J. Anat. (1998) **193**, pp. 347–361, with 8 figures Printed in the United Kingdom

Review

Algorithms for radiological image registration and their clinical application

D. J. HAWKES

Computational Imaging Science Group, Division of Radiological Sciences, United Medical and Dental Schools of Guy's and St. Thomas' Hospitals, London, UK

(*Accepted 21 October 1997*)

ABSTRACT

This paper reviews recent work in radiological image registration and provides a classification of image registration by type of transformation and by methods employed to compute the transformation. The former includes transformation of 2D images to 2D images of the same individual, transformation of 3D images to 3D images of the same individual, transformation of images to an atlas or model, transformation of images acquired from a number of individuals, transformations for image guided interventions including 2D to 3D registration and finally tissue deformation in image guided interventions. Recent work on computing transformations for registration using corresponding landmark based registration, surface based registration and voxel similarity measures, including entropy based measures, are reviewed and compared. Recently fully automated algorithms based on voxel similarity measures and, in particular, mutual information have been shown to be accurate and robust at registering images of the head when the rigid body assumption is valid. Two approaches to modelling soft tissue deformation for applications in image guided interventions are described. Validation of complex processing tasks such as image registration is vital if these algorithms are to be used in clinical practice. Three alternative validation strategies are presented. These methods are finding application outside the original domain of radiological imaging.

Key words: image registration; image transformation; radiological imaging.

INTRODUCTION

Radiological images provide information on the size, shape and spatial relationships between anatomical structures and any pathology, if present. They also provide distributions of physical and physiological attributes of the tissues such as the X-ray attenuation coefficient from X-ray computed tomography (CT), proton density or proton relaxation times from magnetic resonance (MR) imaging and blood flow or glucose metabolism from positron emission tomography (PET). Traditionally these images are interpreted by the radiologist viewing the images as film transparencies on a back illuminated light box. Most imaging modalities involve some digital manipulation and computation and so, increasingly, these images are displayed on a workstation. Subjective judgments of the relative size, shape and spatial relationships of visible structures and physiology inferred from intensity distributions are used for diagnosis, planning therapy and monitoring disease progression or response to therapy.

The radiologist, physician and surgeon now have an impressive array of different imaging modalities available for routine clinical investigation including CT, a wide range of different types of MR images, nuclear medicine PET or single photon emission tomography (SPECT) scans, ultrasound, conventional X-ray radiographs and video images from such devices as endoscopes and microscopes. These are usually viewed separately, often using different media or display technology. While this may often be

Correspondence to Dr D. J. Hawkes, Radiological Sciences, 5th Floor, Thomas Guy House, Guy's Hospital, London Bridge, London SE1 9RT, UK. email: d.hawkes@umds.ac.uk

adequate there are situations where it would be useful to know the precise spatial relationship between 2 structures, each visible on only 1 modality. Alternatively interpretation may require combining information on structure with images of function, for example from nuclear medicine which might lack sufficient spatial resolution to be interpreted in isolation. Computer assisted image registration provides a tool for combining these different sources of information. Recent breakthroughs in computation have led to the development of very robust and accurate registration algorithms. Detailed reviews are provided by Brown (1992), van den Elsen et al. (1993) and Maintz (1996).

Sources of spatial information may include atlas or computer models which have been generated from one prototypical individual or a combination of several individuals. These atlases assist in the identification and delineation of anatomical structures, particularly in the brain. They also can assist computer assisted image interpretation and can allow identification of boundaries between structures which provide little or no image contrast. A key step in using these atlases is establishing correspondence between atlas and image. A related problem is the registration of images of a group of individuals in, for example, cohort studies of brain function using PET or MR. One experiment on an individual may not reveal sufficient information either due to low signal in the resulting images or to wide variations between individuals. Bringing brain images from a cohort of studies into correspondence may reveal more useful information on brain function by improving the statistical significance of the findings or establishing intersubject variability (e.g. Friston et al. 1997).

Another important role of medical imaging is the monitoring of changes in shape, size and function of anatomical structures and pathology in a particular individual. Computer assisted image registration in combination with careful quality control of imaging devices provides the means for monitoring changes much more sensitively than by visual interpretation alone (Hill et al. 1994*b*; Hajnal et al. 1995; Freeborough & Fox, 1997).

Recently there have been significant advances in using images for navigation in certain types of surgery, particularly in the brain, skull base, the maxillofacial region, temporal bone and spine (Maciunas, 1993; Taylor et al. 1995). Inherent in this process is registration of the images, defined in a coordinate system related to the original scanning devices, and the physical space of the patient lying on the operating table. This correspondence or registration can be provided by marker pins or devices, such as the stereotactic frame, which is fixed to the patient's skull during both scanning and the operation. Recently frameless image guided surgery has been introduced, in which anatomical features are used for registration. This may be done by identification of corresponding landmarks but could also be provided by intraoperative images such as video, X-ray imaging or ultrasound. These images are therefore used primarily to provide spatial localisation rather than for their visual appearances. This localisation involves using the methodologies of image registration. The same principles apply to images used to plan and guide other therapies, in particular radiotherapy.

This paper briefly reviews the range of registration methods. The bulk of the paper describes recent work at our laboratory to devise accurate and robust algorithms for registration of different types of images and briefly describes some of the applications that will ensue.

METHODS

Clarification of terminology is appropriate at this stage. The task of image matching can be broken down into the following processes: establishing the type of transformation (e.g. 3D to 3D rigid body), computing the transformation from, for example, common features in the image pair to be registered, and transformation of one of the images (the source image) into the coordinate system of the other (the destination image). Together these processes result in registration in which points in each image will correspond to the same points in the patient. There is also a process of interpretation of the registered images, in which information contained in the 2 images is combined or 'fused'. Corresponding information can be used to check the transformation, while complementary information can be used to deduce useful new information either by qualitative interpretation or by improving measurement. The resulting combination of information is sometimes termed the process of 'data fusion'.

Image transformations

The process of registration involves computation of a transformation between coordinate systems.

2D to 2D

If the geometry of image acquisition is tightly controlled, 2D images may be registered purely via a rotation, 2 orthogonal translations and correction of any scaling errors. Although computationally straightforward clinically relevant examples of this are rare as controlling the geometry of image acquisition is usually very difficult. One example that we have developed is the registration of X-ray radiographs of the hand with ^{99m}Tc methyl diphosphonate planar nuclear medicine images for the diagnosis of suspected scaphoid injury (Hawkes et al. 1991). In this example a purpose built holding device constrains the hand to be in identical positions in the 2 images.

3D to 3D

Of more widespread applicability is the accurate registration of multiple 3D images such as MR and CT volumes. The assumption is usually made that the internal anatomy of the patient has not moved or distorted and hence the 6 degrees of freedom of rigid body motion (3 translations and 3 rotations) will bring the images into registration. Careful calibration of each scanning device is required to determine image scaling, i.e. the size of the voxels in each modality.

Registration of images to an atlas or model and registration of images between individuals

Establishing correspondence of an atlas with a set of images or a number of images from a cohort of individuals requires far more complex transformations. The transformation will need to reflect the variation in anatomy between the atlas and the individual patient. There will be changes in shape and size as well as grosser changes in topology. This remains an area of active research with several approaches under investigation including multiscale deformation, in which a rigid body transformation is computed at a coarse scale followed by multiple rigid body transformations for arrays of volume elements at progressively finer scales with interpolation between these elements (Collins et al. 1994), interpolation based on radial basis functions such as the thin plate spline (Bookstein, 1991), and models based on the physical attributes of a deforming substrate such as used in fluid models, spring models etc. (Christenson et al. 1995). In general deformations are very poorly constrained with a very large number of degrees of freedom. The problem can be made more tractable by using statistical shape models based on the analysis of the variations observed across a population of individuals (Cootes et al. 1994; Ruff et al. 1997). This can dramatically decrease the potential number of degrees of freedom necessary to describe the transformation. For example principle component analysis using the 'point distribution model' (Cootes et al. 1992) has shown that only 5 modes (or additional degrees of freedom) are required to capture 89% of the variance in the fetal liver shape (Ruff et al. 1997).

