Knowledge-Based Deformable Matching for Pathology Detection

Thesis Proposal

Mei Chen

CMU-RI-TR-97-20

The Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

May 1997

This research was partly sponsored by: ASRI-AGH/NIST National Medical Practice Project.

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. government.

Abstract

Image registration is the process of matching corresponding features between images. Medical images present a challenge for registration because the same anatomical features vary considerably in appearance across individuals, and any pathology may aggravate the variation.

We propose a registration method that uses knowledge of anatomy to cope with variations in the appearance of corresponding features and to help detect pathologies. The same anatomical feature may have a different density, shape, and size across individuals, but there exists a normal range of variation. Various clinical conditions may affect the gross anatomical morphology, but certain pathologies have characteristic impacts on the anatomy. Traditional registration methods do not seek guidance from this domain knowledge, and they fail to yield accurate registrations because medical images have complicated shapes, non-uniform textures, and ill-defined boundaries. Our algorithm will attempt to apply domain knowledge in the registration process to cope with the variations in appearance. We are especially interested in the registration of images showing pathologies because of its medical importance. In pursuit of this, we have developed a hierarchical registration algorithm that permits variations in appearance. We will infuse domain knowledge to discern normal and abnormal variations in appearances.

The hierarchical registration algorithm has three levels. First a 3-D transformation (rotation, uniform scaling, and translation) globally aligns the different image sets. Secondly a smooth 3-D deformation approximately matches the anatomical structures in the image sets. This attempts to address the inherent differences in the appearances of their anatomical structures. Finally, the match is refined by a free-form deformation to adjust to the small variations overlooked by the smooth deformation. We have developed this algorithm and conducted experiments on real data sets with and without pathologies.

The results from the above registration algorithm are promising, but the algorithm has some limitations. We intend to extract and encode domain knowledge (e.g. using Principal Component Analysis), and apply it to guide the registration process (only permitting normal variations or characteristic impact caused by certain pathologies). To handle intensity variation, we plan to use mutual information as the matching criterion. We will establish a standard validation method and test database to quantitatively evaluate the accuracy of our algorithm and related work.

The system will help in detecting lesions, observing the development of pathologies over time, indexing and retrieval in medical databases, conducting cross-patient analysis, and studying the functionality of different parts of the brain.

Table of Contents

Knowledge-Guided Deformable Matching for Pathology Detection

Mei Chen (meichen@cs.cmu.edu)

1. Motivation

Image registration, the process of matching corresponding features across multiple image sets, has long been an essential and active area in computer vision research. It involves two interdependent problems: identifying the relevant features, and matching them across different data sets. These tasks are difficult when the features have considerable variations in appearance. One such example is human anatomy. Because of genetic and environmental factors and because of disease, biological structures have a large range of variation. People share a common anatomy, but no two persons have identically shaped body structures. Since the manual registration of anatomical structures is labor-intensive and due to lack of consistency, many researchers are investigating automated and accurate registration methods. There have been encouraging results, both in intensity-based and feature-based deformable registration that attempt to address the variation between people. However, due to the complexity of the problem and the absence of encoded knowledge to guide the registration, it remains a challenge.

This proposal presents a method for *Knowledge-guided deformable matching.* This is a registration process that has knowledge enabling it to cope with variations in the appearance of corresponding features. The primary application we focus on is neuroscience. Images that reveal internal brain structures are crucial for correct diagnosis and accurate radiotherapy or surgical planning. A major problem in the cross-patient analysis of brain images is morphological variability. Aside from the normal variations, various neurological conditions affect the gross anatomical shape of the brain. For example, pathologies like bleeding, or tumors, can cause a shift of brain structures called mass effect, and there could be large intensity differences around the lesion. A *knowledge-guided deformable matching system* intends to accurately register normal brain anatomy and to indicate abnormalities that change the brain structure. It will provide a precise and convenient way for doctors to analyze variations between normal brains, to detect lesions that make brain structures deviate from the norm, and to plan surgeries. It will also facilitate image-based retrieval of similar cases in medical databases, assist doctors in comparing different pathologies' impact on brain morphology to observe the development of a pathology over time, and to study the functions of different parts of the brain.

2. Problem Definition

Considering the human head as a three dimensional volume, the task of registration is to extract and match the corresponding structures from different volumes. Registration may be performed on either a single imaging modality, or on multi-modal data, including: *computerized tomography* (CT), *positron emission tomography* (PET), and *magnetic resonance imaging* (MRI). In this document, we concentrate on the registration within T1 weighted MRI data.

MRI images are good at revealing soft tissue structures.

In neurosurgery, the three principal axes of the head are called *axial*, *sagittal*, and *coronal* (see Figure 1). A set of MRI scans is generally a series of consecutive parallel cross-sections

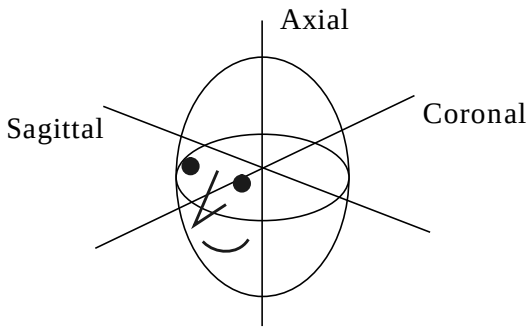


Figure 1. The three principal axes of a head volume.

along one of these axes. Figure 2 shows examples of a person's MRI scans along the three axes. In practice, there is no enforced standard on the acquisition of image data, the axis

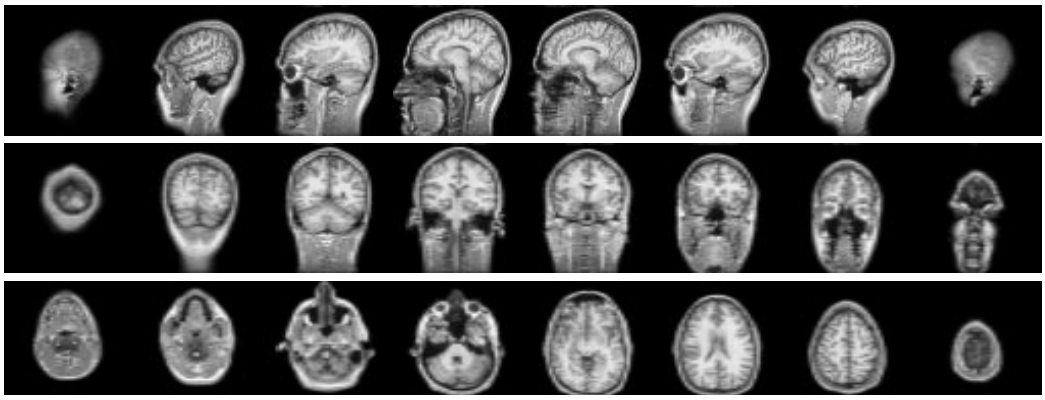


Figure 2. From top to bottom, MRI scans along the sagittal, coronal, and axial axes.

along which the cross-sections are scanned may be at an angle to the principal axes, and the spacings between consecutive cross-sections may be non-uniform. Moreover, a data set can focus on a sub-section of the head if so desired. These complications are illustrated in Figure 3, which depicts an example scan pattern.

Registration within one imaging modality can be classified into inter-person registration and intra-person registration (such as the registration of the same person's data over time). The case we discuss is the inter-person registration of MRI data between an atlas and a patient.

The atlas is a set of 123 coronal MRI scans of a normal brain where each voxel measures 0.9375 x 0.9375 x 1.5 mm^3. The anatomical structures were hand segmented and labelled (courtesy of the Brigham and Women's Hospital of the Harvard Medical School) (see Figure 4). Our task is to precisely segment and label the anatomical structures and detect possible pathologies in the patient volume, using information derived from the atlas. Figure 5 depicts an ideal scenario. Note that the shape, size and location of the labelled anatomical structures

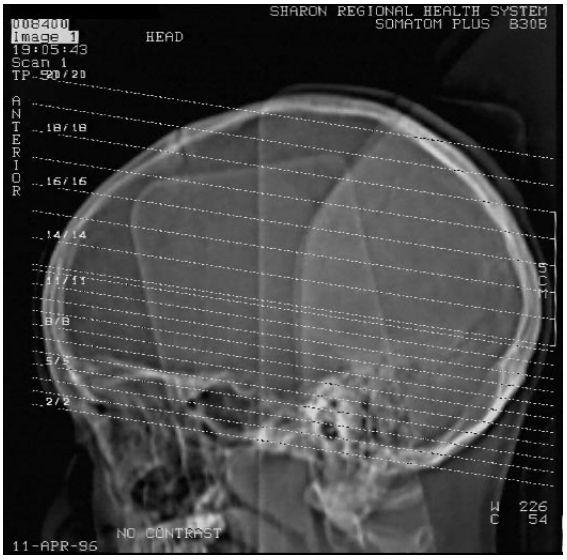


Figure 3. A typical axial scan pattern. Note that the scanning direction is at an angle from the axial axis, the inter-scan spacings are non-uniform and the scans do not cover the entire head volume.

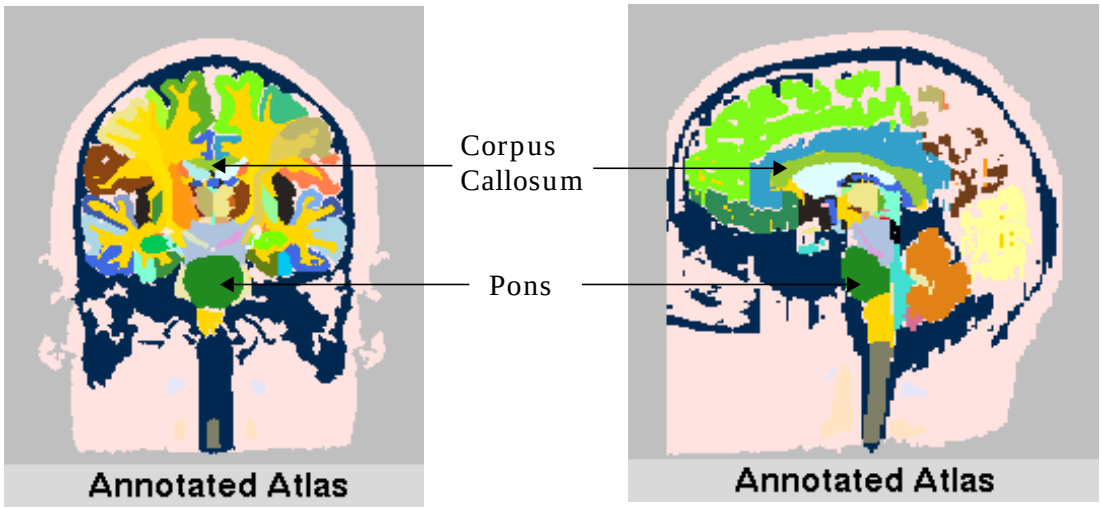


Figure 4. Coronal (left) and sagittal (right) cross-sections of the segmented and labelled atlas. Several anatomical structures are annotated for illustration.

differ between the atlas' and the patient's brains. It is this type of variation in the structures that makes anatomical registration difficult. The current atlas only provides information about a single individual's anatomy, but has no information on the variations across a population. In our work, we intend to encode in the atlas the knowledge of normal variations across individuals, and abnormal but characteristic variations imposed by lesions.

