# An Automatic Registration Method for Frameless Stereotaxy, Image Guided Surgery, and Enhanced Reality Visualization

W.E.L. Grimson[1] G.J. Ettinger[1,3] S.J. White[3] T. Lozano-Pérez[1]
W.M. Wells III[1,2] R. Kikinis[2]

February 22, 1996

**Abstract**

There is a need for frameless guidance systems to help surgeons plan the exact location for incisions, to define the margins of tumors and to precisely identify locations of neighboring critical structures. We have developed an automatic technique for registering clinical data, such as segmented MRI or CT reconstructions, with any view of the patient on the operating table, using a series of registration algorithms, which we demonstrate on the specific example of neurosurgery. The method enables a visual mix of live video of the patient with the segmented 3D MRI or CT model, supporting enhanced reality techniques for planning and guiding neurosurgical procedures, and to interactively view extracranial or intracranial structures non-intrusively. Extensions of the method include image guided biopsies, focused therapeutic procedures and clinical studies involving change detection over time sequences of images.

[1]Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge MA
[2]Department of Radiology, Brigham and Womens Hospital, Harvard Medical School, Boston MA
[3]The Analytic Sciences Corporation, Reading MA

# 1 Motivating Problem

Many surgical procedures require highly precise localization on the part of the surgeon, in order to extract targeted tissue while minimizing damage to adjacent structures. The difficulty of this 3D localization is exacerbated by the fact that it often requires isolating a structure deeply buried within the body. While methods exist (e.g. MRI, CT) for imaging and displaying the 3D structure of the body, the surgeon must still relate what she sees on the 3D display with the patient's actual anatomy.

As current methods often involve a surgeon simply utilizing traditional two dimensional slices of MRI or CT imagery and mentally transforming that information to the actual patient, there is a clear need for registered visualization techniques, in which 3D reconstructions of internal anatomy are exactly overlaid with the surgeon's view of the patient, so that she can directly visualize important structures, and plan and act accordingly.

In this paper, we describe a method for executing this automatic registration and visualization technique. While this problem is particularly relevant in procedures that involve minimal invasion of the body, such as biopsies, or endoscopic procedures, we use the problem of visualization and guidance in neurosurgical procedures as a motivating example.

## 1.1 An Ideal Solution

Ideally, one would prefer a system that could automatically register 3D data sets, and track changes in the position of a data set over time, without requiring the attachment of any devices to the patient. Such an ideal system should support:

- real-time, adaptive, enhanced reality patient visualizations in the operating room;
- dynamic image-guided surgical planning;
- image guided surgical procedures, such as biopsies or minimally invasive therapeutic procedures; and
- registered transfer of *a priori* surgical plans to the patient in the operating room.

While our group is actively developing all aspects of such a system, this paper focuses on one key component of such a system, the registration of different data sources to determine relevant coordinate frame transformations.

## 1.2 Contributions to the Ideal Solution

We have created a system that performs the registration of clinical image data with the position of the patient on the operating table at the time of surgery, using methods from visual object recognition. The method does not require the use of previously attached frames or landmarks. The method has been combined with an enhanced reality visualization technique [14, 4, 41], in which we display a composite image of the 3D anatomical structures with a view of the patient. This registration enables the transfer to the operating room of preoperative surgical plans, obtained through analysis of the segmented 3D preoperative data [8], where they can be graphically overlaid onto video images of the patient. Such transfer allows the surgeon to apply carefully considered surgical plans to the current situation, and to mark internal landmarks used to guide the progression of the surgery.

Thus, the specific problem we consider is:

- **Input:** a video view of the patient, together with a segmented MRI or CT model of the anatomy of the patient, each defined in its own coordinate system.
- **Output:** a transformation aligning the model with the patient, and a transformation describ-

ing the position and orientation of the video camera relative to the patient.

Extensions of our method include adaptively re-registering the video image of the patient to the 3D anatomical data, as the patient or the video source moves, as well as other surgical applications such as image guided biopsy, or focused therapeutic procedures (e.g. laser disc fusion or tumor ablation). We have also demonstrated the use of our system in clinical settings, by registering data sets acquired over extended time periods, thereby enabling the detection of changes in anatomy over time.

## 2 An Example Scenario

The following describes a scenario demonstrating the use of our methods.

**(1)** A patient requiring surgical therapy is scanned by a 3D, high resolution, internal anatomy scanner, such as Magnetic Resonance Imaging (MRI) or Computed Tomography (CT).

**(2)** The current scan is segmented into different tissue types, and graphical models of desired structures (e.g. tumor, ventricles, skin surface, bone, white matter, etc.) are generated.

**(3)** The patient is placed in the operating room, which is equipped with:

- a laser range scanner for obtaining depth data of the patient's skin surface where the surgery is to be performed,
- enhanced reality visualization equipment, such as a video or digital camera, mixer and display monitor; or a head-mounted display with trackable landmarks,
- the operating table may contain landmarks that will remain viewable and fixed relative to the patient during surgery,
- landmark tracking equipment.

**(4)** Prior to draping, the patient is scanned by a laser range scanner. The 3D locations of any table landmarks are also calculated to identify their location relative to the patient.

**(5)** The current MRI or CT scan is automatically registered to the patient skin surface depth data obtained by the laser range scanner. This provides a transformation from MRI/CT to patient.

**(6)** The position and orientation of a video camera relative to the patient is determined, by matching video images of the laser points on an object to the actual 3D laser data. This provides a transformation from patient to video camera.

**(7)** The registered internal anatomy is displayed in enhanced reality visualization [14, 4, 41] to "see" inside the patient. In particular, the two previously computed transformations can be used to transform the 3D model into the same view as the video image of the patient, so that video mixing allows the surgeon to see both images simultaneously.

**(8)** The patient is draped and surgery is performed. The enhanced reality visualization does not interfere with the surgeon, nor does it require her to do anything different from that to which she is accustomed. Rather, the system provides her with additional visualization information to greatly expand her limited field of view.

**(9)** The location of table landmarks can be continually tracked to identify changes in the position of the patient's attitude, relative to the visualization camera. Visualization is maintained by updating the MRI/CT to patient transformation.

**(10)** Viewer location can be continually tracked to identify any changes in the position of the viewer. In the case of a stationary video camera, this is straightforward, though in the case of head-mounted displays such tracking is both more relevant and more challenging. Visualization updates are performed by updating the patient to viewer transformation.

**(11)** In general, the surgical procedure is executed with an accurately registered enhanced visualization of the entire relevant anatomy of the patient. The hope is that by providing this information to the surgeon, procedures can be more efficiently and effectively executed, with reduced side effects to the patient.

# 3 Details of Our Approach

Part 1 of this scenario is standard practice. Methods currently exist for dealing with Part 2 [8, 13]. Here we focus on parts 3–8 of the above scenario. Parts 9–11 are part of our planned future work. For parts 3–8, the key step is the registration of data obtained from the patient in the operating room with previously obtained data and surgical plans.

We use a multi-stage matching and verification of a 3D data set acquired at the time of the surgical procedure with 3D clinical data sets acquired previously. The central ideas are to use a laser striping device to obtain 3D data from the skin of the patient at run time, and to use a sequence of recognition techniques from computer vision to match this data to segmented skin data from the MRI or CT reconstruction. These matching techniques allow us to accurately register the clinical data with the current position of the patient. As a consequence we can display a superimposed image of the 3D structures overlaid on a view of the patient.

The steps of our method are outlined below.

## 3.1 Model input

We obtain a segmented 3D reconstruction of the patient's anatomy, for example using CT or MRI. Current segmentation techniques are generally semi-automatic. This process is typically done by training an intensity classifier on a user selected set of tissue samples, where the operator uses knowledge of anatomy to identify the tissue type. Once initial training is completed, the rest of the scans can be automatically classified on the basis of intensities in the scanned images, and thus segmented into tissue types [8, 13]. Automatically removing gain artifacts from the sensor data can be used to improve the segmentation [42]. Similarly, methods that correct for distortions due to magnetic susceptibility differences between different materials [35] can further improve the accuracy of the segmentation.

We refer to this 3D anatomical reconstruction as the model, which is represented relative to a model coordinate frame. For simplicity, the coordinate system's origin is taken as the model centroid.

## 3.2 Data input

To achieve the registration, we first obtain a set of 3D data points from the patient's skin surface using a laser range scanner, which operates by scanning a laser beam through an optical mechanism that results in a controlled plane of light. A video camera is placed at an angle to this plane such that a portion of the plane is in the camera field of view. When an object is placed in this visible region such that it intersects the laser plane, points in the camera image illuminated by the laser unambiguously correspond to fixed 3D scene points. The 3D measurements of the currently used scanner (Technical Arts 100X) are accurate to within $0.08mm$.

We refer to this 3D information as the data, which is represented in a coordinate frame attached to the laser, and which reflects the position of the patient in a coordinate frame that exists in the operating room. Our problem is to determine a transformation that will map the model into the data in a consistent manner.

### 3.3 Matching data sets

We find this transformation by matching the two data sets using the following steps:

**(1)** First, we separate laser data of the patient's head from background data. Currently we do this with a simple user interface (Figure 1), in which the view from the laser's video camera is displayed with the laser data overlaid on top of that view. The user can use a simple mouse interface to block out laser points coming from the skin of the patient. This laser data can be further edited by displaying the data, overlaid on the MRI model by means of the rough transformation specified above. Using any of the three orthogonal views of the two data sets, the laser data can be furthered edited using a simple set of mouse interfaces. Note that this process need not be perfect, we simply want to remove gross outliers from the data. There are also methods for extracting such data automatically.

**(2)** To initiate the matching, we have several options. First, we have developed a simple graphical interface (see Figure 1) that can be used to roughly align the laser data with the 3D model. The model is displayed in its initial orientation on a series of three orthogonal 2D views, together with the laser data. The user can refine the initial transformation by rotating and translating the data in any of the three orthogonal views. This initial alignment does not need to be very accurate: rotational errors on the order of 10–20 degrees, and translational errors on the order of centimeters are permissible, since the subsequent matching stage is quite reliable at removing these misalignments.