Registration for image guided interventions and 2D to 3D registration

Image guided interventions usually rely on determining the rigid body transformation between image space and the physical space of the operating room. This will be achieved, for example, by identifying sufficient corresponding anatomical landmarks to establish the 6 degrees of freedom of the 3D rigid body transformation. Intraoperative CT or ultrasound provides one or more 2D slices and the registration task is to establish the pose of the slices with respect to the preoperatively acquired 3D volume. X-ray or video images are perspective 2D projections. Establishing the pose of these projections with the 3D preoperative volume, so-called 2D–3D match, will allow 3D information to be projected on to the 2D image for guidance purposes. Alternatively registration of 2 (or more) 2D projections with a 3D volume will allow, via triangulation, reconstruction of 3D locations from the 2D projections in the coordinates of the 3D volume. In general the rigid body assumption is assumed to be a reasonable approximation. Matching a perspective projection to a volume will require establishing between 6 and 10 degrees of freedom dependent on the assumptions and constraints of the perspective projections. If the focal length of the video system or the distance between source and detector of the X-ray set is fixed, then a 'one-off' calibration will determine the 4 parameters of perspective projection leaving the 6 parameters of the rigid body transformation to be determined in the registration process.

Tissue deformation in image guided interventions

In image guided interventions tissue can distort and deform between preoperative scans and the intervention. Anatomical structures may move in relation to each other. Intraoperative data may provide updated information on location and deformation of anatomical structures. This new information might be used as a basis for predicting deformation of adjacent tissues. The general problem, as in matching to an

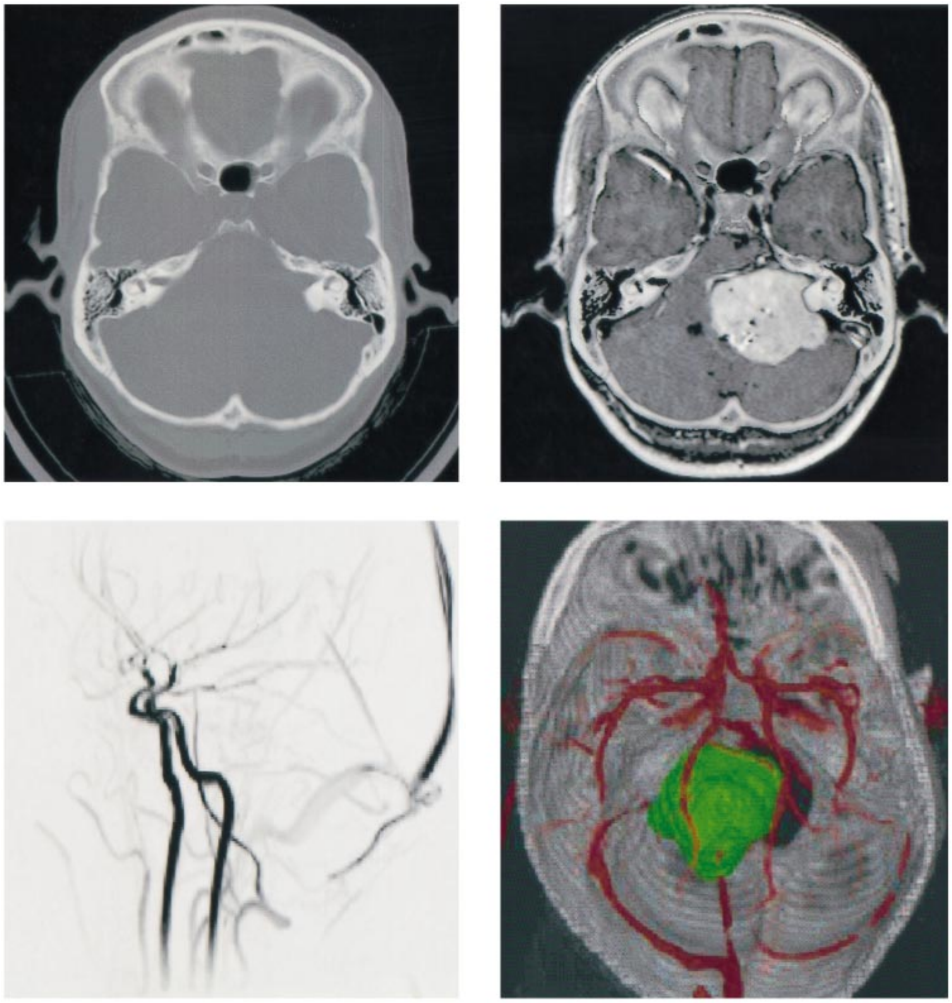

Fig. 1. Top left: a transaxial slice of an x-ray CT volume. Top right: a transaxial slice of this volume registered and overlaid over the corresponding transaxial slice of the gadolinium enhanced MR volume. Bottom left shows a lateral projection of the vasculature from an MR angiogram and bottom right shows a rendering derived from all 3 image volumes. The bone in grey is derived from CT, the tumour in green from gadolinium enhanced MR and the vasculature in red from the MR angiogram.

atlas, is very poorly constrained but in this case we may use the physical constraints of the tissues involved. For example bony structures will usually remain rigid and soft tissues will obey the laws of physics when deforming. Algorithms to exploit physical constraints will be discussed later.

Computation of transformations

Corresponding landmark based registration

One of the most intuitively obvious registration procedures is based on identification of corresponding point landmarks in the 2 images. Identification and location of 3 noncolinear landmarks will be sufficient to establish the transformation between 2 3D image volumes. The algorithm for direct computation of the transformation is well known and straightforward (Arun et al. 1987) involving alignment of centroids of the 2 sets of points followed by rotation to minimise the sum of the squared displacements between source and destination points. This is achieved by simple matrix manipulation using the method of singular valued decomposition.

The point-like features may be pins or markers fixed to the patient and visible on each scan. These may be attached to the skin or screwed into bone. The latter can provide very accurate registration but are more invasive and uncomfortable for the patient. Skin markers in the other hand can easily move by several millimetres due to the mobility of the skin and are difficult to attach firmly. Care must be taken to ensure that the coordinate of each marker is computed as accurately as possible and that the coordinate computed in each modality corresponds to the same point in physical space. Subvoxel precision is possible, for example using the intersection of 2 lines (Colchester et al. 1996), the apex of a 'V' (van den Elsen & Viergever, 1991) and the centre of gravity of spherical or cylindrical markers with a volume much larger than the voxel sizes (Maurer et al. 1995). Automating marker identification is possible.

Alternatively corresponding internal anatomical landmarks may be identified interactively on each image. These may correspond to truly point-like anatomical landmarks such as the cochlea, structures in which points can be unambiguously defined such as bifurcations of blood vessels, the centre of the globes of the eyes or surface curvature features that are well defined in 3D. Several groups, including our own, regularly register clinical images using corresponding anatomical landmarks that have been identified interactively by a skilled user (Evans et al. 1989; Hill et al. 1991, 1994*a*). Registration errors are reduced by increasing the number of fiducial markers. If the error in landmark identification is randomly distributed about the true landmark position, the error in coordinates computed from the resulting transformation reduces as the square root of the number of points identified for a given distribution of points. Typical intra- and inter-observer RMS errors on corresponding point identification are 4 mm when registering PET and MR images of the head and 1 mm when registering MR and CT images of the head, when the points themselves are well defined. Expected misregistration errors of about 2 mm at the centre rising to about 4 mm at the periphery are to be expected when registering MR and PET images of the head using 12 anatomical landmarks. For registering MR and CT images, including the skull base, typical misregistration errors will be about 1 mm at the centre rising to about 2 mm at the periphery for 12–16 landmarks (Hill et al. 1994*a*). Finding these landmarks automatically and reliably is difficult and remains a research issue.

Figure 1 shows an example CT volume registered with a gadolinium-enhanced MR image by picking point landmarks. The spatial relationships between bone and tumour, an acoustic neuroma, are clearly seen. Also shown is a volume rendered image of the combined datasets registered with a further MR image acquired to demonstrate the relationship between the vasculature and tumour (Ruff et al. 1993). The vasculature is shown in red and tumour in green from the different MR acquisitions and bone is shown in grey from CT. The slices are viewed according to radiographic convention from below while the rendering is generated as if viewing the skull and its contents from above. Figure 2 shows example aligned and overlaid MR images (grey) and PET 18-FDG (2-deoxy-2[18-f]fluoro-D-glucose) images (green) of the brain (2a) and aligned and overlaid CT images (grey) and PET 18-FDG (green) images of the pelvis (2b) showing separation of residual bladder activity from the recurrent carcinoma.