The variations between different image sets stem from two sources. The *extrinsic* source is the scanning process: different scanning axes, different resolutions, or intensity inhomogeneities. The *intrinsic* sources are differences between different people's anatomical structures, or the existence of pathology. Variations from both sources need to be addressed to bring different volumes into alignment.

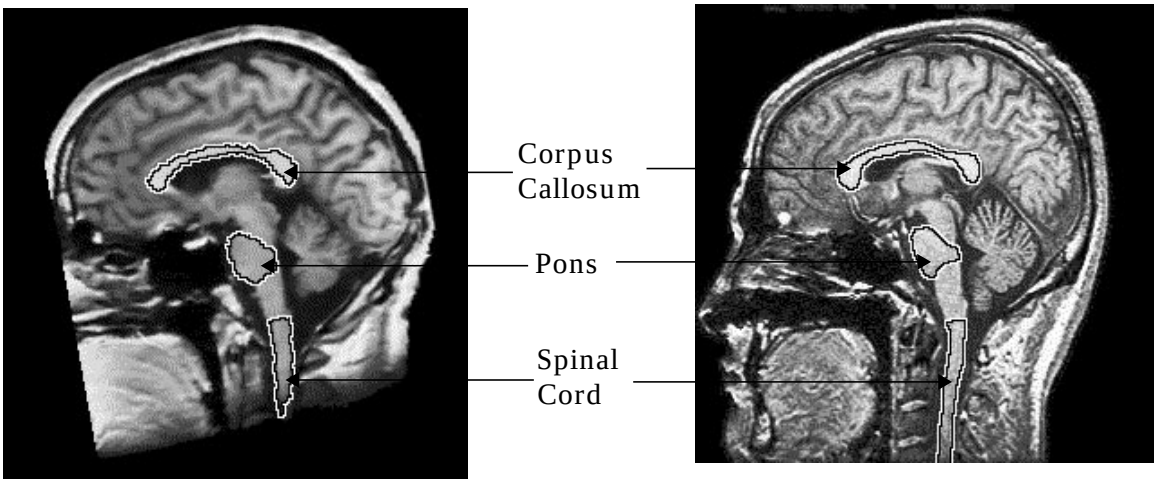


Figure 5. An ideal scenario in which anatomical structures in the patient volume (right) are precisely segmented and labelled using information derived from the atlas volume (left)

3. Current Approach for Matching

Variations caused by extrinsic sources affect the orientation, scale, and intensity consistency of the volumes, while variations resulted from intrinsic sources are manifested as differences in the shape, size, texture, and location of the corresponding anatomical structures across individuals. Because of their different natures, we decompose the registration problem to address them separately.

We adopt a *voxel-based* approach which assumes no prior discrimination among voxels in a volume, that is, all voxels are treated identically with no classification. An alternative is a *feature-based* approach, in which *features* such as boundaries of the anatomical structures are extracted and employed in the registration. Because the anatomical structures in the human brain have complex shapes, non-uniform textures, and ill-defined boundaries, we wish to avoid additional errors incurred by inaccurate feature detection. The varied approaches of anatomical registration are reviewed and compared in section 7.

3.1. Highlights of the Approach

Here are the major points of the current matching method, we will elaborate on them through the document.

- **Preprocess the volumes to improve accuracy.**
  - Apply a *maximum-connected-component* method to extract the head volumes from the background to reduce distractions caused by background noise (4.1.1).
  - *Normalize the intensities* of the volumes to the same range so they are comparable (4.1.2).
- **Hierarchical deformable registration**
  - Apply *global transformation*, i.e. 3-D rotation, uniform scaling, and translation, to the atlas to grossly align it with the patient volume. An iterative optimization algorithm is used to determine the transformation parameters. This process adjusts for

variations caused by extrinsic factors except for the intensity inhomogeneities (3.2).

- Employ a smooth deformation represented by the warping of a coarse *3-D grid* to approximately align the corresponding anatomical structures in the atlas to that of the patient (3.3.1)
- Employ a free-form deformation which allows each voxel to move independently to fine tune the alignment (3.3.2).

3.2. Registration via Global Transformation

We address the variations from extrinsic sources by applying a global transformation to the atlas volume to grossly register it with the patient volume. This global transformation is composed of 3-D rotation, translation, and uniform scaling. There are a total of 7 degrees of freedom.

3.2.1. Representation of the Global Transformation

Figure 6 shows the coordinate systems employed in the global transformation. The origins of the coordinate systems in the atlas volume and the patient volume are placed at their centroids (center of mass), respectively. The Z axis coincides with the axis along which the volume was scanned.

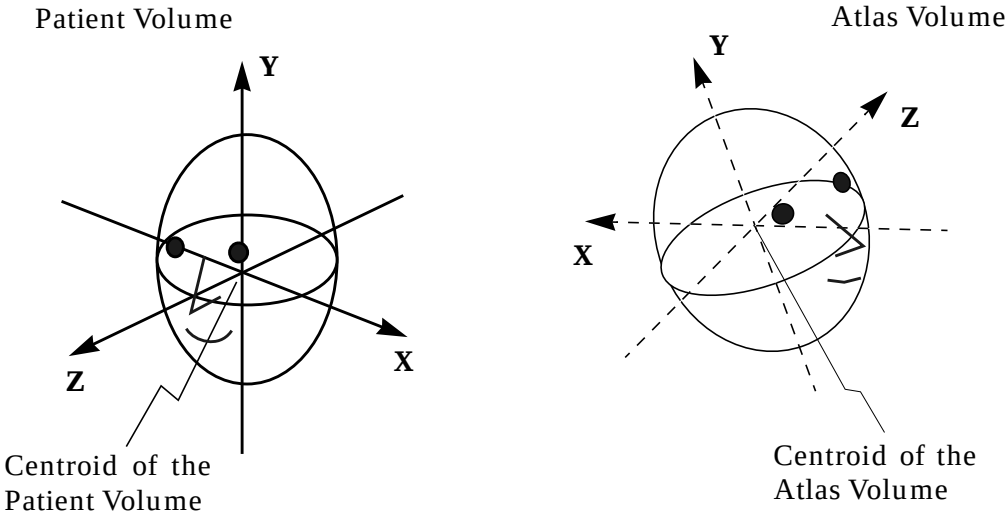


Figure 6. The coordinate system used in the global transformation

The three dimensional rotation is represented by a quaternion.

$$quaternion = q_0 + iq_x + jq_y + kq_z \quad (1)$$

A quaternion can be thought of as a complex number with three imaginary parts. A 3-D rotation by an angle θ about an axis defined by the unit vector $(\omega_x, \omega_y, \omega_z)$ can be represented by a unit quaternion.

$$unitquaternion = \cos\frac{\theta}{2} + \sin\frac{\theta}{2}(i\omega_x + j\omega_y + k\omega_z) \quad (2)$$

Thus the imaginary part of the unit quaternion gives the direction of the rotation axis in 3-D space, whereas the angle of rotation can be recovered from the real part or the magnitude of the imaginary part of the quaternion [5].

The three dimensional translation is represented by vector $(t_x, t_y, t_z)^T$, which represents the displacement of the origin of the atlas coordinate system with respect to the origin of the patient coordinate system.

Uniform scaling is denoted by a scalar, *s*.

The order in which we apply these transformations to the atlas volume is a rotation about the centroid of the atlas, then scaling then translation. Rotation aligns the atlas volume to the same orientation as the patient volume. Scaling adjusts the size of the atlas volume grossly to that of the patient. Translation removes any position difference not accounted for by the alignment of the centroids of the volumes.

3.2.2. Determining the Global Transformation

There are eight parameters to be determined for the global transformation, $(q_0, q_x, q_y, q_z, t_x, t_y, t_z, s)$. Because the atlas and the patient volumes are innately different, we should not expect to find a transformation that exactly registers all the measurements in one coordinate system to the measurements in the other. We can only pursue a transformation that minimizes the error.

We define the quality of a registration in terms of the sum of squared differences (SSD) between the corresponding voxel intensities in the two volumes:

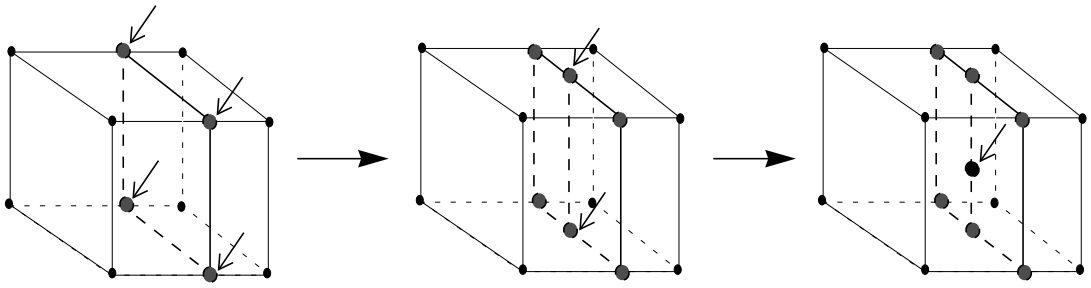
$$SSD(Transformation) = \sum_{(x, y, z)} (Patient_{(x, y, z)} - Atlas_{Transformation(x, y, z)})^2 \quad (3)$$

The SSD is a function of the transformation. To minimize the SSD, we apply the Levenberg-Marquardt iterative non-linear optimization algorithm to determine the transformation parameters (see Appendix 1). The derivative of the SSD with respect to the transformation indicates the change in SSD induced by the changes in the transformation. The Levenberg-Marquardt algorithm adopts gradient descent to iteratively adjust the transformation parameters to reduce the SSD. The iteration continues until the changes in the transformation parameters are negligible, i.e. below a user defined threshold, and at which point the transformation is considered to be recovered. We employ coarse-to-fine processing and stochastic sampling to help prevent the optimization process from being trapped in local minima (see section 4).