**insert Figure 1 here**

**(3)** If we have a rough alignment provided by the user, this is usually sufficient to move to the pose refinement stage described below. In some cases, however, we may not want to rely on user intervention (for example in the batch mode registration of MRI scans acquired over time, described in a later section). In this case, we need an automatic way of efficiently finding good initial alignments. We have developed several schemes for accomplishing this.

If we do not want to rely on operator intervention, we can instead sample a set of evenly spaced directions on the view sphere. For each view, we extract a sampled set of visible points of the model, using a z-buffer. In particular, given a pixel size for the z-buffer and given an estimate for the view direction, we project all of the model points into a plane orthogonal to the view direction, where the plane is tesselated in pixels of the given side. Within each pixel, we keep only the point closest to the viewer. This gives us a temporary model, which we can use for matching. While it is possible in this method for points to leak through the z-buffer from the back of the model, this does not affect the performance of the method. This gives us a temporary model, which we can use for matching. The point on the view sphere serves to estimate the rotation needed to roughly align the model and the data. For each such model and its associated alignment, we execute the matching process described below.

Yet another method for removing the operator intervention takes advantage of data itself to find initial transformations. In particular, as a preprocessing step, we can hash all pairs of MRI points based on distance between them. Furthermore, at each MRI skin point, we can estimate the surface normal to the skin by a local fit of the neighbouring data. To find an initial transformation, we can select any two laser data points, where for stability reasons we select two points that are widely separated. Using the distance between these two laser points, we can access the hash table to find possible matching MRI points. For each such pair in the hash table, if we consider the hypothesis that the laser points match the MRI points (there are two such matches), then this determine 5 of the 6 degrees of freedom associated with the coordinate frame transformation. The missing parameter is the rotation about the axis connecting the two points. To solve for this parameter, we can estimate the normal to the skin surface at the laser point by fitting a plane to the neighbouring data. Then the rotation we need is the rotation about the axis between the points that will rotate the laser normal to align with the MRI normal. Note that such a rotation may not exist, in which case we can discard this pair. Similarly, after solving for the rotation, we can check that application of this rotation to the normal at the other point also causes it to agree with its matching normal. If not, the pair is discarded. By cycling over all the possible pairings of MRI points to laser points, as defined by the entries in the hash table, we can collect the set of feasible initial transformations. These can be ranked on the basis of the RMS fit of the transformed laser data to the MRI data, and the resulting rank ordered list of hypotheses can be further processed using the methods described below, stopping when a sufficiently accurate fit is found.

A related approach is to use Interpretation Tree Search [16, 17, 15] to match triples of visible sampled model points to the three selected laser points. This method basically searches over all possible ways of matching three laser points (selected to be widely separated from one another) to three points selected from the sample MRI model. For each pairing of triples of model and data points, the method tests whether the pairwise distances between model points and laser points are roughly the same. If all such tests are valid, the match is kept, and we compute the coordinate frame transformation that maps the three laser points into their corresponding model points. These transformations form a set of hypotheses. Note that due to the sampling of the model data, the actual object points corresponding to the selected laser points may not exist, so these hypothesized transformations are at best approximations to the actual transformation. For efficiency, we hash pairs of points on distance, and only retrieve likely candidates for testing.

In the example shown in Figure 2, there are 481 laser sample points, and the skull model has 35,265 sample points. Given an initial sampled viewpoint and a coarsely sampled z-buffer, there are 409 model points in the sampled view. In principle, there are

$$\binom{481}{3}\binom{409}{3} \approx 2.02e14$$

possible hypotheses for the aligning transformation, but using simple distance constraints (and allowing for some amount of error in the distance measurements), there are only 16,945 possible hypotheses that remain for further testing.

**(4)** We can use the Alignment Method [22] to filter these hypotheses. In particular, we filter out those hypotheses for which the transformed view vector becomes unviewable (e.g. lies below the operating table). Then, for each remaining hypothesis, we transform all the laser points by the hypothesized transformation, and verify that the fraction of the transformed laser points that do not have a corresponding model point within some predefined distance is less than some predefined bound. We discard those hypotheses that do not satisfy this verification. For efficiency, we use two levels of sampling of the laser points, first verifying that a coarsely sampled set of laser points are in agreement, then further verifying, for those that pass this test, that all the laser points are in agreement.

An example of the two data sets (model and laser data) after a verified alignment is shown in Figure 2. For this example, there is only one surviving hypothesis, out of the 16,945 initial hypotheses. Note that some of the laser points are partially buried in the CT model (displayed as partially transparent), indicating that the initial alignment is close but not sufficiently accurate.

**(5)** For each verified hypothesis (whether obtained automatically or by initial user alignment), we refine the alignment of the two data sets by minimizing an evaluation function that measures the amount of mismatch between the two data sets.

**(5.1)** Evaluate the current pose by summing, for all transformed laser points, a term that is itself a sum of the distances (weighted by a Gaussian distribution) from the transformed point to all nearby model points [40]. This Gaussian weighted distribution is a method for roughly interpolating between the sampled model points to estimate the nearest point on the underlying surface to the transformed laser point. More precisely, if $\ell_i$ is a vector representing a laser point, $m_j$ is a vector representing a model point, and $\mathcal{T}$ is a coordinate frame transformation, then the evaluation function for a particular pose (or transformation) is

$$E_1(\mathcal{T}) = -\sum_i \sum_j e^{-\frac{|\mathcal{T}\ell_i - m_j|^2}{2\sigma^2}}. \tag{1}$$

This objective function is similar to the posterior marginal pose estimation (PMPE) method used in [40], and to the elastic net constraints of [9]. One can visualize this objective function as if we placed a Gaussian distribution of some spread $\sigma$ at each model point, then summed the contributions from each such distribution at each point in the volume. Then the contribution of each transformed laser point towards the evaluation function is simply the summed value at that point. Because of its formulation, the objective function is generally quite smooth, and thus facilitates "pulling in" solutions from moderately removed locations in parameter space. As well, it bears some similarity to the radial basis approximation schemes used for learning and recognition (e.g. [5, 10]).

**(5.2)** In earlier versions of the system, we used Powell's method [34] to iteratively minimize this evaluation function. Powell's method is basically a conjugate gradient descent method. Starting at some initial point, it searches along a line in the parameter space for the point that minimizes the objective function along that line. Then from that point it picks a new direction and mimizes along that line, and so on. Ideally, such methods should pick "conjugate" directions along which to minimize, that is, directions with the property that movement along a new such direction will not spoil the minimization along previously used directions. Powell's method is a simple means for choosing these conjugate directions, without directly estimating the gradient of the objective function.

It is more efficient, however, to use the Davidon-Fletcher-Powell quasi-Newton method [34]. This requires an estimate of the gradient of the objective function, which we describe below. To estimate the gradient of the object function, we need the partial derivatives of that function with respect to the transformation parameters $p$, this is given by

$$\frac{\partial E_1}{\partial p} = \sum_i \sum_j \frac{-1}{\sigma^2} e^{-\frac{|\mathcal{T}\ell_i - m_j|^2}{2\sigma^2}} \left\langle \mathcal{T}\ell_i - m_j, \frac{\partial \mathcal{T}\ell_i}{\partial p} \right\rangle, \tag{2}$$

where $\langle \cdot, \cdot \rangle$ denotes a dot product. For each of the translation components of the transformation, the partial derivative is straightforward, for example:

$$\frac{\partial \mathcal{T}\ell_i}{\partial t_x} = u_x$$

where $t_x$ is the $x$ component of the translation vector, and $u_x$ is a unit vector in the $x$ direction. The other two translation components are similar.

For the rotation components, we can treat the rotation as an instantaneous rotation about three orthogonal axes. For each axis $r$ with associated angle of rotation $\theta$, applying Rodrigues' formula suffices to show that the partial derivative is given by the following cross product:

$$\frac{\partial \mathcal{T}\ell_i}{\partial \theta} = r \times R\ell_i$$

where $R$ is the associated rotation matrix. Thus the gradient of the evaluation function is straightforwardly computed using the appropriate expressions for the partial derivatives in equation (2).

The DFP method essentially iteratively builds up a good approximation to the inverse Hessian matrix, which can then be applied to the data to solve for the parameters yielding a minimum of evaluation function [34]. Solving this minimization problem yields an estimate for the pose of the laser points in model coordinates.

**(5.3)** Execute this refinement and evaluation process using a multiresolution set of Gaussians. Thus initially a broad based Gaussian is used to allow influence over large areas, resulting in a coarse initial alignment, but one which can be reached from a wide range of starting positions. Subsequently, more narrowly tuned Gaussian distributions can be used to refine the pose, while focussing on only nearby model points to derive the pose.

**(5.4)** Using the resulting pose of this refinement, repeat the pose evaluation process, now using a rectified least squares distance measure. In particular, perform a second sampling of the model from the current viewpoint, using a much more finely sampled z-buffer. Relative to this finer model, evaluate each pose by measuring the distance from each transformed laser point to the nearest model point, (with a cutoff at some predefined maximum distance). Evaluate the pose by summing the squared distances of each point. Minimize using the DFP method to find the least-squares pose solution. Here the evaluation function is

$$E_2(\mathcal{T}) = \left[ \frac{1}{n} \sum_i \min \left\{ d_{\max}^2, \min_j |\mathcal{T}\ell_i - m_j|^2 \right\} \right]^{\frac{1}{2}} \tag{3}$$

where $d_{\max}$ is some preset maximum distance, and where $n$ is the number of laser points $\ell_i$. This objective function is essentially the same as the maximum a posteriori model matching scheme of [40]. It acts much like a robust chamfer matching scheme, or an iterative closest point matching scheme, similar to that used by [23, 28, 3]. To use DFP on this evaluation function, we again need to estimate the gradient, which is given by:

$$\begin{aligned} \frac{\partial E_2}{\partial p} = & \ \tfrac{1}{n} \left[ \tfrac{1}{n} \textstyle\sum_i \min \left\{ d_{\max}^2, \min_j |\mathcal{T}\ell_i - m_j|^2 \right\} \right]^{\frac{-1}{2}} \times \\ & \times \textstyle\sum_i \left\langle \mathcal{T}\ell_i - m_j, \frac{\partial \mathcal{T}\ell_i}{\partial p} \right\rangle, \end{aligned} \tag{4}$$

where we assume that for small changes, the closest point does not change, and hence the min operation can be ignored. The partial derivatives in this case are identical to those used in the previous case.

