Frameless registration of MR and CT 3D volumetric data sets

Rakesh Kumar Kristin Dana P. Anandan Neil Okamoto Jim Bergen*
Paul Hemler Thilaka S. Sumanaweera Petra A. van den Elsen John Adler †

Abstract

In this paper we present techniques for frameless registration of 3D Magnetic Resonance (MR) and Computed Tomography (CT) volumetric data of the head and spine. We present techniques for estimating a 3D affine or rigid transform which can be used to resample the CT (or MR) data to align with the MR (or CT) data. Our technique transforms the MR and CT data sets with spatial filters so they can be directly matched. The matching is done by a direct optimization technique using a gradient based descent approach and a coarse-to-fine control strategy over a 4D pyramid. We present results on registering the head and spine data by matching 3D edges and results on registering cranial ventricle data by matching images filtered by a Laplacian of a Gaussian.

1 Introduction

We present techniques for the registration of 3D Magnetic Resonance (MR) and Computed Tomography (CT) volumetric data of the head and the spine. CT provides a clear description of bony structures whereas MR provides more detail of soft tissues. In frameless radio-surgery [Schweikard94], the surgeon typically detects and locates the tumor precisely in MR. Registering the CT and MR data sets allows the location of the tumor to be mapped to the CT coordinate system. The CT data can then be used for automatic real-time registration with stereo x-ray video data during radio-surgery. Since the radiation gun used in radio-surgery is calibrated to the stereo x-ray coordinate system, registration of MR to CT coordinates and CT to x-ray coordinates enables precise location of the tumor during surgery. The techniques we present are general in nature and can be applied to other 3D multi-sensor data. In fact, these 3D registration techniques are an extension of 2D registration techniques developed by our group for registering images taken from infrared and visible light sensors [Dana93].

The early techniques for registering MR and CT head data sets relied on the placement of external frames or markers [Van den Elsen94] on the patients. The external frames are bolted on to the patient and are painful to the patient and cumbersome to the surgeon. Although marker based techniques do not suffer from these problems, they typically use only a few markers and are not as accurate as frameless registration techniques which attempt to register the entire data sets. Moreover, registration techniques based on external frames and markers are not easily extended to other parts of the body where the bone is buried deep inside the anatomy.

1.1 Our Approach

Most previous work on MR-CT registration demonstrates results using head data. The skull in the head data provides an excellent frame of reference for registration. Also, it is relatively easy to segment the thin layer of skin covering the skull from the CT and MR data respectively. This segmentation allows the use of point-to-surface fitting or surface-to-surface fitting schemes [Pelizzari89, Hemler94] for registration of 3D CT and MR head data sets. These techniques can be described as "hat fitting head" methods, since they rely purely on the shape of the skin surface. Grimson et.al. [Grimson94] have used a similar approach for registering MR/ CT data to 3D laser range data. Szelski et.al. [Szelski94] have recently extended this approach to handle non-rigid deformations. They are able to use their technique to match specific data from a patient to generic data from a human atlas map.

The "hat fitting head" approach, cannot be easily extended to other areas of the body. The skin at other parts of the body is connected to non-rigid tissue and cannot be used as a reliable feature for estimating a global affine or rigid transformation. The segmentation problem is significantly more difficult when applied to other anatomical structures of the body where the bone does not lie on the outer boundaries of the image data. Moreover, the surface fitting approach does not use all the available data and hence may be limited in accuracy. For instance, if the head MR and CT data are acquired close in time, the ventricles may be used for registration. In Section 4 we present 3D registration results where we used only the volume data which contains the ventricles. In order to register the head data using just the ventricles with the surface fitting approach, the more difficult problem of segmentation of ventricles would have to be solved first.

Our registration approach recovers information concerning the spatial structures in the data from multiple parts of the data sets (e.g. for the head data sets not only the outer skin surface, but bone, and ventricle data as well), and uses all of this data for registration. The use of structures from many parts

*Rakesh Kumar, Kristin Dana, P. Anandan, Neil Okamoto and Jim Bergen are with the David Sarnoff Research Center, CN5300, Princeton, NJ-08530.

†Paul Hemler, Thilaka S. Sumanaweera, Petra A. van den Elsen and John Adler are with the Stanford University Medical Center, Stanford University, Palo Alto, California.

0-8186-6410-X/94 $4.00

of the data volume leads to a better conditioned estimation problem and more accurate results. In order to recover the variety of spatial structures, we apply different "feature extractors" for the different parts of the data sets. These features range from simple normalized laplacian of gaussian responses to measures of edge presence, strength, orientation, curvature and other simple local properties of the intensity surface. Each feature extractor typically consists of a linear filter applied to the data set at one or more scales, and also possibly a simple non-linear operator that extracts a measure of the feature at each voxel. The resulting representations are iconic and fit naturally into our gradient-based registration techniques. The filters can be tuned to different sections of the data sets based on anatomical knowledge and the typical responses of various anatomical structures in the different sensors. In [Van den Elsen92] ridge like spatial structures associated with the intensity surface in the bone area around the skull were used for registration. Our approach provides a natural way to incorporate a variety of filters all operating simultaneously with matching done at various frequency scales in a coarse-to-fine manner. For matching the feature maps, we do not attempt to solve a combinatorially explosive feature matching problem rather we treat the 3D feature maps as signals and do a direct signal-to-signal registration of the maps. This approach also does not require us to solve the difficult segmentation problem, which is required for the hat fitting head approaches.