After applying the transformation, the voxel coordinates in the atlas may not have integral coordinates. In this case tri-linear interpolation is employed to determine its intensity from its eight bounding neighbors (see Figure 7). Notice that the result of the tri-linear interpolation is not affected by the order of the three linear interpolations along the three axes. If the voxel falls outside of the range of the volume, its effect is ignored.

3.3. Registration via Deformation

After the global transformation, the extrinsic variations between the atlas and the patient volumes are reduced. However, the existence of intrinsic variations impedes accurate segmentation and correspondence of individual anatomical structures.



Linear interpolations among 8 neighbors along one axis results in 4 intermediate val-

Linear interpolation among the 4 intermediate values along another axis then yields 2 intermediate val-

Linear interpolation between the 2 intermediate values along the third axis gives the voxel intensity

Figure 7. Tri-linear interpolation gives the intensity of a voxel with non-integral coordinates by doing linear interpolations among its 8 bounding neighbors along each of the three axes

Figure 8 displays the mid-sagittal planes of both the atlas and the patient after registration via global transformation. Note that their anatomical structures are roughly aligned, but they differ in shape, size, and intensity distribution.

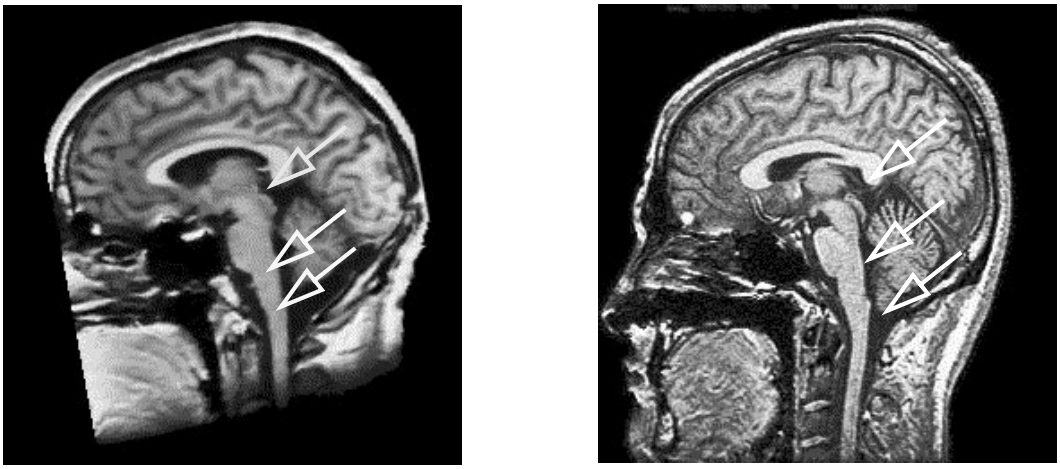


Figure 8. The mid-sagittal planes of the atlas volume (left), and the patient volume (right) after registration via global registration. The arrows indicate the corresponding points if they are overlapped, note that the details do not match

The brain morphology is intricate, and the variations across individuals are large, but the shape and density distribution of anatomical structures are still distinct enough for them to be distinguished. We intend to deform the volumes to align their internal structures.

Two issues need to be addressed: the formulation of the deforming force, and the representation of the deformation. Our current atlas does not provide more information than its own anatomy expressed in image content, so at present the deforming force is simply the intensity difference between spatially corresponding voxels. The deformation process allows voxels to search for their counterparts in the patient volume. A more intelligent registration algorithm that has control over the deformation force will be devised once we embed relevant knowledge into the atlas.

3.3.1. Smooth Deformation

Since we use voxel intensity differences as the deforming force, the most intuitive way to represent the deformation would be the 3-D displacement of each voxel. This allows each voxel to deform freely, but it can only succeed when the voxels' initial positions are close to their sought positions. The intrinsic variations across individuals make the global transformation unable to precisely align the atlas to the patient to provide a good initial alignment. We need a more robust representation of intermediate deformation.

- **Representing the Smooth Deformation**

Our solution is to enforce a smooth deformation by placing a 3-D control grid is each volume. Vertices of the control grid are control points. The 3-D displacements of the surrounding control points determine the deformation of all the voxels enclosed in each grid cell by interpolation. Therefore although the control points can still deform freely, the voxels inside each grid cell are forced to deform smoothly. Since the number of control points is orders of magnitude lower than the number of voxels, there are fewer parameters to estimate, making the deformation process more stable. The simplest 3-D grid is composed of rectangular prisms, in which case the control grid is a regular sub-sampling of the voxel grid. Figure 9 shows an example of a control grid cell and an arbitrary deformation of it. Szeliski employed similar representation in 2-D registration [9], and 3-D registration of surfaces, [8]. Figure 10 is a 2-D illustration of the smooth deformation process.

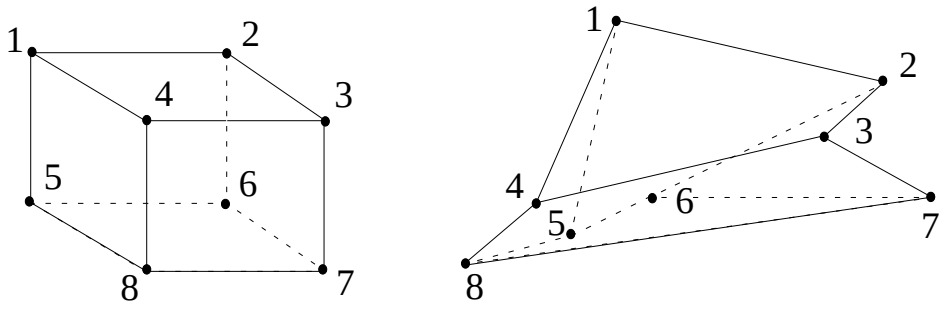


Figure 9. A control grid cell with its vertices as control points (left), and an arbitrary deformation of the cell due to movements of control points (right).

- **Estimating the Smooth Deformation**

Although we represent the smooth deformation with 3-D displacements of the control points, we estimate the deformation force using the intensity differences throughout the volume. For a voxel $[x, y, z]_{patient}$ in the patient volume with intensity $I_{patient}(x, y, z)$, suppose the control grid cell it belongs to is the *i* th in the array of all control grid cells $Cell_{patient}[i]$. The corresponding grid cell in the atlas volume is $Cell_{atlas}[i]$. From the relative position of $[x, y, z]_{patient}$ with respect to the 8 control points of $Cell_{patient}[i]$ in the patient volume, we can tri-linearly interpolate its corresponding location in the atlas volume $[x, y, z]_{atlas}$ using the 8 control points of $Cell_{atlas}[i]$. Should $[x, y, z]_{atlas}$ have non-integral coordinates, tri-linear interpolation is applied to calculate the point's intensity from its 8 bounding neighbors in the atlas (see Figure 7). In the case that $[x, y, z]_{atlas}$ falls outside the atlas volume, it is ignored.

If the smooth deformation is perfect, then the atlas volume completely aligns with the patient volume, and $I_{atlas}(x, y, z)$ should equal to $I_{patient}(x, y, z)$. But due to noise and small shape differences between the atlas and patient, $I_{atlas}(x, y, z)$ and $I_{patient}(x, y, z)$ differ; the best

deformation minimizes these differences. We take a similar approach to that of solving for the global transformation in 3.2.2. We use the sum of squared differences (SSD) over the voxels in the volumes as the error function, and employ the Levenberg-Marquardt iterative nonlinear optimization method to determine the smooth deformation parameters that best match the two volumes.

The only difference between this process and the one in 3.2.2. is that the parameters we compute are not the global transformation parameters for the full volumes, but the collection of local smooth deformation parameters for each control grid cell. Although each control grid cell is deformed independently, the deformations are linked via shared control points across the 3-D control grid. They are evaluated together and they collectively define the smooth deformation.

The number of smooth deformation parameters is 3 times the number of control points, i.e. 3-D displacements of each control point. Since each voxel is only affected by the 8 control points of its control grid cell, the derivative matrix of the SSD with respect to all the smooth deformation parameters will have large number of zero entries. We employ sparse matrix representations to improve computational efficiency.

3.3.2. Free-Form Deformation

The smooth deformation is effective and robust at deforming the atlas to approximately match its anatomical structures to that of the patient, even when the global transformation gives poor initial alignment. However, since only the control points are allowed to deform freely, it can not account for any details smaller than the size of a control grid cell. To further refine the deformation to adjust to finer details, we apply a free-form deformation.

- **Representing the Free-Form Deformation**

After the smooth deformation has approximately aligned the anatomical structures in the atlas and the patient, we use the 3-D displacement of each voxel to represent the free-form deformation. Therefore each voxel can deform freely, allowing the deformation to fine tune to details. This procedure is similar to the approach discussed in [13]. A comparison is given in section 7. Since the deformation parameters are the 3-D displacement of each voxel

- **Estimating the Free-Form Deformation**

The free-form deformation to match the atlas to the patient is similar to optical flow. Consider the atlas and the patient as two consecutive frames in an image sequence, the intensity in the atlas is then a function of voxel location and time, $I_{atlas}(location,time)$. The underlying assumption is that intensity is conserved under the free-form deformation:

$$\frac{\partial I_{atlas}(location,time)}{\partial(location,time)} \bullet (deformation,1) = 0 \qquad (4)$$

The first element of the dot product is the change of intensity with respecct to location and time, and the second element is the change of location during one unit of time, i.e. the deformation. Setting the dot product to zero expresses the intensity conservation assumption. From (4) we have

$$\frac{\partial I_{atlas}(location,time)}{\partial location} \bullet (deformation) = -\frac{\partial I_{atlas}(location,time)}{\partial time} \qquad (5)$$

This gives

$$\nabla I_{atlas} \bullet deformation = I_{atlas} - I_{patient} \tag{6}$$

To prevent the deformation parameters from going out of bounds when the atlas gradient ∇I_{atlas} is close to zero, we add a stabilizing factor α :

$$deformation = \frac{I_{atlas} - I_{patient}}{\left\|\nabla I_{atlas}\right\|^2 + \alpha} \nabla I_{atlas} \tag{7}$$

Tri-linear interpolation is used to compute I_{atlas} when the voxel coordinates after deformation are not integers. The estimation of deformation is done for each voxel. Because optical flow computed in this way is noisy and under-constrained, we apply 3-D Gaussian smoothing to the 3-D deformation parameters after each iteration.

