Anomaly Detection through Registration

Mei Chen, Takeo Kanade, Dean Pomerleau, Henry A. Rowley
School of Computer Science, Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA 15213, USA
({meichen, tk, pomerlea, har}@cs.cmu.edu)

Abstract

We study an application of image registration in the medical domain. Based on a 3-D hierarchical deformable registration algorithm, we developed a prototype for automatic registering a standard atlas to a patient's data to create a customized atlas. The registration algorithm can also be applied to detect asymmetry in the patient data to help indicate the existence and location of any pathology. We have conducted experiments on 11 MRI scans of normal brains, 3 MRI and 1 CT scan of brains with pathologies.

1. Motivation

Human anatomy presents a challenge to image registration algorithms. Because of genetic and environmental factors and because of diseases, biological structures have a large range of variation in appearance. In neuroscience, 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 shapes of the brain. For example, pathologies like bleeding or tumors can cause a shift of brain structures called mass effect. Traditionally, doctors have been using manual registration to detect lesions that make brain structures deviate from the norm, to analyze variations between normal brains, and to plan surgeries. Since the manual registration of anatomical structures is labor-intensive and prone to be inconsistent, many researchers are investigating automatic registration methods. While there have been some encouraging results, due to the complexity of the problem, it remains unsolved.

We present a 3-D *hierarchical deformable registration algorithm*, and a prototype for Anomaly Detection thrOugh REgistration, A*DORE*. ADORE is designed to automatically and accurately match an atlas (a hand-segmented image set of normal anatomy) to a patient's data to create a customized atlas, as well as to indicate any pathologies in the patient's data. It will facilitate image-based retrieval of similar cases in medical databases, assist doctors in detecting pathologies, comparing different pathologies' impact on brain morphology, observing the development of a pathology over time, and studying 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, e.g. *computerized tomography* (CT) and *magnetic resonance imaging* (MRI). MRI is good at revealing soft tissue structures, while CT is good at uncovering bony structures.

In neurosurgery, the three principal axes of the head are called *axial*, *sagittal*, and *coronal* (see Figure 1). An MRI or CT scan consists of a series of parallel cross-sections 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 the image data, so the axis 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.

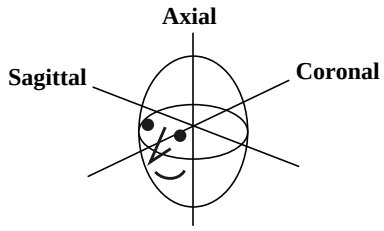


Figure 1. The three principal axes of a head volume.

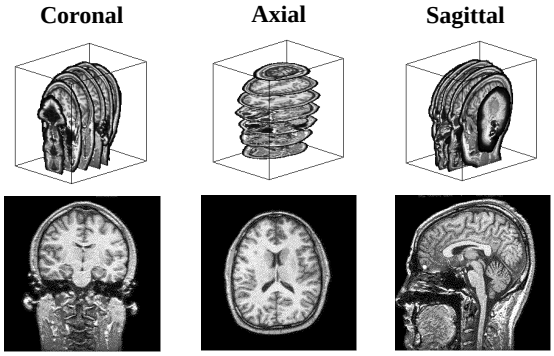


Figure 2. MRI scans along the principal axes.

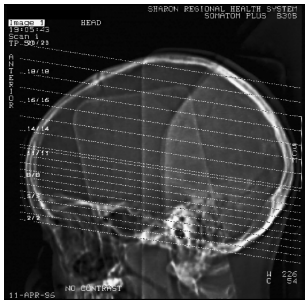


Figure 3. A typical axial scan pattern (indicated by the white lines). Note that the scanning direction is at an angle to the axial axis, the inter-scan spacings are non-uniform, and the scans do not cover the entire head volume.

One problem we will discuss is matching an atlas to a patient's data to create a customized atlas. The brain atlas we use is a set of 123 coronal T1 weighted MRI scans of a normal brain where each voxel measures 0.9375 x 0.9375 x 1.5 mm^3 . 144 anatomical structures were hand segmented and labelled (courtesy of the Brigham and Women's Hospital of the Harvard Medical School) (see Figure 4). By matching the atlas to a patient's data to create a customized atlas, we can segment and label the anatomical structures in the patient's data. Figure 5 depicts the scenario in which the labels of anatomical structures in the customized atlas are used to segment the corresponding anatomical structures of the patient. Note that the original atlas (Figure 4, right) is warped into a customized atlas for the patient (Figure 5, left). We could choose to warp the patient data to match the atlas. However, since the patient's data will be of more interest for diagnosis and analysis, we prefer to keep it unchanged

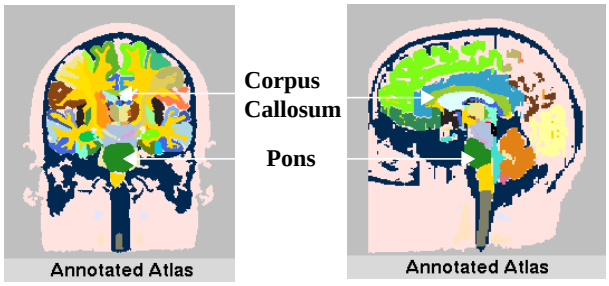


Figure 4. Coronal and sagittal cross-sections of the atlas. Two anatomical structures are annotated for illustration.

and warp the atlas instead.