Any registration technique has three chief characteristics [Brown92]: (1) geometric transformation model (2) match measure and (3) search or matching method. We discuss our proposed approach in this context. The geometric transformation model we initially assume is a global rigid or 3D affine transformation. We will assume that the geometric distortions in the MR data have been removed by the technique developed at Stanford [Sumanaweera93]. The 3D affine model is reasonable for head data because the structures are relatively rigid. In the case of registering the entire spine, our goal is to extend the model to a piece-wise rigid or affine model. For instance, a separate rigid transformation model may be computed for the volume around each vertebra. In this paper, we present results for this single vertebra registration.

The match measure we use is the sum of squares of differences (SSD) between the transformed multisensor data sets. The matching technique used is the same regardless of the filter applied to the data. To estimate the registration parameters, the filtered data can be matched by hierarchical direct correlation. However, this can lead to a very expensive computational search problem when solving for 6 or 12 unknown parameters over the large medical data sets. Instead we use an iterative regression-based technique [Bergen92]. The search is implicitly done by a direct optimization technique using a gradient based descent approach and a coarse-to-fine control strategy. In section 2 we describe our matching method in more detail. In section 3, we describe how we use edge detection as an intensity transformation and in Section 4 we present results for registering MR and CT data sets of the head and the spine using edge data and normalized Laplacian of Gaussian responses.

2 Hierarchical Parameter Estimation for Registration

Any registration technique can be decomposed into two parts: (1) what to register and (2) how to register. This section is about "how to register". We assume for now that the two data sets have been transformed (filtered), so that the filtered data sets can be directly correlated. Let the two filtered data sets be referred to as the reference data set (I_r) and the inspection data set (I_i) respectively. We wish to solve for a global set of transformation (motion) parameters which can be used to resample or warp the inspection data set so that it exactly aligns with the reference data set. The global set of transformation parameters could be any parametric transformation (e.g. 3D affine or 3D rigid). Note we have also used variations of this general pyramid based registration technique for estimating quasi-parametric and non-parametric transformations in 2D data sets [Bergen92]. The registration technique we present is non-linear and requires an initial estimate.

The registration technique we present is based on minimizing the SSD error between the two data sets. First, a 4D pyramid [Burt88] is constructed from each of the two input data sets, and then the motion parameters are estimated in a coarse-fine manner. Within each pyramid level the sum of squared difference (SSD) measure is used as a match measure. This measure is minimized with respect to the motion field parameters. At each pyramid level, the solution from the previous pyramid level is used as an initial estimate. Using this initial estimate, the inspection data set is resampled. The new incremental estimate is computed by minimizing the SSD error between the resampled inspection data set and the reference data set. The estimated increment is then composed with the previous estimate to achieve a new estimate. This whole process is iterated many times at each pyramid level and then across pyramid levels to achieve the final estimation.

The SSD error measure for estimating the flow field within a region and pyramid level is:

$$E(\{\mathbf{u}\}) = \sum_{\mathbf{x}} (I_i(\mathbf{x}) - I_r(\mathbf{x} - \mathbf{u}(\mathbf{x})))^2 \qquad (1)$$

where $\mathbf{x} = (x, y, z)$ denotes the spatial position of a point. $\mathbf{u}(\mathbf{x}) = (u(x, y, z), v(x, y, z), w(x, y, z))$ denotes the motion at that point. The sum is computed over all the points within the region and $\{\mathbf{u}\}$ is used to denote the entire flow field within that region. The motion field due to a 3D affine or rigid motion can be represented as:

$$\begin{bmatrix} u(\mathbf{x}) \\ v(\mathbf{x}) \\ w(\mathbf{x}) \end{bmatrix} = A \begin{bmatrix} x \\ y \\ z \end{bmatrix} + T \qquad (2)$$

where A is a 3x3 matrix and T is a 3x1 matrix. If the motion is rigid, then A must be an orthonormal matrix and T is the translation matrix. The motion field for

the 3D affine transformation can be compactly written as:

$$\mathbf{u}(\mathbf{x}) = \mathbf{X}(\mathbf{x})\mathbf{p} \tag{3}$$

where $\mathbf{p}$ denotes the twelve parameter vector:

$$\mathbf{p} = (T_1, A_{11}, A_{12}, A_{13}, T_2, A_{21}, A_{22}, A_{23}, T_3, A_{31}, A_{32}, A_{33})$$

and

$$\mathbf{X}(\mathbf{x}) = \begin{bmatrix} 1 & x & y & z & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & x & y & z & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & x & y & z \end{bmatrix}$$

In this form, it is immediately obvious that the motion field is linear in terms of the unknown twelve parameter vectors. However, the filter magnitude functions I_r and I_i are non-linear in terms of position; hence, the overall objective function is also non-quadratic.