Once the atlas is deformed to align with the patient, we call it a *customized atlas*. We can apply the labels in the *customized atlas* to directly segment and label the patient volume.

4. Matching Implementation Details

4.1. Preprocessing

To improve the robustness of the algorithm, we apply two types of preprocessing: background separation and histogram normalization.

4.1.1. Background Separation

Since we take a voxel-based approach with no feature extraction, each voxel in the data is assumed to play an equal role in the registration process. There is usually background noise in MRI data due to the scanning process. Therefore it is important to separate the background to ensure only relevant data contributes to the registration procedure.

We first automatically threshold the volume to eliminate some of the dim background noise. The threshold we use is the first valley of the intensity histogram (see Figure 11). Since the region containing the head is connected, we apply a maximum-connected-component algorithm to the binary volume to find the largest connected component. Any holes caused by dark regions inside the head are filled in. We only consider data within that component in later processing. However, if there is spurious noise connected to the head, it will stay. If some internal structures are detached from the main head region, they may be lost in this process.

4.1.2. Histogram Normalization

We adopted an intensity-based method to avoid errors introduced by unreliable feature extraction processes, but intra-scan and inter-scan intensity variation make this approach problematic. The operating conditions and status of the MRI equipment frequently affects the observed intensities. Our current solution is to perform histogram normalization on each volume prior to the registration process. To prevent outlier voxels with extreme intensities

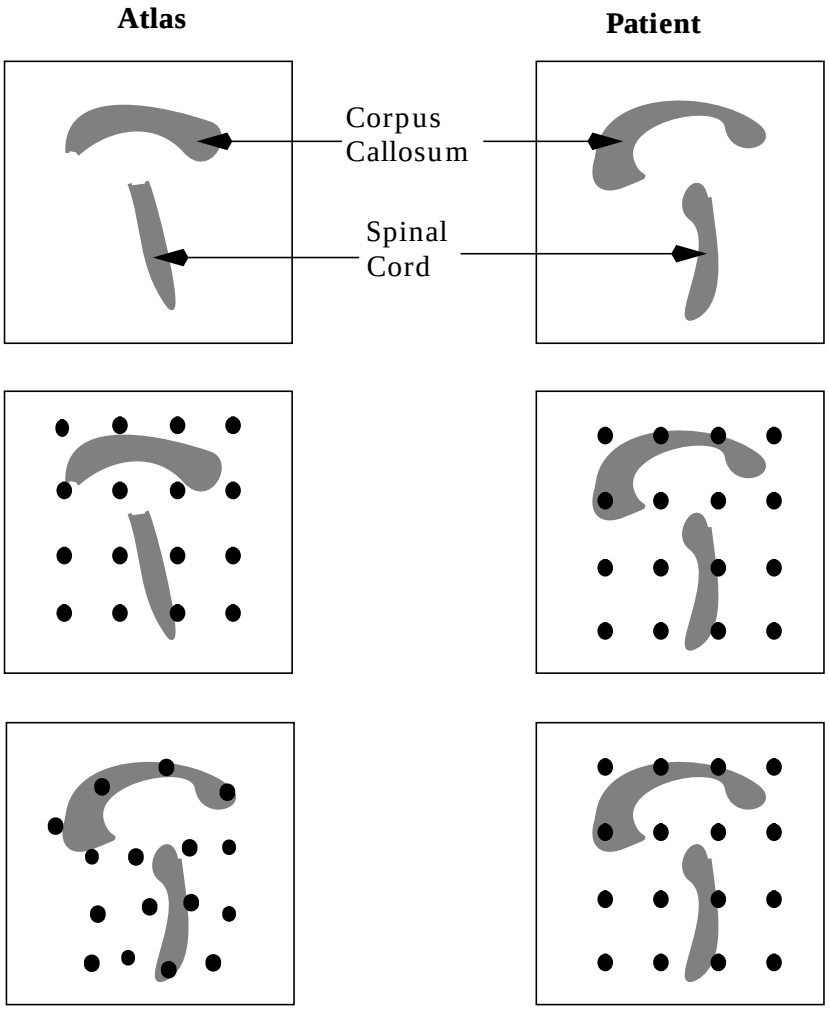


Figure 10. A 2-D illustration of the smooth deformation process. The first row is the original data from the atlas and the patient. The second row is the original data overlayed with the regular control grids. The third row shows the data overlaid with control grids after the smooth deformation. Note that the atlas is deformed to match the patient, and its control grid is deformed accordingly.

from stretching the histogram, we first set the intensity of the darkest 2% of the voxels to 0, and the intensity of the brightest 2% of the voxels to 255. We then linearly scale voxel intensities in between to the range of 0 to 255.

4.2. Efficient and Effective Processing

The volumes we deal with typically have 8 million voxels. Therefore it is imperative to carry out the registration in an efficient, yet accurate, manner.

4.2.1. Coarse-to-Fine Processing

A coarse-to-fine strategy is used not only to improve efficiency, but also to help the opti-

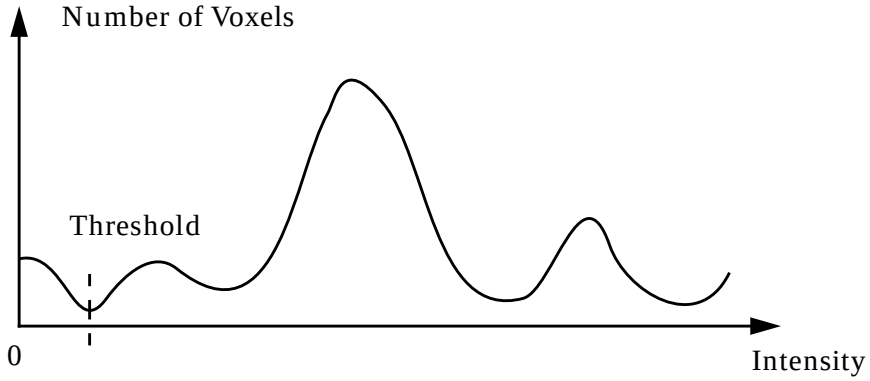


Figure 11. An illustration of thresholding based on the intensity histogram.

mization procedure to first focus on global patterns and gradually shift to the details. We employ two kinds of coarse-to-fine processing: an image pyramid (used in the global transformation and the smooth deformation), and a control grid pyramid (used in the smooth deformation).

The image pyramid is a hierarchy of data volumes generated by successive subsampling. At the lower resolutions there are fewer details present, so the minimization process has less tendency to become trapped in local minima. By passing results from a lower resolution on to a higher resolution, we help the higher resolution registration start closer to the global minimum. Optimization at the higher resolutions refines the result given by the finer details.

The control grid pyramid is a hierarchy of coarse-to-fine control grids. By starting with coarser control grids the deformation can focus on global patterns before plunging into details. The result from the lower resolution control grid initializes the registration with finer control grids close to the optimum. Smooth deformation with finer control grids allows for a more precise registration.

4.2.2. Stochastic Sampling

Instead of going through each voxel of the volumes during each iteration of the optimization process, we sample a random set of voxels at each iteration. Since the problem is extremely over-constrained, i.e. the number of data points far exceeds the number of parameters, the reduction in the number of data points will increase the variance in the estimation, which we believe is negligible. Moreover, the stochastic nature of the sampling helps the minimization process to escape from local minima which arise because of false matches between the high frequency components in the atlas and the patient volume. Because these false matches are based on features that are small, the local minima are narrow in parameter space. Therefore stochastic sampling serendipitously combines efficient computation with an effective way to escape from local minima [25]. Note that in the free-form deformation we have 3 times as many parameters as the total number of voxels, so stochastic sampling was not appropriate.

4.2.3. Parallel Processing

In a voxel-based approach, each voxel contributes independently to the whole evaluation, so the process is voxel-wise parallelizable. To reduce overhead, we employ parallel process-

ing at one level above the voxel representation. During the registration via global transformation and during the free-form deformation, we parallelize the processing of each cross-section of the volumes; whereas in the registration via smooth deformation, we parallelize the processing of each control grid cell in the volumes. Because a voxel in the patient volume can map to any position in the atlas, it is difficult to partition the volumes so each part can be processed independently. Therefore, we choose a shared-memory multi-processor computer so each processor can have access to the full volumes.

5. Matching Experimental Results

We tested our algorithm between the atlas and 6 patient volumes without any pathology, and three patient volumes with pathologies. The results are varied. We present results from matching the atlas with one *normal* volume and one *pathology* volume as examples to analyze the effectiveness of our approach. Table 1 provides information on these data sets.

	Scan Axis	Number of Scans	Scan Spacing (mm)	Scan Dimension (pixels)	Pixel size (mm by mm)	Pixel Representation
Atlas	Coronal	123	1.5	256 x 256	0.9375x0.9375	2 bytes
Normal	Sagittal	124	1.5	256 x 256	0.9375x0.9375	2 bytes
Pathology	Axial	46	Unknown	256 x 256	Unknown	1 byte

Table 1: Information on Data Sets.

5.1. Result from Registration via Global Transformation

Figure 12 displays corresponding cross-sections of the atlas, the normal, and the pathology volumes after registration via global transformation. The atlas is centered and aligned to the same orientation as the other volume, and scaled to the same size. The highlights are labels of anatomical structures in the atlas volume. Note that when transferred to the other volume, they do not match the corresponding structures in that brain, due to intrinsic variations between people and shifts caused by the pathology. The extrinsic variations between the volumes are reduced, but the intrinsic variations still remain to be addressed.

5.2. Result from Registration via Smooth Deformation

Figure 13 shows corresponding cross-sections of the atlas volume and the other volumes after registration via smooth deformation. The atlas is warped to align with the other volumes (compare with the left images in Figure 12). Now the labels of anatomical structures in the atlas match approximately with the corresponding structures in the other volumes. This demonstrates that the registration via smooth deformation can account for some of the intrinsic differences.