Surface based registration

Corresponding surfaces may be identified and used for registration. In these algorithms corresponding surfaces are delineated in the 2 imaging modalities and a transformation computed that minimises some measure of distance between the 2 surfaces (Borgefors, 1986; Pelizzari et al. 1989; Jiang et al. 1992). At registration this measure should be minimum. This method uses more of the available data than landmark identification and robust and accurate methods have been reported for some applications. Unfortunately the technique is highly dependent on identification of corresponding surfaces, yet different imaging modalities can provide very different image contrast between different structures. The process of delineation is hard to do accurately. Computer assisted segmentation will almost always require some manual editing or adjustment. The surface may also exhibit natural symmetries to certain rotations leading to poorly constrained transformations. Other features such as lines and tubes as well as combinations of features have also been used (Meyer et al. 1995). We have shown how, in principle, adjacent surfaces may be used for registration incorporating knowledge of the spatial relationships of different surfaces (Hill & Hawkes, 1994).

Registration of multiple images of the same patient acquired using the same imaging modality

When images of the same patient are acquired with the same modality then a strong correlation will exist between the voxel intensities in one image and voxel intensities in the other. In these cases cross correlation is a powerful measure of alignment but the variance of the ratio of image intensity and the sum of squares of differences in intensities have also been used successfully (Woods et al. 1992). As it is the small differences in very similar images that may have clinical significance care must be taken to ensure that the computation of the transformation does not itself remove this important information. Rescaling of either size or intensity, for example, must not mask real changes in volumes. If possible all images must be rescaled by reference back to an image of a fixed object or calibration phantom.

Subtle changes may be on the threshold of observation above the noise inherent in the imaging process. It is important therefore to ensure that

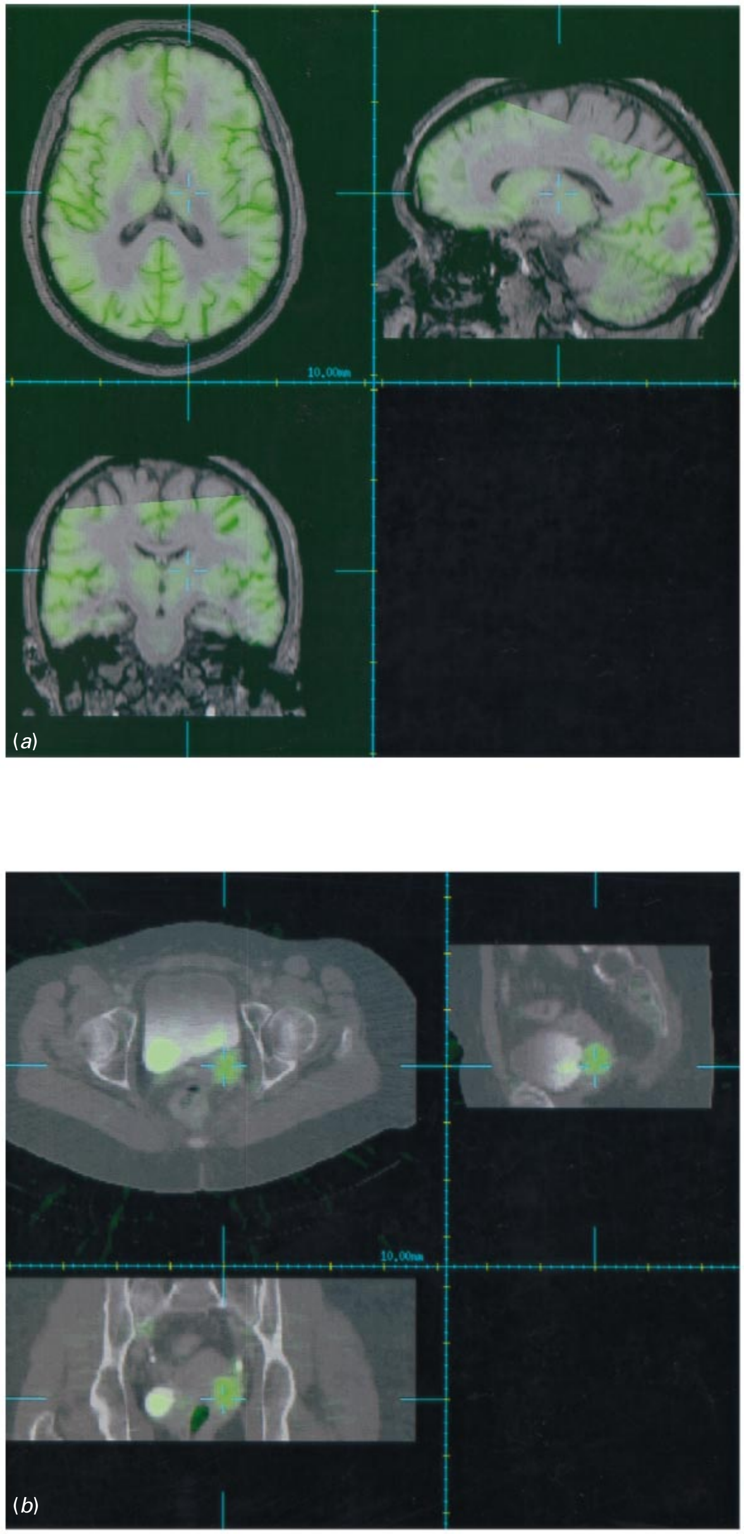

Fig. 2. (*a*) Three orthogonal slices through registered and combined PET 18-FDG and MR volumes of the head, with the PET image overlaid in green on the MR image displayed on a grey scale. (*b*) Superimposition of a transaxial slice from registered CT (grey) and PET 18-FDG (green) image volumes of the pelvis, showing clear separation of residual bladder activity from recurrent cervical carcinoma.

computation of the transformed image is accurate. Both images are discretely sampled in a voxel (or pixel) array. Corresponding intensities in the destination image are computed using the transformation back to the original source image. This process will require a resampling of the source image. A variety of resampling and interpolation schemes have been used ranging from nearest neighbour interpolation, tri-

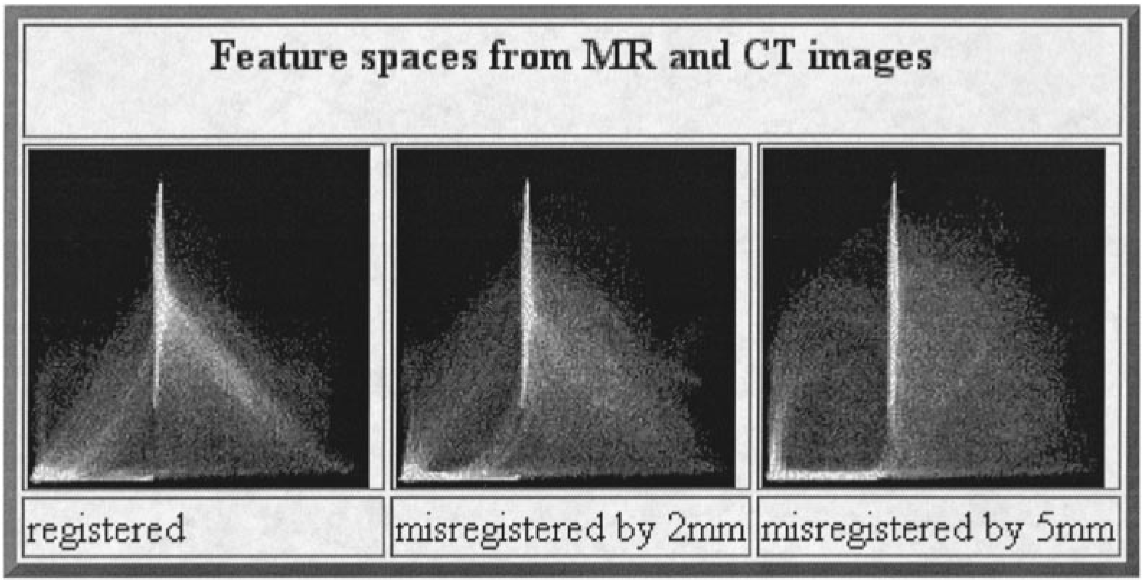

Fig. 3. Joint intensity histogram or feature space for MR and CT images of the head. The vertical axes correspond to the MR image intensity and the horizontal axes to CT intensity, with the left plot corresponding to registration, the middle to 2 mm misregistration and the right to 5 mm (from Hill et al. 1994*c*).

linear interpolation, through to full sinc interpolation. The latter will introduce minimal additional artefactual signal but it is computationally expensive. In general some compromise between accuracy and speed of computation has to be made. Recent work has claimed registration to subvoxel accuracy (Hajnal et al. 1995) with volume changes estimated to be accurate to about 0.1% of total brain volume (Freeborough & Fox, 1997).