To estimate the parameter vector $\mathbf{p}$ we use the Gauss-Newton optimization technique. Let $\mathbf{p}_i$ denote the current estimate of the 3D affine parameters. After using the flow field represented by these parameters in the resampling step, an incremental estimate $\delta\mathbf{p}$ can be determined. To achieve this, we insert the parametric form of $\delta\mathbf{u}$ into Equation 1, and obtain an error measure that is a function of $\delta\mathbf{p}$.

$$E(\delta\mathbf{p}) = \sum_{\mathbf{x}} \left(\Delta I + (\nabla I)^T \mathbf{X}\delta\mathbf{p}\right)^2 \tag{4}$$

where $\Delta I = (I_i(\mathbf{x}) - I_r(\mathbf{x} - u(\mathbf{x}))$ and ∇I is the spatial gradient of the intensity volume I_r. Minimizing this error with respect to $\delta\mathbf{p}$ leads to the equation:

$$\left[\sum \mathbf{X}^T(\nabla I)(\nabla I)^T\mathbf{X}\right]\delta\mathbf{p} = -\sum \mathbf{X}^T(\nabla I)(\Delta I). \tag{5}$$

The incremental estimate δp is computed and added to the current estimate $\mathbf{p}_i$ to obtain a new estimate $\mathbf{p}_{i+1}$. The inspection data set is resampled using this new estimate and the whole process iterated for a fixed number of iterations or until there is convergence.

In the case of estimating a global rigid transformation, at each iteration we solve for three incremental rotation parameters $(\Omega_x, \Omega_y, \Omega_z)$ and the incremental translation parameters $(\delta T_x, \delta T_y, \delta T_z)$. The global rotation parameters may be represented by a 3x3 orthonormal matrix or by quaternions. The incremental motion estimate at any point is given by:

$$\delta\mathbf{u} = \mathbf{\Omega}\mathbf{x} + \delta\mathbf{T} \tag{6}$$

where the incremental rotation Ω is represented by a skew-symmetric matrix:

$$\mathbf{\Omega} = \begin{bmatrix} 0 & -\Omega_z & -\Omega_y \\ \Omega_z & 0 & -\Omega_x \\ \Omega_y & \Omega_x & 0 \end{bmatrix}$$

The above expression for $\delta\mathbf{u}$ is linear in terms of the incremental rotation and translation parameters. We substitute the above expression for $\delta\mathbf{u}$ into equation (1) to obtain an error expression equivalent to the equation (4) derived for the 3D affine case. This expression is minimized to solve for the rotation and translation increments. The rotation increment is converted into a quaternion representation which can then be composed with the current rotation estimate to find the new estimate [Kumar94].

The Gauss-Newton technique is a second order minimization technique which leads to quick convergence. However at times it can diverge from the correct solution. To ensure descent the Gauss-Newton technique minimization is very easily modified to be the Levenberg-Marquardt technique. We found this necessary when matching CT and MR data sets which had been filtered by the 3D Laplacian of a Gaussian.

The above registration technique is a least-squares based approach and is sensitive to large outliers. However, as reported earlier [Bergen92] this sensitivity is reduced by performing the least-squares estimation over a pyramid. The pyramid based approach locks on to the dominant motion in the scene. In the results section for head data, we show how the estimation is not affected by the non-rigidity of the ears. However, in the case of large outliers, instead of minimizing the square of the error function in equation (1) a robust version of the error function may be minimized [Kumar94].

3 Three Dimensional Edge Based Registration

3.1 Edge-based Registration

Although there are intrinsic differences in the imaging modalities of CT and MR, certain image features can be commonly identified. For example, discontinuities in certain structures lead to intensity discontinuities, or edges, in both imaging modalities. In data sets of the human head, skin and bone boundaries give rise to intensity boundaries in both CT and MR. In data sets of the human spine, the boundaries of the vertebrae are visible in both data sets. Not all intensity discontinuities are common between the two data sets. Because of intrinsic differences in the sensors, edge strength cannot be assumed to be the same. In addition, the transformation to match edge strengths varies between anatomical structures and can be quite complex. Therefore we rely only on the presence of a common subset of edges for edge-based registration. Preparing the data for registration requires the following steps: edge detection, edge pruning, and edge normalization.

3.2 Edge Detection

Edge detection is done in three dimensions. [Hemler94] detects 2D edges in each slice of the data volume and stacks these edge images to obtain the 3D edge volume. This approach unnecessarily depletes the information used for registration because edges along one axis are ignored. We use an edge focusing approach similar to [Bergholm87] but extended to three dimensions. Edge focusing coordinates blurring and edge detection to produce well localized edges that are likely to correspond to the true volume structure. Finally, we combine edge focusing and volume pyramids into an efficient edge detection approach.

To detect edges we first compute the gradient vector. The gradient in the each direction is obtained by filtering the data with the 3x3x3 central difference filter. The local maximum in the gradient direction is identified as the location of the edge element or edgel. Local maxima are located to subvoxel accuracy by fitting a quadratic function to the gradient magnitude in the gradient direction.