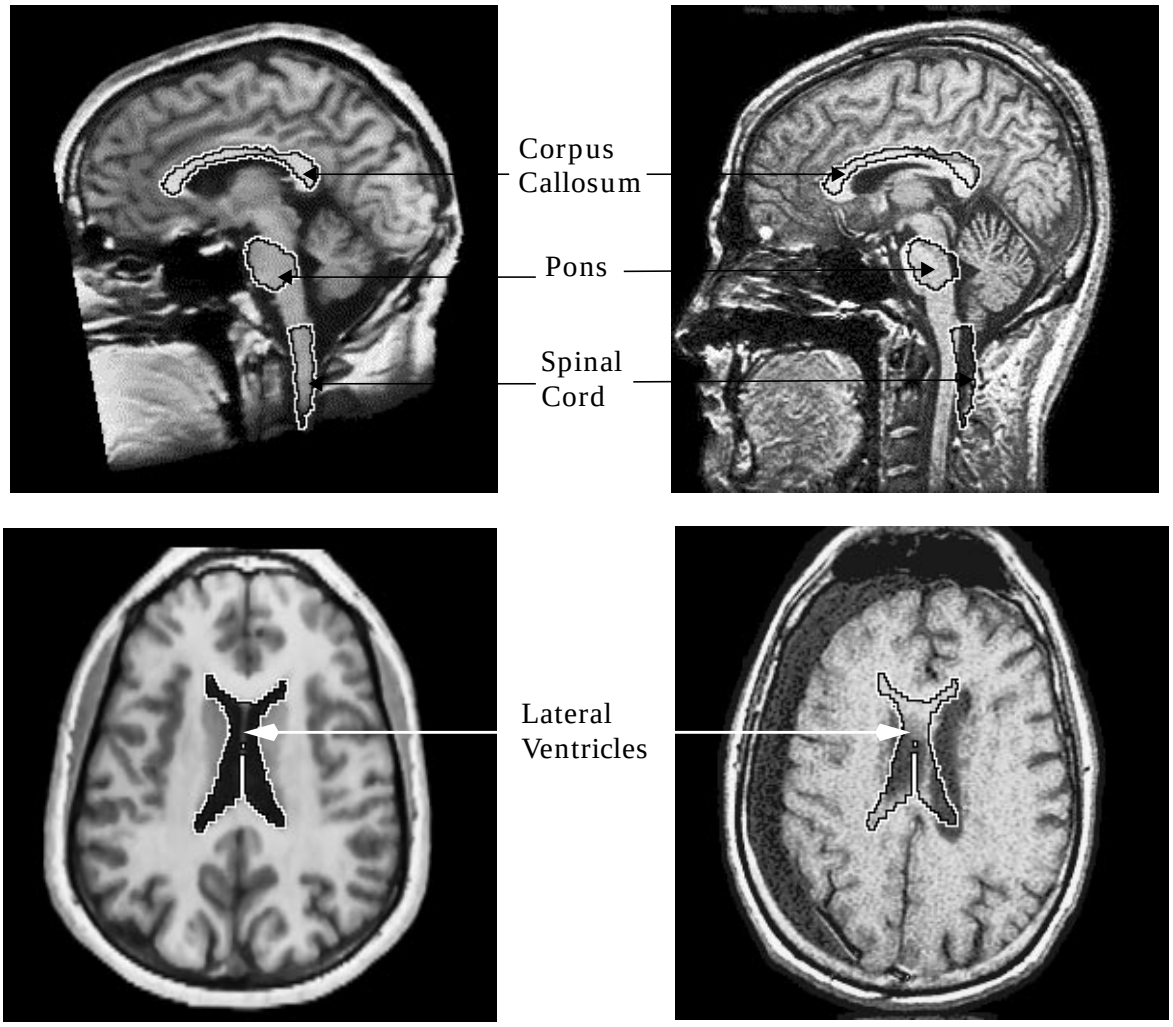


Figure 12. Top: Corresponding cross-sections of the atlas volume (left), the normal volume (right) after registration via global transformation. Bottom: atlas (left), pathology volume (right). Note the labels of anatomical structures from the atlas do not match the corresponding structures in the other volumes.

5.3. Result from Free-Form Deformation

Figure 14 shows cross-sections of the atlas and the other volumes after registration via free-form deformation. The atlas is warped to align with the other volumes more precisely (see images in Figure 13). Note that the labels of the anatomical structures in the atlas match well with the corresponding structures in the other volumes. This demonstrates that the free-form deformation can refine the result from the smooth deformation to address the variations in fine details.

6. Evaluation and Analysis of Matching

The main requirements of a registration procedure for clinical use are simplicity, accuracy,

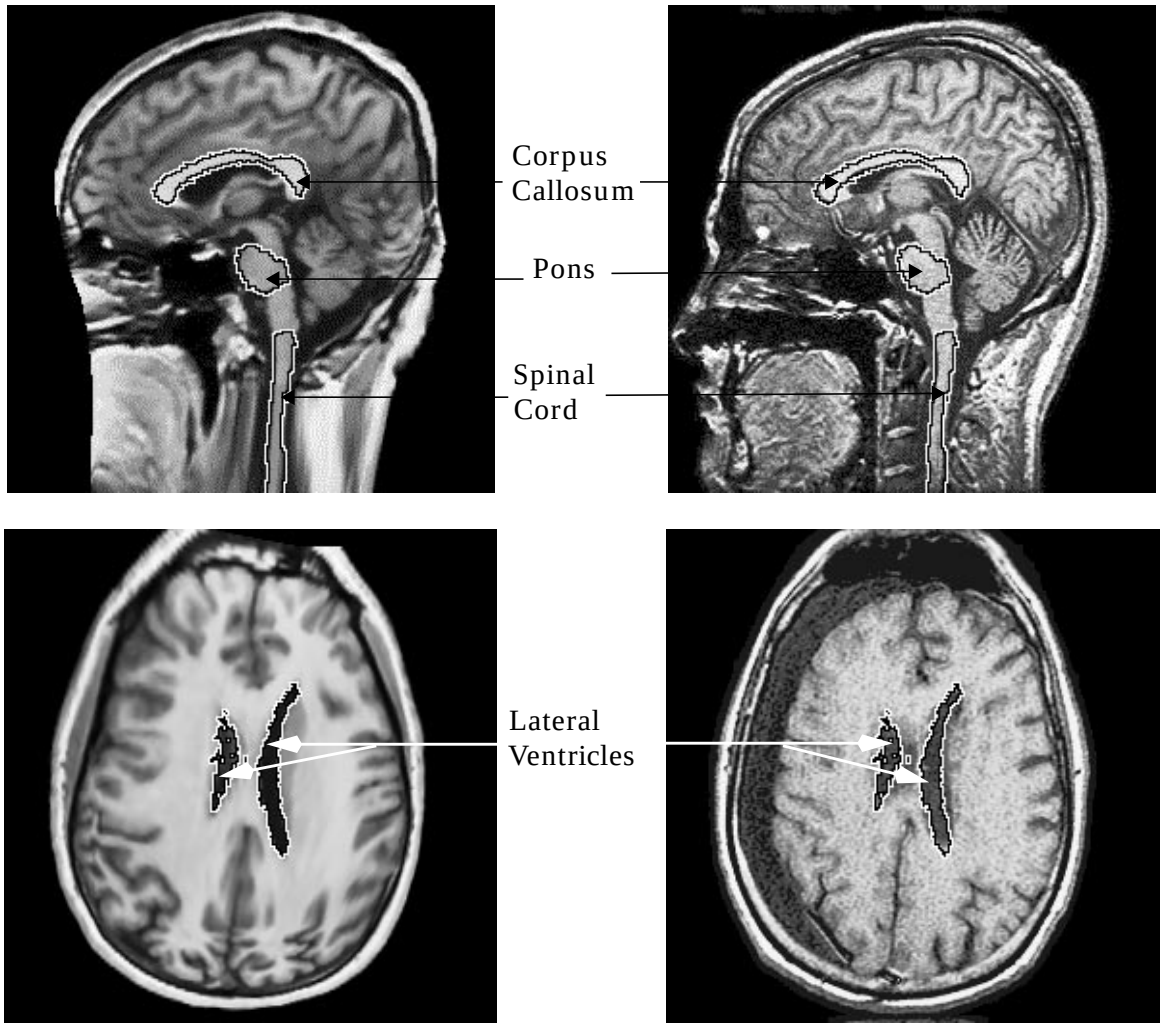


Figure 13. Top: Corresponding cross-section of the atlas (left) and the normal volume (right) after registration via smooth deformation. Bottom: atlas (left), pathology volume (right). Note the atlas is warped to match the patient. Labels of anatomical structures in the atlas match approximately with the corresponding structures in the patient.

and speed. Operating the matching application should be simple, so that untrained medical people are able to perform registrations. The average error should be less than one voxel size. The algorithm should be fast enough to use even during a surgical intervention for the registration of on-line data. We will evaluate our approach accordingly.

6.1. Simplicity

Our current approach, from preprocessing, registration via global transformation, registration via smooth deformation, to registration via free-form deformation, is completely automatic. The role of the user is to specify the principal scanning axis. This can be standardized so the user can simply select among different settings.

