Utilizing Segmented MRI Data in Image-Guided Surgery

W.E.L. Grimson[1] G.J. Ettinger[1] T. Kapur[1] M.E. Leventon[1]
W.M. Wells III[1,2] R. Kikinis[2]

July 9, 1996

Abstract

While the role and utility of Magnetic Resonance Images as a diagnostic tool is well established in current clinical practice, there are a number of emerging medical arenas in which MRI can play an equally important role. In this article, we consider the problem of image-guided surgery, and provide an overview of a series of techniques that we have recently developed in order to automatically utilize MRI-based anatomical reconstructions for surgical guidance and navigation.

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

1 Introduction

In recent years, Magnetic Resonance Imaging (MRI) has become a commonplace medical diagnostic tool [2], especially for cases involving soft tissue, such as in the brain. Two factors combine to make MRI a very valuable clinical tool: fine scale spatial resolution allows for the detection and delineation of detailed structures, and the range of responses of tissues to the MR stimulus allows for the visual differentiation of a range of internal tissues. While the diagnostic characteristics of MRI are of clear value and import, there are a number of newly emerging medical arenas in which MRI scans can play an equally critical role. In this article we focus on one of them: image-guided surgery.

Our goal is to build an end-to-end image-guided surgical aid. Such a system takes as input an MR scan, then processes the MRI data to provide reconstructions of an individual patient's anatomy, and finally utilizes those reconstructions to interactively guide a surgeon's execution of a procedure. Our goal is to give a surgeon the equivalent of "x-ray vision", that is, to allow the surgeon to view a patient and at the same time display in exact alignment with that view all desired internal structures, so that the surgeon in essence may "look inside" a patient before executing each stage in a surgical procedure. By providing such registered visualizations, along with an interactive ability in which the surgeon can probe structures and points within the operating field and see a full 3D visualization of the corresponding points in the MR scan, we hope to provide the surgeon with an enhanced ability to plan, to navigate and to localize throughout a surgical procedure.

Furthermore, by building a complete, end-to-end system, we are forced to directly consider the questions of what information must be extracted from the MRI scans, and how best to present that information to the surgeon. Thus, we provide in this paper an overview of our system, highlighting the following stages of processing:

- Automatic intensity-driven labelling of MRI voxels by tissue type;
- Automatic or semi-automatic shape-driven segmentation of labelled voxels into distinct anatomical structures;
- Rendered displays of those anatomical structures;
- Automatic registration of rendered anatomical structures to the current position of the actual patient, and display of such rendered, registered structures in concert with live video views of the patient;

- Automatic tracking of surgical instruments and probes relative to the patient and relative to the registered anatomical reconstructions, enabling full visualization of the surgical site, either in 3D reconstruction, or in MRI slices.

In the following sections, we detail our current approach to each of these problems. We also demonstrate each stage, with a sequence of examples from actual surgical cases. As our initial system has been used primarily for neurosurgical cases, we use such examples as the focus of our discussion throughout the paper.

2 Intensity-driven labelling of tissue type in MRI

The first stage in our system is to transform the input MR image data, by providing a tissue class label (e.g. white matter, grey matter, csf, bone, skin, etc.) for each voxel of data. Intensity-based classification of MR images has proven problematic, however, even when advanced techniques such as non-parametric multi-channel methods are used, primarily due to spatial inhomogeneities in the equipment's sensitivity. The MRI signal is derived from radio-frequency (RF) signals emanating from the scanned tissue. Diagnostic MRI machines use RF coils (antennas) designed to have uniform RF sensitivity, or *gain*, throughout their working volume. Although images derived from such coils appear visually uniform, there are often significant departures from the ideal that disturb intensity-based segmentation methods. One example is differentiating white matter and gray matter in the brain. The spatial inhomogeneities in RF gain are often of sufficient magnitude to cause the distributions of intensities associated with these two tissue classes to overlap, thereby defeating intensity based classification.

This section describes a statistical method that uses knowledge of tissue properties and gain inhomogeneities to correct the gain artifact of MRI. The result is a method that provides both more accurate segmentation of tissue types and better visualization of MRI data. Additional details may be found in [40].

2.1 Description of Method

If either the RF gain or the tissue type is known at an image location, then it is relatively easy to use models of the imaging process to infer the other parameter at that location, given the measured signal. It is problematic, however, to determine either the gain or the tissue type without knowledge of the other. The strategy used by our method is to cycle between estimating the tissue type and the RF gain throughout the image.

The MRI gain artifact addressed in this section is a spatially-varying factor that multiplies the intensity data. Application of a logarithmic transformation to the data allows the artifact to be modeled as an additive *bias* field.

2.1.1 Bias Field Estimator

We describe a Bayesian approach to estimating the bias field that represents the RF gain artifact in log-transformed intensity data. Similar to other approaches to intensity-based segmentation of MRI [12] [5], the distribution for observed values is modeled as a normal distribution:

$$p(Y_i \mid \Gamma_i, \beta_i) = \mathrm{G}_\sigma(Y_i - \mu(\Gamma_i) - \beta_i) \ , \tag{1}$$

where

$$\mathrm{G}_\sigma(x) \equiv \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{1}{2}(\frac{x}{\sigma})^2)$$

is the scalar Gaussian distribution, with variance σ^2 and where

Y_i	is the observed intensity at the i^{th} voxel[1]
Γ_i	is the tissue class at the i^{th} voxel[2]
$\mu(x)$	is the mean intensity for tissue class x
β_i	is the bias field at the i^{th} voxel.

A stationary prior (before the image data is seen) probability distribution on tissue class is used, it is denoted

$$p(\Gamma_i) \ . \tag{2}$$

If this probability is uniform over tissue classes, our method devolves to a maximum-likelihood approach to the tissue classification component.

The entire bias field is denoted by $\beta = (\beta_0, \beta_1, \ldots, \beta_{n-1})^T$, with a zero mean Gaussian prior probability density,

$$p(\beta) = \mathrm{G}_{\psi_\beta}(\beta) \ , \tag{3}$$

where

$$\mathrm{G}_{\psi_\beta}(x) = (2\pi)^{-\frac{n}{2}} |\psi_\beta|^{-\frac{1}{2}} \exp(-\frac{1}{2} x^T \psi_\beta^{-1} x)$$

is the n-dimensional Gaussian distribution. ψ_β is the $n \times n$ covariance matrix for the bias field. We use this model to characterize the spatial smoothness inherent in the bias field.

[1] The extension to vector measurements per voxel is straightforward.

[2] For example, when segmenting brain tissue $\Gamma_i \in \{\textit{white-matter}, \textit{gray-matter}\}$.

Using the definition of conditional probability, assuming independence of tissue class and the bias field, and assuming independence of voxels, Bayes rule yields the posterior probability on the bias field, given observations,

$$p(\beta \mid Y) = \prod_i \left[\sum_{\Gamma_i} p\left(Y_i \mid \Gamma_i, \beta_i\right) p\left(\Gamma_i\right) \right] \frac{p(\beta)}{p(Y)} \quad . \tag{4}$$

The maximum-a-posteriori (MAP) principle can be used to formulate an estimate of the bias field as the value of β having the largest posterior probability,

$$\hat{\beta} = \arg\max_{\beta} p(\beta \mid Y) \quad . \tag{5}$$

A necessary condition for a maximum of the posterior probability of β is that its gradient with respect to β be zero at the maximum:

$$\left[\frac{\partial}{\partial_{\beta_i}} \ln p(\beta \mid Y) \right]_{\beta = \hat{\beta}} = 0 \qquad \forall_i \quad .$$

Installing the statistical modeling of Equations 1-5 leads to the following matrix expression for the zero gradient condition:

$$\left[Y - \beta - WU - \sigma^2 \psi_{\beta}^{-1} \beta \right]_{\beta = \hat{\beta}} = \mathbf{0} \quad ,$$

where W is a matrix of weights

$$W_{ij} = \frac{p(\Gamma_i = \textit{tissue-class-j}) \mathrm{G}_{\sigma}(Y_i - \mu(\textit{tissue-class-j}) - \beta_i)}{\sum_{\Gamma_i} p(\Gamma_i)\ \mathrm{G}_{\sigma}(Y_i - \mu(\Gamma_i) - \beta_i)} \quad , \tag{6}$$

and where

$$U \equiv (\mu(\textit{tissue-class-1}), \mu(\textit{tissue-class-2}), \ldots)^T$$

is a vector containing the tissue class mean intensities.

Equivalently, we have

$$[\beta = H(Y - WU)]_{\beta = \hat{\beta}} \quad , \tag{7}$$

where

$$H \equiv (I + \sigma^2 \psi_{\beta}^{-1})^{-1} \quad . \tag{8}$$

Although Equation 7 has the superficial appearance of being a linear estimator of the bias field β, given the observations Y, unfortunately, it is not linear, because the weights W are non-linear functions of β (Equation 6).

The result of the statistical modeling in this section has been to formulate the problem of estimating the bias field as a non-linear optimization problem embodied in Equation 7.

2.1.2 EM Algorithm

We use the Expectation-Maximization (EM) algorithm to obtain bias field estimates from the non-linear estimator of Equation 7. The EM algorithm was originally described in its general form by A.P. Dempster, N.M. Laird and D.B. Rubin in 1977 [8]. It is often used in estimation problems where some of the data is "missing". In this application the missing data is knowledge of the tissue classes – if the tissue classes were known, then estimating the bias field would be straightforward.

In this application, the EM algorithm iteratively alternates evaluations of the following expressions:

$$\hat{\beta} \leftarrow H(Y - WU) \tag{9}$$

$$W_{ij} \leftarrow \frac{p(\Gamma_i = \textit{tissue-class-j})\mathrm{G}_\sigma(Y_i - \mu(\textit{tissue-class-j}) - \hat{\beta}_i)}{\sum_{\Gamma_i} p(\Gamma_i)\ \mathrm{G}_\sigma(Y_i - \mu(\Gamma_i) - \hat{\beta}_i)} \quad . \tag{10}$$

Equation 10 (the *E-Step*) is equivalent to calculating the posterior tissue class probabilities (a good estimator of tissue class) when the bias field is known. Equation 9 (the *M-Step*) is equivalent to a MAP estimator of the bias field when the tissue probabilities W are known. The linear operator H is applied to the difference between the measurement and a prediction based on the tissue class model. Thus, a good estimator of tissue class is produced as a side effect of estimating the bias field.

The iteration may be started on either expression. Initial values for the weights will be needed to start with Equation 9, and initial values for the bias field will be needed to start with Equation 10.

It is shown in [8] that in many cases the EM algorithm enjoys pleasant convergence properties – namely that iterations will never worsen the value of the objective function. Provided that the bias estimates are bounded, our model satisfies the necessary conditions for guaranteed convergence.

Filtering To use the method, we must determine the operator H (Equation 8), which acts as a linear filter that incorporates models of the signal intensity, via σ, and the bias field, via ψ_β. Ideally, this would be done by estimating the covariance ψ_β, but given the size of this matrix, this is impractical. H can be estimated by several other means. For example, it is the optimal filter for estimating the bias when the tissue classes are known constants rather than random variables (the "complete-data" situation with the EM algorithm). In this case the measurement model would be $p(Y_i \mid \beta_i) = \mathrm{G}_\sigma(Y_i - \mu(\Gamma_i) - \beta_i)$. H is additionally the optimal *linear* least squares estimator (LLSE) for arbitrary models when ψ_β describes the second-order statistics of the bias field.

H is also essentially equivalent to the LLSE for discrete random *processes* with given second order statistics (auto-correlation functions). This estimator is characterized by the Wiener-Hopf equations. The application of Wiener filtering is often approached via Fourier transforms methods, yielding a filter frequency response in terms of the power spectra of the signal and noise. These power spectra may be estimated using techniques of spectral estimation.