When using edge focusing techniques, edges are detected over the entire volume only at coarse resolution where there are very few false edges (since the signal to noise ratio of the data set is relatively high). At finer resolutions, the neighbors of these edges are searched in order to find better localized edges (see [Bergholm87]). The edge focusing approach requires filtering the volume with low-pass filters of increasingly high cut-off frequencies so that the data goes through incremental changes in scale space. When the steps in scale space are small, there is a high probability that an edgel will change position by at most one voxel. The initial scale is chosen to give a reasonable trade-off between noise reduction and loss of volume structure due to blurring. [Bergholm87] uses Gaussian blurring with the standard deviation σ decreasing in increments of 0.35 and a typical initial scale corresponding to $\sigma = 4$.

Our approach to edge detection uses Gaussian pyramids [Burt88] for the edge focusing process. Specifically, we use volumetric pyramids (a 3D extension of image pyramids) to step through scale. The primary advantage of using pyramids in edge focusing is the computational efficiency obtained from subsampling. Since pyramid levels are relatively large steps in scale space, the assumption that the edgel can move only one voxel is no longer valid and a larger neighborhood of voxels that must be searched. In practice, a 5x5x5 search range at each level yields reasonable results. This increased search neighborhood leads to an increase in "noise edges". In practice, this increase seems to have a minimal effect on registration. (A hybrid approach can be used where the data undergoes incremental changes in scale within each pyramid level. This approach would provide computational efficiency without increased noise edges.) The efficiency obtained by using pyramids with edge focusing is especially important since the MR and CT data sets can be quite large.

A significant benefit of using 3D edge focusing is the reduced dependence on thresholds. Coarse-to-fine tracking is used instead of thresholding for most of the noise reduction. Although we often find it useful to use some threshold when tracking edges through scale space, the final edges are less dependent on the value of the threshold. Specifically, the threshold only controls the noise around "true" edges but not throughout the volume.

3.3 Edge Pruning

Some of the edges that are obtained during edge detection have no match in the other data set and should not be used in computing the match measure. Other edges may be strong in one data set and weak in the other set. These edges are also not suitable for matching because the match would not be robust to noise. However, there exists a common subset of edges in each of the data sets that are suitable for matching. In the CT and MR data sets of the human head, we have identified three edge types that are suitable to use for registration: the boundaries of the skull, skin and brain. In the CT and MR data sets of the human spine, the vertebrae boundaries are used for registration.

The determination of whether a particular edge corresponds to one of the desired types of boundaries is made by analyzing the local intensity distribution around the edges. If the calibration parameters of the sensors are fixed, different types of anatomical structures map to specific intensity distributions. Given this knowledge, we can determine if an edge belongs to one of the acceptable categories of boundaries (e.g., skin-air, bone-tissue, etc.) by verifying that the intensities on either side of a 3D edge are roughly within the expected ranges. The remaining edges are then pruned away as being unreliable for registration.

In order to compute the intensity distributions on either side of the 3D edge, we use the tangent plane to the edge boundary surface to split a cube of 5x5x5 voxels around the edge. This tangent plane is simply the plane perpendicular to the 3D gradient vector. On each side of this plane, the intensity is summed. Note, voxels in the intensity transition region are not included in the summation. If the summed intensities on each side of the boundary plane are within the predetermined intensity range for each edge type, the edgel is classified as belonging to that type.

In general, this classification approach works very well for CT data, because the mapping from the structures of interest to the intensity ranges is unique. Also, the ranges can be specified with small uncertainties. On the other hand, in MR data, different structures (e.g., bone and air) map to similar intensity ranges. In addition, the uncertainty in the range of each structure is also larger, requiring greater tolerance. However, since we do not require a distinction between skin-air and bone-tissue boundaries, but rather a determination whether the edge is in either of these categories, the approach of analyzing intensity distributions on either side of the edge works reasonably well even with the roughly specified intensity ranges. Pruning is important to get accurate registration results because there are often many edges that are not common between the data sets. Retaining only the common edges by pruning prevents the optimization-based registration algorithm from falling into a local minimum.

3.4 Edge Normalization

Since the strength of corresponding edges between the data sets cannot be assumed to be the same, it cannot be used as part of the match measure. A local linear transformation of the gradient magnitude may bring the edge strengths to closely matching ranges. However, since the edge strength mapping between the data sets varies with anatomical structure it is not clear that this is the correct approach. We have chosen to use binary edges so that the only likeness assumed is the presence of an edge.

The iterative gradient based approach for registration is a descent method which requires information to lock onto when the edges are misaligned. This approach requires blurring of binary edges before regis-

tration. Finite difference approximation to the gradient is a poor approximation when applied to binary edges. As the edge volume is blurred, the finite difference approximation to the gradient operator becomes more accurate. In our current implementation, we blur edges with a Gaussian 5x5x5 tap filter. In future implementations, we plan to blur the edges just along the gradient direction. Since the edges may be at subpixel locations, it is convenient to apply blurring filter at the edge detection stage.

4 Results

In this section, we present results obtained for registering two sets of head data and one set of spine data. A fair amount of research and experimental effort is put into setting the parameters of the MR and CT sensors when acquiring these data sets. The reader is referred to the papers [Hemler94] and [Van den Elsen92] on the details of how the imaging protocols for the MR and CT machines were chosen.