The expectation is that this second objective function is more accurate locally, since it is composed of saturated quadratic forms, but it is also prone to getting stuck in local minima. Hence we add one more stage.

**(5.5)** We observe that while the above method always gets very close to the best solution, it can get trapped into local minima in the minimization of $E_2$. To improve upon this, we take the pose returned by the above step, and perturb it randomly, then repeat the minimization. We continue to do this, keeping the new pose if its associated RMS error is better than our current best. We terminate this process when the number of such trials that have passed since the RMS value was last improved becomes larger than some threshold.

**(5.6)** The best found solution is a pose, and a measure of the residual deviation of the fit to the model surface. An example is shown in Figure 2.

**(5.7)** An alternative to using all of the laser data is to instead utilize laser data that is "more distinctive". We do this by estimating the curvature, along the plane of the laser, at each laser data point, and keeping only a predefined percentage of the most highly curved laser points. The motivation is to restrict our matching to points that carry more information about the local structure of the surface. Initial experiments show that using only highly curved points improves the residual RMS error.

**(5.8)** Once the final solution is found, the residual error at each point can be measured, and laser points with large residual errors can be removed from the data set. This automatically deletes possible outliers. The final stages of the registration process can then be rerun using the remaining data, to obtain a tighter fit to the surface.

**insert Figure 2 here**

We collect such solutions for each verified hypothesis, over all legal view samples, and rank order them by smallest RMS measure. The result is a highly accurate transformation of the MRI data into the coordinate frame of the laser sensor.

## 3.4 Camera Calibration

Given a registration, we need to relate it to a view of the patient. A video camera can be positioned in roughly the viewpoint of the surgeon, i.e. looking over her shoulder. By calibrating the position and orientation of this camera relative to the laser coordinate system, we can render the aligned MRI or CT data relative to the camera's view. This rendering can be mixed with the live video signal, giving the surgeon an enhanced reality view of the patient's anatomy [14, 4, 41]. This can be used to plan a craniotomy or a biopsy, or to define the margins of an exposed tumor for minimal excision.

When the laser scanner video camera is used for this purpose, the solution for the camera model is straightforward. Points in the laser scan of any object in the camera field have unique correspondences with points in the image plane. During calibration, scan points from a calibration object are used to generate an approximation of the camera pose. When a patient is scanned, corresponding 3D laser measurements and image points from the skin are used to refine the camera pose model via Powell's method.

When an independent camera is used for video capture from the surgeon's view a similar method is used using objects in the scene with flat surface facets. Images of the laser slices are taken with the video camera. Straight 2D line segments are located in the video images and matched to corresponding straight 3D line segments in the laser data. If three such matching segments are found, they can be used to solve for an approximation to the perspective projection transformation, and thus for the pose of the camera. Using this as a starting point, Powell's method can again be used to optimize the pose estimate to best bring all of the laser line segment data into projective alignment with the corresponding video data. Thus, one can use simple faceted objects in the surgical setting directly to calibrate the camera, and this process can be repeated throughout the surgical procedure as needed (e.g. if the position of the camera relative to the operating table is perturbed).

If a small number of fiducial points are known in both laser coordinates and image coordinates, then methods exist for automatically solving or updating the camera calibration parameters (e.g. [29]).

Figure 3 shows the alignment of the CT model of the skull with the actual image of the skull in a calibrated video camera, demonstrating the type of visual overlay provided by this system.

**insert Figure 3 here**

## 3.5 Visualization

We can combine the camera calibration and the registration of the data sets, to achieve a visualization of the data. In particular, we can apply the transformation between the data sets to bring the MRI or CT model into alignment with the patient, in the coordinate frame of the laser system. We can then project that model into the coordinate frame of the video camera, by applying the computed camera model. This gives us a virtual view of the MRI model, as seen by that camera. This can then be mixed with an actual video view of the camera, and used as a visualization tool by the surgeon.

## 4 Testing and Application of the Method

As a first test, we have run a series of controlled experiments, in which we have registered a CT reconstruction of a plastic skull with laser data extracted for a variety of viewpoints. We have run the system both with an initial estimate of body and nose axes, and where we sample over a range of views on the viewing sphere. In all cases, the system finds a "correct" registration, with typical residual RMS errors of 1.6 millimeters. The issue of how reliable RMS error is as a measure of success is discussed in the next section. Examples are shown in Figure 4. For example, using the skull data of Figure 2, and sampling over a set of views, leads to the statistics shown in Table 1.

**Insert Figure 4 here**

**Insert Table 1 here**

As a second test, we have run trials matching laser data against an MRI scan of one of the authors, an example of which is shown in Figures 5 and 6. Figure 5 shows the result of registering laser data against the MRI skin surface. Figure 6 is a visualization in which the surface of the brain has been rendered and registered with the view of the patient, and the result projected and mixed with a video image of the patient. We have run the system both with an initial pose estimate, and by sampling a range of views on the viewing sphere. In all cases, the system finds a correct registration (Figure 6), with typical residual RMS errors on the order of $1.5mms$. The resolution of this MRI scan is $0.9375 \times 0.9375 \times 1.5mms$.

**insert Figure 5 here**

**insert Figure 6 here**

We have also run a series of trials with neurosurgery patients. An example registration of the laser data against MRI models of a patient is shown in Figure 7. The tumor and the ventricles of the patient are highlighted. The RMS errors in such cases were typically about $1.5mms$.

**insert Figure 7 here**

We have been using this registration and visualization method to transfer presurgical plans to the patient. In our current method, we use our registration method to provide a visual overlay of the view of the patient with internal structures selected by the surgeon. By viewing this overlay on a live video monitor, the surgeon can trace with a marker the outline of key structure on the patient's scalp [14]. This enables the surgeon to mark locations for planned surgical steps, prior to the placement of the patient in the OR. To date, we have used this procedure on roughly 20 neurosurgical patients at Brigham and Women's Hospital in Boston, MA. (See Section 6.)

Besides applications for surgical planning and guidance, the method has other applications, including the registration of multiple clinical data sets such as MRI versus CT. As a demonstration of this, we have registered sequences of MRI scans of the same patient, taken over a period of several months, and used differences in the registered scans to visualize and measure changes in anatomy [11]. These scans were part of a NIH study of multiple sclerosis (MS) at Brigham and Women's Hospital aimed at determining the optimal frequency for performing MR imaging of MS patients. In this testing, we automatically registered 20 sets of 50 MRI scans each.

## 5 Computational Evaluation

We have examined a series of issues relating to the computational performance of our system. The results of our testing are described in this section.

### 5.1 Minimization error

It is worth commenting on the interpretation of the RMS error. First, the issue of what is a correct registration is somewhat difficult to evaluate on the basis of RMS error. Certainly solutions with low RMS are likely to be good fits of the two data sets, by definition, but this does not ensure that they are "correct". Rather, RMS is a measure of the "success" of the method at finding a good fit between data sets, while the correct registration is one in which homologous anatomical points are mapped onto one another. To test this latter issue, we need to perform phantom studies of some sort, an issue we are currently investigating.

Second, we are actually mostly interested in how accurately we have measured the transformation. While the RMS measure is a useful indicator of fit, the actual transformation, when applied to model points, may be much more accurate. This is discussed later in this section.

Third, the RMS error will clearly be a function of the density of the model points. For efficiency, we often run the method with a subsampling of the MRI model. To explore the effect of this sampling on the residual error, we ran experiments on two patients, in which we varied the sampling of the MRI and recorded the final RMS error. This is graphed in Figure 8. One can see that as more model points are included the RMS error declines to roughly 1 mm. Note that the voxel size is $0.9375 \times 0.9375 \times 1.5 mms$ so that the expected RMS error just due to the discrete size of the voxels is $.578mm$. When combined with the fact that we are measuring distance to the nearest vertex of an isosurface triangle, not to the actual surface itself, one can see that the method is close to the limit of RMS accuracy. The key issue here is how to strike a good balance between the efficiency one gains by using a sample of the full MRI model and the potential cost in goodness of fit introduced by such a sampling. Figure 8 that sampling rates of 2 or 3 probably introduce very little error, while sampling rates of 5 are probably the limit one should tolerate.

One way of dealing with the fact that the MRI data points are only a sampling of the surface is to attempt to account for residual errors in modeling the surface. In particular, when evaluating the terms in equation 3, one still searches for the closest MRI point to the transformed laser point. Rather than summing

$$|\mathcal{T}\ell_i - m_j|^2$$

over all such points, we use the normal component

$$|\,(\mathcal{T}\ell_i - m_j) \cdot N_j|^2$$

where $N_j$ is the normal to the surface at the point $m_j$. Thus, we use the component of the distance vector between the transformed laser point and the nearest MRI point, in the direction of the normal to the MRI surface at that point, rather than simply using the distance of the vector itself. This will reduce the effects of errors due to subsampling the MRI surface. We will return to this point in the testing described below.

Finally, it is worth noting that the sparsity of the laser data clearly impacts on the RMS error. If the laser data were dense, so that we could be guaranteed of getting data from portions of the skin with varying structure, it is likely that the RMS error would be significantly reduced. Our experience in matching MRI scans to MRI scans using this technique supports that observation. We stress, however, that the RMS error is probably less relevant to determining accuracy of the method than the issue of repeatability discussed below.