However H is obtained, if it is the optimal linear filter for estimating β when the tissue class is known, then Equation 7 describes the (non-linear) MAP estimator when the tissue class is unknown.

In practice, it may be difficult to obtain *the* optimal linear filter. H may be instead chosen as a good engineering approximation of the optimal linear filter. In this case, Equations 9 and 10 are still a good estimator for the missing data case, and the good convergence properties of the EM algorithm still obtain. This is the approach we have taken in our implementation, where the filter was selected empirically. The justification here is the good results obtained with the method.

2.1.3 Implementation

The results described in Section 2.2 were obtained using a preliminary implementation of the method coded in the C programming language. This single-channel implementation accommodates two tissue classes, and uses an input segmentation to limit the region of interest (ROI) to be classified and gain-corrected. Subsequent implementations accomodate multi-channel data and more general tissue models [40].

The algorithm of Section 2.1.2 is initiated on the "E step", Equation 10, with a flat initial bias field. A moving-average lowpass filter was used for the operator H in Equation 9. This filter only operates in the input ROI, and adjusts the effective averaging window at boundaries to confine influence to the ROI. This filter is shift-invariant, except at the boundary regions. Window widths of 11 – 15 were used. A uniform distribution was used for the prior on tissue class.

Operating parameters were selected manually. These were found to be stable with respect to the type of acquisition. The method is relatively insensitive to parameter settings – for example, errors in the specifications of the class means can be accommodated by shifts in the resulting bias field estimate.

In a typical case, the program was run until the estimates stabilized, typically in 10 – 20 iterations, requiring approximately 1 second per iteration on a Sun Microsystems Sparcstation 2.

2.2 Results

In this section we describe the application of our method to the segmentation of the brain into white matter and gray matter in spin-echo and fast-spin-echo images. Examples are shown for gradient echo coronal images and surface coil images. The method has also been applied to spine phased-array and surface-coil images.

All of the MR images shown in this section were obtained using a General Electric Signa 1.5 Tesla clinical MR imager. An anisotropic diffusion filter developed by Gerig et al. [11] was used as a pre-processing step to reduce noise.

2.2.1 Gradient Echo Brain

This example describes white matter / gray matter segmentation in a coronal gradient echo acquisition. Figure 1 shows the input (proton-density) image, acquired using a conventional "birdcage" head coil. The brain tissue ROI was generated manually. Figures 2 and 3 shows the initial and final segmentations, respectively. Figure 4 shows the final bias field estimate, with range -15 to -5. The largest value of the input data was 85. Note the dramatic improvement in the right temporal area – in the initial segmentation the white matter is completely absent in the binarization.

2.2.2 Surface Coil Brain

Here we show the results of the method on a sagittal surface coil brain image. The five-inch receive-only surface coil was positioned at the back of the head. Figure 5 shows the intensity image, after having been windowed by a radiologist for viewing in the occipital area. Figure 6 shows the final gray matter probability and Figure 7 shows the final bias field estimate. The brain ROI was generated manually. Figure 8 shows a corrected intensity image. Here the gain field estimate has been applied as a correction, in the brain tissue only. Note the dramatic improvement in "viewability" – the entire brain area is now visible, although the noise level is higher in the tissue farthest from the surface coil.

2.3 Accuracy Assesment

This section describes an assessment of the accuracy of the method (EM segmentation) in segmenting white matter and gray matter in a single slice of an axial T2-weighted spin-echo image. Two comparisons were performed. The method was compared to manual segmentation performed by experienced raters, and to a method of supervised multivariate classification. The images, manual segmentations, and supervised segmentations are described in [23]. An ROI was obtained by selecting those pixels that were labelled as brain tissue by four of the five raters in the manual segmentations. The amount of difference between segmentations was calculated as the percentage of pixels in the ROI having different labels.

2.3.1 Comparison to Manual Segmentation

The percentage of difference was calculated for comparisons within a group consisting of the manual segmentations and the results of EM segmentation. For the manual segmentations, the average percentage of difference from the other segmentations ranged between 16.9 % and 21.13 %, and for the EM segmentation it was 21.04 %. Thus the EM segmentation is consistent with the manual segmentations.

2.3.2 Comparison to Supervised Classification

The method was also compared to segmentations performed by the same expert raters using a supervised segmentation method. In this test percentages of difference were calculated for comparisons of the supervised segmentations and the EM segmentations with the manual segmentations described above. For the supervised segmentations, the average percentage of difference from the manual segmentations ranged from 21.71 % to 24.15 %, and for the EM segmentation it was 21.04 %. The EM segmentation is seen to have less average difference from the manual segmentations than the supervised segmentations.

3 Shape-driven segmentation of labelled MRI into anatomical structures

The result of the previous stage is an initial mapping of MRI voxels into tissue type. The next stage is twofold: we need to extract out specific anatomical structures, primarily by finding connected regions of the data with the same tissue label; and we need to account for the

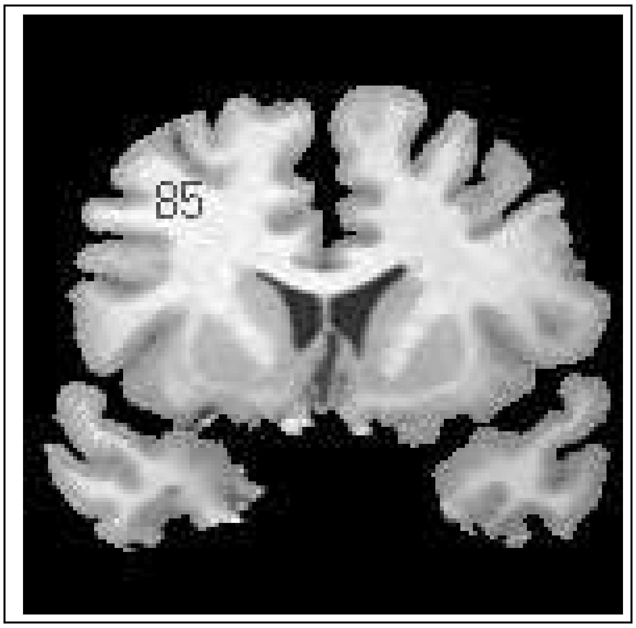

Figure 1: Original Gradient Echo Image

Figure 2: Initial White Matter Segmentation

Figure 3: Final White Matter Segmentation

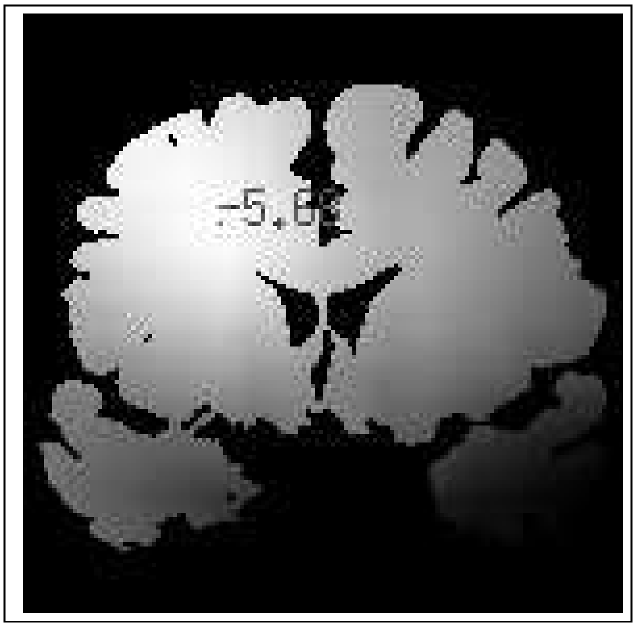

Figure 4: Estimated Bias Field

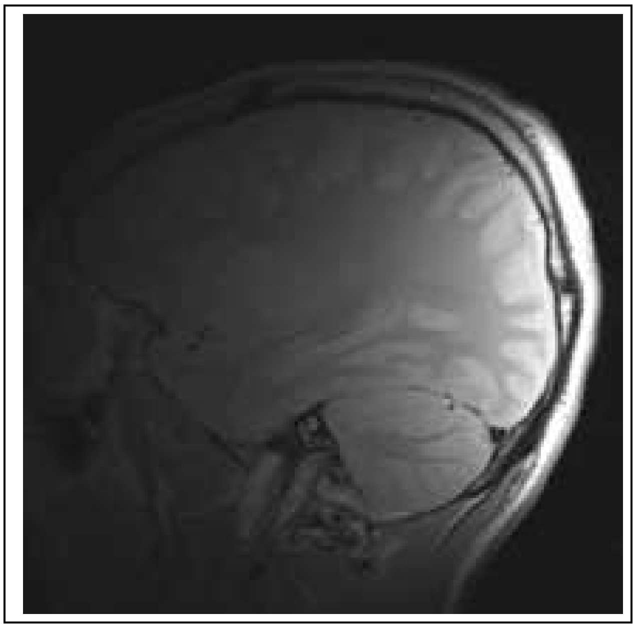

Figure 5: Input Image

Figure 6: Final Gray Matter Probability

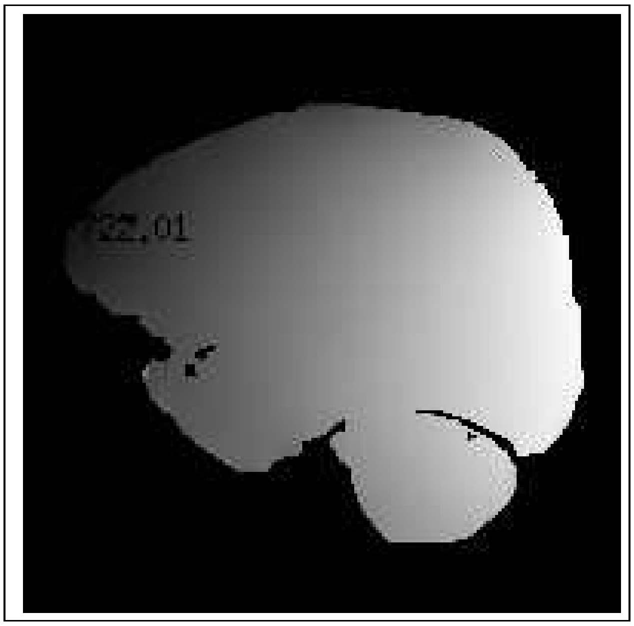

Figure 7: Final Gray Matter Bias

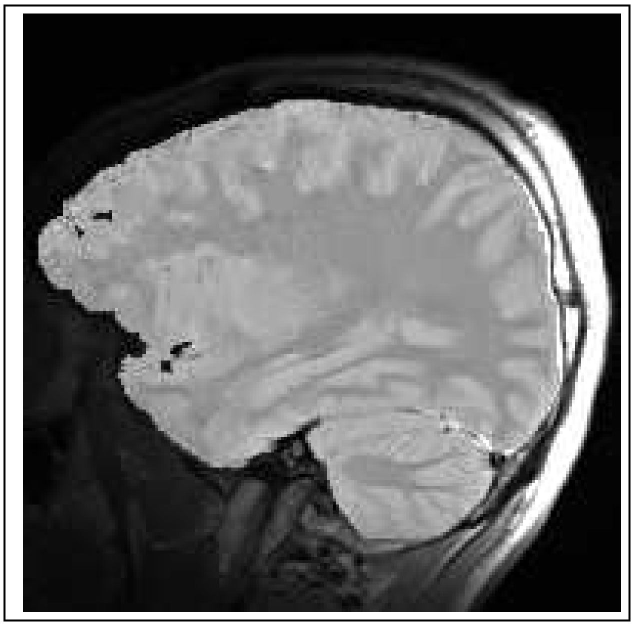

Figure 8: Corrected Image

fact that occasionally different tissue types have similar intensities, by incorporating anatomical knowledge into the labelling process. We have developed automated techniques for incorporating knowledge of anatomy into the segmentation process. These methods work well for normal tissue or for cases where pathologies have predictable appearances, such as Multiple Sclerosis lesions. With other pathologies, such as tumors, or structures which vary significantly from the trained cases, segmentation requires semi-automated tools for incorporation of anatomical expertise. Semi-automated segmentation tools developed at Brigham and Women's Hospital allow for efficient verification and updating of segmented structures. These tools consist of visualizers for validating segmentation results and, if necessary, semi-automated refinement tools. Both the automated techniques and semi-automated tools are described below.