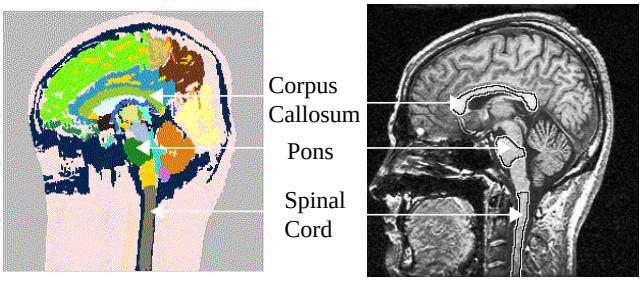


Figure 5. The scenario in which anatomical structures in a patient volume (right) are precisely segmented and labelled using information from the customized atlas (left).

Aside from creating the customized atlas, we will investigate the problem of using registration to indicate any pathology in the patient's data. Figure 6 displays the corresponding axial cross-sections of the atlas and a brain with pathology. The atlas exhibits an approximate symmetry about the central line, whereas the symmetry is destroyed in the patient's data due to the existence of the pathology. Note that the shape, size, location, and intensity of the anatomical structures are different in the two data sets.

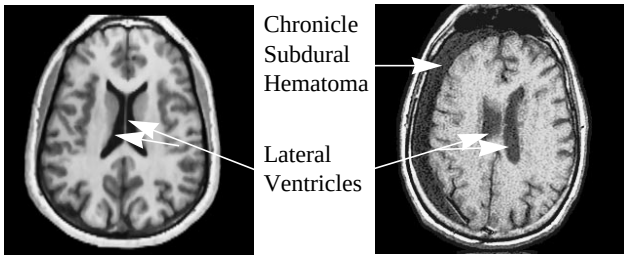


Figure 6. The corresponding axial cross-sections of the atlas (left) and a brain with pathology (right). Note the exhibition of symmetry in the atlas, and the loss of symmetry in the patient. Also note the difference in the shape, size, location, and intensity between the corresponding anatomical structures.

The variations between different data sets stem from two sources. The *extrinsic* source is the scanning process, which results in 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 image volumes into alignment.

3. Current Approach for Matching

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

We adopt a *voxel-based* approach which assumes no prior segmentation among voxels in a volume. 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. Alternative approaches to anatomical registration are reviewed and compared in section 7.

3.1. Highlights of the Approach

The major points of the current matching method are listed below. We will elaborate on them in the following sections.

- **Preprocessing**
  - Apply a *maximum-connected-component* method to extract the data of interest (the head volume), from the background to reduce distortions 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 a *global transformation* (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 (3.2). This process adjusts for variations caused by all extrinsic factors except for intensity inhomogeneities (which was dealt within preprocessing).
  - Employ a *smooth deformatio*n represented by the warping of a grid of 3-D control points to approximately align the corresponding anatomical structures in the atlas to those of the patient. An iterative optimization algorithm is used to determine the deformation parameters (3.3.1). This process partially adjusts variations caused by intrinsic factors.
  - Employ a *fine-tuning deformation* in which each voxel moves independently to improve the alignment. An iterative optimization algorithm is used to determine the deformation parameters (3.3.2). This process further adjusts for variations caused by intrinsic factors.

3.2. Registration via Global Transformation

We address the variations from extrinsic sources by applying a *global transformation* to the atlas volume to grossly align it with the patient volume. This *global transformation* is composed of 3-D rotation, translation, and uniform scaling.

3.2.1 Representation of the Global Transformation

Figure 7 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). The Z axis coincides with the axis along which the volume was scanned.

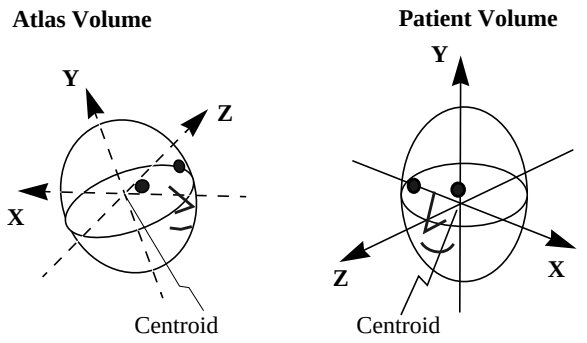


Figure 7. The coordinate system.

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.

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

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 [3].

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

The uniform scaling is denoted by a scalar *s*.

The order in which we apply these transformations to the atlas volume is first 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 T*, $(q_0, q_x, q_y, q_z, t_x, t_y, t_z, s)$. Because the atlas and the patient volumes are innately different, we should not expect to find a transformation that exactly matches them. We can only pursue a transformation that minimizes the error.

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

$$SSD(T) = \sum_{(x,y,z)} \left(I_{Patient}(x,y,z) - I_{Atlas}(T(x,y,z)) \right)^2 \quad (3)$$

$I_{Patient}(x,y,z)$ is the intensity of voxel $[x,y,z]_{patient}$ in the patient volume, and $I_{Atlas}(T(x,y,z))$ is the intensity of voxel $[T(x,y,z)]_{atlas}$ in the atlas volume. $T(x,y,z)$ is (x,y,z) after applying the *global transformation T*. The SSD is a function of this transformation *T*. We apply the Levenberg-Marquardt non-linear optimization algorithm to iteratively adjust the transformation parameters to reduce the SSD. The iteration continues until the changes in the transformation parameters are below a user defined threshold, at which point the *global transformation* is considered to be recovered. We employ multi-resolution processing and stochastic sampling for efficiency and to help prevent the optimization process from being trapped in local minima (see section 4.2).

After applying the transformation, the voxel coordinates in the atlas may not have integral coordinates. Tri-linear interpolation is employed to determine a voxel's intensity from its eight bounding neighbors (see Figure 8). 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.