**insert Figure 8 here**

## 5.2 Repeatability

A key issue with any registration algorithm based on minimizing an evaluation function is the question of local minima. Since most realistic objective functions will have complex associated energy landscapes, it is important to test that the algorithm avoids getting trapped in local minima, and instead is reliable at finding the best global minimum.

To explore this issue, we ran the following test. We took several cases involving real patients, and found a transformation close to what appears to be the correct alignment. We then ran a series of trials in which we perturbed that transformation and used the result as an initial alignment to the registration algorithm. We then let the registration method converge to a solution, which we recorded. In typical trials, we perturbed the transformation by $n$ millimeters of translational offset, and $2n$ degrees of rotational offset. We have run trials where $n$ ranged from 1 to 10, with the method successfully finding correct solutions. A detailed example of a set of trials is given in Table 2, using a perturbation of 5 mms and 10 degrees, and using a sampling of one in every 5 MRI points. One can see that factoring out the normal component of the residual error reduces the overall RMS error, and that typically the RMS error is on the order of the size of a voxel.

**Insert Table 2 here**

More importantly, to test the repeatability of the system in finding the same transformation, we took each cluster of transformations and applied the following test. We computed an entry vector into the skull, based on the surgeon's expected trajectory. We then applied each transformation out of the clustered set to that vector, collecting a set of possible entries. Starting at the skin surface, we traced a set of points along each of the transformed entry vectors, for example using a vector 6 centimeters long and sampling every centimeter along the ray. This defines a cluster of points at each sampling, based on the deviation in the computed transformations applied to the expected entry vector. The key question is how much spread there is in these points, since that will determine the deviation that the surgeon might be expected to see from the ideal location. To measure this, we found the centroid of each cluster of points, then computed statistics on those clusters, which are described in in Table 3. We ran tests in which the normal component of residual error was used and in which the full residual error was used. As well, we varied the threshold on RMS error below which a solution was accepted. As one might expect, while this may occassionally cause the system to miss a solution (as shown by the success rate) it does reduces the deviation in the computed entry vectors.

**Insert Table 3 here**

It is worth noting the range of variation in the results. In particular, in the case of the skull and Patient 070194, the laser data primarily covered the face, where there is considerable variation in structure. This led to very low deviations in entry vector, because the structure served to distinctively identify the best registration. The remaining variation (on the order of 10-30 microns) is probably mostly due to the fact that we are randomly sampling every fifth point in the MRI skin surface, which can introduce a small amount of residual perturbation.

In the case of Patient 071594 and Patient 082294, the laser data primarily covered the cheek and side of the face (other cases we have run cover the back or the top of the head), where there is much less distinctive data available. In these cases, the residual deviation is higher. In one case, where some of the laser data came from the nose of the patient, the typical deviation was 100 microns. In the second case, the laser data was almost entirely from the check of the patient, and here the typical deviation was on the order of 1.6 millimeters. In all cases, one can see from the 95 percentile statistics that extreme deviations are very rare.

One conclusion from this data is that a good way to utilize the system is to ensure that the laser system acquires data from the face if the patient, ensure reliable registration, then using the visualization camera in whatever orientation is best suited for the surgeon.

Finally, we note that while using normal residual error reduces the RMS error associated with solutions, it does not improve the reliability of the system to find the same solution.

### 5.3 Capture Radius

Another way of testing the method is to run a capture radius test. In this trial, we took the data from one of our patients, and found a good registration for that patient's data. With that as a starting point, we randomly perturbed the solution by a translational offset of $n$ mms and a rotational offset by $2n$ degrees (note that the rotational effects will result in an overall translation of data by more than $n$mms). For each value of $n$ we ran 20 trials, gathered the resulting transformations and performed a cluster analysis similar to the cases above. Table 4 and 5 summarize the results. Note that during these trials we set $d_{\max}$ to 10 mms so that we rely on the Gaussian based minimization $E_1$ doing a good job of pulling the registration into close range. In Table 4 and 5 we list two sets of results, one where we keep all legal transformations as part of our cluster, and one where we restrict ourselves to transformations with good RMS values. The latter case increases the number of trials with no solution, but improves the accuracy of the found solutions. The observation is that a capture radius of 5mms and 10degress is very reliable, and can easily be obtained automatically by our indexing method, or can easily be obtained by our simple user interface. If we are willing to take more than one trial, should the first trial result in an unacceptable registration, as measured by the RMS error, then the capture radius easily extends to twice this range.

**Insert Table 4 here**

**Insert Table 5 here**

## 6 Clinical Testing

We have tested our system on a number of actual neurosurgery patients[1] at Brigham and Women's Hospital for the purpose of accurately planning craniotomy locations. Each patient had a MRI scan performed approximately one day prior to surgery. Clinicians at the Surgical Planning Lab segmented the MRI scan into skin, brain, tumor, ventricles, and other structures of interest. Once the patient's head was shaved in preparation for surgery he was brought to the Surgical Planning Lab for laser registration by our system. The registration procedure consisted of the following steps:

**(1)** Place the laser scanner over the patient's head such that the camera is positioned at the expected angle of approach to the tumor.

**(2)** Scan the patient's head with the laser.

**(3)** Edit the laser data, as overlaid on the video image of the patient, by roughly outlining the region of interest, namely the patient's head.

**(4)** Coarsely align the laser data with the segmented MRI skin model, using refinements through interactive mouse rotation and translation of the laser data overlaid on the MRI model.

**(5)** Execute the automated registration starting with the coarse initial alignment.

**(6)** Verify the registration by a visual animation of the transformed MRI skin overlaid on the video

[1] Twelve patients tested as of April, 1995.

image as well as a color coded examination of residual errors of transformed laser data overlaid on the MRI skin model.

**(7)** Render the internal brain structures in their registered locations relative to the camera, such that they can be mixed with the video signal to produce an enhanced reality visualization.

**(8)** The surgeon draws the position of the tumor and other structures of interest directly on the scalp while looking at the visualization. The selection and opacity of the different anatomical structures is dynamically controlled to provide the surgeon with as much geometric information as possible from the given viewpoint.

Each of the steps in the process can be repeated to refine the surgical plan. Elapsed time for steps 2 through 6 is approximately five minutes. Previously performed *manual* alignment aimed at achieving such visualizations could take as much as 45 minutes with an accuracy of 10-30 mms. Our registration thus achieves an order of magnitude improvement in both efficiency and accuracy, two factors which are critical to the neurosurgeons. Feedback from the surgeons has been highly positive as they have found this easily accessible form of 3D geometric knowledge to prepare them well for the surgeries. Current work involves migrating the use of the laser registration system into the OR for neurosurgery guidance as well as for planning and guiding other types of surgery.

## 7 Other Applications

As mentioned earlier, the method has applications for surgical planning and guidance, including tumor excision and biopsy. The method has broader application, however, including the registration of multiple clinical data sets such as MRI versus CT. As a demonstration of this, we have taken two MRI scans of the same patient, obtained several months apart. These scans are part of an ongoing NIH study of multiple sclerosis (MS) at Brigham and Womens Hospital aimed at determining the optimal frequency for performing MR imaging of MS patients. Under this study patients with varying disease stages are imaged at different frequencies to identify changes in MS lesion activity. To support this analysis, one needs to register the MRI scans from different points in time, and compare them to detect relevant changes. We have applied our technique to this task, using the surface of the intracranial cavity (ICC) as the basis for the registration. An original MRI ICC surface is used as a model, and a second MRI ICC surface is aligned with the first by sampling points on that surface (taking 1 out of every 30 points at random). Two views of the resulting overlay are shown in Figure 9. The resolution of these scans is $0.9375mm \times 0.9375mm \times 3mm$.

**insert Figure 9 here**

Given this alignment, we can transform the second data set into the coordinate frame of the first data set, then resection the data to obtain 2D slices equivalent to those of the first data set (with an interpolation to estimate the values in the resectioned slices). Given these new sections, we can then compare individual slices of the first data set to the resectioned second data set, and do image differencing to find noticeable changes. An example of this, highlighting the growth of a lesion in the patient is shown in Figure 10.

**insert Figure 10 here**

## 8 Related Work

Registration of data sets is an important problem in many medical applications, and hence has seen considerable work in the past. Two recent reviews of registration methods include [39, 1]. Of particular relevance are three different approaches. First, Pelizzari and colleagues [31, 27, 26, 18, 38, 33, 32, 30] have developed a method that matches retrospective data sets, such as MRI or CT or PET, to one another. Similar to our approach, this work uses a least squares minimization of distances between data sets, although their system uses a different distance function for minimization, in particular using for each data point, the distance from that data point to the model surface, as measured along a ray from the data point to the centroid of the model. Typical reported RMS errors are on the order of 3-5mms. As opposed to our system, however, this approach does require some operator intervention to set a decent initial starting position, which our system does not. It also apparently requires some operator intervention to steer the system towards the correct solution, suggesting that local minima are a potential problem. Our system avoids this difficulty by randomly perturbing near final solutions to find better nearby minima.

A second related approach is that of [6, 7, 24, 25, 36]. This method also does a least-squares minimization of a distance function to match data sets. Here, the distance is weighted by an estimate of the inverse variance of the noise in the measurements, and a Levenberg-Marquardt method is used to find the minimum. It appears that at present the method requires a reasonable initial starting position, though the authors observe that sampling over the view sphere could remove this restriction. Once a first estimate of the solution is found, points with large errors are removed (they are considered to be outliers) and the minimization process is repeated to refine the pose. It is unclear whether the removal of outliers is sufficient to keep the method from getting trapped into local minima.