3.1 Automated structural segmentation

In this stage of our process, we address the issues of finding connected regions of data with the same tissue label and accounting for the possible overlaps in intensity values across different tissue types. The EM Segmentation, described in Section 2, generally classifies the MRI data into four classes roughly corresponding to white matter, grey matter, csf and skin. We say "roughly" since there are other non-brain tissues in the scan that naturally overlap in intensity with the four classes mentioned above, and as a result are incorrectly classified into one of the four classes. We reduce these misclassifications by using mathematical morphology operations [17, 19, 32, 33] such as erosion and dilation, as well as connectivity analysis. The procedure we use to extract the brain from the labelled scan uses anatomical information about the relation of the brain to other non-brain tissue in the scan. It consists of the following steps:

- Perform an erosion operation with a spherical (in real space which usually implies elliptical in image space due to the anisotropy of the voxels) structuring element with radius corresponding to the thickness of the connectors between brain and the cranium (determined empirically, and held constant over scans), so that it eliminates connections from the brain to any misclassified non-brain structure.

- Find the largest 3D connected component with tissue labels corresponding to the brain.

- Dilate the brain component obtained in the previous step by a structuring element slightly larger in size to the one used in the erosion, conditioned on the brain labels in the input image. Since the dilation is conditioned on the original image, no boundaries are expanded in this process. This corresponds approximately to restoring the boundaries of the brain component that were distorted in the erosion step.

The result of this process is an improved segmentation of the tissue types, which incorporates topological information into the results of the pure intensity classification. We have used this process to segment over 200 brains thus far. Figure 9 shows two views of the segmented white matter surface for two different normal controls.

In addition to visual inspection of our results by neuroanatomy experts, we have performed quantitative performance analysis of our system against segmentations produced manually by experts. Our results are summarized in the following three tables. Table 1 shows a simple comparison of the number and percentage of pixels that are classified differently by our segmenter as compared to expert segmentation, based on a 3-class segmentation of the data into white matter, grey matter and csf. Note that while the total number of voxels classified differently in the two cases appears large, in fact compared to the total volume in each data set, only a small fraction (typically 1-2%) are actually different. To test the extent to which this difference in classification affects the overall segmentation, we ran a second test. Tables 2 and 3 present the result of a test in which the edges of the brain tissue generated by our segmentations and the edges generated by the manual segmentations are compared. This comparison is performed in two steps: first by finding the mean distance (city-block) from each edge in our segmentation to edges in the manual segmentation (which basically penalizes false positives) as shown in Table 2, and then by finding the mean distance from an edge in the manual segmentation to edges in our segmentation (which penalizes false negatives) as shown in Table 3. We also include some other statistics on the edges, such as the percent of edges in our segmentation that coincide with (or are one and two pixels away from) edges in the manual segmentation. We repeat this test to compute the number of edges in the manual segmentation that coincide with, or are within one or two pixels of the edges of our segmentation. The idea behind these tests is to convey the "goodness of fit" between the brain boundaries produced by our segmentation and ones produced manually. As can be seen by the numbers in these tables, we are usually within 2-3% of manual segmentations, which is considered acceptable performance in the applications with which we are dealing [23].

Although we did not test the distribution of false positives and false negatives in a detailed manner, we observed from sampling of the data that these errors appeared to be randomly distributed, rather than aggregated in particular places.

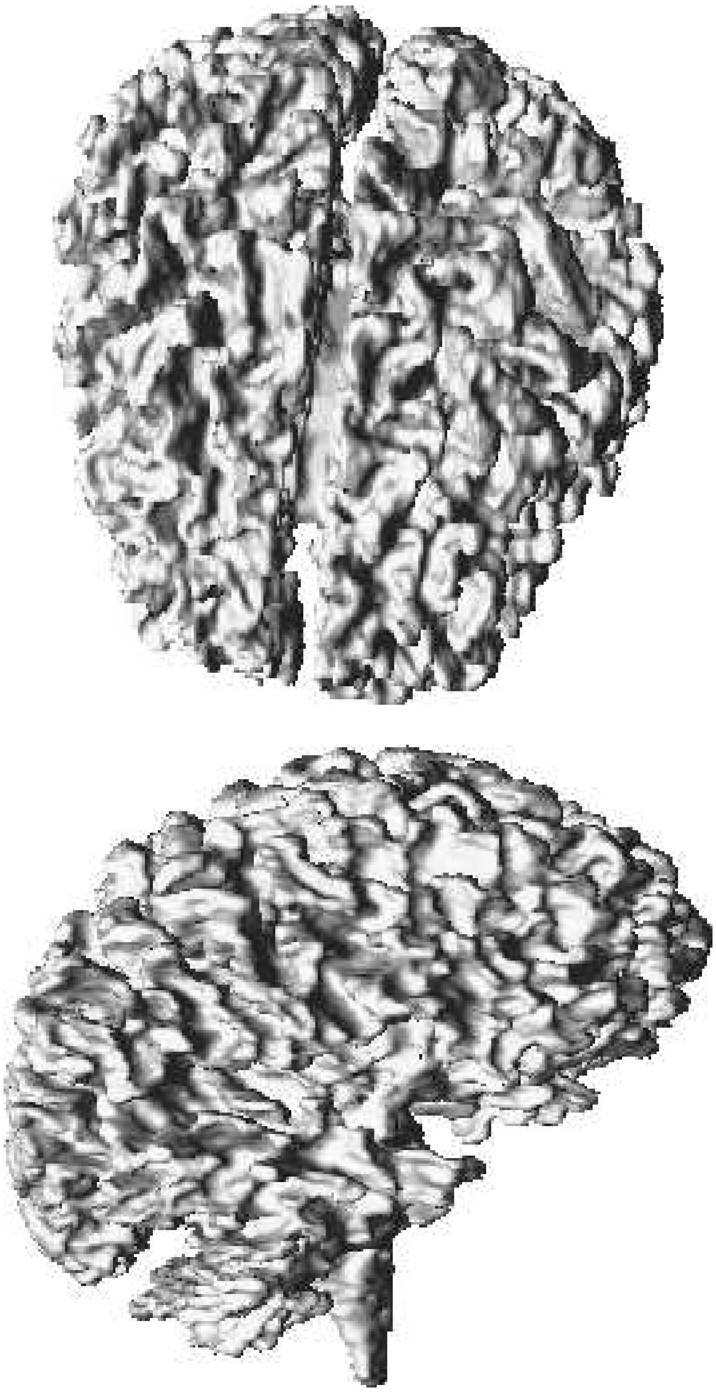

Figure 9: Top: Top view of a reconstructed white matter surface. Bottom: Side view of the same reconstructed white matter surface.

Case No.	Number of Different Voxels	Percentage Difference
1	120463	1.72
2	129648	1.85
3	153631	2.19
4	124335	1.77
5	150297	2.14
6	133743	1.91
7	192502	2.75
8	221291	3.16
9	188528	2.69
10	171776	2.45
11	176724	2.52
12	144990	2.07
13	136753	1.95
14	184274	2.63
15	142237	2.03
16	258752	3.69
17	194908	2.78
18	197774	2.82
19	101356	1.45
20	103202	1.19

Table 1: Difference in classification between our method and manual segmentation. The middle column shows the number of pixels that were classified differently by our 3-class (white matter, grey matter, csf) segmentation method as compared with manual segmentation for each of 20 cases of size 256x256x124 voxels.

Case No.	Mean d	% at d=0	% at d=1	% at d=2	% at d>2
1	1.10	19.84	69.05	7.65	3.47
2	1.23	8.19	73.49	12.67	5.65
3	0.84	31.49	61.27	5.10	2.13
4	1.07	20.91	66.45	7.29	5.35
5	1.11	11.00	75.19	9.27	4.54
6	0.85	28.20	65.20	4.82	1.78
7	0.76	38.58	55.20	4.42	1.80
8	0.81	34.74	57.45	5.04	2.78
9	0.52	57.18	39.12	2.29	1.41
10	1.21	22.34	63.32	6.78	7.56
11	0.77	42.97	47.00	5.14	4.89
12	2.22	15.27	61.86	8.13	14.74
13	1.67	17.96	58.14	8.43	15.47
14	1.36	13.12	68.58	8.84	9.47
15	1.59	13.63	63.66	10.33	12.38
16	6.44	37.79	31.49	4.58	26.14
17	1.29	42.25	40.81	5.32	11.63
18	1.26	41.16	40.10	5.94	12.80
19	1.15	16.53	71.84	8.29	3.35
20	1.54	30.51	64.55	2.19	2.04

Table 2: Measuring false positives: The second column shows the mean distance d (in pixels) between a pixel that falls on the boundary of the brain in the manual segmentation and the nearest brain-boundary pixel in our segmentation. The rest of the columns show the percentage of brain-boundary pixels in the manual segmentation at different distances from the nearest brain-boundary pixels in our segmentation.

Case No.	Mean d	% at d=0	% at d=1	% at d=2	% at d>2
1	1.20	10.37	72.57	12.39	4.67
2	1.34	21.41	69.02	4.46	5.11
3	1.51	7.44	66.06	15.97	10.54
4	1.69	10.97	72.20	11.81	5.03
5	1.08	12.49	78.51	6.56	2.43
6	1.46	6.46	69.14	16.03	8.37
7	1.44	6.67	65.49	18.50	9.34
8	1.32	8.34	67.57	19.73	4.36
9	2.09	5.33	48.32	28.47	17.88
10	1.10	14.91	69.71	11.58	3.80
11	1.34	8.76	63.61	22.36	5.27
12	1.25	14.82	71.71	9.71	3.76
13	1.17	14.21	70.97	10.74	4.07
14	1.17	13.08	75.40	8.00	3.52
15	1.00	17.91	72.09	7.48	2.53
16	1.48	9.00	55.81	26.81	8.39
17	2.26	8.87	56.99	24.13	10.02
18	1.43	11.28	56.98	25.07	6.68
19	1.42	15.12	68.81	10.04	6.03

Table 3: Measuring false negatives: The second column shows the mean distance d (in pixels) between a pixel that falls on the boundary of the brain in our segmentation and the nearest brain-boundary pixel in the manual segmentation. The rest of the columns show the percentage of brain-boundary pixels in our segmentation at different distances from the nearest brain-boundary pixels in the manual segmentation.

3.2 Semi-automated segmentation tools

Semi-automatic segmentation methods are used in cases where the automated segmentation techniques are inaccurate or where tissues to be segmented were not modeled for the automatic methods. One of the sources of inaccuracies in the automatic segmentation methods is the selection of the morphology parameters. In about 10% of the scans we have segmented, due to the variation in the size of the connecting elements from the brain tissue to the cranium, the empirically determined radius of the erosion kernel does not adequately model the width of the connectors between the brain and non-brain structures, and therefore the brain tissue is not isolated at the end of the segmentation. Such scenarios are currently detected by manual inspection, and are corrected by the use of deformable models, such as snakes [42], customized to this application.