4.1 3D MR-CT registration of head using only ventricle data

We present results (Figures 1-4) on registering parts of the head data using just the volumes around the ventricle area without including bony regions. For this data set, we did not know the calibration parameters (such as voxel size) of either the MR or the CT data sets. We therefore did the registration assuming a global 3D affine transformation to compensate for differences in voxel size. The 3D MR and CT volumes we used were each 50 x 80 x 12 voxels. The boxed areas in the MR images (in Figures 1 and 2) show the volume used. We filtered each of the data sets with a 3D Laplacian filter and built a 4D pyramid. These 4D pyramids were further processed by a nonlinear adaptive local gain control filter to match the gray level values of the two data sets. These filtered 3D pyramids were then matched using the technique described in Section 2. The goal of the matching process was to derive a 3D affine (12 parameters) which warps the CT data so that it exactly aligns with the MR data. An initial 3D affine estimate was provided by clicking on 6 pairs of corresponding points in the two data sets.

The CT 3D data was warped using the computed 3D affine transform to align with the MR data. We demonstrate the 3D registration results by showing the registration obtained for one slice in the 12 slice data set. Fusion is a process which combines two images to form a single image by selecting salient features from each image [Burt93]. We use it here to show the quality of our registration results. If the registration is poor then "ghost" structures would be seen in the fused image. Figures 1 and 2, show slice 3 of the input MR and CT data respectively. Figure 3 and 4 show the MR image warped with the estimated affine parameters and the fused CT and warped MR image respectively. The fused image in Figure 4 shows both the CT bone outline and MR inner detail. and indicates that our 3D registration results are quite good.

4.2 3D MR-CT registration of head data using skin and bone edges

In this subsection results are presented on registering the entire head data by matching edges detected in the MR and CT data sets obtained from a single patient[1].

We scaled both the data sets in order to have equal size voxels (1.55mm) in each of the three dimensions. The resulting scaled CT and MR data were of dimensions (155 x 155 x 100) and (161 x 161 x 100) voxels respectively. We then detected 3D edges to subpixel accuracy in the CT and MR data sets respectively (see Section 3). The 3D edge sets were pruned using simple classification to the edges which are either on the skin-air interface or bone-tissue interface. Note in MR we cannot distinguish between the skin-air interface edges from the bone-tissue interface edges. We do not use this classification for matching; it is only used to prune the large number of edges detected. The pruned edges are then blurred by a 5 tap filter. The resulting blurred edge maps can be seen in Figure 5. These blurred CT and MR edge maps were then registered using the technique described in Section 2. The registration algorithm was given an initial estimate of zero for all parameters and was run in a coarse to fine manner from level 3 to level 0 of the 4D pyramid with 10 iterations at each level. The MR data was resampled using the computed 3D affine parameters determined by the registration algorithm. In Figure 5, the reader can view from top row to bottom row: the original CT data, the CT edge maps, the original MR data, the MR edge maps and the resampled MR data for two slices. Figure 6 shows the registration result in the axial view for 2 slices. This figure shows a rectangular window of the warped MR data spliced into the original CT data. The reader can note the quality of the registration by seeing how well the anatomical features align in the spliced images. For instance for slice 43, one can see an alignment of the ventricles (black area in the center of the image) between the CT and the warped MR. As can be seen in the figure, the original CT and MR were quite different from corresponding slices. In fact, the final registration warped the MR data by 38 voxels along the z-axis. The gray values in the CT data range from 0 to 2500 approximately. The gray values for the soft tissue (brain matter) in CT however lie in a very narrow range from 925 to 1075. In Figure 6,

[1] "The CT scan was obtained on a Philips Tomoscan 350, at tube settings of 120 kV and 200 mA. The scan was made without gantry tilt. The set contains 100 contiguous slices (no gap between the slices) with slice thickness of 1.55 mm. Field of view is 240 mm, and each slice is a 256x256 matrix of pixels, rendering the pixel size 0.9375 mm. MRI was performed on a Philips Gyroscan S15/ACS, a 1.5 Tesla machine. A transverse T1 weighted 3D FFE series was acquired with one acquisition (TR/TE 30/9 msec). Slice thickness is 1 mm. The slices are contiguous, there is no gap present. Each slice is a 256x256 matrix of pixels, and field of view is 250 mm, rendering the pixel size 0.9765625 mm. Flip angle is 30 degrees, water-fat shift 0.50 pixels, and scan percentage is 80%. The upper 10 and lower 10 slices (with very low signal-to-noise ratio) of this originally 200 contiguous slices containing data set were removed, resulting in 180 slices" [Van den Elsen92].

the CT data is shown with the gray values that were clipped to lie between 925 and 1075. Therefore, the appearance of the bone in the CT images is saturated and the bone appears somewhat thicker than it is. Figure 7 shows the registration results by resampling the CT and warped MR along the sagittal view. We then spliced in a triangular wedge of warped MR data into the CT data. In Figure 7, we do window into the CT gray value range and hence the bone is not saturated, however no inner details of the soft tissue (brain matter) are visible. From the spliced image shown in Figure 7, one can note how well the anatomical structures (skin and bone) align after registration.