Voxel similarity measures and entropy as a measure of alignment

In the last 4 y there has been significant progress, worldwide, in using relationships between voxel intensity values to align images acquired from different modalities. This work stems from the observation that while images from different modalities exhibit complementary information there is usually a high degree of shared information between images of the same structures. For example the human observer is able to fuse very different images such as MR and CT of the same structure provided intensity look-up tables are adjusted appropriately. The first successful application of voxel similarity methods was the variance of intensity ratios proposed by Woods et al. (1993) for registration of MR and PET images of the brain. Alignment corresponds to the transformation which minimises the variance of corresponding PET intensities for all voxels within defined ranges of MR intensity. The approach works very well if all but the brain volume has been removed from the MR images. Van den Elsen et al. (1994) proposed a method based on intensity correlation for MR and CT images. The CT image intensities are remapped using a triangular look-up-table so that both air and bone are dark with soft tissue bright, as in the corresponding MR image. Successful registration of images of the head and spine were reported.

There has been significant interest in measures of alignment based on the information content or entropy of the registered image. An important step to understanding these methods is the formation of the joint intensity histogram which is an estimate of the joint probability distribution of image intensities (Hill et al. 1994*c*). An example of such plots is shown in Figure 3 for MR and CT images of the head of the same patient and Figure 4 for MR and PET images. The joint intensity histogram is formed by accumulating or binning the occurrences of pairs of intensities in the 2 images for each trial alignment. Figures 3 and 4 show histograms formed at alignment and at 2 misalignments. It can be seen that the histograms disperse as misalignment increases and that each image pair has a distinctive joint intensity histogram signature at alignment.

Minimising the entropy of this distribution, the joint entropy, has been proposed by Studholme et al. (1995) as a measure of misalignment. As an illustrative example consider 2 images of the same individual, each containing 2 eyes. Misaligned, the combined images will contain 4 eyes while at alignment there will

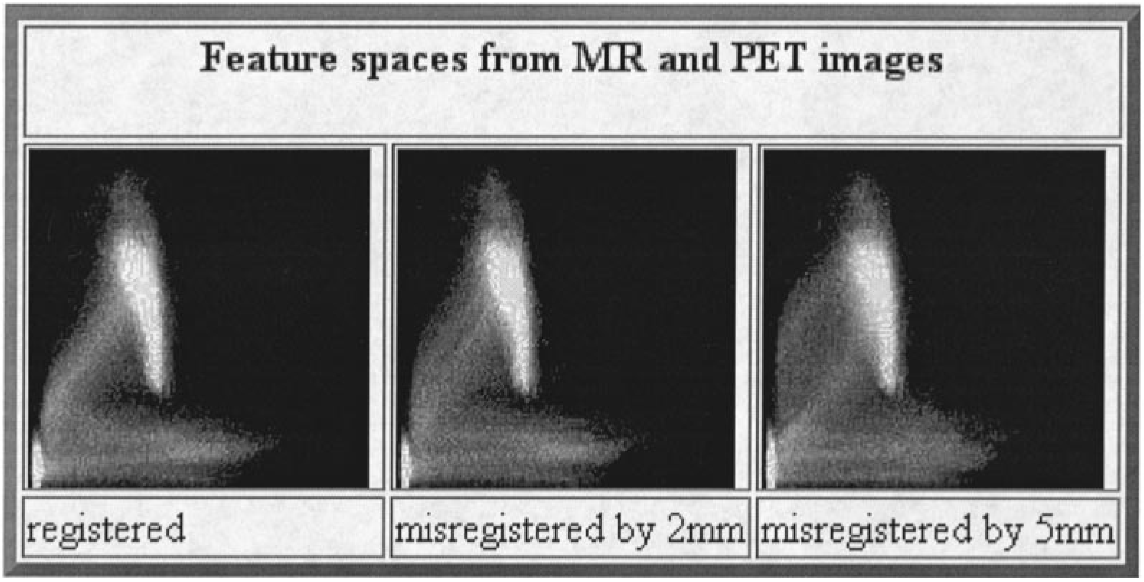

Fig. 4. Same plots for MR and PET images of the head as in Figure 3 with the MR intensities plotted on the vertical axes and the CT intensities plotted on the horizontal axes (from Hill et al. 1994*c*).

only be 2 eyes. There is, therefore, less information in the combined images at registration.

The Shannon–Wiener entropies, $H(M)$ and $H(N)$ of images M and N may be defined by

$$H(M) = -\sum_{m \in M} p\{m\} \log(p\{m\})$$

and

$$H(N) = -\sum_{n \in N} p\{n\} \log(p\{n\})$$

where $p\{m\}$ is the probability that a voxel in image M has value m and $p\{n\}$ is the probability that a voxel in image N has a value n. The individual entropies $H(N)$ and $H(M)$ may be estimated from the histograms of image values. The joint entropy $H(M,N)$ of the overlapping region of image M and N may be defined by

$$H(M,N) = -\sum_{n \in N} \sum_{m \in M} p\{m,n\} \log(p\{m,n\})$$

where $p\{m,n\}$ is the probability that a voxel in the overlapping region of image M and N has values m and n respectively. The joint probability may be estimated from the joint histogram of image values.

While joint entropy can provide a useful measure of alignment it has proved not to be very robust as other misalignments can result in much lower joint entropy. As an example, alignment of just the air surrounding the patient will produce a global minimum of entropy. An appropriate measure would be the difference in information between the overlapping volume of the combined image with respect to the information in the overlapping volumes of the 2 original images. Such a measure is provided by mutual information proposed independently by Collignon et al. (1995) and the MIT group (Wells et al. 1996). Mutual information is given by

$$I(M;\, N) = H(M) + H(N) - H(M,N).$$

At alignment we postulate that the joint entropy is minimised with respect to the entropy of the overlapping part of the individual images. Mutual information is a measure of how one image ‘explains’ the other. It makes no assumption of the functional form or relationship between image intensities in the 2 images.

While these entropy measures provide a measure of misalignment, they do not allow direct computation of the best estimate of alignment. An optimisation process is therefore required to compute a registration estimate. Any optimisation process involving 6 degrees of freedom will require computation of potentially a very large number of estimates. We have implemented a multiresolution scheme for registration in which a multiresolution stack of each image volume is generated from the finest to the coarsest resolution. Mutual information is computed at the coarsest resolution from the joint intensity histogram. A simple search strategy, at discrete steps in each of the 6 directions of the 6 dimensional search space of rigid body motion, computes the optimal registration at this resolution and then the search continues at the next finer resolution. This simple procedure has the advantage that large scale misalignments are retrieved rapidly at the coarsest scale with smaller adjustments computed at finer scales. We have tested the ro-