For segmenting tissues which are not modeled for the automatic methods, such as tumors, an interactive segmentor is used by clinicians. This tool, developed by the Surgical Planning Lab at Brigham & Women's Hospital, uses real-time ray tracing running on a Thinking Machines Inc. CM-2 to render segmented tissues as parameter values are dynamically varied. The parameters of interest are the intensity thresholds used to define intensity-based tissue classifiers. These thresholds are interactively modified and both 2D slices and 3D renderings of the resultant classifiers are viewed in real time. If necessary, further structural correction of the segmentation is performed by interactively editing the 2D slices. Such editing consists of highlighting regions of interest with the mouse and specifying the desired label for the contained pixels. Such structural correction may be necessary due to overlaps in intensity values across tissue classes or partial voluming artifacts.

Figure 10 shows a case where, as part of the surgical planning, the automated segmented grey matter of the patient was overlaid with manually segmented pathology.

3.3 Combining segmentations from multiple sensing modalities

We often combine segmented tissues obtained from different modalities into a single segmented tissue map in order to generate a complete structural segmentation of a patient. For example, the detailed vasculature from high resolution MR angiograms is combined with soft tissue segmentation from standard MR images, bone structures from CT is combined with the soft tissue labeling from MRI, or functional data from PET or SPECT is combined with structural information from MRI. In order to perform the multi-modality fusion we register the different images to a single coordinate frame. The registration is performed using either:

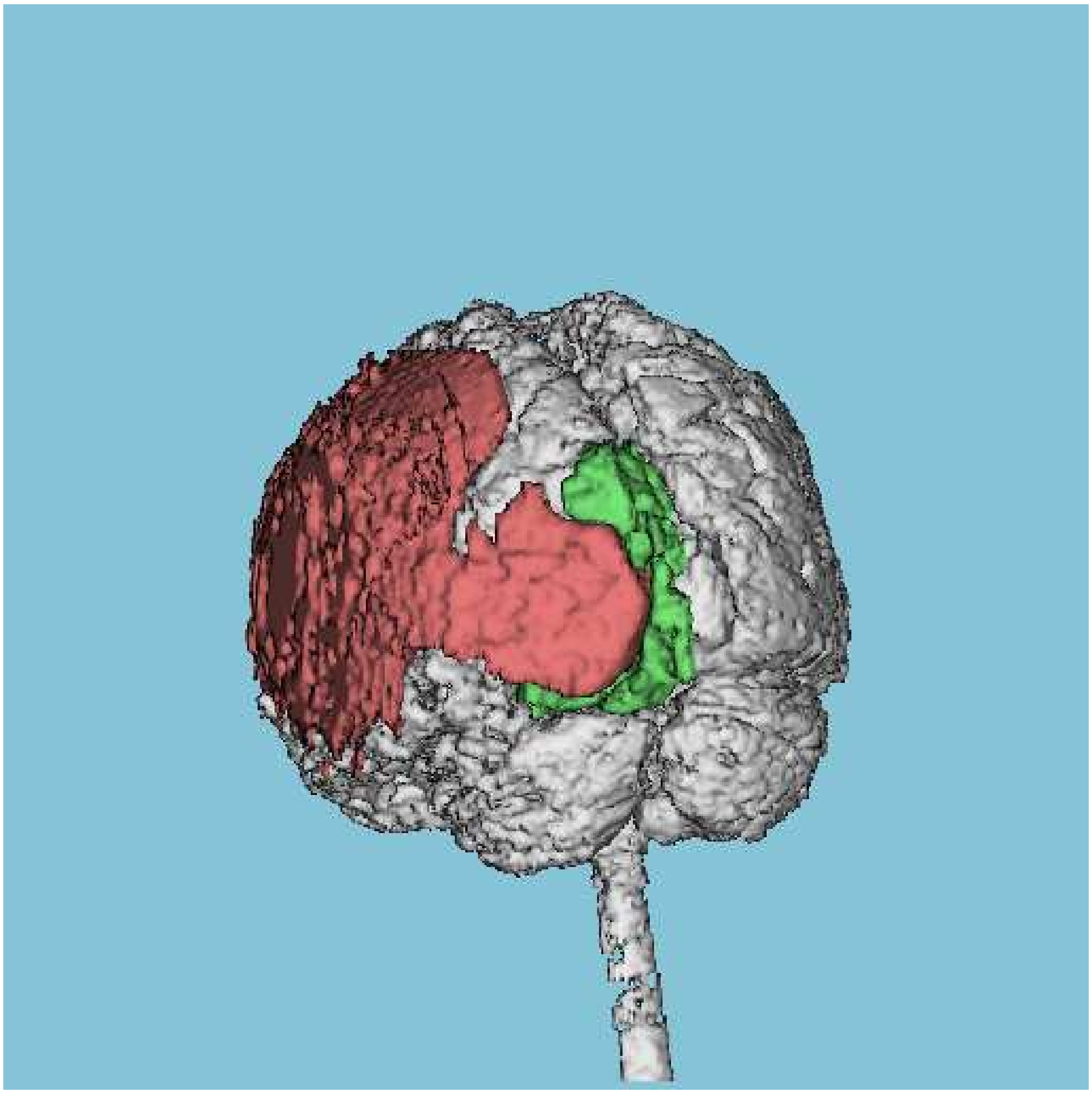

Figure 10: Side view of a reconstructed brain surface of a patient overlaid with manually segmented pathology – subdural haematoma (red) and tumor (green).

- specific 3D surfaces which are expected to remain relatively rigid across scanners, such as scalp or intra-cranial cavity, or
- the whole 3D volume.

Surface registration is performed using the techniques described in Section 5. Volume registration is performed by using an information-theoretic approach which maximizes a mutual information metric [41]. Mutual information, $I(u(x), v(T(x)))$, of a reference data set $u(x)$ and a test data set $v(T(x))$ undergoing a transformation T, is defined as:

$$h(u(x)) + h(v(T(x))) - h(u(x), v(T(x))).$$

$h(x)$, the entropy of random variable x, is defined as:

$$-\int p(x) ln(p(x)) dx.$$

This mutual information metric encourages solutions for T in which u explains v well, while at the same time maximize the complexity of the matched portions of the two data sets. In order to search for the transformation which maximizes I, optimization techniques such as gradient descent are used. Since this technique is based solely on the intensities in the images rather than particular features and does not require an *a priori* model of the relationships between the intensities of the different images, it has proven to be a general approach applicable to a wide range of multi-modality fusion applications (in addition to [41], see also [7, 27]).

The use of these techniques is illustrated in the following surgical example. The patient had a skull base meningioma which consisted of both intra- and extra-cranial parts. The 3D models were made from CT images and two sequences of MR images (see Figure **??**). In this 3D model, the skull was constructed from CT images, the tumor and the optic nerves from gradient echo MR images and the vascular tree from MR angiograms. The tumor is seen in extra-cranial and intra-orbital areas as well as in the middle cranial fossa. A top view of the anatomical structures appears in Figure **??**.

4 Graphical display of segmented MRI

At this stage, we have extracted a set of 3D models, represented both by surface and volume data, in a coordinate frame attached to the MR scanner. For image-guided surgery we leverage the segmented MRI data by providing the surgeon with spatial feedback on locations of key

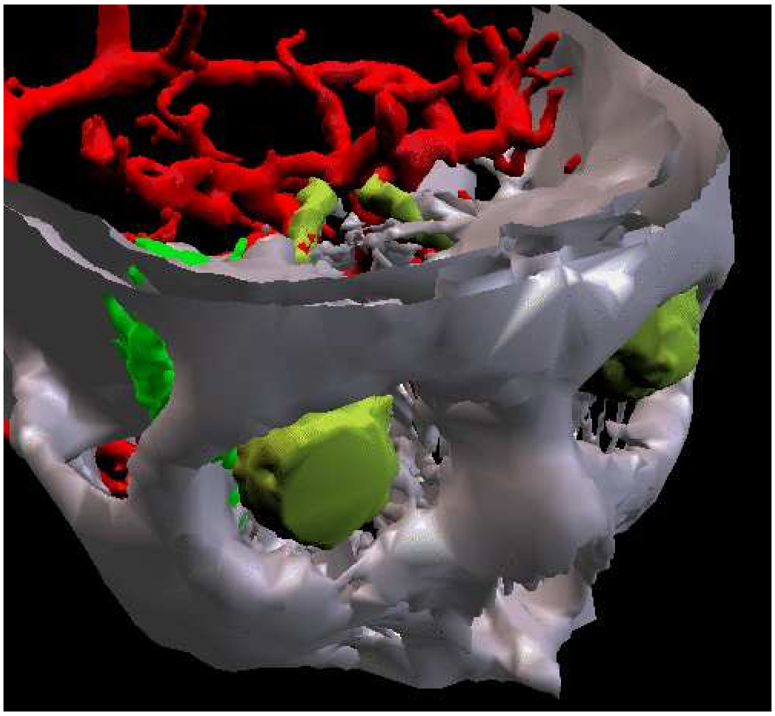

Figure 11: Right frontal view of the 3D model. The skull (colored white) is derived from CT images, while the vascular tree (red) was derived from an MR angiogram. Models of the tumor (green) and optic nerve with the ocular bulb (yellow) were derived from an gradient echo MR sequence.

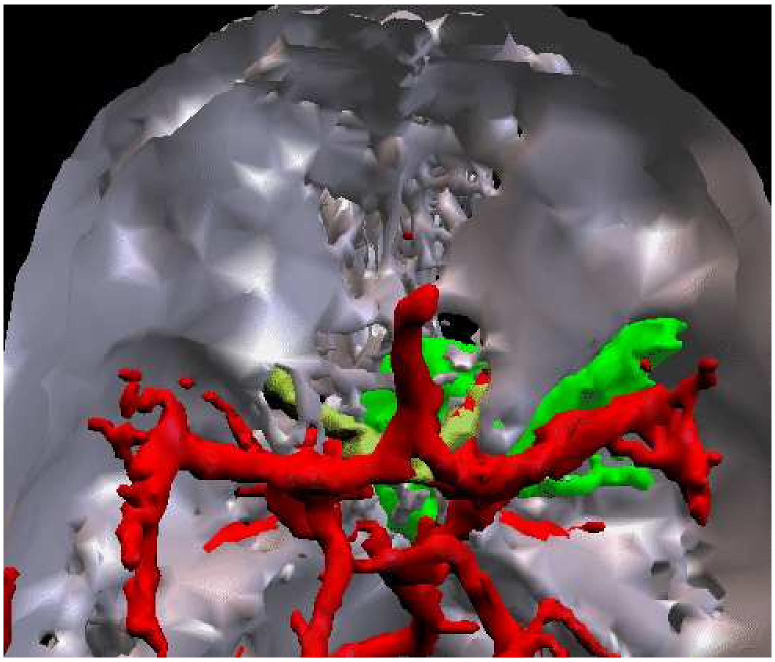

Figure 12: Top view of the 3D models for case two, showing the spatial relationship of the circle of Willis (a part of the intracranial vascular tree), the optic nerves and the bony structure. The tumor (colored green) is in the middle cranial fossa and in the sphenoid sinus (one of the paranasal sinuses). This spatial information was found useful during surgical planning.

anatomical structures. Such feedback is accomplished by overlaying the anatomical structures in the surgeon's field of view (see Section 5) or by pinpointing the location of medical instruments relative to the segmented structures (see Section 6). In both cases we need to render the 3D segmented MRI structures on a display viewable to the surgeon. Such displays may be computer monitors, as we are currently using, or line-of-sight displays such as head-mounted goggles, transparent panels, or surgical microscopes.