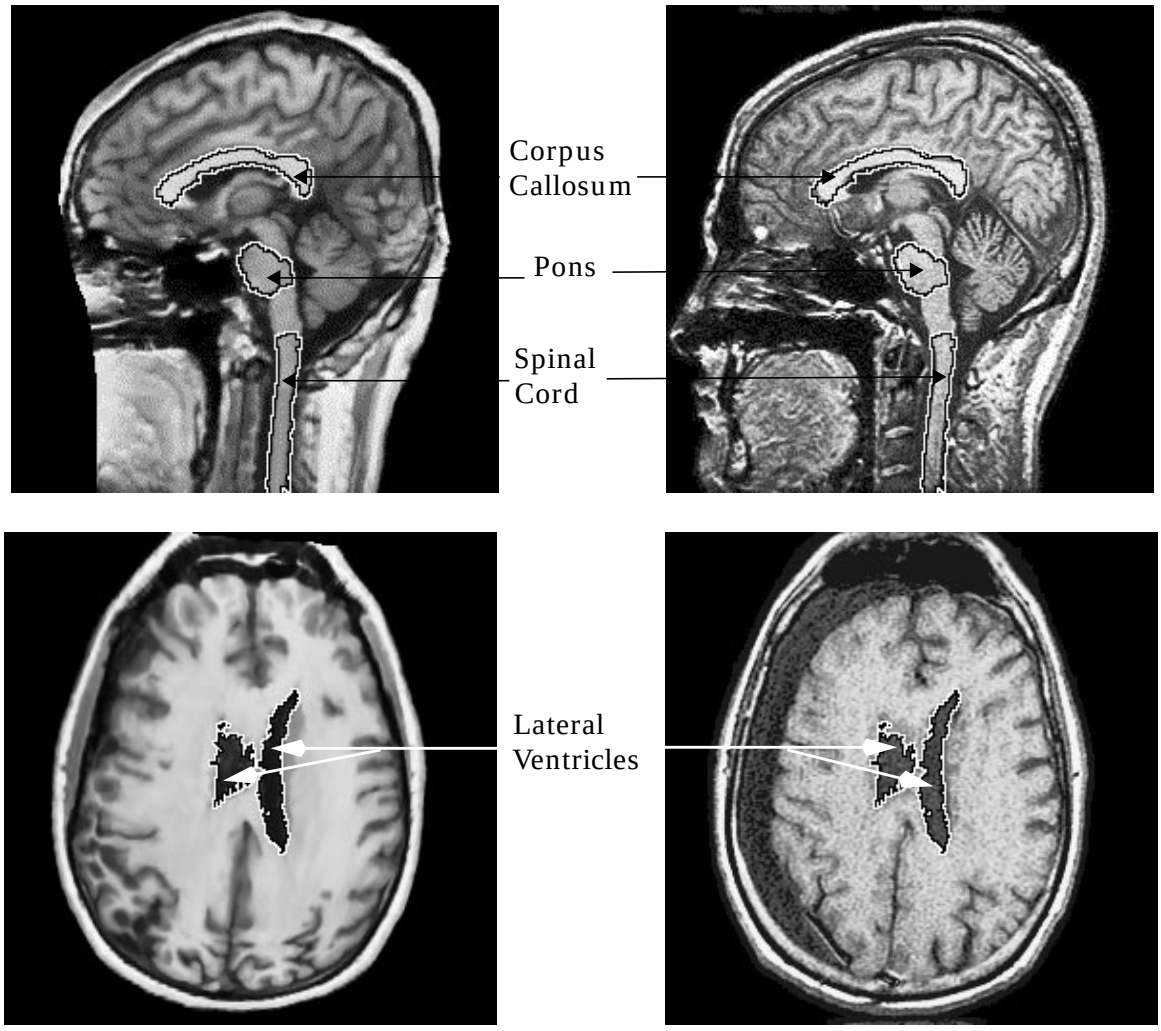


Figure 14. Top: Corresponding cross-sections of the atlas (left) and the normal volume (right) after registration via free-form deformation. Bottom: atlas (left), pathology volume (right). Note the atlas is warped to match the patient more precisely. Labels of anatomical structures in the atlas match well with the corresponding structures in the other volumes.

6.2. Speed

Currently it takes 12 minutes to match 256x256x124 volumes on an SGI computer with four 194 MHz R10000 processors. There are parameters that can be tuned to improve efficiency, such as the number of stochastic samples, the number of levels in the pyramids, and the number of control points.

6.3. Accuracy

Objective estimation procedures for the error of a matching method are hard to design. Validation techniques that have been proposed include correspondence measurements of landmarks identified by human observers [11], [13], comparison of a matching method to a

technique with a higher accuracy, such as a stereotactic frame or marker-based method [6]. One way to evaluate our registration precision is to compare our segmentation and labelling of the patient volume with an expert's segmentation and labelling of the patient volume, or of another atlas volume. Note that because of the inconsistency in manual registration, we will need to reconcile atlases provided by different experts to create *standard atlases*. Since this information is currently unavailable, we will first artificially transform and deform the atlas volume to create a "patient" volume, and try to segment and label this volume using information from the atlas, so as to assess the accuracy of our approach. Another approach is to

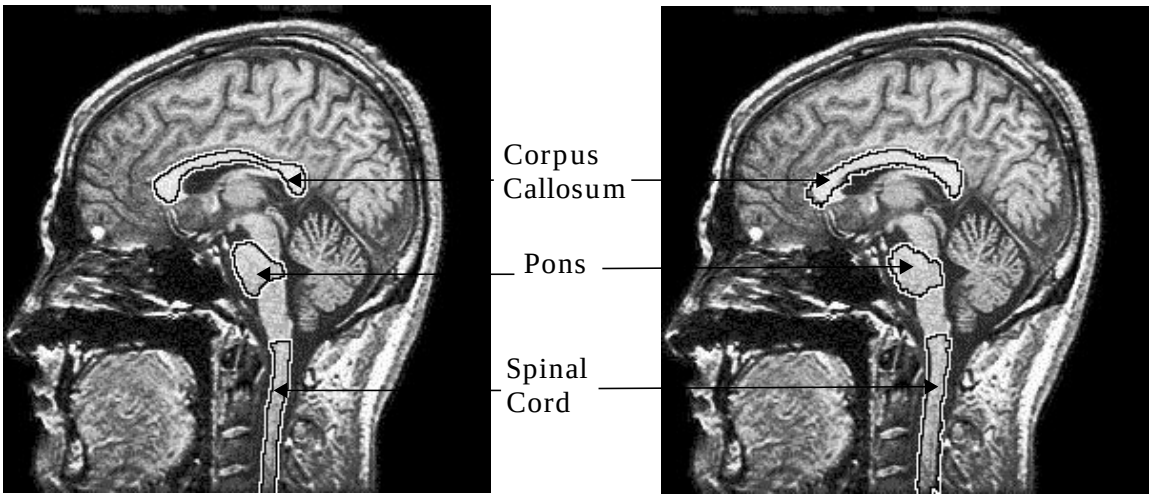


Figure 15. The normal volume with labels from hand segmentation (left), compared to the result of our algorithm (right).

match the atlas to a patient volume to create a customized atlas for the patient, then match this customized atlas back to the atlas to create a customized atlas for the atlas. The amount of overlapping between labels in the atlas and labels in its customized atlas shows the consistency of the algorithm. To appraise the algorithm's ability to deal with intensity inhomogeneities, we can add noise to the artificial "patient" volume to imitate that from the scanning mechanism.

7. Related Work

The registration of medical images by optimization in transformation space has been an active research area--the comprehensive survey article by van den Elsen et al. [14] lists 161 citations. The primary division of registration approaches is into methods using external and internal image properties for matching, because this distinction is important for the clinical protocol. *External image properties* are introduced by artificial objects that are "added" to the patient, such as head frames or skin markers. *Internal image properties* are patient related characteristics of the image data set, such as the intensity difference between anatomical structures.

External, marker-based registration methods have the advantage that any two modalities can be matched, as long as a marker can be constructed that is detectable in the images. They yield high accuracy matching with respect to rigid transformations [6]. But to ensure accuracy, the markers are required to be attached rigidly to the patient (by driving screws into the patient's skull), so as to indicate precisely the patient's position and orientation, which is

invasive and inconvenient. Registration methods using internal image properties have the advantage of being fully retrospective, which means that one does not need to know prior to the acquisition of the images whether matching will be required, and non-invasive. They also have the potential to deal with non-rigid transformations between different image sets. This makes them more favorable in anatomical registration across individuals, where a rigid transformation will not suffice.

Two popular schools of registration using internal image properties are *feature-based* and *voxel-based*. *Feature-based* methods attempt to extract the anatomical structures in different data sets, and find the correspondence between them. They have the characteristic of being efficient in representation and independent of imaging modalities. However, feature-based registration is critically dependent on the quality of the feature extraction, which is not trivial especially since anatomical structures tend to have complex shapes and ill-defined boundaries. Human interaction is generally necessary to help select and extract features or to guide the matching procedure. Consequently, it is subject to user subjectivity, time-consuming, and inconvenient. [15] describes an automated approach to register CT and MR brain images using features extracted by geometrically invariant operators in scale space, but it is restricted to rigid transformations to match brain images from the same patient. The extension of this method to enable non-rigid transformations is claimed to be straightforward, although computationally expensive.

Many researchers apply *deformable models* to enable feature-based non-rigid registration. [7] gives a comprehensive survey. The initial placement of the deformable model must be very close to the sought feature to guarantee a successful result for elastically deformable models, therefore human intervention is generally necessary. The approach in [16] employs a 3-D elastic warping transformation to register 3-D images. The transformation is driven by an external force field defined on a number of distinct anatomical surfaces, which were acquired in an interactive manner.

As an alternative, *voxel-based* algorithms obviate the need for an explicit segmentation, although the representation is not as concise. The most intuitive voxel-based approach is based on voxel *intensities*. Bajcsy et al. develop a system that elastically deforms a 3D atlas to match anatomical brain images [11],[10]. The atlas is modelled as a physical object and is given elastic properties. After an initial global alignment, the atlas is deformed to match corresponding regions in the brain image volume in response to forces derived from image features. The technique is sensitive to the initial positioning of the atlas--if the initial rigid alignment is off by too much, then the elastic matching may perform poorly. The presence of neighboring features may also cause matching problems--the atlas may warp to an incorrect boundary. Finally, without user interaction, the atlas can have difficulty converging to complicated object boundaries. Although their approach is similar to our current one, they assume the intrinsic variations between people can be modelled by an elastic deformation whereas we only enforce smooth deformation in the intermediate stage. Also their method is more computationally expensive and requires interactive and time-consuming preprocessing, while ours can be fully automated.

Christensen et al. presented a method very close to ours, but they used a fluid dynamic model for the deformation [2], [3]. It constrains neighboring voxels to have similar deformations, while allowing large deformation for small sub-volumes. It takes 1.8 hours to match 128x128x100 volumes on a 16384-processor MasPar, while our algorithm takes 12 minutes to match 256x256x124 volumes on an SGI with four 194 MHz R10000 processors.

In [13], Thirion takes a similar approach as ours, except that he assumes the volumes are already globally aligned, and he applies optical flow from the beginning. To reduce computational expense, he used the gradient of the patient volume instead of the deformed atlas,

because the computation of the latter is more expensive, requiring tri-linear interpolation of each voxel's gradient. However, this quicker method can cause errors when the deformed atlas does not resemble the patient closely. Because optical flow relies heavily on the constant brightness assumption, it is prone to failure when there are large intensity variations between different image sets. We implemented his algorithm and tested it on the same data sets that we tested our method. Figure 16 shows a comparison of the patient volumes segmented and labelled using Thirion's method and our approach. Note that our results are more accurate in these cases. For volumes with intensities more similar to that of the atlas, the two methods performed equally well.