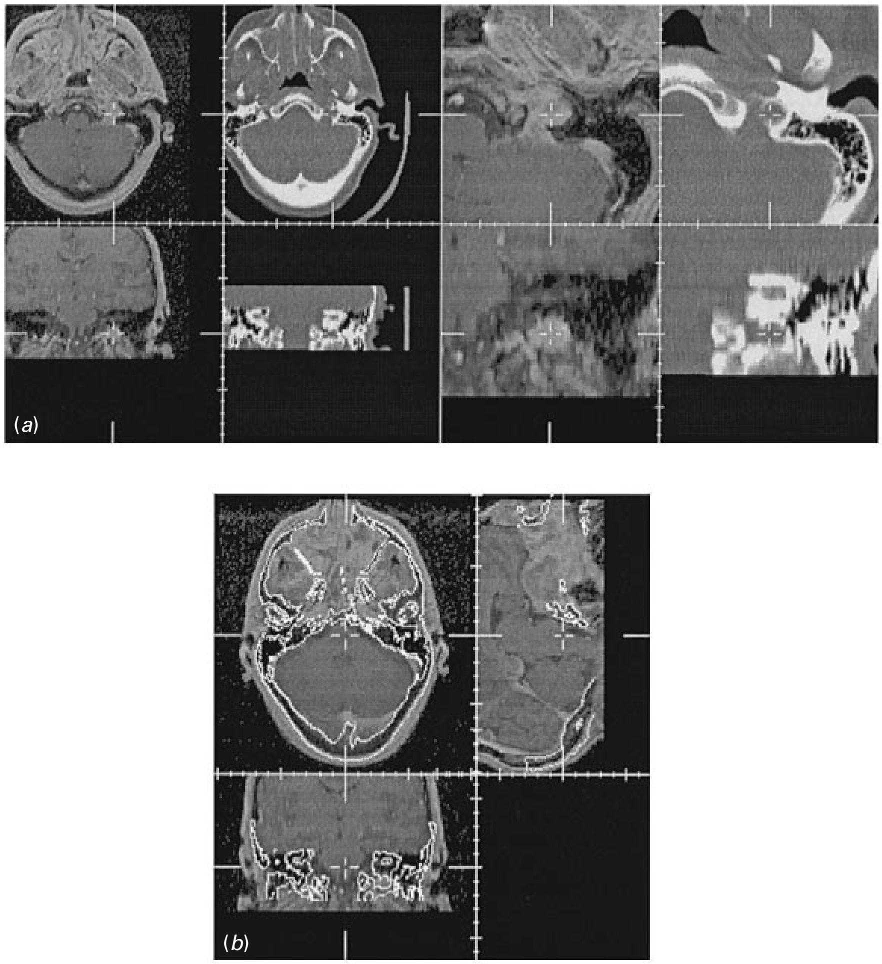

Fig 5. Transaxial and coronal slices through CT and MR volumes of the head, registered fully automatically by multiresolution maximisation of mutual information. (*a*) Panels on the left show side by side display of the aligned images and those on the right are zoomed versions of the same images. (*b*) Alternative display with a contour corresponding to the surface of bone, obtained by intensity thresholding the CT volume, superimposed on the MR images. Note the limited axial extent of the CT volume yet registration is visibly of high quality.

bustness and precision of the method to multiple random starting misalignments on large numbers of clinically acquired MR PET and MR CT image pairs of the head. Results are reported in detail elsewhere (Studholme et al. 1996, 1997; Studholme, 1997) but in summary we found that the system was robust for initial misalignments of up to 30 mm and 30° provided sufficient axial sections of image data were available (30 mm or more of axially overlapping images at registration). Registration takes between 15 and 45 min on a medium range Unix workstation, depending on the image size and resolution. Figure 5 shows example MR and CT images registered fully automatically by this algorithm.

2D to 3D registration

We have also explored the use of voxel intensity based methods in 2D to 3D image registration. Our task is to establish the pose of digital fluoroscopic x-ray images in relation to a previously acquired CT volume, for image guidance in interventional procedures undertaken on or near to the spine. In, for example, percutaneous laser discectomy it might be advantageous to perform the intervention under x-ray guidance. Currently multiple slices of CT data are acquired during a procedure. This is time consuming and can deliver a high radiation dose. If the preoperative CT could be mapped into the x-ray image then the procedure might safely be done under digital fluoroscopy guidance, while preserving the 3D guidance ability of the previously acquired CT images.

The method is based on digitally reconstructed radiographs (DRR) first proposed for stereotactic neurosurgical applications (Lemieux et al. 1994). Trial estimates of pose are computed by integrating along rays through the CT volume to simulate the process of perspective projection in x-ray imaging. The DRR is

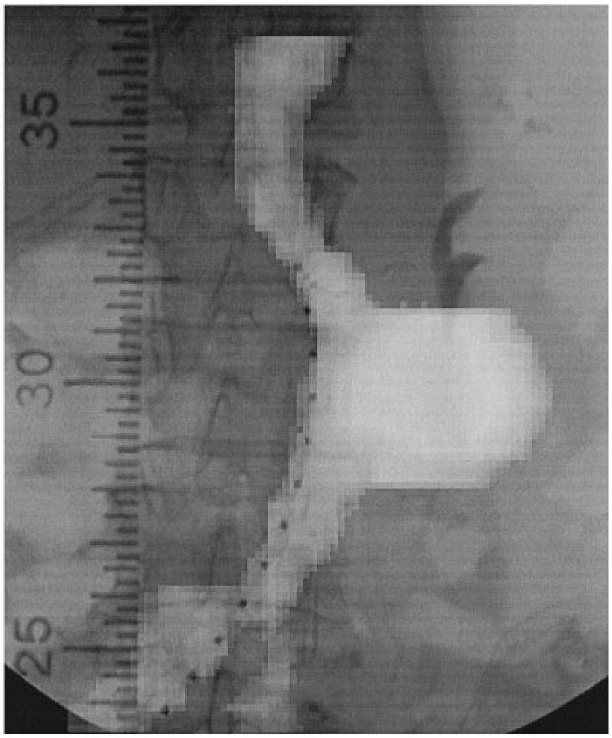

Fig. 6. A rendered abdominal aorta and aneurysm generated from CT angiography, registered and superimposed on a digital fluoroscopic image obtained during an interventional procedure to place an aortic stent.

compared with the true x-ray image using normalised cross correlation as a measure of similarity (Penney et al. 1997). Normalised cross correlation is used because of the similarity between the physical processes of x-ray and CT image formation. Recently a more robust measure based on pattern intensity has been proposed (Weese et al. 1997). The x-ray set is first calibrated using an object of known dimensions (steel markers at the apices and centres of the 6 faces of a 60 mm Perspex cube) to determine the 4 perspective parameters. The confounding effect of soft tissue movement is minimised by removing the soft tissue from the CT image by intensity thresholding. A multi-resolution gradient descent method is used to determine alignment. An approximate starting estimate is provided by a small number of user defined points. Results on a spine phantom show registration accuracy better than 1 mm and 1°, which is sufficient for the clinical application. Clinical evaluation of this process in percutaneous laser discectomy is about to start. Figure 6 shows an example of a 2D to 3D registration in which a rendered image from a CT volume acquired of a large aortic aneurysm is registered and overlaid on a digital fluoroscopic x-ray image acquired during an interventional procedure to place an aortic stent.

Image registration with soft tissue deformation in image guided surgery

This section briefly describes work in progress in our laboratory to use information on spatial location, derived from intraoperative images, to update 3D models of the patient's anatomy and pathology that have been derived from preoperative images. We are pursuing 2 approaches, one based on interpolation using radial basis functions and the other on simplified computational models of physical characteristics of the tissue. Both approaches assume that each bony structure can be represented by a rigid body. Each separate bone (for example each vertebra of the spine) is treated as a separate rigid body each with its own transformation.

This work is based on deformable models which use both 'a priori' information, such as physical constraints, as well as image derived information. A recent review of the literature is provided by McInerney & Terzopoulos (1996). Little et al. (1997) in our laboratory have recently shown how the constraints of rigid bodies can be incorporated into interpolation based on radial basis functions. This results in an interpolating solution that is a summation of a linear term corresponding to the rigid bodies and a basis function which smoothly tends to zero at the surface of the rigid bodies. The resulting transformation is exact at rigid bodies, given the rigid body transformation, and provides smooth interpolation elsewhere. Initial results on MR images of the spines of volunteers are promising but considerable further work on validation is required before this methodology can enter clinical practice. An example of this transformation on a sagittal MR image of the head of a volunteer who was scanned in 2 positions is shown in Figure 7. Soft tissue structures close to the spine and cranium are accurately registered.