We use surface rendering techniques to display the segmented MRI structures. This procedure consists of first extracting bounding surfaces from the segmented MRI volume using the marching cubes algorithm [25]. This algorithm generates a set of connected triangles to represent the 3D surface for each segmented structure. These surfaces are then displayed by selecting a virtual viewing camera location and orientation in the MRI coordinate frame and using standard computer graphics techniques to project the surface onto the viewing camera. This rendering process removes hidden portions of the surface, shades the surface according to its local normal, and optionally varies the surface opacity to allow glimpses into internal structures. A sample rendering is shown in Figure 13.

5 Registering MRI scans to views of a patient

The capability to render 3D anatomical structures provides a useful tool for the surgeon to visualize relationships between the structures. These visualizations, though, are in the MRI coordinate frame which is generally not the same reference frame that the surgeon has of the patient on the operating table. In order to provide direct guidance to the surgeon, the MRI visualization must be performed in the patient's reference frame–i.e., the MRI rendering must be performed such that the surgeon's view of the MRI structures is the same as the surgeon's view of the actual patient. Our goal is to overlay the MRI data onto an actual camera's view of the patient such that the MRI data is correctly aligned to the patient. For this purpose, the camera viewing the patient is positioned such that it is looking over the surgeon's shoulder.

In order to perform this alignment process we register skin surface data from the MRI scan to positional information for the patient's skin surface on the operating table. For neurosurgery, and other head surgeries, we use skin surfaces of the rigid scalp of the patient. The positional data of the patient is acquired with a laser striping device which achieves high positional accuracy ($<$ 1 mm) while avoiding direct contact with the patient. We then apply a 3D surface registration algorithm to match the laser data to the MRI data, as depicted in Figure 14 and described below. Further details are in [14]. Related surface-based registration approaches include [9, 15,

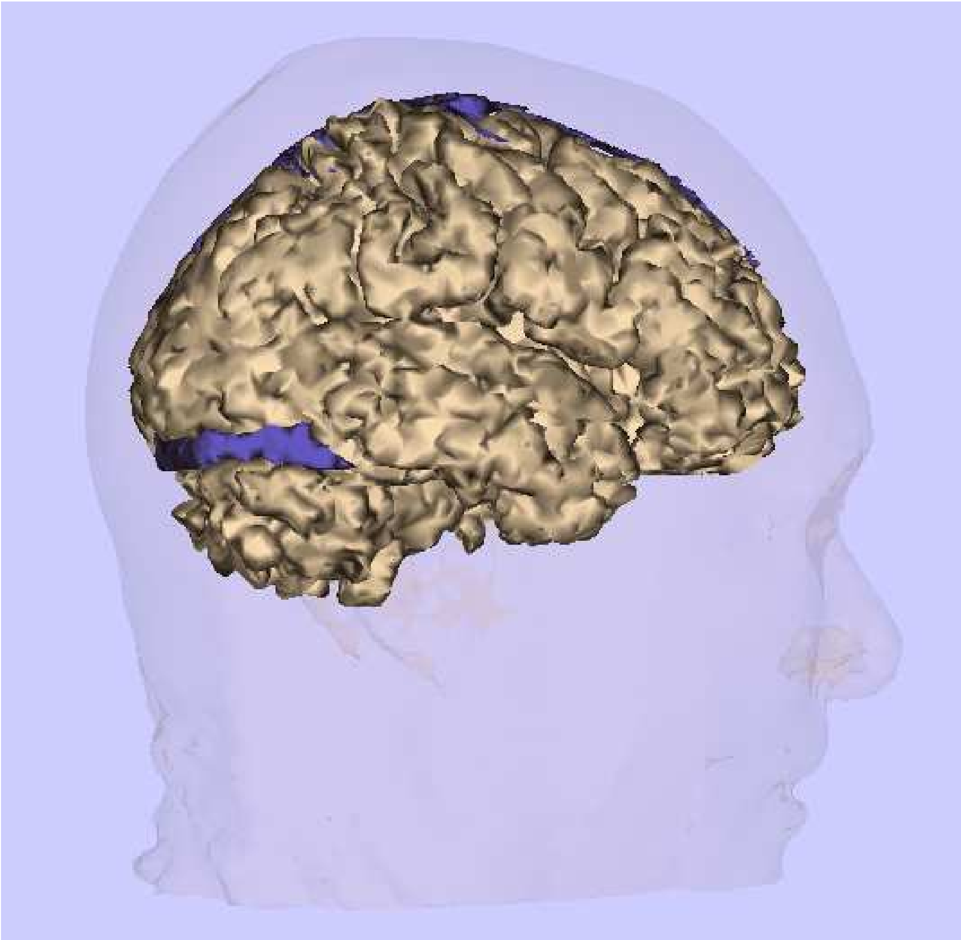

Figure 13: Sample MRI rendering of segmented brain and vessels along with transparent skin.

29, 35, 36, 43].

5.1 Initial Match

Before commencing the matching process we preprocess the laser data to separate data of the subject's head from background data. Currently we do this with a simple user interface in which the view obtained from the video camera used in conjunction with the laser is displayed, with the laser data overlaid two dimensionally on top of that view. The user can thus use a simple mouse interface to highlight laser points coming from the skin of the subject. Note that this process need not be perfect–the matching process is designed to deal robustly with outliers.

To initiate the matching, we have two options: interactive alignment and coarse constrained search. Under interactive alignment we use a graphical interface to manipulate the MRI and laser data. We use two displays to view our alignment. In one display we view the 3D MRI skin surface superimposed on the laser camera's video image of the subject. This display supports 3D translation and rotation of the MRI head data to achieve close alignment with the video image of the head. The second display consists of three orthogonal projections of sampled points from the MRI skin surface overlaid with the projection of the laser points. To facilitate interpretation of sidedness within each projection, the intensity of the points is made proportional to the normal depth within each projection. In this display we interactively translate and rotate the laser points such that they roughly align with the underlying MRI points. This initial alignment does not need to be very accurate: rotational errors on the order of 20 degrees, and translational errors on the order of centimeters are permissible, since the subsequent matching stage is quite reliable at removing these misalignments.

If we do not want to rely on operator intervention, we can instead use an automated constrained search procedure to find the initial alignment. In this method, we first select a pair of widely separated laser points, and at each point, we estimate the surface normal, by a local least squares fit. We then search over all possible pairs of MRI surface points, at some subsampling, and use those pairs whose distance roughly agrees with the distance between the chosen laser points. For each pair of MRI points and the pair of laser points, we solve for the transformation that aligns the points and the surface normals at the two points, provided such a transformation exists. For those pairings of points with legal transformations, we apply the transformation to all of the laser points, and measure the least squares distance between the transformed laser data and the MRI skin model. We use this measure to rank order the possible alignments. For efficiency purposes, we hash the pairings of MRI points by distance, in a coarse-to-fine manner,

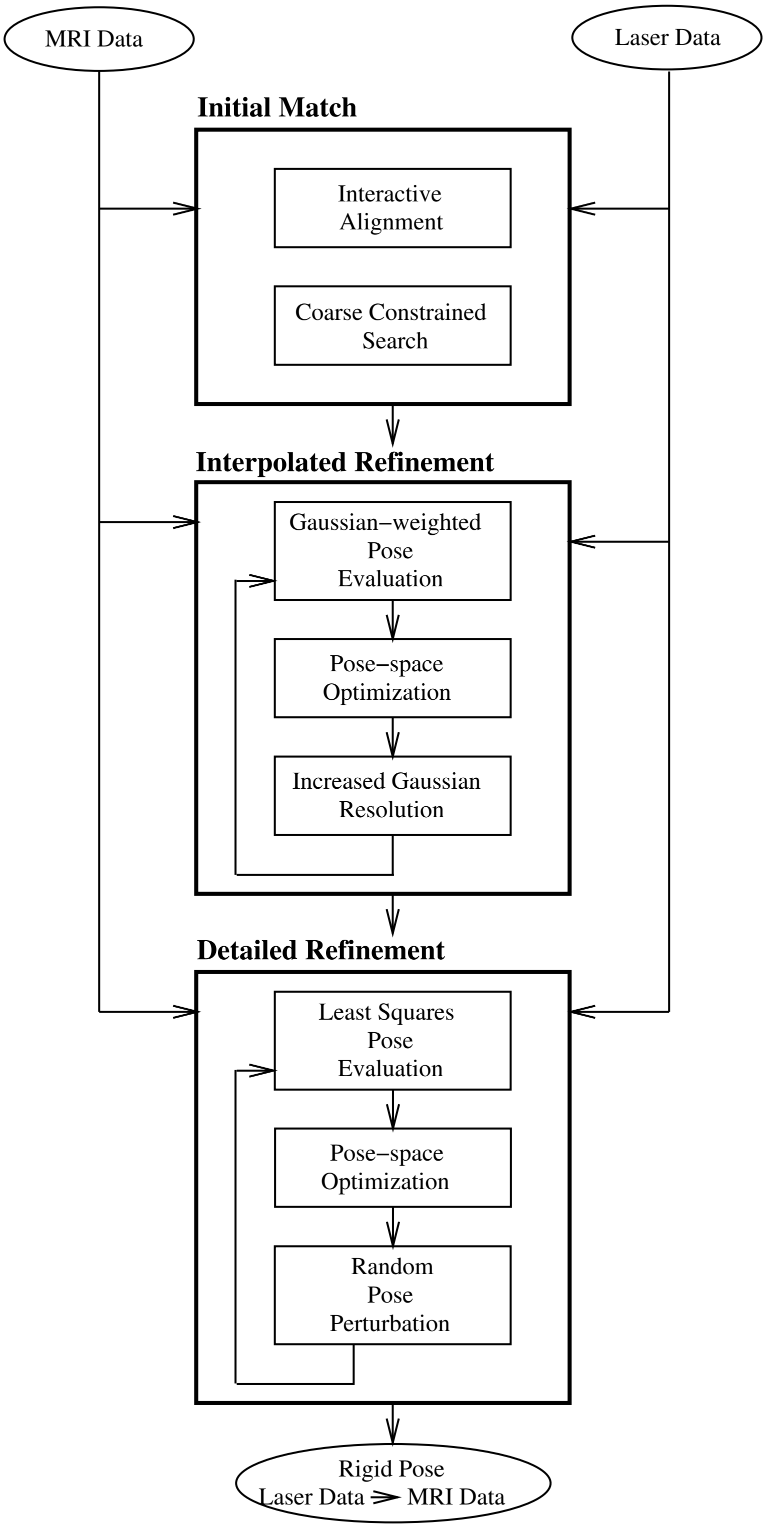

Figure 14: Registration algorithm outline.

thereby saving considerably on the computation required. We keep the n best transformations for use at the next stage. Note that we can accomplish the same constrained search by using triples of points without normal information, if desired.

For each selected alignment transformation, we execute the refinement processes described below.

5.2 Interpolated Refinement

We first refine the alignment of the two data sets by minimizing an evaluation function that measures the amount of mismatch between the two data sets. In particular, we sum, for all transformed laser points, a term that is a sum of the distances from the transformed laser point to all nearby MRI points, where the distance is weighted by a Gaussian distribution [38]. This Gaussian weighting roughly interpolates between the sampled MRI points to estimate the nearest point on the underlying surface to the transformed laser point. More precisely, if vector ℓ_i is a laser point, vector m_j is an MRI point, and $\mathcal{T}$ is a coordinate frame transformation, then the evaluation function for a particular transformation is

$$E_1(\mathcal{T}) = -\sum_i \sum_j e^{-\frac{|\mathcal{T}\ell_i - m_j|^2}{2\sigma^2}}. \tag{11}$$

Because of its formulation, the objective function is quite smooth, and thus facilitates "pulling in" solutions from moderately removed locations in parameter space.