A third approach is that of [2, 19, 20, 21, 12, 37] who perform automatic rigid registration of 3D surfaces by matching ridge lines which track points of maximum curvature along the surface. The ridge lines are characterized by five intrinsic parameters (curvature, torsion, maximum curvature of the surface, angle between curve and surface normal, and angle between curve tangent and direction of maximum curvature) which are used to hash the ridge points of the model data set into a five-dimensional hash table. During matching, the hash table is used to efficiently find model ridge points which are similar to ridge points from a second data set. A rigid transform is computed for each pair of matched ridge points and the results are collected in a six-dimensional transform accumulator. The most accessed cell in the accumulator is selected as the best transform. This method is not directly suitable for dealing with sparse data, such as the laser input.

# 9 Conclusions

We have described a method that automatically registers segmented clinical reconstructions (such as MRI or CT) of a patient with a live view of that patient, enabling a surgeon to visualize internal structures before executing procedures. The method enables a visual mix of live video of the patient with the segmented 3D MRI or CT model, supporting enhanced reality techniques for planning and guiding neurosurgical procedures, and to interactively view extracranial or intracranial structures non-intrusively. Extensions of the method include image guided biopsies, focused therapeutic procedures and clinical studies involving change detection over time sequences of images. The method has been applied to a series of neurosurgery cases, and has demonstrated reliability in accurately locating internal structures.

The heart of the method is a multi-stage, multi-resolution, registration method that can be used in fully automatic or semi-automatic mode, and which reliably finds solutions whose accuracy at finding a target ranges from 10 microns, in cases where highly varied structures are sensed, to 1.5 millimeters, in cases where much smoother structures are sensed. These results suggest that the best use of this visualization technique is to register based on data taken from a highly structured viewpoint, such as the face of a neurosurgery patient, but to use a visualization viewpoint from whatever view is desired by the surgeon.

## References


[1] Ayache, N., "Medical Computer Vision, Virtual Reality and Robotics" *Image and Vision Computing*, **13**(4):295–313, May 1995.

[2] Ayache, N., J.D. Boissonnat, L. Cohen, B. Geiger, J. Levy-Vehel, O. Monga, P. Sander, "Steps Toward the Automatic Interpretation of 3-D Images", In *3D Imaging in Medicine*, edited by H. Fuchs, K. Hohne, S. Pizer NATO ASI Series, Springer-Verlag, 1990, pp 107-120.

[3] Besl, J.P., N.D. McKay, "A method for registration of 3D shapes", *IEEE Trans PAMI* **14**:239–256, 1992.

[4] Black, P. McL., R. Kikinis, W. Wells, D. Altobelli, W. Lorensen, H. Cline and F. Jolesz, "A New Virtual Reality Technique for Tumor Localization" *Program Book of the Annual Meeting of the Congress of Neurological Surgeons*, 1993.

[5] Brunelli,R and T. Poggio, 1991, "HyberBF Networks for Real Object Recognition", *Int. Joint Conf. Artif. Intell.*, Sydney, Australia.

[6] Champleboux, G.; Lavallee, S.; Szeliski, R.; Brunie, L., "From accurate range imaging sensor calibration to accurate model-based 3D object localization", *IEEE Conf. CVPR*:83–89, 1992

[7] Champleboux, G.; Lavallee, S.; Sautot, P.; Cinquin, P., "Accurate calibration of cameras and range imaging sensor: the NPBS method", *IEEE Int. Conf. Rob. Autom.*:1552–1557, 1992

[8] Cline, H.E. and W.E. Lorensen and R. Kikinis and F. Jolesz, 1990, "Three-Dimensional Segmentation of MR Images of the Head Using Probability and Connectivity." *JCAT* **14**(6):1037-1045.

[9] Durbin, R., R. Szeliski, and A. Yuille, "An analysis of the elastic net approach to the travelling salesman problem," *Neural Computation* **1**(3):348–358, Fall 1989.

[10] Edelman, S. and T. Poggio, 1992 "Bringing the grandmother back into the picture: a memory-based view of object recognition", *Intl. J. Pattern Recog. Artif. Intell.* **6**:37–61.

[11] Ettinger, G., E. Grimson, T. Lozano-Perez, W. Wells, S. White, R. Kikinis, "Automatic Registration for Multiple Sclerosis Change Detection", *IEEE Workshop Biomed. Image Anal.*, Seattle, 1994.

[12] Feldmar, J. and N. Ayach, "Rigid and affine registration of smooth surfaces using differential properties," EECV, Stokholm, Sweden, 1994, pp. 397–406.

[13] Gerig, G. and W. Kuoni and R. Kikinis and O. Kübler, 1989, "Medical Imaging and Computer Vision: an Integrated Approach for Diagnosis and Planning," *Proc. 11'th DAGM Symposium*, Hamburg FRG, Springer, pp. 425–443.

[14] Gleason, P.L., R. Kikinis, D. Altobelli, W. Wells, E. Alexander III, P. McL. Black and F. Jolesz, "A New Virtual Reality Technique for Non-Linkage Stereotactic Surgery" *Abstract Booklet of the Annual Meeting of the World Society for Stereotactic and Functional Neurosurgery, Ixtapa, Mexico*, 1993.

[15] Grimson, W.E.L., 1990, *Object Recognition by Computer: The role of geometric constraints*, MIT Press, Cambridge.

[16] Grimson, W.E.L. and T. Lozano-Pérez, "Model-Based Recognition and Localization from Sparse Range or Tactile Data", *Int. J. Rob. Res.*, **3**, No. 3, 3 – 35, 1984.

[17] Grimson, W.E.L. and T. Lozano-Pérez, "Localizing Overlapping Parts by Searching the Interpretation Tree", *IEEE Trans. PAMI*, **9**, No. 4, 469 – 482, 1987.

[18] Grzeszczuk, R.; Tan, K.K.; Levin, D.N.; Pelizzari, C.A.; Hu, X.; Chen, G.T.; Beck, R.N.; Chen, C.T.; Cooper, M.; Milton, J.; et al., "Retrospective fusion of radiographic and MR data for localization of subdural electrodes", *J. Comput. Assis. Tomogr.* **16**(5):764–773, 1992.

[19] Gueziec, A., "Large Deformable Splines, Crest Lines and Matching", *INRIA TR 1782*, October 1992.

[20] Gueziec, A., N. Ayache, "Smoothing and Matching of 3-D Space Curves", *Second Eur. Conf. Comp. Vis.*, May 1992, pp 620-629.

[21] Gueziec, A., N. Ayache, "New Developments on Geometric Hashing for Curve Matching", *IEEE Conf. CVPR*, June 1993, pp 703-704.

[22] Huttenlocher, D. and S. Ullman, 1990, "Recognizing Solid Objects by Alignment with an Image," *Int. J. Comp. Vis.* **5**(2):195–212.

[23] Jiang, H., R.A. Robb and K.S. Holton, 1992, "A New Approach to 3-D Registration of Multimodality Medical Images by Surface Matching", in *Visualization in Biomedical Computing – SPIE*: 196–213.

[24] Lavallee, S.; Brunie, L.; Mazier, B.; Cinquin, P., "Matching of medical images for computed and robot assisted surgery", *Proc. 13th Int. Conf. Engineering in Medicine and Biology* **13**(1):39–40, 1991.

[25] Lavallee, S.; Szeliski, R.; Brunie, L., "Matching 3d smooth surfaces with their 2d Projections using 3d distance maps", *SPIE – Geometric methods in computer vision*, 1991, pp. 322–336.

[26] Levin, D.N.; Xioping Hu; Tan, K.K.; Galhotra, S.; Herrmann, A.; Chen, G.T.Y.; Pelizzari, C.A.; Balter, J.; Beck, R.N.; Chin-Tu Chen; Cooper, M.D., "Integrated 3-D display of MR, CT and PET images of the brain", *NCGA '89 Conference PRoceedings*: 179–186, 1989.

[27] Levin, D.N.; Xioping Hu; Tan, K.K.; Galhotra, S.; Pelizzari, C.A.; Chen, G.T.Y.; Beck, R.N.; Chin-Tu Chen; Cooper, M.D.; Mullan, J.F.; Hekmatpanah, J.; Spier, J.-P., "The Brain: integrated three-dimensional display of MR and PET images", *Radiology* **172**(3):783–789, 1989.

[28] Malandain, G., J.M. Rocchisani, "Registration of 3D medical images using a mechanical based method", *IEEE EMBS satellite symposium on 3D advanced image processing in medicine*, Nov. 2-4, 1992, Rennes, France.

[29] Mellor, J. P., "Realtime Camera Calibration for Enhanced Reality Visualization", *CVRMED '95*, April 1995, Nice, France, pp. 471–475.

[30] Neiw, H.M.; Chen, C.T.; Lin, W.C.; Pelizzari, C.A.. "Automated three-dimensional registration of medical images", *Medical Imaging V: Image Capture, Formatting, and Display: SPIE* **1445**:259–264, 1991.

[31] Pelizzari, C.A.; Chen, G.T.Y.; Spelbring, D.R.; Weichselbaum, R.R.; Chin-Tu Chen, "Accurate three-dimensional registration of CT, PET, and/or MR images of the brain", *JCAT* **13**(1):20–26, 1989.

[32] Pelizzari, C.A.; Evans, A.C.; Neelin, P.; Chen, C.T.; Marrett, S., "Comparison of two methods for 3D registration of PET and MRI images", *Proc. 13th Int. Conf. Engineering in Medicine and Biology* **13**(3):221–223, 1991.

[33] Pelizzari, C.A.; Tan, K.K.; Levin, D.N.; Chen, G.T.Y.; Balter, J., "Interactive 3D patient-image registration", *Information Processing in Medical Imaging. 12th International Conference, IPMI '91 Proceedings*:132–141, 1991.

[34] Press, W.H., S.A. Teukolsky, S.T. Vetterling, B.P. Flannery, *Numerical Recipes in C, The Art of Scientific Computing, Second Edition*, Cambridge University Press, 1992.

[35] Sumanaweera; Glover; Binford; Adler, "MR Susceptibility Misregistration Correctoin", *IEEE TR. Med. Imaging* **12**, 1993.