An alternative strategy is to attempt to model soft tissue deformation directly. We have demonstrated plausible soft tissue deformation using a simple multiresolution model in which the tissue is represented on as an array of discrete elements (Edwards et al. 1997). Energy terms associated with these elements include spring energy where the connection between each node is represented by a spring providing forces of compression and tension; stiffness energy associated with bending of the connections of sets of 3 adjacent nodes; membrane energy associated with changes in the area of triangular elements; and combinations of all 3. This method has been tested on 2D CT and MR slices of the brain acquired before and after surgery for placement of electrode mats on

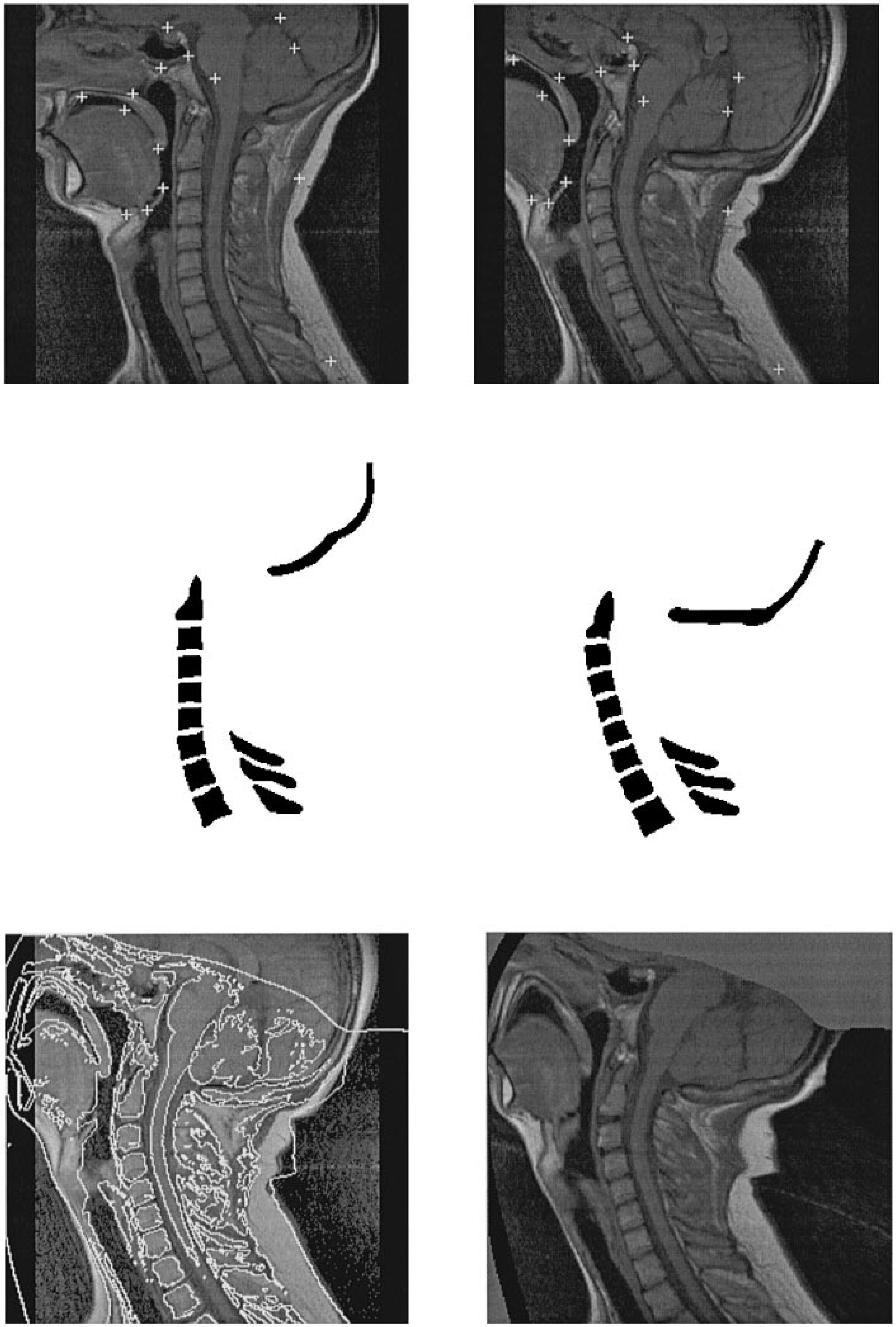

Fig. 7. Example image that has been deformed using a modified radial basis function (from Little et al. 1997). The top pair of images show sagittal MR slices through the head of a volunteer scanned with the head held in 2 very different positions. Landmarks used to constrain the deformation are also shown. The middle pair show corresponding segmentations, hand drawn, of the individual vertebrae and part of the cranium. The bottom right image shows the result of deforming the top left image to match the top right image, while the bottom left image shows an edge map of the deformed image superimposed on the top right image.

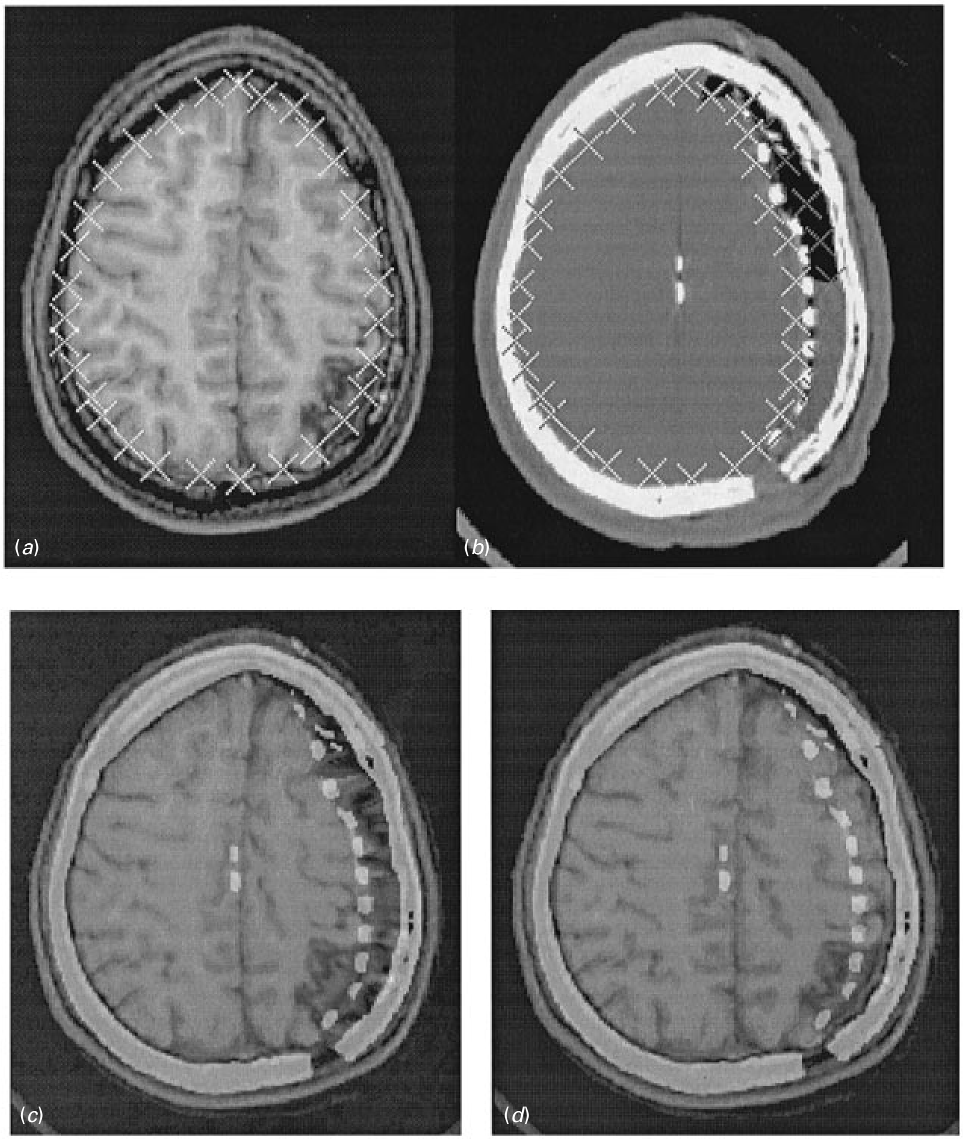

Fig. 8. Slice of a preoperative MR volume of a patient suffering from epilepsy (*a*). The corresponding CT slice is shown (*b*) after craniotomy and placement of the electrode mat. Rigid body registration shows (*c*) significant displacement of the surface of the brain which is corrected, in this case, with a deformation computed from a combined stiffness and membrane model (*d*). (From Edwards et al. 1997.)

the brain surface prior to excision of areas of focal activity in the treatment of epilepsy. We model the scanned slice with a 3 component model, 1 component representing rigid structures (bone), 1 representing fluid (cerebrospinal fluid and air surrounding the patient) which has zero energy associated with deformation and the 3rd representing deforming soft tissues. In this example the image data provided their own internal standard, in this case electrodes placed in the interhemispheric fissure. In this particular example the combination of the stiffness and membrane model proved to be most accurate. Results are shown in Figure 8. The algorithm is currently being extended to 3D and further data for validation are being collected. At present computing time is impractical, taking several hours per transformation, but methods to reduce these are in progress.

