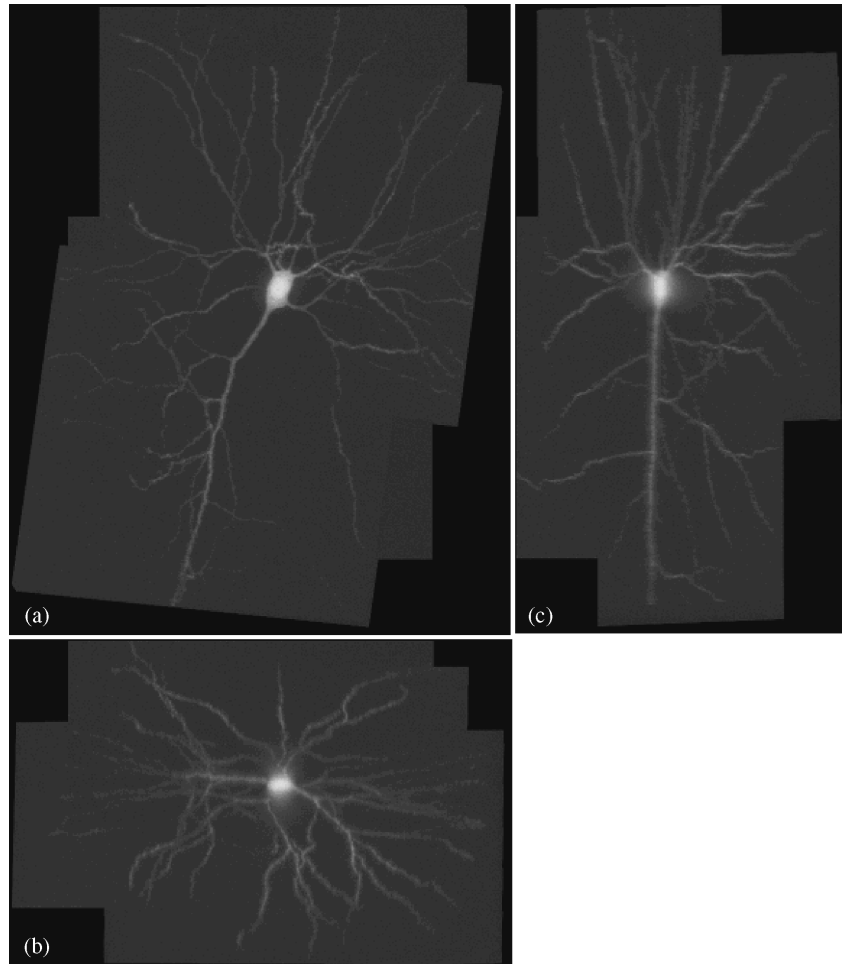
Fig. 6. Mosaic image constructed from the four views in Fig. 2. The 3D result is displayed as a set of three projections in panels (a), (b) and (c) as x–y, x–z and y–z, respectively. The result of the montaging is a high-resolution image for the whole neuron with increased effective imaging depth.

Bradbury, S. & Bracegirdle, B. (1998) *Introduction to Light Microscopy*. BIOS Scientific Publishers Ltd, New York.

Brown, L.G. (1992) A survey of image registration techniques. *ACM Computer Surveys*, **24**, 325–376.

Can, A., Al-Kofahi, O., Lasek, S., Szarowski, D.H., Turner, J.N. & Roysam, B. (2003) Attenuation correction in confocal laser microscopes: a novel two-view approach. *J. Microsc.* **211**, 67–79.

Can, A., Lasek, S., Szarowski, D.H., Turner, J.N. & Roysam, B. (2000) A robust two-view method for increasing the imaging depth and correcting for signal attenuation in confocal microscope images. *Proc. Microsc. Microanal.* **6**, 818–819.

Can, A., Shen, H., Turner, J.N., Tanenbaum, H.L. & Roysam, B. (1999) Rapid automated tracing and feature extraction from live high-resolution retinal fundus images using direct exploratory algorithms. *IEEE Trans. Information Technol. Biomedicine*, **3**, 125–138.

Can, A., Stewart, C.V., Roysam, B. & Tanenbaum, H.L. (2002) A feature-based. Robust, hierarchical algorithm for registering pairs of images of the curved human retina. *IEEE Trans. PAMI*, **24**, 347–364.

Capek, M. & Krekule, I. (1999) Alignment of adjacent picture frames captured by a CLSM. *IEEE Trans. Information Technol. Biomedicine*, **3**, 119–124.

Friedman, J.B. & Finkel, R. (1977) An algorithm for finding best matches in logarithmic expected time. *ACM Trans. Mathemat. Software*, **3**, 209–226.

Hampel, F.R., Rousseeuw, P.J., Ronchetti, E. & Stahel, W.A. (1986) *Robust Statistics: the Approach Based on Influence Function*. John Wiley & Sons, New York.

Haralick, R.M. & Shapiro, L.G. (1992) *Computer and Robot Vision*. Addison-Wesley Co., New York.

Hartley, R. & Zisserman, A. (2000) *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge.

Holland, P.W. & Welsch, R.E. (1977) Robust regression using iteratively re-weighted least-squares. *Commun. Statistics. Theory Meth.* **A6**, 813–827.

James, C.B., James, A.M., Dennis, A.O., Willows, B. & Mark, S.C. (2000) Computer-assisted visualizations of neural networks: expanding the field of view using seamless confocal montaging. *J. Neurosci. Methods*, **98**, 55–163.

Kanatani, K. (1994) Analysis of 3-D rotation fitting. *IEEE Trans. PAMI*, **16**, 543–549.

Lavallee, S. (1995) Registration for Computer-Integrated-Surgery: Methodology, State of the Art. In: *Computer-Integrated-Surgery: Technology and Clinical Applications*, pp. 77–97. MIT Press, Cambridge, MA.

Lorusso, A., Eggert, D.W. & Fisher, R.B. (1997) A comparison of four algorithms for estimating 3-D rigid transformations. *Machine Vision Applications*, **9**, 272–290.

Luck, J., Little, C. & Hoff, W. (2000) Registration of range data using a hybrid simulated annealing and iterative closest point algorithm. *Proc. IEEE Int. Conf Robotics Automation*, **4**, 3739–3744.

Maes, F. (1998) *Segmentation and registration of multi-modal medical images: from theory, implementation, and validation to useful tool in clinical practice. PhD thesis*, Catholic University of Leuven, Leuven, Belgium.

Maintz, A.J.B. & Viergever, M.A. (1998) An overview of medical image registration methods. *Med. Image Anal.* **2**, 1–36.

Oldmixon, E.H. & Carlsson, K. (1993) Methods for large data volumes from confocal scanning laser microscopy of lung. *J. Microsc.* **170**, 221–228.

Pavel, V.B. & Annica, D. (1995) Confocal laser scanning microscopy and 3-D reconstructions of neuronal structures in human brain cortex. *J. Neuroimage*, **2**, 201–207.

Pawley, J. (ed.) (1995) *Handbook of Confocal Microscopy*. Plenum Press, New York.

Rusinkiewicz, S. & Levoy, M. (2001) Efficient variants of the ICP algorithm. *Proceedings, Third International Conference on 3-D Digital Imaging and Modeling*, pp. 145–152.

Stewart, C.V. (1999) Robust parameter estimation in computer vision. *SIAM Rev.* **41**, 513–537.