4.3 3D MR-CT Registration of a Single Vertebra using Edges

In this subsection, registration results are presented for a single vertebra. The registration is done by matching edges detected in the MR and CT data sets. The spine is a non-rigid structure; therefore a single affine transformation would not be appropriate to register the data sets. By isolating a single vertebra from the rest of the spine we are able to assume rigidity and register using an affine model. For the results of this paper we manually isolate a vertebra by choosing a rectangular region around the vertebra as shown in Figure 8. In future work we intend to approximately register the entire spine first, so that choosing corresponding vertebrae will be simplified. The spine data is shown in Figure 9 [2].

We scaled both the data sets in order to have equal size voxels (1.2 mm) in each of the three dimensions. The resulting scaled CT and MR data (after isolating one vertebra) were both of dimensions (60 x 30 x 60) voxels. We then detected 3D edges to subpixel accuracy in the CT and MR data sets respectively (see Section 3). For both CT and MR, we used edge focusing starting at pyramid level 1. In MR we used edge classification to choose only the vertebra boundaries (without edge classification the edges from the spinal cord were also present). In CT, edge classification was not necessary since the vertebra boundary was essentially the only boundary present. The edges were blurred with a 5x5x5 gaussian-like filter. The resulting blurred edge maps are shown in Figure 9. These blurred CT and MR edge maps were then registered using the technique described in Section 2. The registration algorithm was given an initial estimate of all zero values and the algorithm was executed only at level 0 of the pyramid with 30 iterations. Coarser levels of the pyramid were not used to avoid merging data from neighboring vertebrae. The MR data was resampled using the computed 3D affine parameters determined by the registration algorithm. Figure 9 shows the from top row to bottom row: the original CT data, the CT edge maps, the original MR data, the MR edge maps and the resampled MR data for two slices. Figure 10 shows the registration result by illustrating a rectangular window of the warped MR data spliced into the original CT data. The reader can note the quality of the registration by seeing how well the vertebra boundaries on both sides of the spinal cord align in the spliced images. Notice that most boundaries line up quite well. Note that in slice 30 of the MR data the vertebra boundary and the leftmost spinal cord boundary are quite close and it is difficult to distinguish the vertebra boundary. It is clear when viewing Figure 9 that the original CT and MR data are significantly misaligned. In particular, there is a relatively large rotation (approximately 10° about the y and z axes).

[2] The CT data was accquired in axial format with the in-plane resolution of of 0.3125 mm by 0.3125 mm (512 x 512 pixels per slice). A total of 253 slices were accquired at a resolution of 0.5 mm. The MR data was accquired in sagittal format with the in-plane resolution of of 0.78125 mm by 0.78125 mm (256 x 256 pixels per slice). A total of 60 slices were accquired at a resolution of 1.2 mm with 0 mm slice spacing. For further details about the acquisition of this data see [Hemler94]

Acknowledgements

This research is supported by a grant from the Murray Foundation, Princeton, New Jersey. The registration algorithms presented in this paper were developed at the David Sarnoff Research Center, Princeton, New Jersey. The research on developing the appropriate imaging protocols for the MR and CT machines was done at Stanford University. The techniques and complete overall system for frameless radiosurgery are being developed at Stanford University by Dr. Adler and his group. The spine medical data and the first head medical data sets were collected at Stanford University. The volumetric head MRI data (Figure 5) was acquired in the Hospital "Medisch Spectrum" in Enschede, NL, with special thanks to Dr. G. Wilts. The high resolution CT (Figure 5) was acquired in Utrecht University Hospital, NL, by Prof. Dr. F.W. Zonneveld, Drs. D.N. Velis and P.A. van den Elsen coordinated the multimodal acquisition.

References

[Schweikard94] A. Schweikard, Tombropoulos, R., Adler, J. and J.C. Latombe, "Planning for image-guided radiosurgery," Working Notes, AAAI Spring Symposium Series, Stanford University, March 1994.

[Bergen92] Bergen, J., Anandan, P., Hanna, K.J., Hinograni, R., "Hierarchical Model-Based Motion Estimation", Proceedings of European Conference on Computer Vision, G. Sandini, Ed., Springer-Verlag, 1992.

[Bergholm87] Bergholm, F., "Edge Focusing", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 9(6):726-741, November 1987.

[Brown92] Brown, L., "A survey of image registration techniques," ACM Computing Surveys, 24(4):325-376, December 1992.

[Burt88] Burt, P., "Smart Sensing within a Pyramid Vision Machine", Proceedings of the IEEE, Vol. 76, No. 8, Aug., 1988.

[Burt93] Burt, P., "Image Fusion", Proceedings of the International Conference on Computer Vision, Berlin, May, 1994.

[Dana93] Dana, K. and Anandan, P., "Registration of visible and infrared images", SPIE Conference on Architecture, Hardware and Forward Looking Infrared Issues in Automatic Target Recognition, May, 1993.

[Van den Elsen92] P. A. Van den Elsen, Maintz J. B. A, Pol E. J. D and Viergever M. A, "Image fusion using geometrical features," In: Robb R. A (Ed): Proc SPIE Vol 1808 Visualization in Biomedical Computing, SPIE Press, Bellingham, WA, pp. 172–186, 1992.