In order to minimize this evaluation function we use the Davidon-Fletcher-Powell (DFP) quasi-Newton method [30]. This method requires an estimate of the gradient of the objective function, which is easily obtained in closed form. Solving this minimization problem yields an estimate for the pose of the laser points in MRI coordinates.

We execute this minimization stage with a multiresolution set of Gaussians. A broad Gaussian is used to allow influence over large areas, resulting in a coarse initial alignment, which can be reached from a wide range of starting positions. Then, narrower Gaussian distributions are used to focus on only nearby MRI points to derive the pose.

5.3 Detailed Refinement

Starting from the pose obtained with the interpolated refinement stage, we repeat the evaluation process, using a rectified least squares distance measure. We again use the DFP method to minimize the evaluation function:

$$E_2(\mathcal{T}) = \left[\frac{1}{n}\sum_i \min\left[d^2_{\max}, \min_j |\mathcal{T}\ell_i - m_j|^2\right]\right]^{\frac{1}{2}} \tag{12}$$

where $d_{\max}$ is some preset maximum distance used to limit the impact of outliers. This objective function acts much like a robust chamfer matching scheme (e.g. [21]). The expectation is that this second objective function is more accurate locally, since it is composed of saturated quadratic forms.

We observe that while this refinement method gets very close to the best solution, it can get trapped into local minima in the minimization of E_2. To improve upon this problem, 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. The final result is a pose, and a measure of the residual deviation of the fit to the MRI surface.

Using the resultant pose we can now overlay the MRI data on the surgeon's view of the patient. An example of such a visualization is shown in Figure 15. We can provide such registered visualizations prior to initiation of the surgical procedure, thus enabling the surgeon to plan the position of the skin flap and the craniotomy. In fact, since we provide a video mix of the registered internal structures with a live view of the patient, the surgeon can interactively trace structures on the scalp of the patient, by viewing his hand in this live video mix. In this way, the system interactively provides surgical planning support to the surgeon.

Similarly, during the surgical procedure itself, this registered visualization allows the surgeon to view nearby structures as well as the position of the tumor relative to his current working area, thus providing navigation guidance to the surgeon.

Initial computational experiments on the registration method [13] show that RMS errors of the registered points are on the order of the resolution of the MR images–generally 1.5 mm. Susceptibility to local minima has been small as it consistently finds the same solution over a wide range of starting positions–viewed points are within one or two voxels of each other with different starting positions. Finally, the region of convergence is upwards of $10°$ of rotational offset.

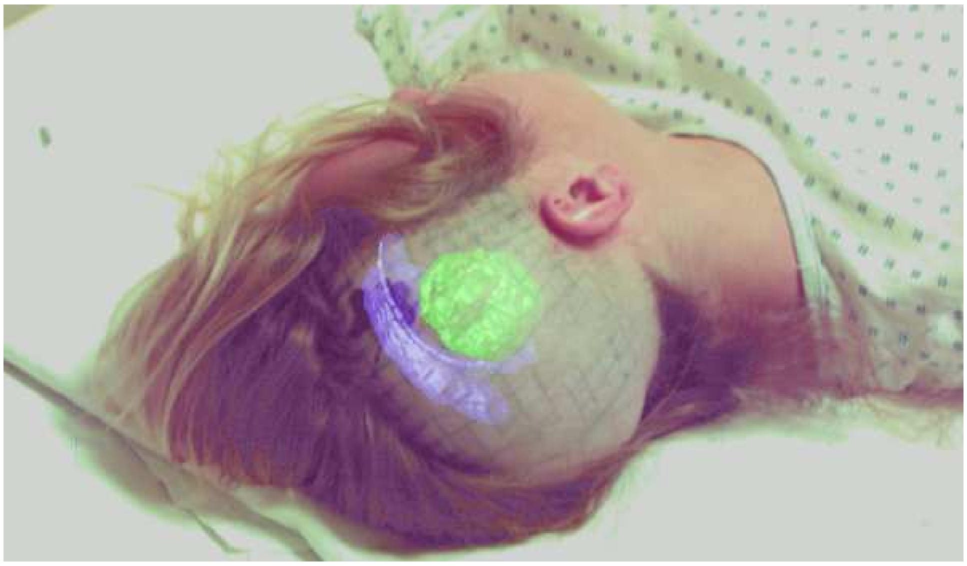

Figure 15: Sample enhanced reality visualization showing internal tumor and ventricle structures overlaid on video view of patient.

6 Registered tracking of surgical probes

Having a precise registration of a patient's anatomical structures to the current position of the actual patient clearly enables full visualizations of the internal anatomy. In addition, however, we want to allow the surgeon to interactively explore the visualization, that is, to be able to identify points in the current operating field and be able to see the corresponding points in the original 2D MRI slices and in 3D reconstructions. To do this, we need a way of tracking a surgical probe and relating its position to the registered anatomical reconstructions.

We have incorporated into our registration and visualization system a Flashpoint (IGT, Inc., Boulder, CO) 3D tracking system. This is a self-contained system which can be used to generate 3D coordinates of LEDs in the system's field of view. The system is based on a straightforward triangulation process, in which a point is observed in three orthogonal linear cameras, whose positions and orientations are known with respect to one another. By identifying the image projection of the same point in each camera, one can back out the projection geometry to determine the position of the point in scene coordinates. To achieve simple and robust identification of the same point in each image, infrared light emitting diodes (IR LEDs) are used, and the pulsing of each diode is synchronized to the imaging process. In this way, there is no possible ambiguity in identifying corresponding image points, and reliable estimation of 3D point positions of the LEDs is possible.

This active triangulation system is highly reliable, with an accuracy of about 1 mm at the 1 m standoff from the 3 linear cameras which we normally use. By mounting three LEDs to a rigid object we can track the object's pose (position and orientation) in three-space. Using two LEDs allow us to solve for five degrees of freedom (all but the twist angle around the axis connecting the two LEDs). Using more than three LEDs provides us with some redundancy, allowing for a least-squares pose solution and wide range of motion in which some LEDs may be blocked from view.

We use the Flashpoint system for tracking head motion (using five LEDs taped to the head) and localizing the position and orientation of the surgical probe (using at least two LEDs mounted to the probe). In order to track the head motion we record the position of the LEDs attached to the patient's head at the time we perform the laser data/MRI registration. This reference position provides a basis for tracking the head. As the surgical probe is positioned, we record the new position of the head-mounted LEDs and compute the transform necessary to return the head to its reference position. This transform is applied to the position/orientation of the surgical probe in order to then apply the laser data/MRI transformation described in Section 5. Since we may use more than three LEDs to track we use Horn's closed form least-squares solution based on quaternions [18] for the tracking transform.

In order to use the Flashpoint system for localizing the surgical probe we perform two calibration steps: calibrate the probe to identify the position of the end of the probe relative to the Flashpoint coordinates of the attached LEDs and calibrate the Flashpoint coordinate system to the laser coordinate system. The probe calibration is performed by measuring the distance from the LEDs to the probe tip. In order to calibrate the Flashpoint coordinate system to the laser coordinate system we use a Flashpoint probe to record points on a calibration gauge which have known laser coordinates. Given the correspondences between Flashpoint and laser points we solve for the transformation between the two coordinate systems. Since the three linear Flashpoint cameras and laser scanner (laser and camera) are all mounted on the same rigid bar, this calibration remains fixed.

6.1 Interoperative Pointer

One type of surgical probe that we have tracked is an interoperative surgical pointer. With this instrument, the surgeon can point at various places or structures in the brain and then see the position relative to other structures in the 3D model, as shown in Figure 16. For example, the pointer can be used to aid in the approach and localization of a tumor and as a tool to measure

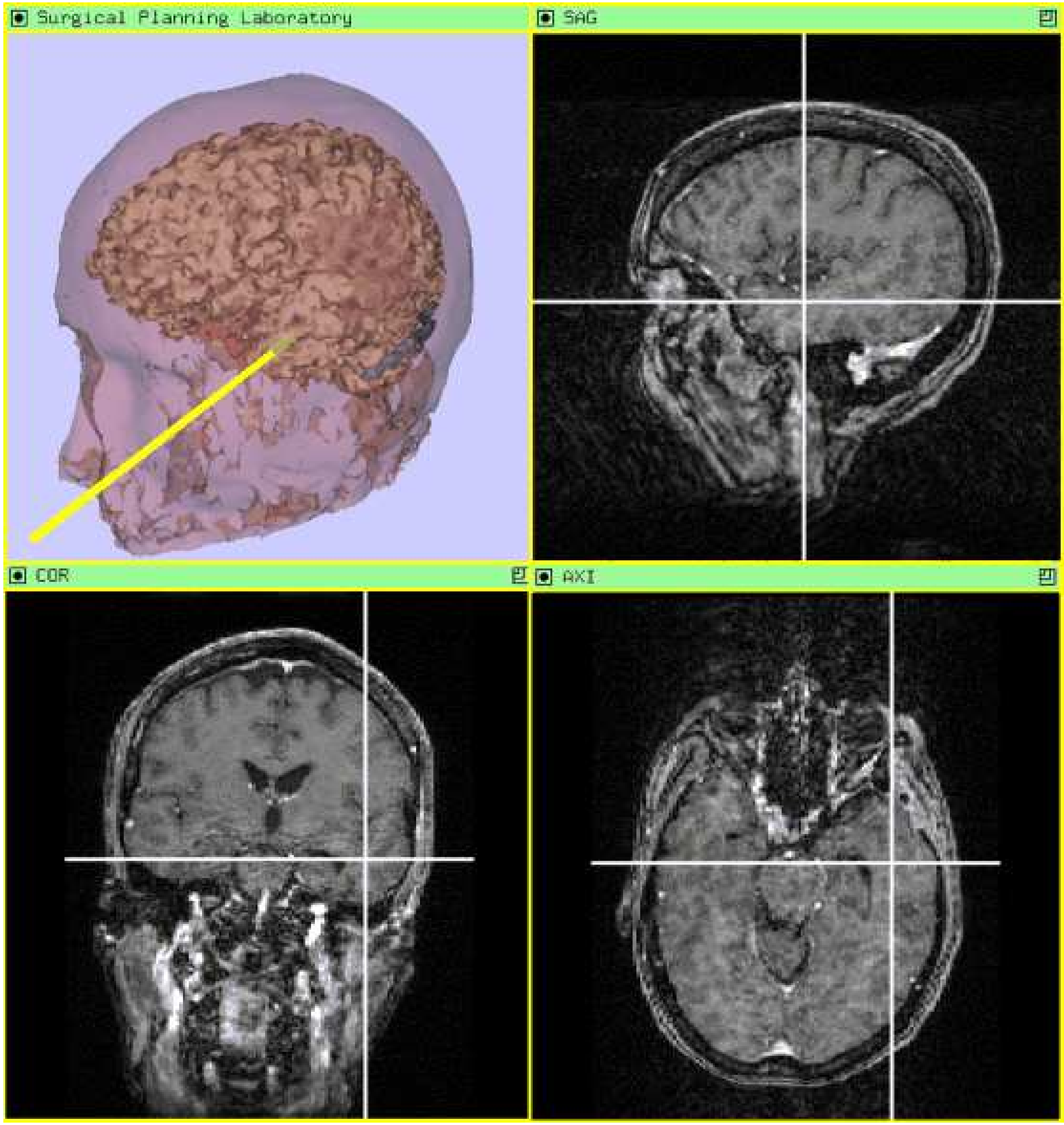

Figure 16: Display of the position of the interoperative pointer as a yellow arrow in the rendered 3D model and as cross-hairs in the three orthogonal slices of the MR scan.

the distance between two points to find a tumor's extent.