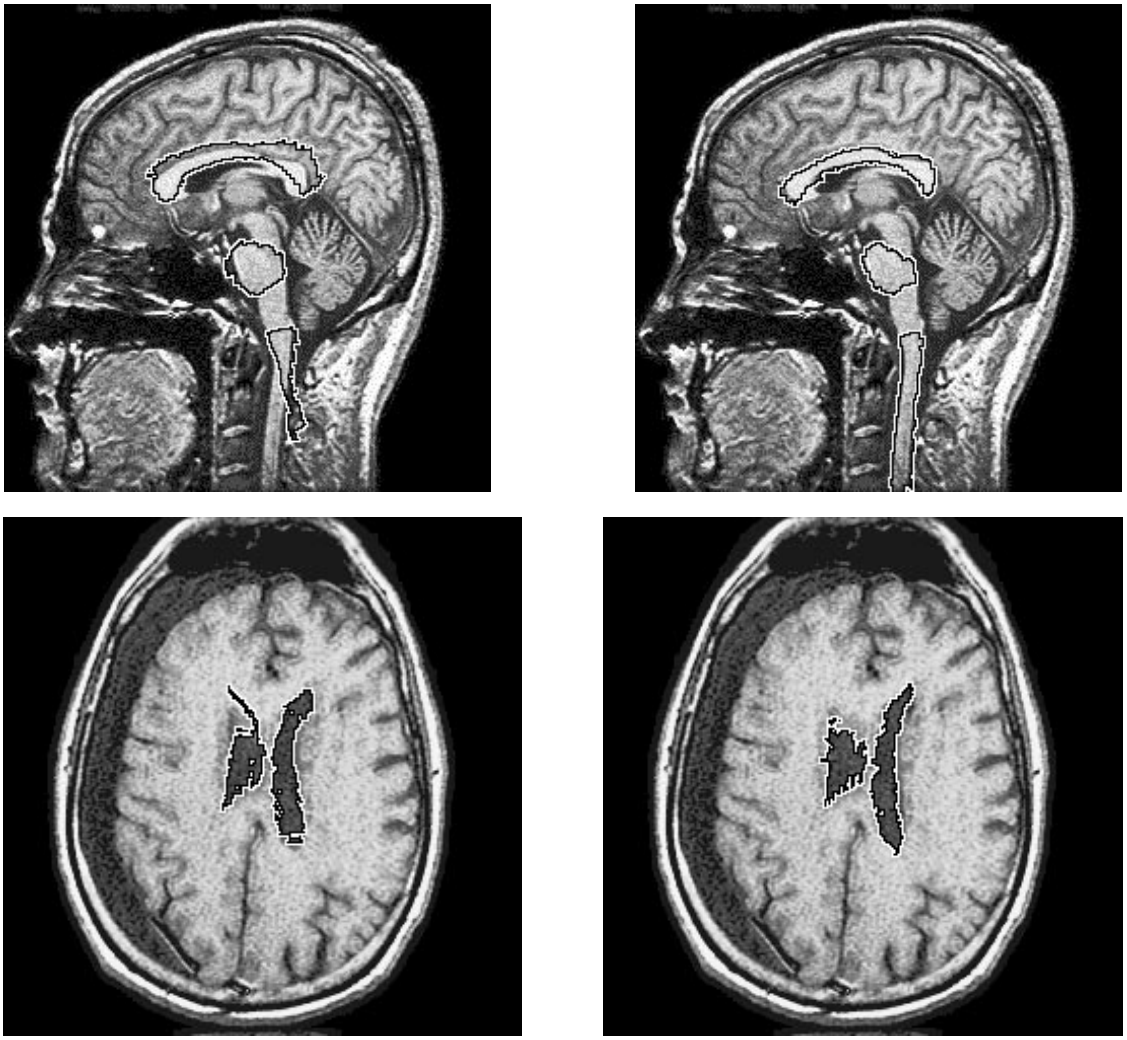


Figure 16. Top: Corresponding cross-sections of the normal volume labelled using the result of Thirion's method (left), and our approach (right). Bottom: The pathology volume labelled using Thirion's method (left), our approach (right). Labels of anatomical structures result from our approach align better with the corresponding structures.

Although voxel intensity-based approaches have shown encouraging results, they are problematic when there are intensity inhomogeneities. Moreover, they only work for multi-modal-data if there exists a linear mapping between intensity values, which is unfortunately almost never the case. Viola and other researchers have investigated registration based on mutual information (MI) [23], [24], [25], [26]. MI is a basic concept from information theory,

which measures the statistical dependence between two random variables, or the amount of information that one variable contains about the other. The MI registration criterion assumes that the statistical dependence between corresponding voxel intensities is maximal if the images are geometrically aligned. Because no assumptions are made regarding the nature of this dependence, the MI criterion is highly data independent and allows for robust and completely automatic registration. Current applications of registration using MI only perform rigid transformations to register image data of the same person from different modalities. The possibility of applying the MI criterion in deformable registration remains to be studied.

To date, most efforts are focused on exploring the information content in the images to achieve registration. Little work has been done in using domain knowledge to guide the process. Also very little effort has been engaged to tackle cases with pathologies present, which are of more clinical importance. We would like to investigate the potential of this areas.

8. Research Plan

The preliminary results from our current approach are promising, and it can lead to many research directions. I plan to pursue the following paths.

- **Use domain knowledge to guide the registration.**

Our current approach deforms the atlas with no constraints from the brain anatomy. If the deformation could be limited to shapes within the normal anatomical variation of structures, or the characteristic deformation caused by pathologies, local minima could be avoided. Knowledge of natural shape variability and pathology-afflicted deformation can be encoded into the atlas as allowed deformations. This knowledge can be acquired from the statistics of training samples.

For deformations described by a large number of possibly highly correlated parameters, principal component analysis (PCA) may offer a promising solution by characterizing the dominant part of the deformation by the few largest *eigendeformations*, reducing the dimensionality of the deformation space substantially. To represent the deformations in a way that is appropriate for applying PCA, Fourier descriptors or modal representations of deformations have been examined in previous work to match curves and surfaces in 2-D and 3-D [17], [19], [20].

Other knowledge of the natural variability of the anatomy, such as the difference in density and texture could also be used to guide the deformation. Knowledge of anatomy, such as symmetry and relative positions of structures, can also provide guidance. Moreover, some anatomical structures are more distinctive and therefore more easily registered. These structures should be given more weight in the estimation of the deformation.

- **Indicate potential areas of pathology.**

For more clinical significance, we want to be able to register volumes with neurological conditions and locate pathologies. Certain pathologies cause mass effect, which shifts the brain structures, and causes anatomical features close by to deform significantly (see Figure 17, left). However, since our deformation process is localized (the smooth deformation is localized to each control grid cell; the free-form deformation is localized to each voxel), it can still match the corresponding anatomical features (see Figure 14). The right image in Figure 17 displays an overlay of the deformation vectors resulted from deforming the pathology volume to its mirror volume. Notice the flow of deformation vectors indicates the direction of the deformation caused by the mass effect. Given that pathologies have a characteristic

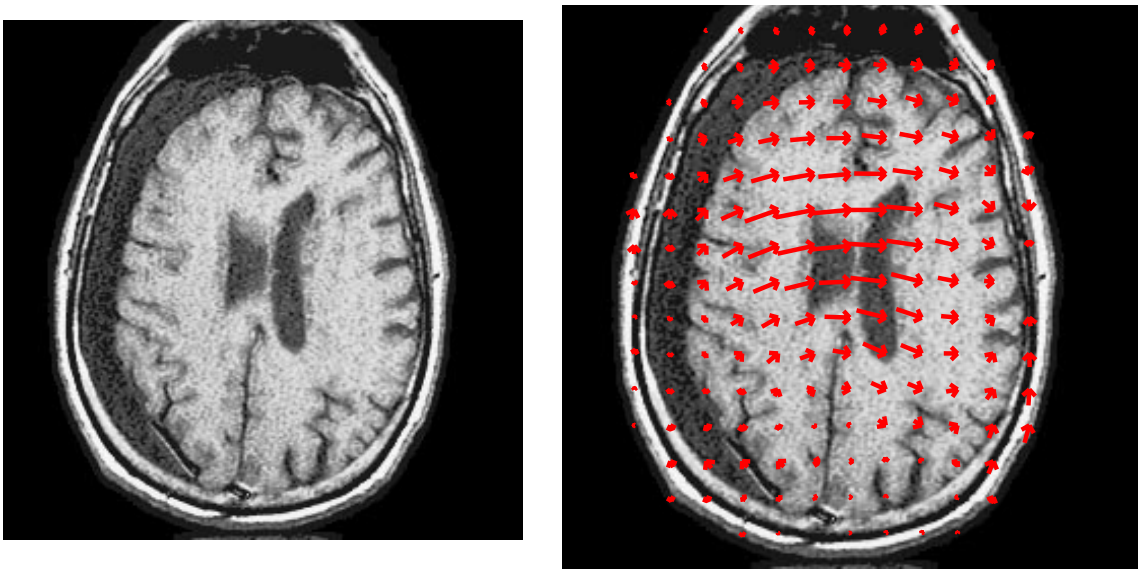


Figure 17. A cross-section of the pathology volume showing *chronic subdural hematoma* (left). The hematoma is the elongated dark area long the left side of the skull. Note it has caused mass effect, squashing the left ventricle. An overlay of the deformation vectors is displayed on the right.

influence on the brain morphology, we can apply this knowledge, along with a textual description of the symptoms (e.g. speech difficulty), to help identify and localize the pathology.

Knowledge of the normal range of variation can also help localizing the pathology. Figure 18 compares the corresponding cross-sections of the atlas after globally transformed (left) and then deformed (right) to match the pathology volume. After deformation, the anatomical structure labelled as *skull* becomes thicker and uneven. If this is out of the normal range of variation, it will imply a pathology. This applies to pathologies that only cause intensity variations as well. After matching the atlas to the data, the area of pathology will have an anatomical label. If the intensity distribution of the pathology does not conform to that of the labelled anatomical feature, it suggests an abnormality.