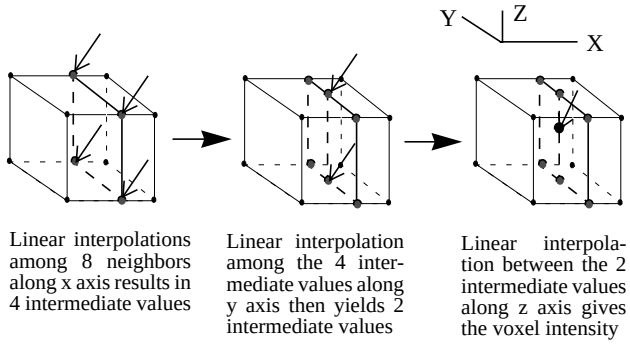


Figure 8. Tri-linear interpolation gives the intensity of a voxel by interpolating among its 8 bounding neighbors.

Figure 9 displays the gray level data of the atlas, a patient's data, and the result of using the global transformation to align the atlas with the patient. Note that the atlas volume is rotated (approximately 90 degrees, from the coronal view to the sagittal view), translated, and scaled to grossly match with the patient's volume. Labels of several anatomical structures in the transformed atlas are directly projected to the patient volume. They only roughly align with the corresponding anatomical structures of the patient, and considerable discrepancies remain in their shape, size, and location.

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 anatomical structures, as shown in Figure 9.

Because brain morphology is intricate, the discrepancy between the atlas and the patient cannot be fully addressed by applying one transformation to the whole atlas volume. Localized deformation is necessary.

3.3.1 Smooth Deformation

Although the variations across individuals can be large, the shape and density of corresponding anatomical structures are still distinct enough for them to be related. Therefore, the intensity difference between spatially corresponding voxels can act as the deforming force. The deformation process causes atlas voxels to be spatially shifted towards their counterparts in the patient volume.

The most intuitive way to represent the deformation would be a 3-D displacement for each voxel. This will allow 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 align

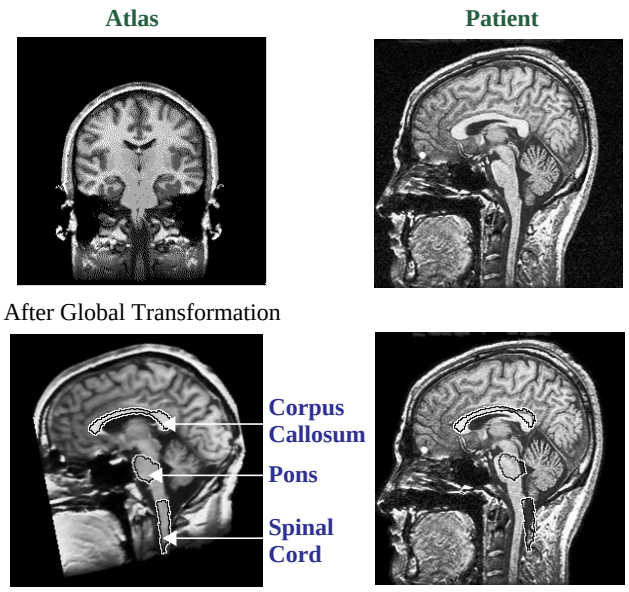


Figure 9. Cross-sections of the atlas volume and a patient's volume, before (top row) and after (bottom row) registration via *global transformation*. Notice the atlas is rotated from a coronal view to align with the sagittal view of the patient. The outlines of anatomical structures in the transformed atlas are directly projected on the patient volume. Note that they only roughly align with their counterparts in the patient data.

the atlas to the patient precisely enough for a good initial alignment. We need a more robust representation for intermediate deformations.

- **Representing the Smooth Deformation**

Our solution is to ensure a *smooth deformation* by placing a 3-D control grid in each volume. The vertices of the control grid are control points. The 3-D displacements of the control points determine the deformation of the voxels they enclose. Therefore although the control points can deform freely, the voxels inside each control 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 10 shows an example of a control grid cell and an arbitrary deformation of it. Szeliski employed similar representation in 2-D registration [7] and 3-D registration of surfaces, [6]. Figure 11 is a 2-D illustration of the s*mooth deformation* process.

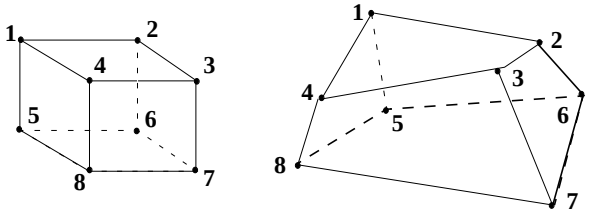


Figure 10. A control grid cell whose vertices are control points (left), and an arbitrary deformation of it due to movements of the control points (right).