6.2 Transcranial Magnetic Stimulation

Another application of tracking a surgical probe is a method for mapping the functional regions of the brain using a transcranial magnetic stimulation (TMS) device. The device, a magnetic coil, when placed on a subject's scalp, stimulates the underlying neurons by generating focused

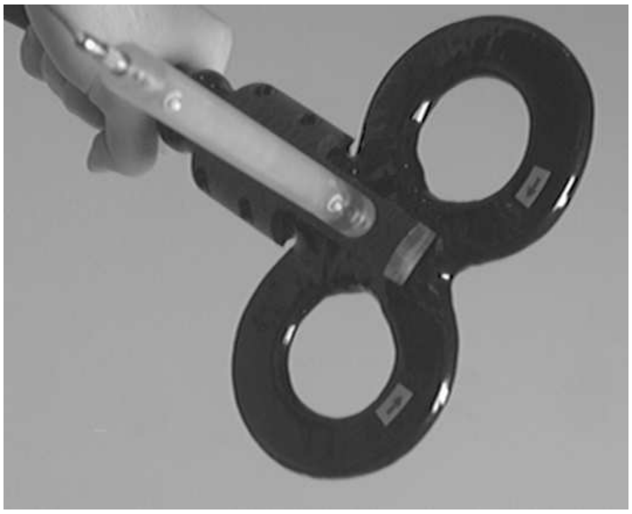

Figure 17: Stimulation coil used for TMS mapping. Note the rod mounted orthogonal to the plane of the coil–two LEDs are fixed on the rod for tracking the position and orientation of the coil.

magnetic field pulses. A brain mapping is then generated by measuring responses of different motor and sensory functions to this stimulation (in a typical example, we record electrical muscle reaction in a sequence of muscle groups). The magnetic coil is tracked by placing two LEDs on a rod mounted to the coil, as shown in Figure 17. Each time the probe is stimulated, the position and orientation of the TMS coil is measured, and the positions of the LEDs placed on the patient's head are obtained.

We combine the registration and tracking data to obtain the functional brain mapping using the following transforms:

- $\mathbf{F_L}$ — transformation from Flashpoint coordinates to laser coordinates; computed from *a priori* calibration.

- $\mathbf{L_M}$ — transformation from laser coordinates to MRI coordinates; computed from dynamic registration procedure.

- $\mathbf{H_r^t}$ — transformation of head from time t to reference position at time 1, computed from calibration of the LEDs at time t to the LEDs at time 1.

We have also collected the following TMS data:

- C_p^t, C_o^t — position and orientation of TMS coil at time t, $t \in [1, T]$, in Flashpoint coordinates.

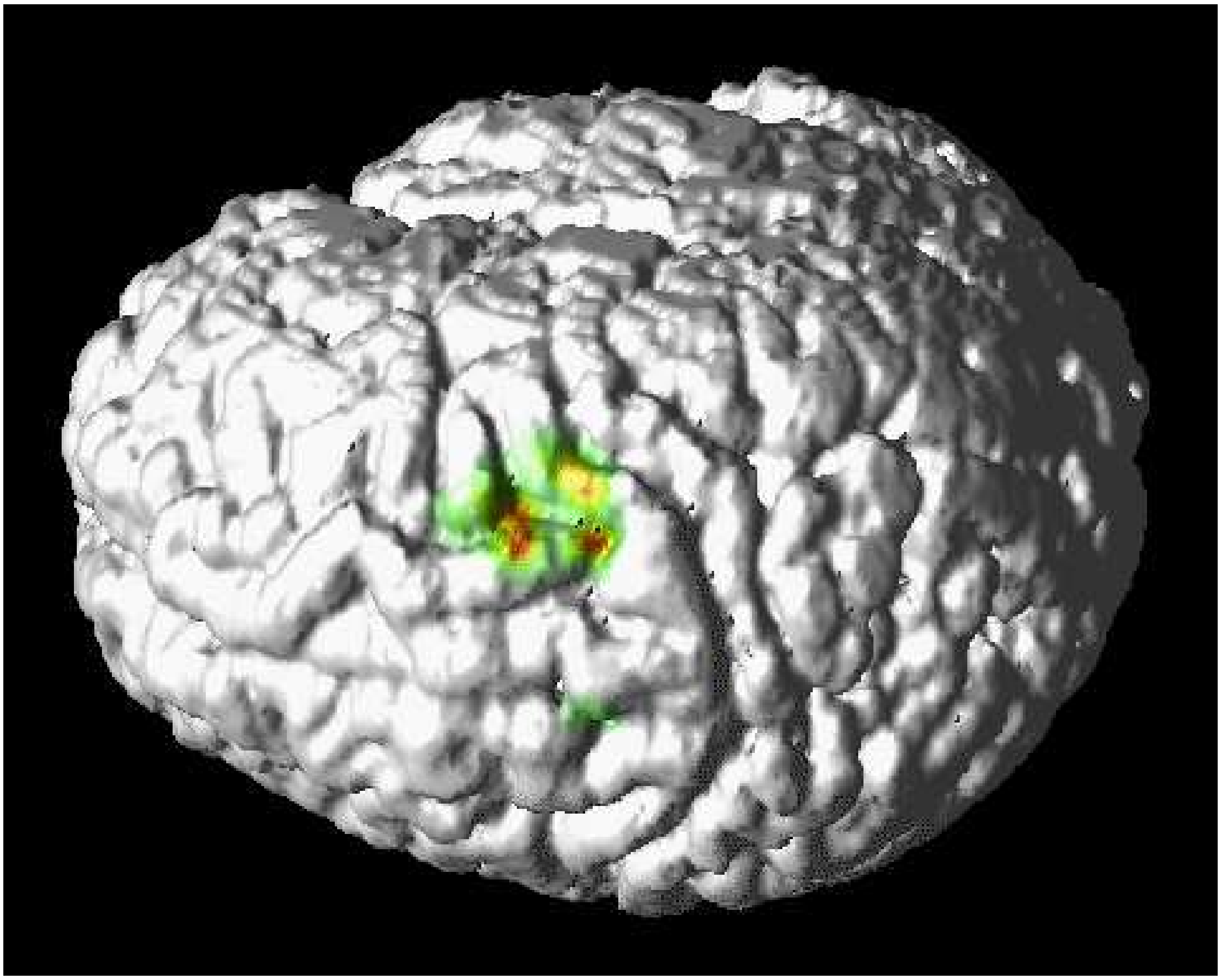

Figure 18: Motor map of the positions and intensities of the responses of the bicep muscle. The color red indicates the highest responses.

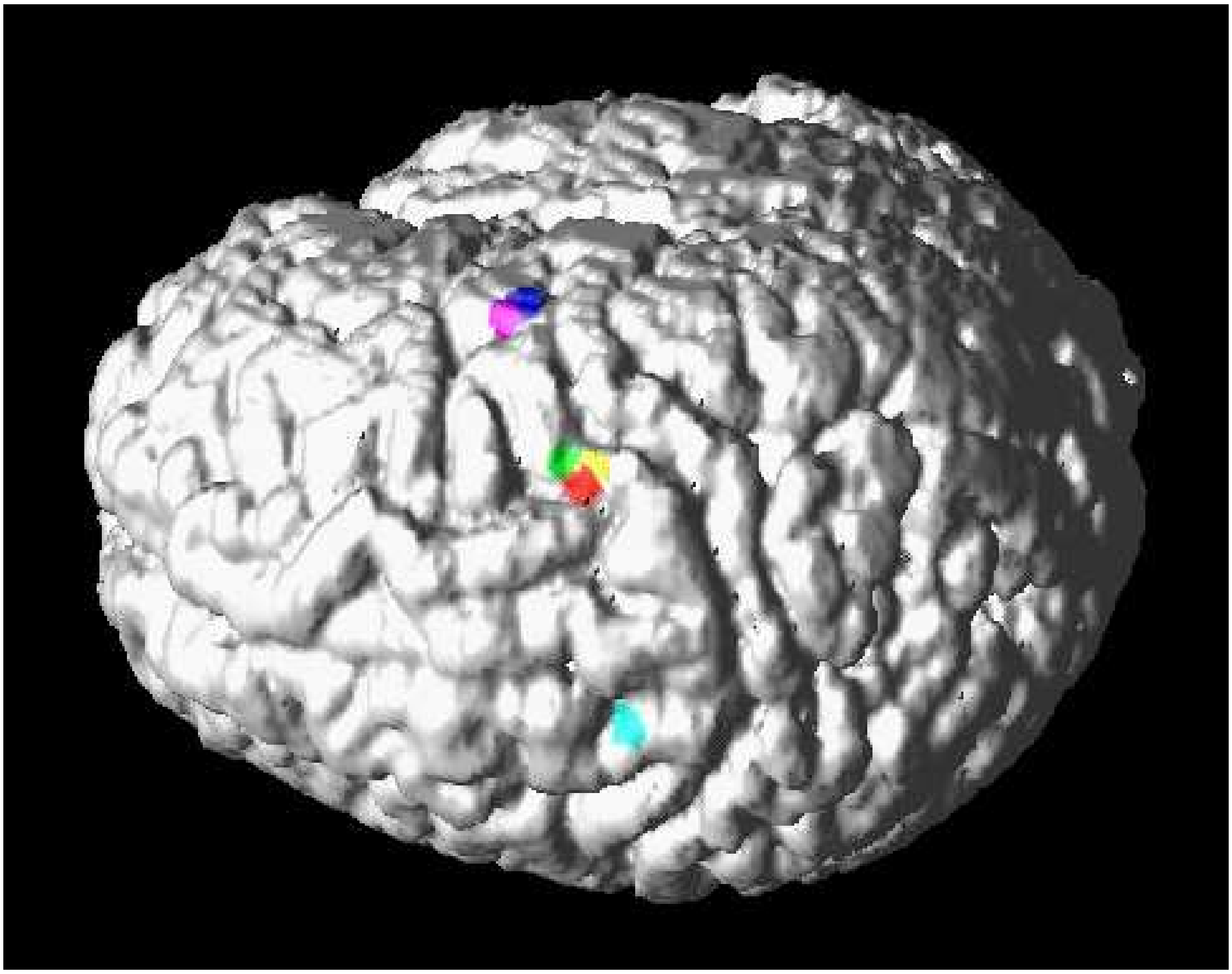

Figure 19: This figure shows the lowest latency responses for each muscle that was mapped. The blue and magenta stimulation points correspond to the left facial and right facial muscles, respectively. Green, yellow, and red correspond to the left biceps, left forearm, and left hand respectively, and cyan indicates the left leg response.

- R^t_j — measured response j to stimulation t. Multiple responses are usually collected such as from several different hand, arm, and leg muscles.

In order to compute the brain mapping we need to map the TMS responses to the brain surface using the measured coil positions/orientations and associated transformations. To perform this mapping, for each stimulation t, we process those MRI surface points, $S[i]$, that are sufficiently close to $\mathbf{L_M F_L H^t_r} C^t_p$ to have been possibly stimulated by the pulse. Note that closeness here can be determined in a number of ways. We can simply use the transformed position of the tip of the probe and gather all MRI points within some predefined distance. Alternatively, if we have a detailed model of the shape of the magnetic field generated by the probe, we can use it, together with information about the orientation of the probe, to select the relevant MRI points. For now, we are simply using Euclidean distance from the position of the probe to select $S[i]$.

For each such $S[i]$ we compute the distance, $d^t[i]$, to the line defined by the point $\mathbf{L_M F_L H^t_r} C^t_p$ and the orientation $\mathbf{L_M F_L H^t_r} C^t_o$. We are currently using a Gaussian weighting function proportional to that distance to "spread" the response R^t_j to the points $S[i]$. The purpose of this (simple) weighting function is to interpolate across the stimulation to obtain a smooth and visible map. If we let $map^t_j[i]$ represent the mapping of response j to stimulation t on the selected surface, then $map^t_j[i] = G(d^t[i], \sigma) R^t_j$, with G being the Gaussian weighting function. We then let $map_j[i]$, the composite mapping from all stimulations, be the maximum $map^t_j[i]$ over all t, which are then normalized over i.