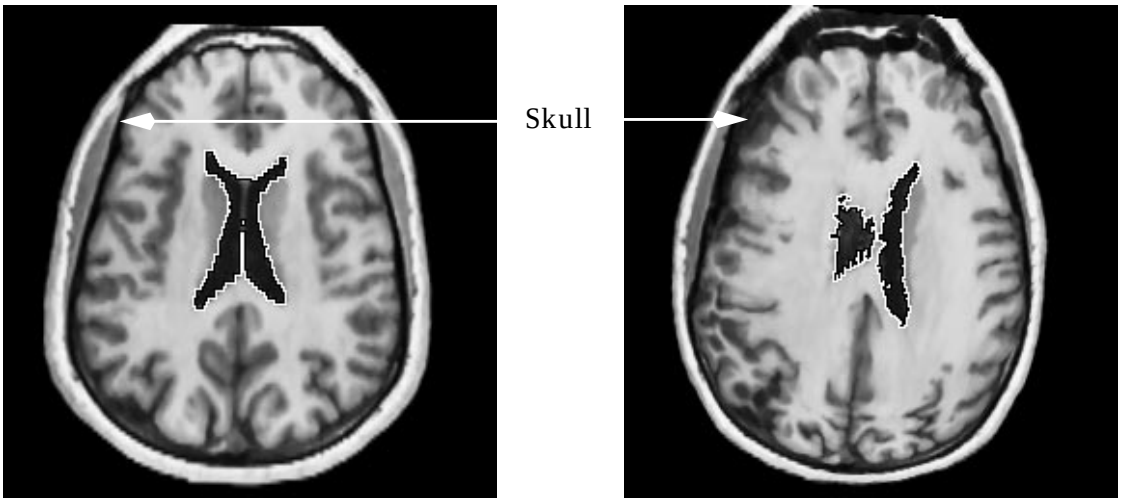


Figure 18. The corresponding cross-sections from the atlas after globally transformed (left) and deformed (right) to match the pathology volume. Note that after deformation the dark region (which is labelled in the atlas as the skull) becomes thicker and uneven.

- **Establish a standard method to quantitatively evaluate the registration results.**

The goal of the current approach is to deform the atlas into a *customized atlas* for the patient. To evaluate the performance, we would need multiple atlases. Using one of the atlases as the patient, the registration process computes its *customized atlas*. The registration accuracy can be quantified by the fraction of voxels whose labels given by an expert agree with that in the *customized atlas*. To acquire more atlases, we need to collaborate with medical experts and other groups working on related research. In addition to the evaluation of our algorithm, we would also like to do a quantitative comparison with related work.

- **Investigate the mutual information criterion for modality independent registration.**

The performance of our approach will decrease if there is large intensity discrepancy between the atlas and the patient volume. And it may fail at aligning data from different imaging modalities. Since mutual information can discern similar patterns despite differences in intensities, we plan to apply the mutual information criterion in the registration. Adopting this criterion will also enable us to register images from different modalities. Previous applications of mutual information have been restricted to rigid transformations. Since we used gradient descent in optimizing deformation parameters, we need an estimate of the derivative of the mutual information with respect to the parameters. Existing approaches for computing this derivative with respect to a large number of parameters are computationally prohibitive.

9. Expected Contributions

- **Pathology detection using deformable matching guided by domain knowledge.**
- **A standard methodology and test database for quantifying medical registration performance.**
- **A modality independent deformable registration algorithm.**

10. Time Table

May -- August, 1997	HP summer internship
by January 1998	Investigate pathology detection via deformable matching. Use knowledge of the normal variations of the anatomy and the characteristic impact of certain pathologies to guide registration.
by February 1998	Establish a standard method for quantitative evaluation of registration results and validate the approach.
by May 1998	Apply the MI criterion in registration to achieve modality independent deformable registration.
by September 1998	Write thesis and defend.

Acknowledgments

Many thanks to Henry Rowley, Shumeet Baluja, John Hancock, Prem Janardhan, Eric Rollins, and David Simon, for their generous help and precious comments. The atlas data was obtained by Yanxi Liu from the Brigham and Women's hospital of Harvard Medical School.

Appendix

The Levenberg-Marquardt algorithm is a non-linear optimization algorithm based on gradient descent. Suppose a vector function $R(p)$ expresses the residual difference at each voxel, where p is the vector of transformation parameters. The derivative of $R(p)$ with respect to p shows how each component of $R(p)$ will change given a change in the transformation parameters Δp.

$$\Delta R(p) = \frac{\partial R}{\partial p}\Delta p \tag{8}$$

Ideally, the goal would be to find the transformation parameters that make $R(p)$ equal to the zero vector:

$$R(p_{new}) = R(p_{old}) + \Delta R(p) = 0 \tag{9}$$

Because of the existence of noise and the innate discrepancies in different volumes, $R(p)$ can not be reduced to the zero vector in practice. The Levenberg-Marquardt algorithm attempts to minimize $R(p)$ by adjusting the transformation parameter vector p. From (8) and (9) we have

$$-R(p_{old}) = \frac{\partial R}{\partial p}\Delta p \tag{10}$$

For a given p, there are far more voxels than the number of transformation parameters, so $\frac{\partial R}{\partial p}$ is not a square matrix. To solve this over-constrained system for p, the Levenberg-Marquardt algorithm employs the pseudo inverse method.

$$\Delta p = \left[\left(\frac{\partial R}{\partial p}\right)^T \frac{\partial R}{\partial p}\right]^{-1}\left[\left(\frac{\partial R}{\partial p}\right)^T (-R(p_{old}))\right] \tag{11}$$

$$\left(\frac{\partial R}{\partial p}\right)^T \frac{\partial R}{\partial p} \tag{12}$$

If matrix (12) is not of full rank, the pseudo inverse will not exist. The Levenberg-Marquardt algorithm adds a stabilizing term λ to the diagonal elements of the matrix

$$\Delta p = \left[\left(\frac{\partial R}{\partial p}\right)^T \frac{\partial R}{\partial p} + \lambda I\right]^{-1}\left[\left(\frac{\partial R}{\partial p}\right)^T (-R(p_{old}))\right] \tag{13}$$

In each iteration of the Levenberg-Marquardt optimization, Δp is computed from the current $R(p_{old})$ and p_{old}. The summation of Δp and p_{old} gives p_{new}. The $R(p_{new})$ is in turn used to compute the new Δp. The iteration goes on until Δp is smaller than a user defined threshold, at which point the transformation parameters that minimize the residual difference are considered to be recovered [27].

References

[1] Christensen et al., “Individualizing Neuroanatomical Atlases Using A Massively Parallel Computer”, IEEE Computer, pp. 32-38, January 1996.

[2] Christensen et al., “Deformable Templates Using Large Deformation Kinematics”, IEEE Transactions on Image Processing, September 1996.

[3] Horn, “Closed-form solution of absolute orientation using unit quaternions”, Journal of the Optical Society of America, Vol. 4, No. 4, pp 629-642, April, 1987.

[4] Lavallee et al., “Image-Guided Operating Robot: A Clinical Application in Stereotactic Neurosurgery”, Computer-Integrated Surgery--Technology and Clinical Applications, Taylor et al., pp. 346-351.

[5] McInerney and Terzopoulos, “Deformable models in medical image analysis: a survey”, Medical Image Analysis, Vol. 1, No. 2, pp 91-108, 1996.

[6] Szeliski, “Matching 3-D Anatomical Surfaces with Non-Rigid Deformations using Octree-Splines”, International Journal of Computer Vision, Vol. 18, No. 2, pp 171-186, 1996.

[7] Szeliski and Coughlan, “Spline-Based Image Registration”, Digital Equipment Corporation, Cambridge Research Lab, Technical Report 94/1.

[8] Bajcsy and Kovacic, “Multiresolution Elastic Matching”, Computer Vision, Graphics, and Image Processing, Vol. 46, pp 1-21, 1989.

[9] Gee, Reivich and Bajcsy, “Elastically Deforming 3D Atlas to Match Anatomical Brain Images”, Journal of Computer Assisted Tomography, Vol. 17, No. 2, pp 225-236, 1993.

[10] Jean-Philippe Thirion, “Fast Non-Rigid Matching of 3D Medical Images”, INRIA, Technical Repart No. 2547, May, 1995.

[11] van den Elsen, Pol, and Viergever, “Medical Image Matching--A Review with Classification”, IEEE Engineering in Medicine and Biology, Vol. 12, No. 1, pp. 26-39, 1993.

[12] Antoine Maintz, van den Elsen, and Viergever, “Evaluation of Ridge Seeking Operators for Multimodality Medical Image Matching”, IEEE Transactions on Pattern Analysis and

Machine Intelligence, Vol. 18, No. 4, April, 1996.

[13] Davatzikos, “Spatial Normalization of 3D Brain Images Using Deformable Models”, Journal of Computer Assisted Tomography, Vol. 20, No. 4, pp. 656-665, 1996.

[14] Szekely et al., “Segmentation of 2-D and 3-D Objects from MRI Volume Data Using Constrained Elastic Deformations of Flexible Fourier Contour and Surface Models”, Medical Image Analysis, Vol. 1, No. 1, pp. 19-34, 1996.

[15] Sclaroff and Pentland, “On Modal Modelling for Medical Images: Underconstrained Shpae Description and Data Compression”, M.I.T. Media Laboratory Perceptual Computing Section Technical Report No. 275.

[16] Sclaroff and Pentland, “Modal Matching for Correspondence and Recognition”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 6, June, 1995.

[17] Moghaddam Nastar and Pentland, “A Bayesian Similarity Measure for Direct Image Matching”, M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 393.

[18] Pokrandt, “Fast Non-Supervised Matching: A Probabilistic Approach”.

[19] Pokrandt, Stein and Hassfeld, “Medical Image Registration: Fast Feature Space Evaluation with Gaussian Entropy Function”.

[20] Wells III et al., “Multi-Modal Volume Registration by Maximization of Mutual Information”, Medical Image Analysis, Vol. 1, No. 1, pp 35-51, 1996.

[21] Viola and Wells III, “Alignment by Maximization of Mutual Information”, Proceedings of the fifth International Conference on Computer Vision, pp 16-23, 1995.

[22] Maes et al., “Multimodality Image Registration by Maximization of Mutual Information”, IEEE Transactions on Medical Imaging, Vol. 16, No. 2, April, 1997.

[23] Press et al., “Numerical Recipes in C”, Cambridge University Press, 1992.