Both these approaches require significant further development and validation before entering clinical practice.

VALIDATION

Any process that entails manipulation of data for clinical purposes must undergo extensive validation. For image registration work this will usually follow a sequence of evaluation on computer generated models (software phantoms), images of physical phantoms of accurately known construction and dimensions and images of patients or volunteers. The process must demonstrate both high robustness and high accuracy. Robustness implies a very low failure rate and if failure does occur that this is communicated to the user.

Assessment of accuracy requires knowledge of a 'gold standard' or 'ground truth' registration. This is difficult to achieve with clinical images but 3 methods have recently been reported. In 1 approach 1 image is simulated from another. Strother et al. (1994) generated PET images from 6 MR images of the head using the characteristics of the CTI-Siemens 953B PET scanner. They tested 5 algorithms including using a stereotactic frame, user identified anatomical landmarks, surface matching and Woods algorithm for MR to MR registration, MR to PET registration and PET to PET registration. Woods algorithm, the only voxel similarity technique tested, performed best.

Hemler et al. (1995) used a cadaver head in which glass tubes had been inserted to provide a gold standard registration of MR and CT images. They compared surface based, stereotactic frame and voxel intensity correlation, reporting that stereotactic frame based registration was the most accurate.

By far the most extensive study to date is that undertaken by Vanderbilt University (West et al. 1997). In this study 7 sets of MR and CT images and 7 sets of MR and PET images were made available to researchers participating in the study. The patients imaged had had marker pins inserted into the bone of their skulls prior to imaging for image guided neurosurgery. These markers provided a very accurate 'gold standard' registration. Prior to distribution the markers were digitally removed from the images. The registrations were undertaken at each site blind to gold standard estimates. A number of surface and voxel based methods were compared and the voxel based methods performed significantly better and required little or no user interaction. The results using mutual information were amongst the most accurate. The median registration errors for the MR and CT images were 1.9 mm or better and for the MR and PET images were 3.2 mm or better using our implementation.

All these validation methods have only been applied to the head and only in situations where the rigid body transformation is a good approximation. Validation of algorithms which incorporate deformations is difficult and remains a research issue. Deformations to account for variations across populations have potentially a very large number of degrees of freedom, although the principal component analysis of the point distribution model shows one way of reducing the dimensionality of the problem. Deformations during surgical interventions will vary significantly from patient to patient. We need to demonstrate that the transformations predicted by either of our 2 approaches are more accurate than those provided by the conventional rigid body assumptions. This could be done using anatomical landmarks, but it is difficult to define landmarks which can be identified accurately. Paradoxically if these landmarks could be identified then they should be used to aid registration. It may be possible to use internal image visible markers inserted in previous interventions. Alternatives include the use of cadavers or animals, both of which pose significant logistic and ethical problems. It might be possible to simulate imaging processes using the widely available Visible Human Dataset (Spitzer et al. 1995). This dataset consists of high resolution MR and CT images all from the same individual, obtained postmortem. All 3 datasets cover the whole of the body.

DISCUSSION AND CONCLUSIONS

Image registration and the subsequent integration of corresponding and complementary information has become an important area of computation in medical imaging. Combining images from different modalities or acquired at different times is proving useful in interpreting lower resolution functional images, such as those provided by nuclear medicine, in determining spatial relationships of structures seen in different modalities, in planning surgery and other therapies such as radiotherapy, in monitoring subtle changes over time and finally to register and update preoperative 3D images in image guided surgery and therapy. Recent breakthroughs in the application of information theory to medical image registration have produced fully automated algorithms for registration of 3D images of the head which are accurate and robust. The same methodology is being applied outside the head. Application to the spine and pelvis for structures near bone are under evaluation. There are similarities in the process of registration using the joint intensity histogram and segmentation or classification using the same representation. A project is underway to explore whether segmentation and registration can proceed as a single iterative process.

We have introduced the concept of an imaging device as a localiser. In image guided surgery and in interventional radiology this often requires a match of 2D slices or projection images to the preoperative 3D volume. We have described one method for registering digital fluoroscopy projection x-ray images to corresponding preoperative CT volumes for applications in interventional radiology. These applications are in their infancy but significant progress is expected in this area in the next few years.

The voxel similarity paradigm may also be applied to matching video images to rendered 3D volumes of MR or CT images (or the surface normal data that is used to generate the rendered images). This opens up the exciting possibility of registering video with MR or CT images directly. This may have a profound effect on the practicality of applying image guidance methodology to a much wider range of interventions.

One of the main limitations of image guided interventions is the, as yet, unpredictable movement and deformation of soft tissue structures during intervention. This is a very difficult area of computation but we have shown 2 examples of approaches that may help to update 3D models derived from preoperative CT using sparse or incomplete information derived from intraoperative projection images. Both methods assume that individual bony structures will move as rigid structures within a deforming matrix. One method computes a smooth interpolation of this matrix based on radial basis functions while the other constructs an explicit discrete model containing some physical attributes of soft tissue, bone and fluid filled regions. These methods are far from clinical use but both methods show promise. Validation of deformation algorithms remains a significant problem with further research required to develop appropriate validation strategies.

The methodologies described in this paper may have application beyond radiology and surgery. It will be interesting, for example, to examine whether the registration methodologies described here could be applied to the accurate registration of autoradiographs to atlases of the mouse embryo obtained in studies of gene expression. This area was the subject of several of the other papers of the Symposium 'Computer Modeling for Anatomists and Clinicians' where this paper was presented. The registration and fusion methodologies described here are being applied outside medicine in, for example, oil exploration to the registration of seismic mappings and other sources of 3D information on geological structure.

ACKNOWLEDGEMENTS

The work at UMDS described in this paper was funded by EPSRC and Philips Medical Systems. The work was undertaken by Dr Derek Hill, Dr John Little, Colin Studholme, Philip Edwards and Cliff Ruff. The technical support of Colin Renshaw, the Radiography staff at Guy's and St. Thomas' Hospitals and King's Healthcare and the Staff at the Clinical PET Centre is acknowledged. The work described here would not have been possible without the support of our clinical and surgical colleagues and in particular Dr Tim Cox, Professor Michael Gleeson, Mr Anthony Strong, Professor Charles Polkey and Professor Michael Maisey. Ron Gaston assisted in the collation of the figures and paper preparation.

REFERENCES

Arun KS, Huang TS, Blostein SD (1987) Least squares fitting of two 3D point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **9**, 698–700.

Bookstein FL (1991) Thin-plate splines and the atlas problem for biomedical images. In *Information Processing in Medical Imaging* (ed. Colchester ACF, Hawkes DJ), pp. 326–342. Heidelberg: Springer.

Borgefors G (1986) Distance transformations in digital images. *Computer Vision, Graphics and Image Processing* **34**, 344–371.

Brown LG (1992) A survey of image registration techniques. *ACM Computer Surveys* **24**, 325–375.

Christenson CA, Rabbitt RD, Miller MI, Joshi SC, Grenander U, Coogan A et al. (1995) Topological properties of smooth anatomic maps. In *Information Processing in Medical Imaging* (ed. Bizais Y, Barillot C, Di Paola R), pp. 101–112. Dordrecht: Kluwer.

Colchester ACF, Zhao J, Holton-Tainter KS, Henri CJ, Maitland N, Roberts PTE et al. (1996) Development and preliminary evaluation of VISLAN, a surgical planning and guidance system using intra-operative video imaging. *Medical Image Analysis* **1**, 73–90.

Collignon A, Maes F, Delaere D, Vandermeulen D, Suetens P, Marchal G (1995) Automated multimodality image registration using information theory. In *Information Processing in Medical Imaging* (ed. Bizais Y, Barillot C, Di Paola R), pp. 263–274. Dordrecht: Kluwer.

Collins DL, Peters TM, Evans AC (1994) An automated 3D non-linear image deformation procedure for determination of gross morphometric variability in human brain. *Visualization in Biomedical Computing 1994*, SPIE **2359**, 180–190.

Cootes TF, Taylor CJ, Cooper DH, Graham J (1992) Training models of shape from sets of examples. In *British Machine Vision Conference 1992* (ed. Hogg D, Boyle R), pp. 9–18. London: Springer.