While we generally are most interested in the mapping of the TMS responses to the cortical surface of the brain, we may perform this mapping onto any underlying segmented surface. The white matter surface is often useful since it highlights some of the major sulci. The skin surface itself is also sometimes useful to examine coverage range of stimulations. The process generates a high resolution, accurate brain mapping which supports surgical planning, surgical guidance, neuroanatomy research, and psychiatric therapy. Figure 18 shows an example of a motor map of a volunteer's bicep muscle. Figure 19 shows the position of the minimum latency response for all muscles that were mapped.

6.3 Alternative tracking methods

We are also currently investigating the possibility of using a more passive system to track both head position and probe position and orientation. This method [28] utilizes some simple visual markers placed on the objects of interest, which are then tracked reliably and rapidly by observation in a single video camera. The advantage of this system is that the passive markers are less intrusive than the LEDs, and that tracking can in principle be done by any camera in any location, rather than relying on the Flashpoint cameras.

7 Experience using the full system

Neurosurgeons at Brigham & Women's Hospital have used our image-guided surgery system for both *a priori* surgical planning and dynamic operating room guidance. Surgical planning applications include localizing craniotomy location in the context of the relationship of the pathology to blood vessels or critical functional brain regions. Our enhanced reality visualization technique has enabled such plans to be directly transferred to the patient by overlaying the internal anatomical structures on views of the patient. In the operating room our system guides the surgeon by providing internal structure visualization from the surgeon's viewpoint as well as by tracking medical instruments relative to the segmented anatomical structures. By displaying the location of the medical instrument on a rendering of the labeled anatomy of interest, the surgeon not only views nearby hidden structures, but also gains better insight as to the identity of visible structures.

Overhead of using the image-guided surgery system in the operating room is about 5-10 minutes for performing the registration and attaching the tracking markers. But this system is replacing a manual registration system, in which video-based alignment was achieved by manually adjusting viewing and rendering parameters, which often took about 45 minutes to perform. In addition, the accuracy of the manual registration method was only estimated to be about 10-30 mm. The order of magnitude improvement of our system in both efficiency and accuracy is a critical factor in the success of this technology. Feedback from the surgeons has been highly positive–specifically, this easily accessible from of 3D geometric knowledge provides critical knowledge for planning and guiding surgery. This technology will be an important factor in accelerating the trend towards minimally-invasive surgeries.

Image-guided surgery is a quickly growing field–other researchers are developing related capabilities. Of note is the VISLAN system [6] which is also designed to be an end-to-end image-guided surgery system with many automated procedures. Other related image-guidance capabilities include [3, 4, 10, 16, 31, 37].

References

[1] L. Axel, J. Costantini, and J. Listerud. Intensity Correction in Surface-Coil MR Imaging. *AJR*, **148**(4):418–420, 1987.

[2] N. Ayache. "Medical Computer Vision, Virtual Reality and Robotics". *Image and Vision Computing*, **13**:295–313, 1995.

[3] G. Barnett, D. Kormos, C. Steiner. "Frameless Stereotaxy Using a Sonic Digitizing Wand: Development and Adaptation to the Picker ViStar Medical Imaging System". R.J. Maciunas, ed., *Interactive Image-Guided Neurosurgery*, American Association of Neurological Surgeons, 1993.

[4] D.R. Bucholtz, K.R. Smith. "A Comparison of Sonic Digitizers Versus Light Emitting Diode-Based Localization". R.J. Maciunas, ed., *Interactive Image-Guided Neurosurgery*, American Association of Neurological Surgeons, 1993.

[5] H. Cline, W. Lorensen, R. Kikinis, F. Jolesz, "3D Segmentation of MR Images of the Head Using Probability and Connectivity." *JCAT* **14**(6):1037–1045, 1990.

[6] A.C.F. Colchester, J. Zhao, K.S. Holton-Tainter, C.J. Henri, N. Maitland, P.T.E. Roberts, C.G. Harris, R.J. Evans, "Development and Preliminary Evaluation of VISLAN, a Surgical Planning and Guidance System Using Intra-Operative Video Imaging". *Medical Image Analysis*, **1**(1):73–90, 1996.

[7] A. Collignon, D. Verndermeulen, P. Suetens, and G. Marchal. "3D multi-modality medical image registration using feature space clustering". In *Proceedings of the First Conference on Computer Vision, Virtual Reality and Robotics in Medicine.* Springer, 1995.

[8] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *J. Roy. Statist. Soc.*, **39**:1 – 38, 1977.

[9] J. Feldmar, N. Ayache. "Locally Affine Registration of Free-Form Surfaces". *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Seattle WA, June 1994.

[10] R.L. Galloway. "Interactive Image Guided Neurosurgery". *IEEE Trans. Biomedical Engineering*, **39**:1226–1231, 1992.

[11] G. Gerig, O. Kübler, and F. Jolesz. Nonlinear Anisotropic Filtering of MRI data. *IEEE Trans. Med. Imaging*, **11**:221–232, 1992.

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

[13] W.E.L. Grimson, G.J. Ettinger, S.J. White, P.L. Gleason, T. Lozano-Pérez, W.M. Wells III, R. Kikinis. "Evaluating and Validating an Automated Registration System for Enhanced Reality Visualization in Surgery". *Proceedings of First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine*, Nice France, April 1995.

[14] W.E.L. Grimson, G.J. Ettinger, S.J. White, T. Lozano-Pérez, W.M. Wells III, and R. Kikinis. "An Automatic Registration Method for Frameless Stereotaxy, Image Guided Surgery, and Enhanced Reality Visualization". *IEEE Transactions on Medical Imaging*, **15**(2), April 1996.

[15] A. Gueziec, N. Ayache. "Smoothing and Matching of 3-D Space Curves". *Proceedings of European Conference on Computer Vision*, May 1992.

[16] B.L. Guthrie, R. Kaplan, P.J. Kelly. "Neurosurgical Stereotactic Operating Arm". *Stereotactic and Functional Neurosurgery*, **54**:497-500, 1995.

[17] R. Haralick, S. Sternberg, and X. Zhuang. "Image Analysis Using Mathematical Morphology". *IEEE Transactions PAMI*, **9**:532–550, 1987.

[18] B.K.P. Horn, "Closed-Form Solution of Absolute Orientation Using Unit Quaternions", *Journal of the Optical Society of America A*, **4**, April 1987, pp. 629–642.

[19] K.H. Höhne and W. Hanson. "Interactive 3D Segmentation of MRI and CT Volumes Using Morphological Operations". *Journal of Computer Assisted Tomography*, **16**:285–294, 1992.

[20] K. Höhne et al. A framework for the Generation of 3D Anatomical Atlases. In *SPIE Vol. 1808, Visualization in Biomedical Computing 1992*, 1992.

[21] H. Jiang, R. Robb, K. Holton, "A New Approach to 3D Registration of Multimodality Medical Images by Surface Matching", *Proceedings of Visualization in Biomedical Computing*, 1992.

[22] R. Kikinis, F.A. Jolesz, W.E. Lorensen, H.E. Cline, P.E Stieg, and P. McL. Black. 3D Reconstruction of Skull Base Tumors from MRI Data for Neurosurgical Planning. *Proceedings of the Society of Magnetic Resonance in Medicine Conference*, 1991.

[23] R. Kikinis, M. Shenton, F. Jolesz, G. Gerig, J. Martin, M. Anderson, D. Metcalf, C. Guttmann, R.W. McCarley, W. Lorensen, and H. Cline. "Quantitative Analysis of Brain and Cerebrospinal Fluid Spaces with MR Imaging". *JMRI*, **2**:619–629, 1992.

[24] K.O. Lim and A. Pfferbaum. Segmentation of MR Brain Images into Cerebrospinal Fluid Spaces, White and Gray Matter. *JCAT*, **13**(4):588–593, 1989.

[25] W.E. Lorensen, H.E. Cline, "Marching Cube: A High Resolution 3-D Surface Construction Algorithm", *Computer Graphics* **21**(3), 1987, pp. 163–169.

[26] R.B. Lufkin, T. Sharpless, B. Flannigan, and W. Hanafee. Dynamic-Range Compression in Surface-Coil MRI. *AJR*, **147**(379):379–382, 1986.

[27] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. "Multi-modality image registration by maximization of mutual information", in *Proceedings of workshop on Mathematical Methods in Biomedical Image Analysis*. IEEE Computer Society Press, 1996.

[28] J.P. Mellor, "Realtime Camera Calibration for Enhanced Reality Visualization", *Proceedings of First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine*, Nice France, April 1995, pp. 471–475.

[29] C. Pelizzari, G. Chen, D. Spelbring, R. Weichselbaum, and C. Chen. "Accurate Three Dimensional Registration of CT, PET and/or MR Images of the Brain". *J. Comp. Assis. Tomogr.*, **13**:20–26, 1989.

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

[31] H.F. Reinhardt, G.A. Hortsmann, O. Gratzl. "Sonic Stereometry in Microsurgical Procedures for Deep-Seated Brain Tumors and Vascular Malformations". *Neurosurgery*, **32**:114-117, 1993.

[32] S. Sandor and R. Leahy. "A 3D Morphological Algorithm for Automated Labelling of the Cortex in Magnetic Resonance Brain Images. *AAAI Spring Symposium Applications of Computer Vision in Medical Image Processing*, Palo Alto, CA., March 1994.

[33] J. Serra. *Image Analysis and Mathematical Morphology.* London Academic, 1982.

[34] M. Shenton, R. Kikinis, and et al. F. Jolesz. Left Temporal Lobe Abnormalities in Schizophrenia and Thought Disorder. *N. Engl. J. Med.*, **327**:604–612, 1992.

[35] R. Szeliski and S. Lavallee. "Matching 3D Anatomical Surfaces with Non-Rigid Deformations Using Octree-Splines". *Proceedings of IEEE Workshop on Biomedical Image Analysis*, Seattle WA, June 1994.

[36] J.P. Thirion. "Extremal Points: Definition and Application to 3D Image Registration". *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Seattle WA, June 1994.

[37] E. Watanabe. "A Potentiometer-Based Localization Arm System". R.J. Maciunas, ed., *Interactive Image-Guided Neurosurgery*, American Association of Neurological Surgeons, 1993.

[38] W.M. Wells, *Statistical Object Recognition*, MIT AI Technical Report 1398, January 1993.

[39] W. Wells, R. Kikinis, E. Grimson, and F. Jolesz. "Statistical Intensity Correction and Segmentation of Magnetic Resonance Image Data". *Proceedings of the First Conference on Computer Vision, Virtual Reality and Robotics in Medicine.* Springer, 1995.

[40] W. Wells, R. Kikinis, E. Grimson, and F. Jolesz. "Adaptive Segmentation of MRI Data". *IEEE Transactions on Medical Imaging, To Appear*, 1996.

[41] W.M. Wells III, P. Viola, H. Atsumi, S. Nakajima, R. Kikinis. "Multi-Modal Volume Registration by Maximization of Mutual Information". *Medical Image Analysis*, **1**(1):35–51, 1996.

[42] A. Witkin, M. Kass, and D. Terzopoulos. "Snakes: Active Contour Models". *International Journal of Computer Vision*, **1**(4):321–331, June 1988.

[43] Z. Zhang. "Iterative Point Matching for Registration of Free-Form Curves and Surfaces". *IJCV*, **13**(2):119–152, 1994.