- **Estimating the Smooth Deformation**

Although we represent the *smooth deformation* using only the 3-D

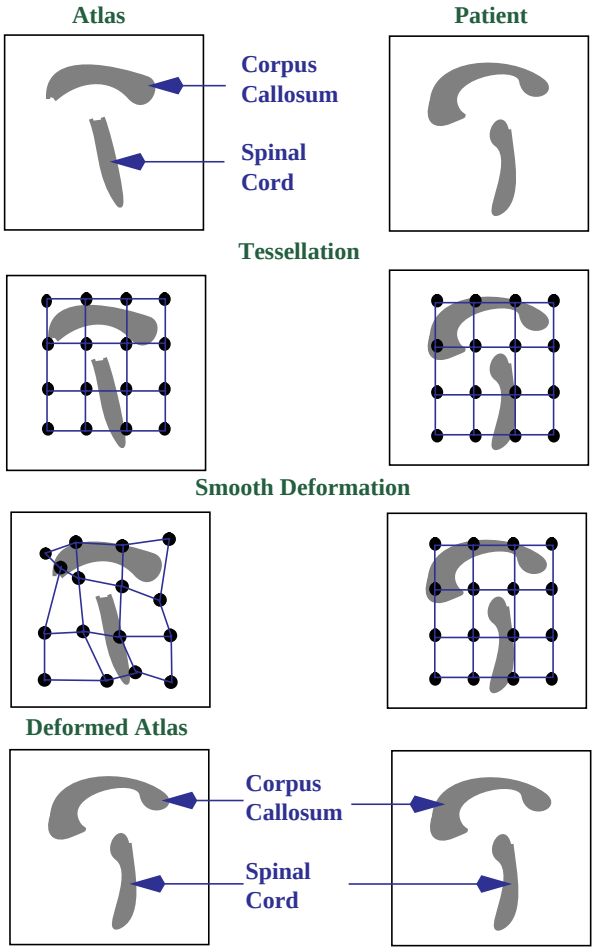


Figure 11. A 2-D illustration of the *smooth deformation.* The first row is the original data. The second row is the original data overlaid with the regular control grids. In the third row the control points of the atlas are shifted towards their counterparts in the patient. In the last row the atlas is deformed to match the patient.

displacements of the control points, we estimate the deformation force using the intensity differences over the whole volumes. 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 ith in the array of all control grid cells $Cell_{patient}[i]$. The corresponding grid cell for 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 the location of its counterpart in the atlas volume $[S(x, y, z)]_{atlas}$ using the 8 control points of $Cell_{atlas}[i]$. Should $[S(x, y, z)]_{atlas}$ have non-integral coordinates, tri-linear interpolation is applied to compute its intensity from its 8 neighboring voxels in the atlas (see Figure 8). In the case that $[S(x, y, z)]_{atlas}$ falls outside the atlas volume, it is ignored.

Under an ideal *smooth deformation*, the atlas volume should completely align with the patient volume, and $I_{atlas}(S(x, y, z))$ should equal to $I_{patient}(x, y, z)$. In practice, $I_{atlas}(S(x, y, z))$ 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) between all voxels' intensities in the volumes as the error function, and employ the Levenberg-Marquardt iterative non-

linear optimization method to determine the *smooth deformation* parameters that best match the two volumes.
The only difference between this optimization process and the one in 3.2.2 is that the parameters we compute are not the global transformation parameters, but the collection of local *smooth deformation* parameters for each control grid cell. The control grid cells do not deform independently, the deformations are linked by shared control points across the 3-D control grid. They collectively define the *smooth deformation*.
Figure 12 shows the effect of the *smooth deformation* on the previous example. Notice that the atlas volume is deformed in 3-D to match the patient volume. The labels of the anatomical structures in the deformed atlas approximately align with the corresponding structures of the patient.

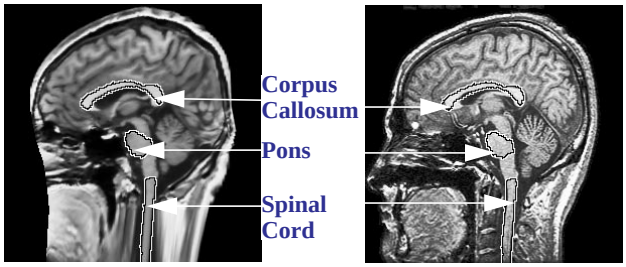


Figure 12. Corresponding cross-sections of the atlas (left) and the patient (right) after registration via *smooth deformation*. Note the atlas is deformed to match the patient. Labels of anatomical structures in the atlas approximately align with the corresponding structures in the patient.

3.3.2 Fine-Tuning 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 independently, it can not account for any details smaller than the size of a control grid cell. To adjust to finer details, we apply a *fine-tuning deformation*.

- **Representing the Fine-Tuning Deformation**

Similar to the smooth deformation, the intensity difference between spatially corresponding voxels acts as the deforming force. The difference is that the 3-D displacement of each voxel in the atlas volume partakes in the representation of the *fine-tuning deformation*. Therefore each voxel can deform freely, allowing the deformation to attend to details. This procedure is similar to the approach discussed in [10].

- **Estimating the Fine-Tuning Deformation**

For a voxel $[x, y, z]_{patient}$ in the patient volume with intensity $I_{patient}(x, y, z)$, its corresponding voxel $[D(x, y, z)]_{atlas}$ in the atlas volume has intensity $I_{Atlas}(D(x, y, z))$. $D_{(x, y, z)}$ is (x, y, z) after applying the current deformation D. Suppose $\delta D(x, y, z)$ is the difference between the current deformation D and the optimum deformation, then we have:

$$I_{Atlas}(D(x, y, z) + \delta D(x, y, z)) - I_{Patient}(x, y, z) = 0 \quad (4)$$

The first order Taylor expansion of the first term in (4) gives

$$I_{Atlas}(D(x, y, z)) + \delta D(x, y, z) \cdot \nabla I_{Atlas}(D(x, y, z)) \quad (5)$$

From (4) and (5), one solution for $\delta D(x, y, z)$ is:

$$\delta D(x, y, z) = \frac{I_{Patient}(x, y, z) - I_{Atlas}(D(x, y, z))}{\left\|\nabla I_{Atlas}(D(x, y, z))\right\|^2} \nabla I_{Atlas}(D(x, y, z)) \quad (6)$$

To prevent the deformation parameters from going out of bounds when the atlas' gradient $\nabla I_{Atlas}(D(x, y, z))$ is close to zero, we add a stabilizing factor α:

$$\delta D(x, y, z) = \frac{I_{Patient}(x, y, z) - I_{Atlas}(D(x, y, z))}{\left\|\nabla I_{Atlas}(D(x, y, z))\right\|^2 + \alpha} \nabla I_{Atlas}(D(x, y, z)) \quad (7)$$

The deformation D is recovered by iteratively solving for δD, until δD is smaller than a user defined threshold. Because the number of deformation parameters is 3 times the number of voxels, the problem is under-constrained. We apply 3-D Gaussian smoothing to the 3-D deformation parameters after each iteration to constrain the deformation. This process is similar to an optical flow algorithm. Trilinear interpolation is used to compute $I_{Atlas}(D(x, y, z))$ when the voxel coordinates after deformation are not integers.
Figure 13 is the result of the previous example after the fine-tuning deformation. Note that the atlas is further warped in 3-D to better align with the patient. Labels of anatomical structures in the atlas match well with their counterparts in the patient.