Cootes TF, Hill A, Taylor CJ, Haslam J (1994) The use of active shape models for locating structures in medical images. *Image and Vision Computing* **12**, 276–285.

Edwards PJ, Hill DLG, Little J, Hawkes DJ (1997) Deformation using a three component tissue model. In *Information Processing in Medical Imaging* (ed. Duncan J, Gindi G), pp. 218–231. Berlin: Springer.

Evans AC, Marret S, Collins L, Peters TM (1989) Anatomical-functional correlative analysis of the human brain using three dimensional imaging systems. *SPIE* **1092**, 264–274.

Freeborough PA, Fox NC (1997) Measurement of global and regional cerebral volume changes by integrating boundary shifts between registered serial 3D MR scans. In *Information Processing in Medical Imaging* (ed. Duncan J, Gindi G), pp. 355–368. Berlin: Springer.

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1997) Comparing functional (PET) images: the assessment of significant change. *Journal of Cerebral Blood Flow and Metabolism* **11**, 690–699.

Hajnal JV, Saeed N, Oatridge A, Williams EJ, Young IR, Bydder GM (1995) Detection of subtle brain changes using subvoxel registration and subtraction of serial MR images. *Journal of Computer Assisted Tomography* **19**, 677–691.

Hawkes DJ, Robinson L, Crossman JE, Sayman HB, Mistry R, Maisey MN et al. (1991) Registration and display of the combined bone scan and radiograph in the diagnosis and management of write injuries. *European Journal of Nuclear Medicine* **18**, 752–756.

Hemler PF, Van den Elsen PA, Sumanaweera T, Napel S, Drace J, Adler JR (1995) A quantitative comparison of residual errors for three different multimodality registration techniques. In *Information Processing in Medical Imaging* (ed. Bizais Y, Barillot C, Di Paola R), pp. 251–262. Dordrecht: Kluwer.

Hill DLG, Hawkes DJ, Crossman JE, Gleeson MJ, Cox TCS, Bracey EECML et al. (1991) Registration of MR and CT images for skull base surgery using point-like anatomical features. *British Journal of Radiology* **64**, 1030–1035.

Hill DLG, Hawkes DJ (1994) Medical image registration using knowledge of adjacency of anatomical structures. *Image and Vision Computing* **12**, 173–178.

Hill DLG, Hawkes DJ, Gleeson MJ, Cox TCS, Strong AJ, Wong WL et al. (1994*a*) Accurate frameless registration of MR and CT images of the head: applications in planning surgery and radiation therapy. *Radiology* **191**, 447–454.

Hill DLG, Hawkes DJ, Studholme C, Summers PE, Taylor MG (1994*b*) Accurate registration and transformation of temporal image sequences. *Proceedings Society of Magnetic Resonance* **2**, 830 (Abstract).

Hill DLG, Studholme C, Hawkes DJ (1994*c*) Voxel similarity measures for automated image registration. *SPIE* **2359**, 205–216.

Jiang H, Robb RA, Holton KS (1992) New approach to 3-D registration of multimodality medical images by surface matching. *SPIE* **1808**, 196–213.

Lemieux L, Kitchen ND, Hughes S, Thomas DGT (1994) Voxel based localisation in frame-based and frameless stereotaxy and its accuracy. *Medical Physics* **21**, 1301–1310.

Little J, Hill DLG, Hawkes DJ (1997) Deformation incorporating rigid structures. *Computer Vision and Image Understanding* **66**, 223–232.

Maciunas RJ (1993) *Interactive Image Guided Neurosurgery*. American Association of Neurological Surgeons.

Maintz JB (1996) *Retrospective registration of tomographic brain images*. *Ph.D*. Thesis. Netherlands: University of Utrecht.

Maurer CR, Fitzpatrick JM, Galloway RL, Wang ML, Maciunas RJ, Allen GS (1995) The accuracy of image-guided neurosurgery using implantable fiducial markers. In *Computer Assisted Radiology* (ed. Lemke H), pp. 1197–1202. Berlin: Springer.

Mcinery T, Terzopoulos D (1996) Deformable models in medical image analysis: a survey. *Medical Image Analysis* **1**, 91–108.

Meyer CR, Leichtman GS, Brunberg JA, Wahl RL, Quint LE (1995) Simultaneous usage of homologous points, lines and planes for optimal 3-D linear registration of multimodality imaging data. *IEEE Transactions in Medical Imaging* **14**, 1–11.

Pelizzari CA, Chen GTY, Spelbring DR, Weichselbaum RR, Chen C (1989) Accurate three dimensional registration of CT, PET and/or MR images of the brain. *Journal of Computer Assisted Tomography* **13**, 20–26.

Penney GP, Little JA, Hawkes DJ, Hill DLG (1997) 2D-3D image registration for use in image guided interventions. In *Medical Image Understanding and Analysis* (ed. Taylor CJ, Noble A, Brady M), pp. 49–52. University of Oxford.

Ruff CF, Hill DLG, Robinson GP, Hawkes DJ (1993). Volume rendering of multimodal images for the planning of skull base surgery. In *Computer Assisted Radiology*, *CAR'93* (ed. Lemke HU, Inamur K, Jaffe CC, Felix R), pp. 574–582. Berlin: Springer.

Ruff CF, Hughes SW, Hawkes DJ (1997) Volume estimation from sparse planar images using deformable models. *Proceedings of the British Machine Vision Conference* (ed. Clark AF), pp. 90–99. University of Essex: BMVA.

Spitzer VM, Whitlock D, Scherzinger AL, Ackerman MJ (1995) The visible human (male and female). *Radiology* **197**, 553.

Strother SC, Anderson JR, XU XL, Liow JS (1994) Quantitative comparisons of image registration techniques based on high resolution MRI of the brain. *Journal of Computer Assisted Tomography* **18**, 954–962.

Studholme C (1997) *Measures of 3D image alignment*. Ph.D. Thesis: University of London.

Studholme C, Hill DLG, Hawkes DJ (1995) Multiresolution voxel similarity measures for MR-PET registration. In *Information Processing in Medical Imaging* (*IPMI'95*) (ed. Bizais Y, Barillot C, Di Paola R), pp. 287–298. Dordrecht: Kluwer.

Studholme C, Hill DLG, Hawkes DJ (1996) Automated 3D registration of MR and CT images of the head. *Medical Image Analysis* **1**, 163–175.

Studholme C, Hill DLG Hawkes DJ (1997) Automated 3D registration of MR and PET brain images by multi-resolution optimisation of voxel similarity measures. *Medical Physics* **24**, 25–36.

Taylor RH, Lavallee S, Burdea G, Moesges R (1995) *Computer Integrated Surgery*. Boston: MIT Press.

Van den Elsen PA, Viergever MA (1991) Marker guided registration of electromagnetic dipole data with tomographic images. In *Information Processing in Medical Imaging* (ed. Colchester ACF, Hawkes DJ), pp. 142–153. Heidelberg: Springer.

Van den Elsen PA, Pol EJD, Viergever MA (1993) Medical image matching: a review with classification. *IEEE Engineering in Medicine and Biology* March, 26–39.

Van den Elsen PA, Pol EJD, Sumanaweera TS, Hemler PF, Napel S, Adler Jr (1994) Grey value correlation techniques used for automatic matching of CT and MR volume images of the head. *SPIE* **2359**, 227–237.

Weese J, Buzug TM, Lorenz C, Fassnacht C (1997) An approach to 2D/3D registration of a vertebra in 2D x-ray fluoroscopies with 3D CT images. In *CVRMed-MRCAS'97* (ed. Troccaz J, Grimson E, Moesges R), pp. 119–128. Berlin: Springer.

Wells, WM, Viola P, Atsumi H, Nakajima S, Kikinis R (1996) Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis* **1**, 35–51.

West J, Fitzpatrick JM, Wang MY, Dawant BM, Maurer CR, Kessler RM et al. (1997) Comparison and evaluation of retrospective intermodality image registration techniques. *Journal of Computer Assisted Tomography* **21**, 554–566.

Woods RP, Cherry SR, Mazziotta JC (1992) Rapid automated algorithm for aligning and reslicing PET images. *Journal of Computer Assisted Tomography* **16**, 620–633.

Woods RP, Mazziotta JC, Cherry SR (1993) MRI PET registration with automated algorithm. *Journal of Computer Assisted Tomography* **17**, 536–546.