[Van den Elsen94] P. A. van den Elsen and M.A. Viergever, "Marker-guided multimodality matching of the brain," European Radiology, Vol. 4(1), pp45–51, 1994.

[Grimson94] E. Grimson, T. Lozano-Perez, W. Wells, G. Ettinger, S. White and R. Kikinis, "Automated registration for enhanced reality visualization in surgery," Working Notes, AAAI Spring Symposium Series, Stanford University, March 1994.

[Hemler94] Paul F. Hemler, T. Sumanaweera, R. Pichumani, P. A. van den Elsen, S. Napel, J. Drace, J. Adler, "A system for multimodality image fusion of the spine", Working Notes, AAAI Spring Symposium Series, Stanford University, March 1994.

[Kumar94] R. Kumar and A. R. Hanson, "Robust methods for estimating pose and a sensitivity analysis," *Computer Vision, Graphics and Image Processing: Image Understanding (CVGIP-IU)*, Volume 60, Number 3, November, 1994.

[Pelizzari89] Pelizzari, C.A., et. al., "Accurate Three-Dimensional Registration of CT, PET and/or MR Images of the Brain", Journal of Computer Assisted Tomography, 13(1):20-26, Jan./Feb. 1989.

[Sumanaweera93] Sumanweera, T. S., Glover, G.H., Binford, T., and Adler, J.R., "MR Susceptibility Misregistration Correction", IEEE Transactions on Medical Imaging, Vol. 12, No. 2, June 1993.

[Szelski94] Szelski, R. and Lavallee, S., "Matching 3-D anatomical surfaces with non-rigid deformations using octree splines," *Proceedings of the IEEE Workshop on Biomedical Image Analysis*, Seattle, Washington, June 24-25th, 1994.

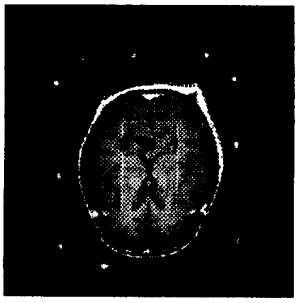

Figure 1: **MR data: Slice 3 of 12.**

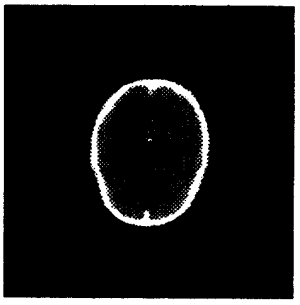

Figure 2: **CT data: Slice 3 of 12**

Figure 3: **Warped CT data: Slice 3 of 12.**

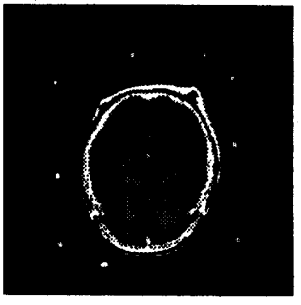

Figure 4: **Fusion of MR and Warped CT data, Slice 3 of 12.**

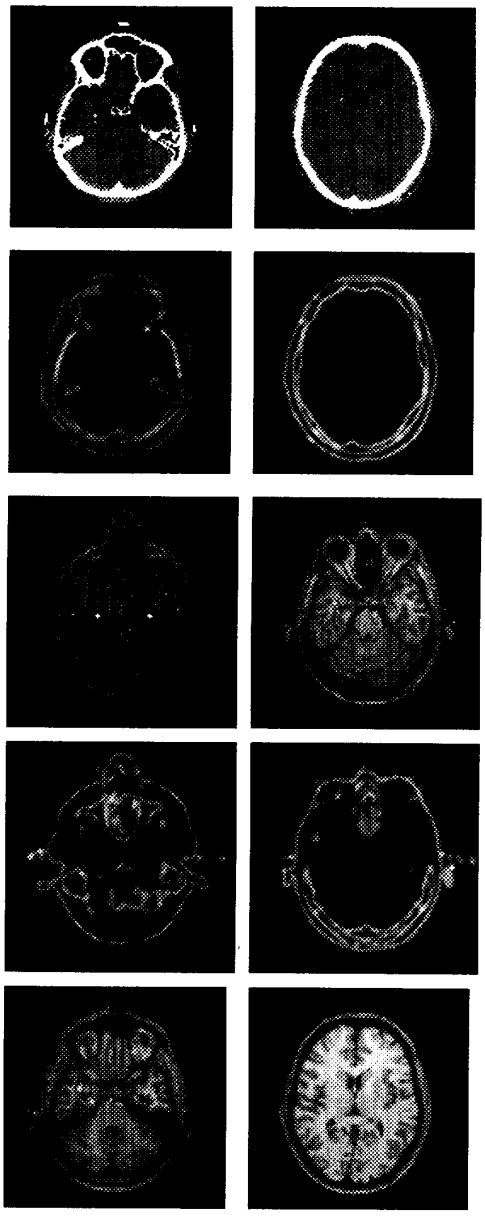

Figure 5: (Top - Bottom) CT head data, corresponding CT edges, MR head data, corresponding MR edges, warped MR data (registered to CT). (Column 1) results for slice 21. (Column 2) results for slice 43.