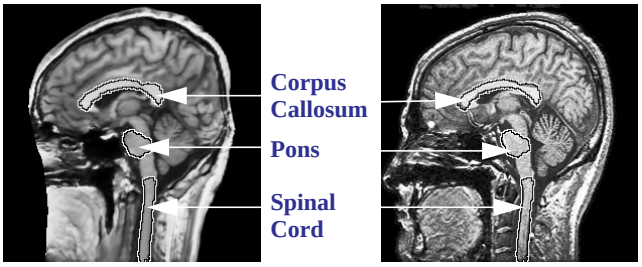


Figure 13. Corresponding cross-sections of the atlas (left) and the patient (right) after registration via *fine-tuning deformation*. Note the atlas is warped more to match the patient. Labels of anatomical structures in the atlas match well with their counterparts in the patient volume.

Once the atlas is deformed to align with the patient, we call it a *customized atlas*. Labels of anatomical structures in the *customized atlas* can be used to directly segment and label the patient volume.

4. Algorithm 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 prior segmentation, each voxel in the data is assumed to play an equal role in the registration process. There is usually background noise in the data, which necessitates background separation 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 is the first valley of the intensity histogram, after it is smoothed using a Gaussian filter (see

Figure 14). Since the region containing the head is connected, we apply a connected-component algorithm to the binary volume and 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.

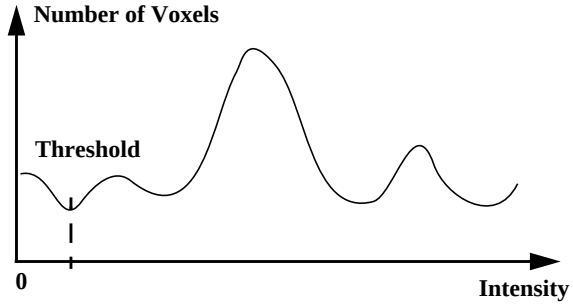


Figure 14. Thresholding based on the intensity histogram.

4.1.2 Histogram Normalization

We adopted an intensity-based method to avoid errors introduced by unreliable feature extraction processes, but inter-scan intensity variations make this approach problematic. The current solution is to perform histogram normalization on each volume prior to the registration process. To prevent outlier voxels with extreme intensities from stretching the histogram, we set the intensity of the darkest 2% of the voxels to 0, the intensity of the brightest 2% of the voxels to 255, and 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 Multi-resolution Processing

A multi-resolution strategy is used not only to improve efficiency, but also to help the optimization procedure to first focus on global patterns and gradually shift to the details. We employ two kinds of multi-resolution 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. Currently it has 3 levels, each level being half the resolution of the next higher level. At the lower resolutions there are fewer details present, so the minimization process has less tendency to become trapped in local minima. By inheriting results from a lower resolution, the higher resolution registration can start closer to the global minimum. Optimization at the higher resolutions uses the finer details to refine the result.

The control grid pyramid is a hierarchy of different resolution control grids. Currently it has 5 levels, in which the numbers of control points along x, y, and z directions are 2x2x2, 3x3x3, 4x4x4, 5x5x5, and 6x6x6. 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

While estimating the *global transformation* and the *smooth deformation*, we do not process every voxel of the volumes during each iteration of the optimization process. Instead, we sample a random set of voxels at each iteration. This improvement in computation efficiency is possible because the optimizations are over-constrained. Moreover, the stochastic nature of the sampling helps the minimization process to escape from local minima [20]. Note that because the *fine-tuning deformation* has 3 times as many parameters as the number of voxels, the problem is highly underconstrained, so stochastic sampling is not appropriate.

4.2.3 Parallel Processing

In our approach, the computation at each voxel is identical, so the process is voxel-wise parallelizable. To reduce overhead, we employ parallel processing at a higher level than the voxel representation. During the registration via *global transformation* and during the *fine-tuning 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. Because a voxel in the atlas volume can map to any position in the patient's volume, it is difficult to partition the volumes so each part can be processed independently. Therefore, we used a shared-memory multi-processor computer in which each processor has access to the full volumes.

5. ADORE

Based on the 3-D *hierarchical deformable registration algorithm*, we developed a prototype, ADORE (Anomaly Detection thrOugh REgistration). ADORE is designed to create a customized atlas for the patient, and use registration to indicate pathologies in the patient data. Experiments on actual medical data were carried out to evaluate the performance in matching and pathology indication.

5.1. Registration Performance

We conducted experiments on eleven MRI data sets. The performance was successful on five data sets, reasonable on four data sets, and poor on two data sets. Consistent satisfactory performance is impeded by two factors: one is that our approach is intensity-based, and suffers when the patient volume has a considerably different intensity distribution from the atlas; another factor is that if the patient data is coarsely sampled or only covers part of the head volume, it will not contain sufficient information for accurate registration.