[36] Szeliski, R.; Lavallee, S., "Matching 3d Anatomical surfaces with non-rigid deformations using octree-splines", *SPIE* **2031**, 1993.

[37] Thirion, J.P., "Extremal points, definition and application to 3D image registration,", *IEEE Proc. CVPR*, 1994, Seattle, pp. 587–592.

[38] Turkington, T.G.; Jaszczak, R.J.; Greer, K.L.; Coleman, R.E.; Pelizzari, C.A., "Correlation of SPECT images of a three-dimensional brain phantom using a surface fitting technique", *IEEE Trans. Nuclear Science* **39**(5):1460–1463, 1991.

[39] van den Elsen, P, E.J.D. Pol, M. Viergever, "Medical Image Matching – A Review with Classification", *IEEE Engineering in Medicine and Biology* **12**(4):26–39, March 1993.

[40] Wells, W. M., 1993, *Statistical Object Recognition*, Ph.D. Thesis, MIT. (MIT AI Lab TR 1398)

[41] W. Wells, R. Kikinis, D. Altobelli, W. Lorensen, G. Ettinger, H. Cline, P.L. Gleason and F. Jolesz, 1993, "Video Registration using Fiducials for Surgical Enhanced Reality" *Proc. 15th Annual Conf. of the IEEE Engineering in Medical and Biology Society*, IEEE.

[42] Wells, W.M., and R. Kikinis and F.A. Jolesz, and W.E.L. Grimson, 1993, "Statistical Gain Correction and Segmentation of Magnetic Resonance Imaging Data", in preparation.

| Coarse model | Hypotheses | Fine model | Solutions | RMS |
|---|---|---|---|---|
| 406 | 88012 | - | 0 | - |
| 409 | 16945 | 9428 | 1 | 1.71 |
| 500 | 97862 | - | 0 | - |
| 502 | 95978 | - | 0 | - |
| 497 | 114602 | - | 0 | - |
| 494 | 101477 | - | 0 | - |
| 429 | 62656 | - | 0 | - |
| 426 | 58521 | - | 0 | - |

Table 1: Statistics on registration from multiple views. Shown are the views for which the laser and camera are visible to one another. Coarse model is the number of model points to initially match, against 481 laser points. Hypotheses indicates the number of acceptable matches found by Interpretation Tree search. In only one case did a hypothesis survive Alignment verification, in which case a finer model of 9428 points was refined relative to the laser data, as shown.

| Patient | Mean RMS Error |
|---|---|
| skull wo normals | 1.705 |
| skull w normals | 0.585 |
| 070194 wo normals | 1.495 |
| 071594 wo normals | 1.819 |
| 071594 w normals | 1.371 |
| 082294 wo normals | 1.514 |
| 082294 w normals | 1.007 |

Table 2: Table of mean RMS values for set of perturbation trials.

| Patient | RMS cutoff | Entry Vector Deviation | 95 Percentile Deviation | % Success Rate |
|---|---|---|---|---|
| skull wo normals | 5.0 | 0.011–0.009 | 0.645–.625 | 100 |
| skull w normals | 5.0 | 0.213–0.199 | 1.085–1.113 | 96 |
| 070194 wo normals | 5.0 | 0.031–0.034 | 1.017–1.090 | 98 |
| 070194 wo normals | 2.0 | 0.030–0.033 | 1.013–1.087 | 96 |
| 071594 wo normals | 5.0 | 0.104–0.113 | 0.982–01.004 | 100 |
| 071594 w normals | 5.0 | 2.129–2.458 | 18.911–19.939 | 100 |
| 071594 w normals | 1.37 | 1.260–1.166 | 18.459–13.783 | 56 |
| 082294 wo normals | 5.0 | 1.680–1.630 | 6.384–6.440 | 100 |
| 082294 wo normals | 1.52 | 1.582–1.419 | 6.032–5.727 | 96 |
| 082294 w normals | 5.0 | 1.160–0.908 | 2.846–2.851 | 100 |
| 082294 w normals | 1.01 | 0.723–0.742 | 3.274–3.107 | 76 |

Table 3: Table of deviations in entry tube, for set of perturbation trials. Entry Vector Deviation is median deviation, from start to end of 6 cm tube. 95 Percentile Deviation lists the deviation below which 95 percent of the samples lie, giving an indication of the spread of the deviation. Success rate is the percentage of trials in which an RMS value was obtained below the desired cutoff.

| Perturbation | Found | Median Deviation | 95 Percentile Deviation |
|---|---|---|---|
| 1.0 | 20 | 0.969 | 3.440 |
| 2.0 | 20 | 0.921 | 2.083 |
| 3.0 | 20 | 0.814 | 3.031 |
| 4.0 | 19 | 1.080 | 3.202 |
| 5.0 | 13 | 0.990 | 6.620 |
| 6.0 | 14 | 1.223 | 2.674 |
| 7.0 | 11 | 2.273 | 4.993 |
| 8.0 | 14 | 0.846 | 3.425 |
| 9.0 | 9 | 1.937 | 5.280 |
| 10.0 | 7 | 1.368 | 3.009 |
| 11.0 | 8 | 0.941 | 4.341 |
| 12.0 | 5 | 1.009 | 5.519 |
| 13.0 | 4 | 0.621 | 2.814 |

Table 4: Table of deviations in entry tube, for set of perturbation trials. Perturbation is $n$ mms of translation and $2n$ degrees of rotation. Median Deviation is median deviation, from start to end of 6 cm tube. 95 Percentile Deviation lists the deviation below which 95 percent of the samples lie, giving an indication of the spread of the deviation. Found is number of trials out of 20 for which a solution with an RMS error of less than 1.6 mms is found.

| Perturbation | Found | Median Deviation | 95 % Deviation |
|---|---|---|---|
| 1.0 | 20 | 0.969 | 3.440 |
| 2.0 | 20 | 0.921 | 2.083 |
| 3.0 | 20 | 0.814 | 3.031 |
| 4.0 | 20 | 1.080 | 4.402 |
| 5.0 | 19 | 1.662 | 22.318 |
| 6.0 | 19 | 1.363 | 17.634 |
| 7.0 | 18 | 3.754 | 131.091 |
| 8.0 | 18 | 2.073 | 16.714 |
| 9.0 | 17 | 4.375 | 27.773 |
| 10.0 | 15 | 3.044 | 137.169 |
| 11.0 | 13 | 2.528 | 85.483 |
| 12.0 | 13 | 11.435 | 127.461 |
| 13.0 | 11 | 13.493 | 81.886 |
| 14.0 | 9 | 8.554 | 197.988 |
| 15.0 | 7 | 9.768 | 136.693 |
| 16.0 | 9 | 10.808 | 136.934 |
| 17.0 | 11 | 114.949 | 179.150 |
| 18.0 | 12 | 8.680 | 106.366 |
| 19.0 | 10 | 69.834 | 180.454 |

Table 5: Table of deviations in entry tube, for set of perturbation trials. Perturbation is $n$ mms of translation and $2n$ degrees of rotation. Median Deviation is median deviation, from start to end of 6 cm tube. 95 Percentile Deviation lists the deviation below which 95 percent of the samples lie, giving an indication of the spread of the deviation. Found is number of trials out of 20 for which a solution with an RMS error of less than 5.0 mms is found.

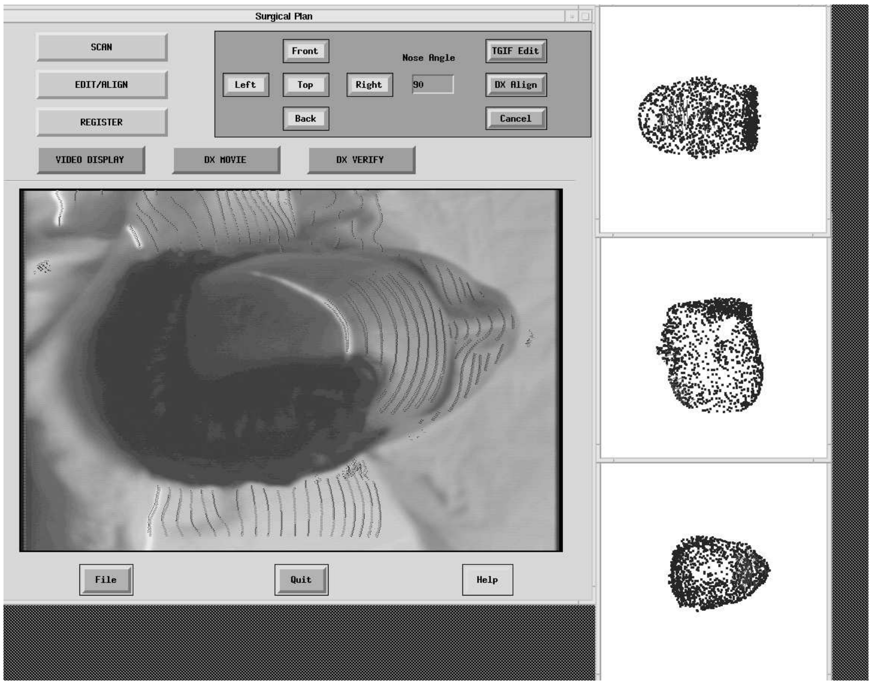


Figure 1: Example of graphical interface used to obtain initial alignment.