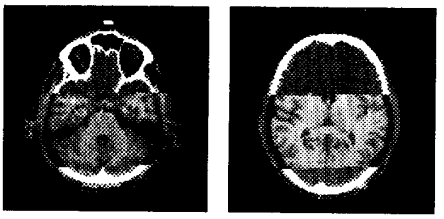

Figure 6: Splice of registered MR and CT head data (slices 21 and 43) to show registration.

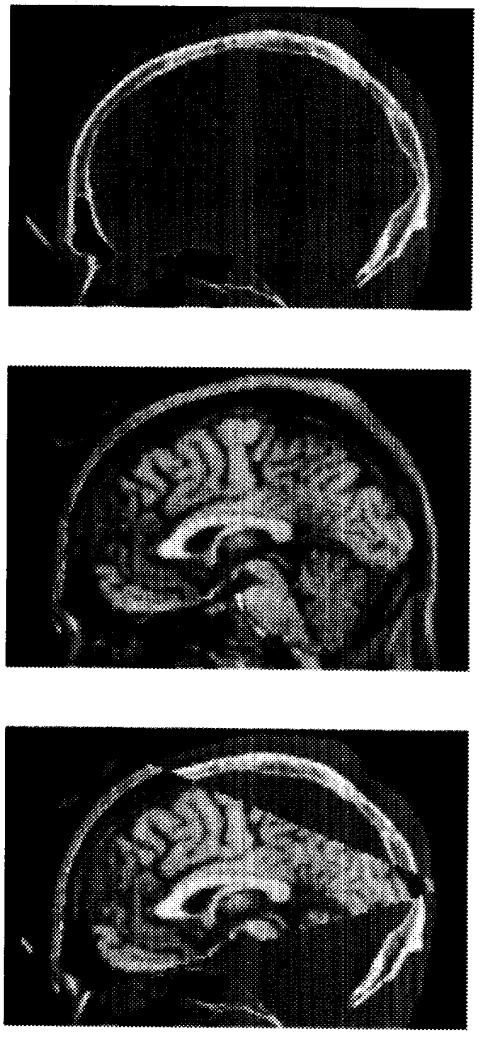

Figure 7: (Left) CT data, sagittal view. (Middle) registered MR data, sagittal view. (Right) splice of CT and MR to show registration.

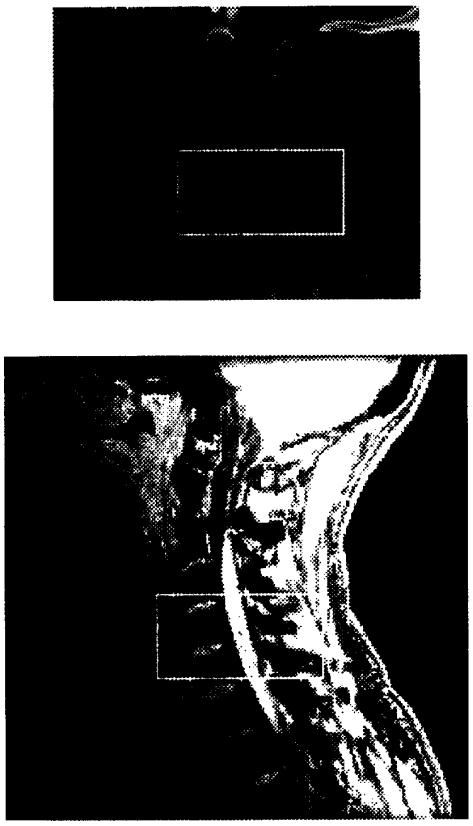

Figure 8: Section of data that was used for the registration in Figure 10.

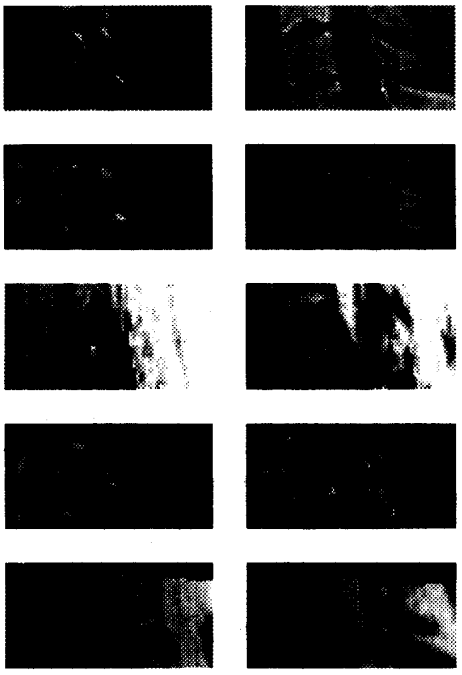

Figure 9: (Top - Bottom) CT spine data, corresponding CT edges, MR spine data, corresponding MR edges, warped MR data (registered to CT). (Column 1) results for slice 20. (Column 2) results for slice 30.

Figure 10: Splice of registered MR and CT spine data (slices 20 and 30) to show registration.