The result of matching the atlas to one normal brain was presented in Section 3. Figure 15 shows the progressive results of the 3-D *hierarchical deformable registration* between the atlas and a brain with pathology. Note that the *global transformation* only addresses the extrinsic variations between the atlas and the patient using 3-D rotation, scaling, and translation. The atlas is rotated approximately 90 degrees from the coronal view to match the axial view of the patient. Labels of anatomical structures in the atlas can not align with the corresponding structures in the patient data. The intrinsic variations are adjusted roughly during the *smooth deformation* via 3-D warping of the control grid. Labels of anatomical structures in the atlas approximately align with the corresponding structures in the patient data. The deformation is refined using the *fine-tuning defor-*

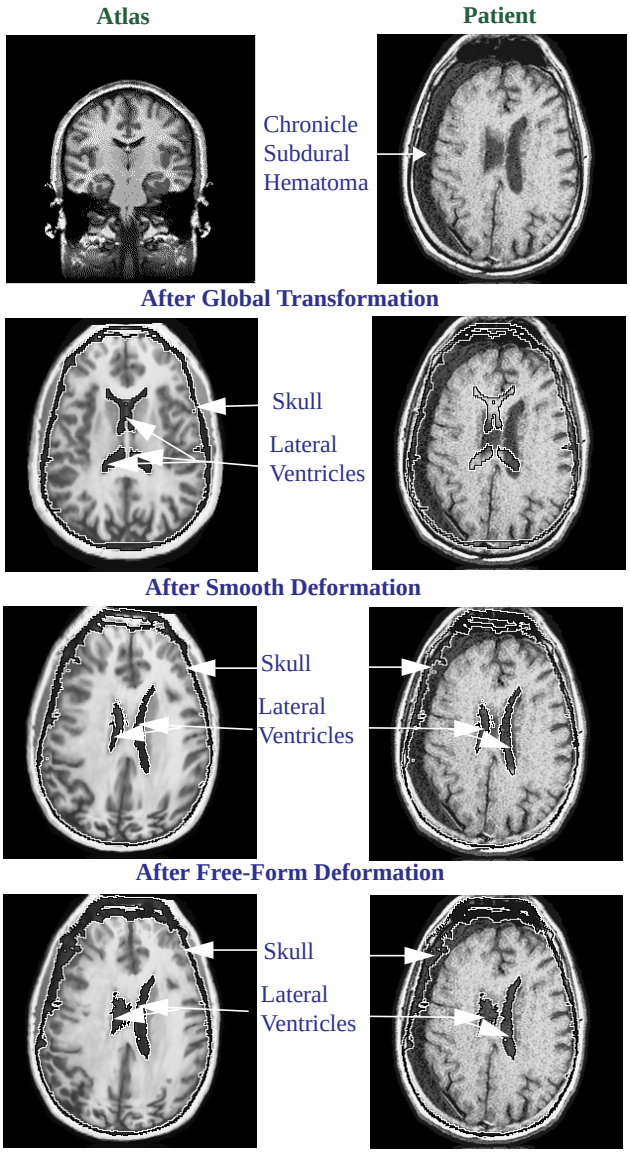


Figure 15. The progressive result of the 3-D hierarchical deformable registration between the atlas and a patient with pathology. The atlas was warped in 3-D to match the patient. The alignment of the corresponding anatomical structures improve along the registration hierarchy.

mation, where each voxel moves independently to align with its counterpart. Labels of anatomical structures in the atlas align well with the corresponding structures in the patient data. Figure 16 is a close-up on the registration of the lateral ventricles along the deformation hierarchy. The improvement of alignment accuracy is significant. The contours of the lateral ventricles appear to change during the registration process because the transformation and deformations move the 3-D volumes with respect to the fixed cross-section we are observing.

5.2. Anomaly Detection

It is of great clinical significance to create the customized atlases and to indicate pathologies for patients with neurological conditions. Certain pathologies cause mass effect, which shifts the brain

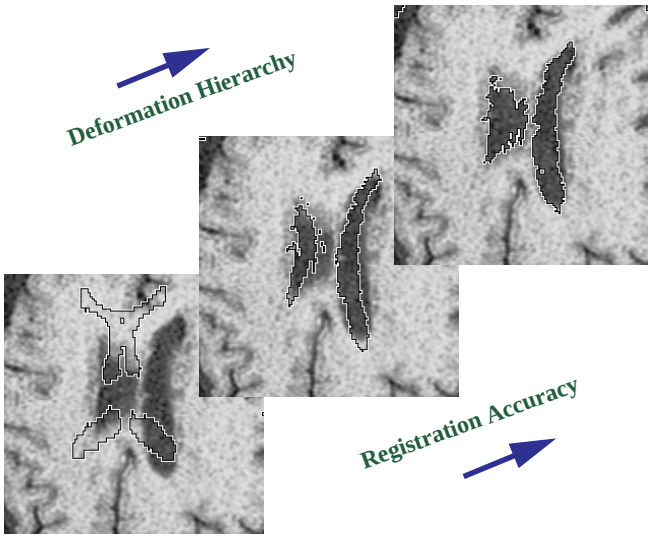


Figure 16. A close-up on the registration of the lateral ventricles at each stage in the registration hierarchy. The improvement in accuracy is significant.

structures, and causes anatomical features close by to deform significantly (see Figure 6). However, since our deformation process allows local deformations, it can still create the customized atlases with anatomical structures matching their counterparts in the patients' data, such as the case in Figure 15.