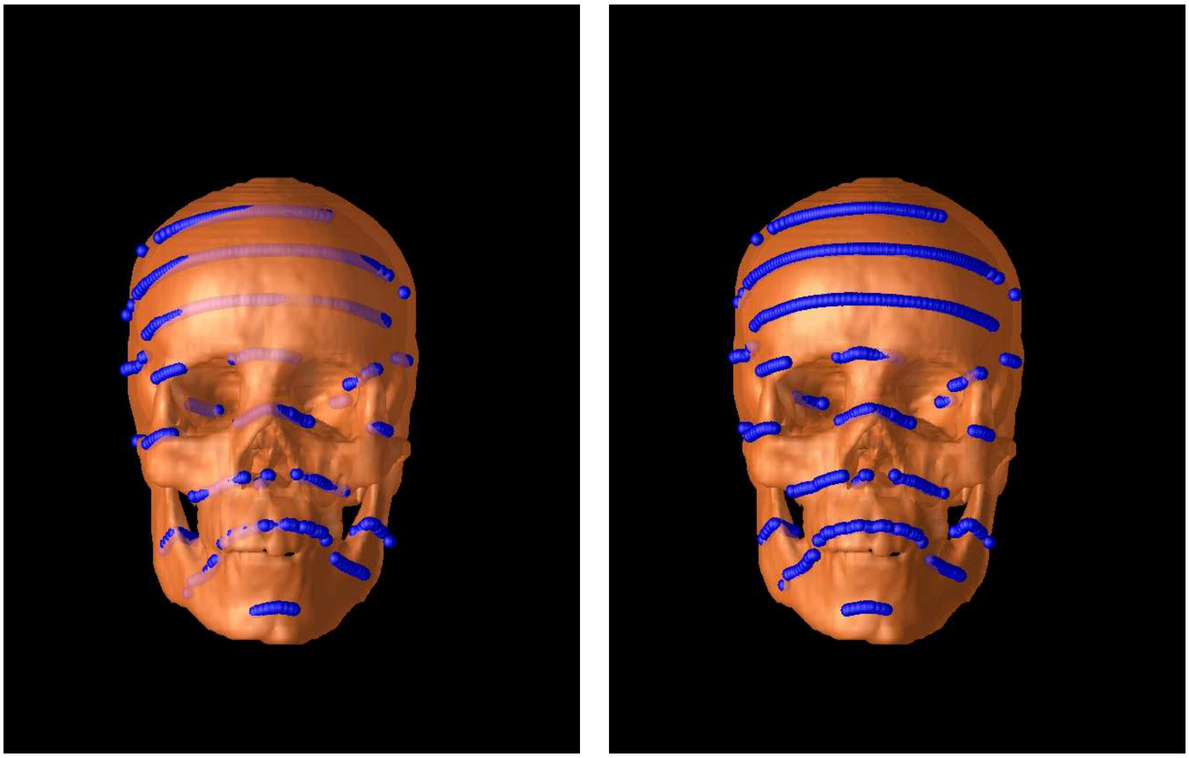

Figure 2: (left) Example of laser data (shown as large dots) overlaid on CT model of a plastic skull, after an initial alignment of the two point sets. Note the transparent laser points which are actually lying inside the skull. (right) Example of registered laser data (shown as large dots) overlaid on CT model.

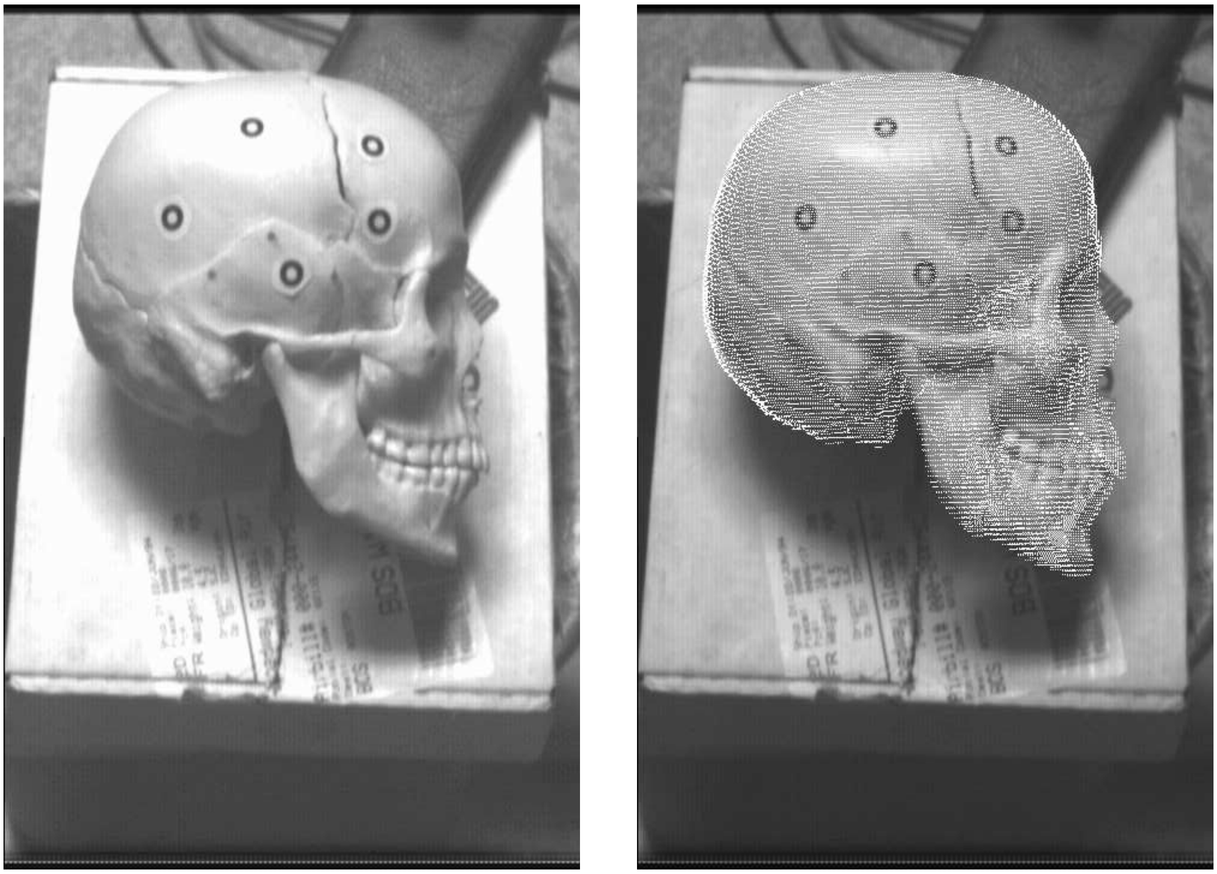

Figure 3: Example of video image, and the overlay of a registered 3D (CT) model of an object with the real object in that image (figures courtesy of J.P. Mellor).

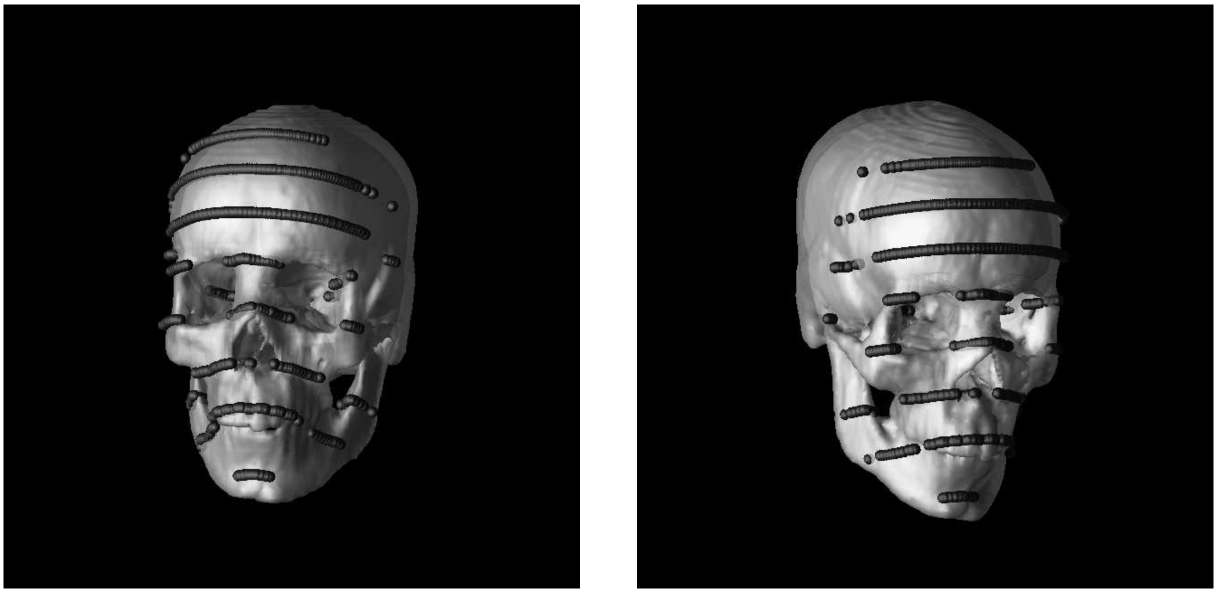

Figure 4: Examples of registered laser data (shown as large dots) overlaid on CT model.

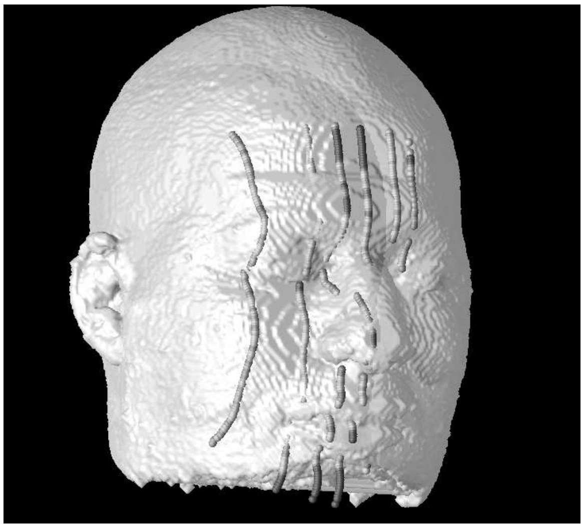

Figure 5: Example of registered laser data (shown as large dots) overlaid on MRI model.

Figure 6: Visualization image of brain merged with video view.