For the experiments on anomaly detection, we matched the patient's mirror volume (with left and right exchanged) to itself. Since a normal brain is approximately symmetric about the central line (visible from the axial and coronal view), not much deformation is needed to align its mirror volume to itself. A significant amount of deformation will indicate the absence of symmetry, and the possible existence of pathologies. In Figure 17, the left column displays an axial cross-section of each patient's volume, the right column is the same cross-section overlaid with the deformation vectors of matching the patient's mirror volume to itself. For the normal brain (first row), there is very little deformation, because it is approximately symmetric. For the brain with chronicle subdural hematoma (second row), there is significant and uniform deformation. This is because the pathology caused mass effect and destroyed the symmetry of the brain, forcing the anatomical structures (such as the lateral ventricles) to shift considerably. The direction of the deformation vectors reveals the source of the mass effect, which is the possible location of the pathology. For the brain with a tumor (third row), aside from the uniform deformation caused by the shifting of the lateral ventricles, there is a swirl pattern in the deformation vectors. This is because the tumor is so close to the central line that its "counterpart" in the mirror volume managed to shift back (partly) to match with itself. Note that unlike the other data sets, this is a CT scan. The characteristics of the deformation vectors combined with domain knowledge can provide indications of the existence and location of anomalies.

6. Evaluation and Analysis of Matching

It is difficult to quantitatively evaluate registration results in medical images, due to the lack of ground truth. We give a qualitative assessment of our algorithm in terms of simplicity, accuracy, and speed.

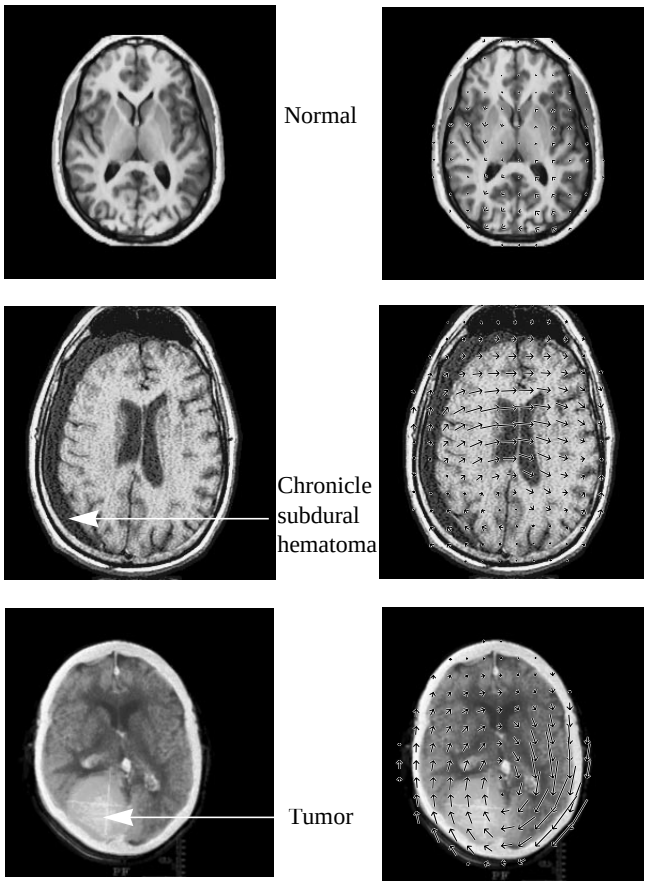


Figure 17. An axial cross-section of each patient volume (left), and the same cross-section overlaid with corresponding deformation vectors (right). The pattern, magnitude and direction of the deformation vectors indicate the possible existence and location of the pathologies.

6.1. Simplicity

Our current approach, from preprocessing, registration via *global transformation*, registration via *smooth deformation*, to registration via *fine-tuning 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.

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

Quantitative evaluation of a matching method is hard to perform due to the lack of ground truth. Figure 18 compares ADORE's segmentation and labelling of a patient volume with an expert's result, and it is noticeably less accurate at fine details. This is because ADORE adopts an intensity-based approach, but the same anatomical structure in different data sets can have different intensities.

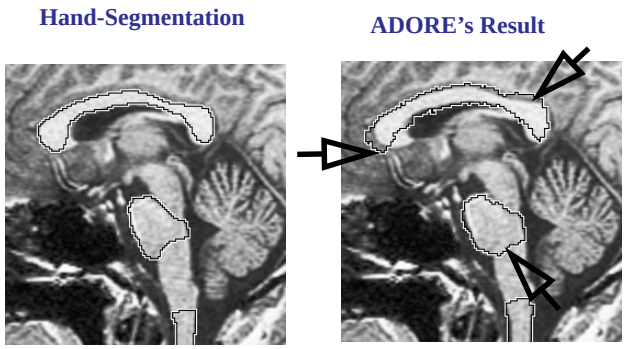


Figure 18. Labels from hand segmentation compared to the result of ADORE's.

7. Related Work

The registration of medical images via optimization in transformation space has been an active research area--the comprehensive survey article by van den Elsen et al. [11] 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 differences 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 both modalities. They yield high accuracy matching with respect to rigid transformations [4]. But to ensure accuracy, the markers must be rigidly attached to the patient (by driving screws into the patient's skull), so as to indicate precisely the patient's position and orientation. Obviously, this is invasive and inconvenient.

Registration methods using internal image properties have the advantage of being non-invasive and fully retrospective, which means that one does not need to know prior to the acquisition of the images whether matching will be required. They also have the potential to deal with non-rigid transformations between different image sets. This makes them more favorable for 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 correspondences between them. They have the characteristic of being efficient in representation and independent of imaging modality. However, feature-based registration is critically dependent on the quality of the feature extraction, which is not trivial 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.

Many researchers use *deformable models* for *feature-based non-rigid registration*. [5] 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 [9],[8]. The atlas is modelled as a physical object and is given elastic properties. 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 an intermediate stage. Without user interaction, their atlas can have difficulty converging to complicated object boundaries. Also their method is more computationally expensive and requires interactive and time-consuming preprocessing.