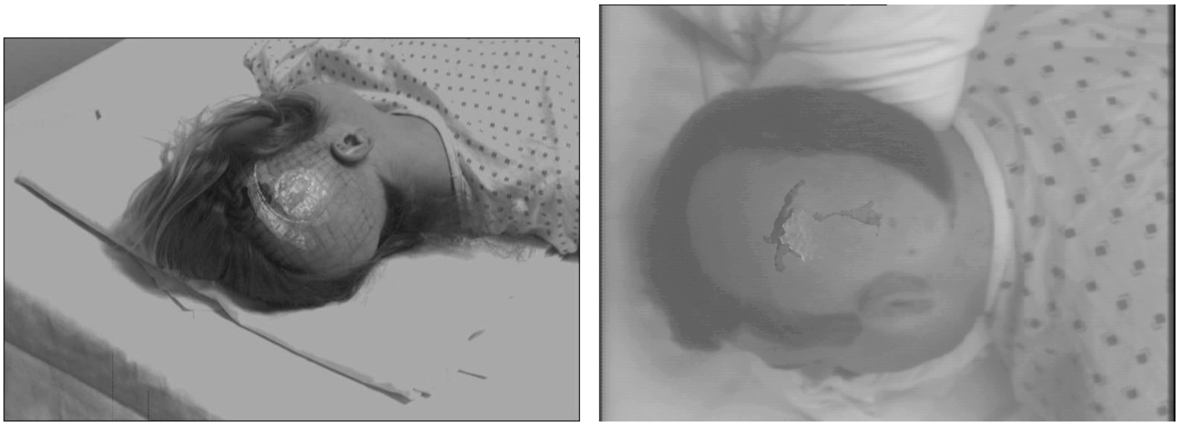

Figure 7: Examples of combining the registration of MRI to laser data and the calibration of a video camera relative to the laser to provide an enhanced reality visualization of patients. In each case, the tumor and the ventricles are displayed in registration with the patient.

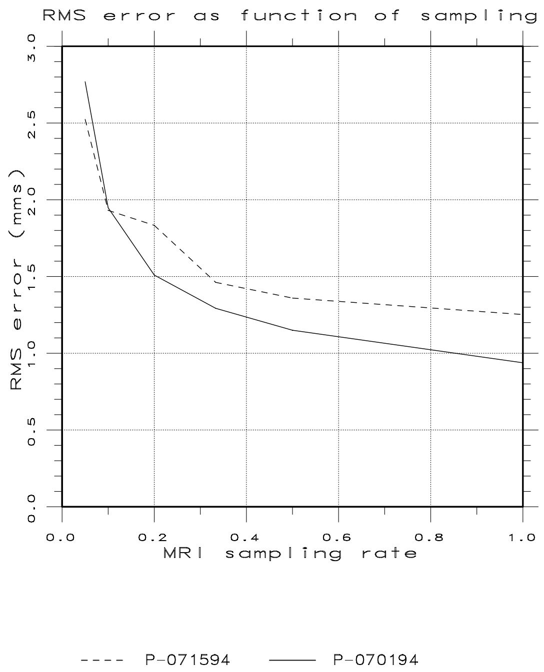


Figure 8: RMS error as a function of the sampling of the MRI data. Patient 071594 had an MRI model of 123,725 points, which were matched against 956 laser points. Patient 070194 had an MRI model of 109,641 points, which were matched against 949 laser points. The graphs show the RMS error (in mms) as a function of the fraction of the model actually sampled.

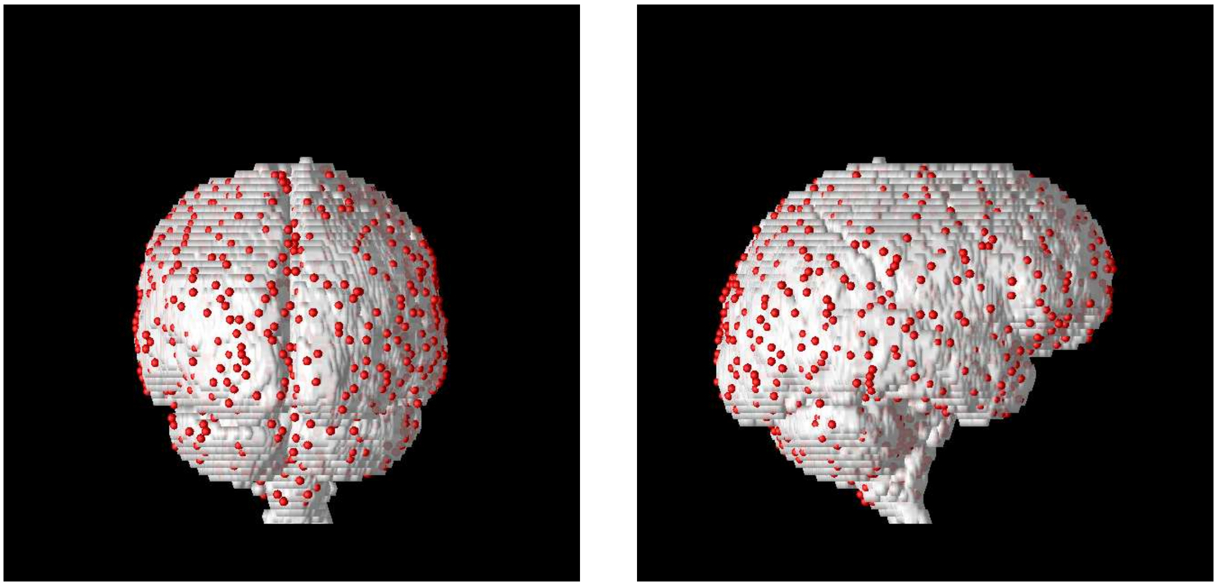

Figure 9: Example of registering MRI data sets from different times. A subsample set of surface points from one image are matched and overlaid on the surface of the second image.

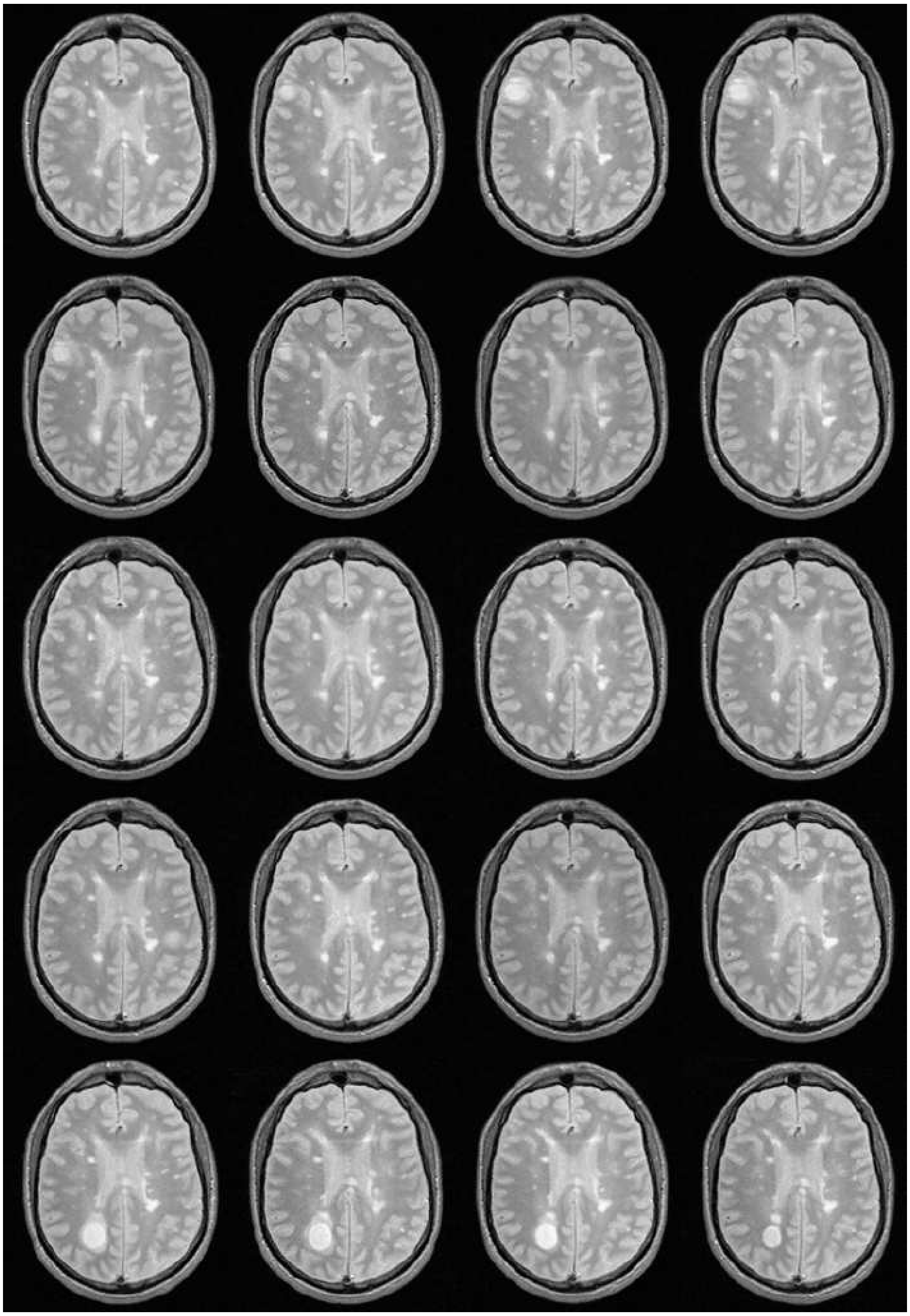

Figure 10: A series of MRI slices of a patient taken several weeks apart, over a period of eight months. Shown are the results of reslicing and normalizing each subsequent MRI scan, relative to the first scan, so that, in principle, the same slice of the anatomy is shown. One can easily see the changes in lesions in the upper left and lower left, over time.