Christensen et al. presented a method very close to ours, but they used a fluid dynamic model for the deformation [1], [2]. 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 [10], 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 computation time, he used the gradient of the patient volume instead of the deformed atlas, because the computation of the latter is more expensive, requiring trilinear 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.

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) [18], [19], [20], [21]. 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 MI to registration 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, and to tackle pathological cases which are of more clinical importance.

8. Research Directions

Our experiments demonstrated the strength and promise of the hierarchical deformable registration algorithm, and the automatic registration and pathology detection prototype, ADORE. However, for them to be applied to real medical practice, much research remains to be done.

8.1. Improve Registration Performance

Although the histogram normalization step can address considerable intensity discrepancies between data sets, the performance of our approach still decreases if the discrepancy is large. It may also fail to align data from different imaging modalities. Since mutual information can discern similar patterns despite differences in intensities, and has been successfully applied to multi-modality registration via rigid transformations, it is promising to apply this criterion to deformable registration.

8.2. Use Domain Knowledge as Guidance

Our current approach has no constraints on the deformations. If the deformation could be limited to shapes within the normal anatomical variation of structures, or the characteristic deformation caused by pathologies, potential local minima could be avoided.

Knowledge of natural shape variability and pathology-afflicted deformation can be used to define allowable deformations, as well as to assist pathology detection. Figure 19 compares the corresponding cross-sections of the atlas after global transformation (left) and then local deformation (right) to match a patient. After deformation, the anatomical structure labelled as *skull* becomes thicker and uneven. Since this is out of the normal range of variation, it implies the possible existence of 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.

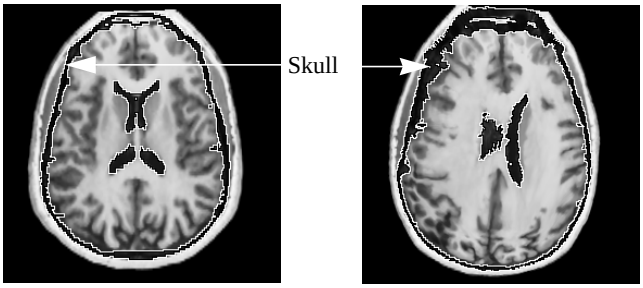


Figure 19. The corresponding cross-sections from the atlas after global transformation (left) and local deformation (right) to match a patient volume. Note that after deformation the region labelled as the skull becomes thicker and uneven.

Domain knowledge can be acquired from the statistics of training samples. The challenge is how to formulate a representation of the knowledge that can be incorporated into the registration procedure. For information described by a large number of possibly highly correlated parameters, principal component analysis (PCA) may offer a promising solution by characterizing the few largest eigendeformations, reducing the dimensionality substantially. To represent the information in a way that is appropriate for applying PCA, Fourier descriptors or modal representations have been examined in previous work [17], [19], [20].

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.

8.3. Quantitative Evaluation

Quantitative evaluation of registration results is important because that will allow researchers to compare different approaches, and to assess the trade-off between computational efficiency and accuracy. One way to quantitatively evaluate the performance of ADORE would be to employ hand-labelled multiple atlases. Using one of the atlases as the patient, the registration process creates its customized atlas. The registration accuracy could be quantified by the fraction of voxels whose labels given by an expert agree with the labels in the customized atlas. We are currently investigating potential sources for additional atlases to make this type of evaluation possible.

9. Conclusion

The goal of this research is to perform automatic, fast and accurate registration of volumetric data in 3-D space, as well as anomaly detection using registration and domain knowledge. One important application domain is medical image registration. Anatomical structures vary considerably in appearance across individuals or within one individual over time, and any pathology may aggravate these variations.

We adopted an intensity-based approach for registration, which assumes no prior segmentation of the voxels in a volume. To address the appearance variations of anatomical structures, we have developed a three-level hierarchical deformable registration algorithm. First, a 3-D *global transformation* (rotation, uniform scaling, and translation) aligns the different image volumes. This procedure corrects for variations induced during the image acquisition process. Secondly, a *smooth deformation*, represented by the warping of a 3-D grid of control points, approximately matches the corresponding anatomical structures in the image volumes. This step partially adjusts for the inherent variations in the appearances of the anatomical structures across individuals. Several grid resolutions are used to progressively change the emphasis from global alignment to alignment of specific anatomical structures. Finally, a *fine-tuning deformation*, which allows each voxel to move independently, refines the matching of corresponding anatomical features. For computational efficiency, we use a hierarchy of image resolutions, stochastic sampling, and parallel processing.

Based on this algorithm, we have developed a prototype for automatic registration and pathology detection, ADORE (Anomaly Detection thrOugh REgistration). We conducted experiments on 12 sets of real image data of the brain. ADORE can match two 16MByte volumes in 12 minutes on an SGI workstation with four 194 MHz R10000 processors, with an accuracy qualitatively comparable to manual registration. By matching an expert-segmented atlas to a patient's data, ADORE can automatically build a customized atlas for different individuals. Moreover, ADORE is able to indicate the possible existence and location of pathologies by measuring the asymmetry of the anatomical features.

Acknowledgments

Many thanks to Shumeet Baluja, Carlos Ernesto Guestrin, John Hancock, Prem Janardhan, Andrew Johnson, 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.

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 Report 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] 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.

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

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

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

[17] Pokrandt, "Fast Non-Supervised Matching: A Probabilistic Approach".

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

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

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

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

[22] Press et al., "Numerical Recipes in C", Cambridge University Press, 1992.