1-1-2011

Fusion and visualization of intraoperative cortical images with preoperative models for epilepsy surgical planning and guidance.

A Wang

S M Mirsattari

A G Parrent

T M Peters

Preprint of paper published as
Fusion and visualization of intraoperative cortical images with preoperative models for epilepsy surgical planning and guidance. Wang A, Mirsattari SM, Parrent AG, Peters TM, in Computer Aided Surgery 2011 ; 16(4) 149-60, PMID: 21668293

Chapter 3

Fusion and Visualization of Intraoperative Cortical Images with Preoperative Models for Neurosurgical Planning and Guidance

3.1 Introduction

In typical image-guided neurosurgery implementations, a preoperative surgical plan is registered to patient space in the operating room so that the surgeon may be guided to the surgical targets. Nevertheless, the lack of correspondence between the intraoperative context and the preoperative plan on the display, such as a computer screen or head-mounted display (HMD), poses a challenge to the surgeon's capability to mentally correlate the two spaces. For example, for epilepsy surgery or tumor resection on the patient's left temporal lobe, electro-cortical stimulation mapping

(ESM) is often performed to locate the important language and/or memory areas so that they can be spared during the surgery. However, as the surgeon stimulates the cortical surface and marks the location of these critical functional areas, the intraoperative context can not be easily updated so that it can be displayed with the preoperative plan. This limits the surgeon's ability to quantitatively examine the intraoperative information with the preoperative plan, such as provided by functional magnetic resonance imaging (fMRI) analysis. Fusion of the intraoperative scene, as captured by a camera during the procedure, with the preoperative plan, is one approach to solve this problem.

Fusion can be achieved by directly registering an intraoperatively acquired photographic image with the preoperative model. Clarkson et al. [1] performed an intensity-based 3D-2D registration between optical images from multiple views to a preoperative 3D surface model while Miga et al. [2] combined surfaces captured with a laser range scanner with texture acquired simultaneously with a digital camera and registered them to a preoperative cortical model. An alternative to accomplish fusion indirectly is by monitoring the camera pose with respect to the intraoperative scene using a tracking system. For example, in several studies [3] [4], endoscopic static and video images were overlaid onto the models of the brain and the heart. The above mentioned methods either require an accurate preoperative and/or intraoperative surface model, two or more optical images, or require that the camera be tracked for accurate registration.

This chapter proposes a fusion method that differs from the above approaches, and consists of the following steps. First, a landmark-initialized intensity-based perspective 3D-2D registration is employed to estimate the position and orientation of the patient's 3D cortical model with respect to the camera coordinate system, using a single optical image. Second, to increase robustness and efficiency of the registration algorithm, an improved multi-stage optimization method is employed. Finally, the photographic image is back-projected onto the 3D cortical model using texture-

mapping with the estimated orientation and position parameters. To evaluate the proposed fusion algorithm, we first conducted phantom experiment, in which fiducial landmarks were used to assess the accuracy. Next, we evaluated the fusion method using a neuro-navigation system and intraoperative data. Finally, we present the clinical application of the proposed fusion method for image and function guided neurosurgery.

The remainder of the chapter is organized as follows: Section 3.2 describes the methods proposed to achieve fusion. Sections 3.3 and 3.4 present evaluation experiments using phantom data and *in-vivo* data with neuro-navigation respectively. Section 3.5 explain the clinical application of the fusion method for image and function guided neurosurgery. Finally, discussion and conclusion appear in section 3.6 and 3.7 respectively.

3.2 Methods and Materials

3.2.1 System Overview

Our system comprises a neuro-navigation component, which employs an optical tracking system (NDI Polaris) and provides the common utility for image-guided neurosurgery. In addition, a consumer grade digital camera (Nikon D60) is used to acquire the intraoperative photographs. These intraoperative photographic images are transferred to the navigation system via a USB cable, post-processed, and overlaid onto the preoperative models and image volumes using the proposed fusion method. In this study, the visualization and navigation system runs on an Intel Pentium IV PC with an NVidia 7800GTS graphics card.

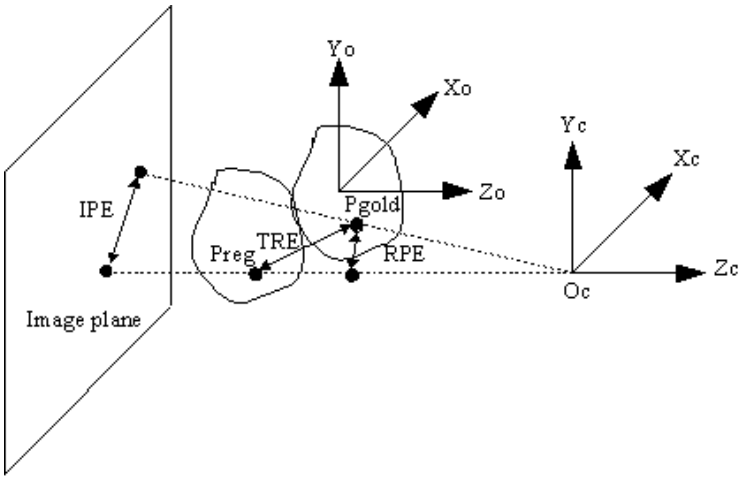

Fig. 3.1: 2D projection of 3D model and error definition.

3.2.2 Problem Definition Using a Pinhole Camera Model

Although a consumer grade digital camera is employed in this work, its optical characteristics can be modeled by a simple pinhole camera model [5]. The pinhole camera model describes the mathematical relationship between the coordinates of a 3D point and its projection onto the image plane of the pinhole camera. Figure 3.1 illustrates the relationship between points in 3D space and the corresponding points on the 2D image plane using a pinhole camera model. O_c is the aperture of the camera or focal point, and using O_c as the origin, a camera centered coordinate system is established. Focal length f is the distance of the image plane to the focal point in the pinhole camera model.

In this model, a point N represented by a homogeneous vector $\mathbf{n} = (x, y, z, 1)^T$ in the 3D scene is projected to intersect with the image plane at a point M represented by a homogeneous vector $\mathbf{m} = (u, v, 1)^T$ in the 2D image and they are related through the following relation:

$$a\mathbf{m} = T\mathbf{n} = PC\mathbf{n}, \tag{3.1}$$

where

$$P = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \tag{3.2}$$

is a 3×4 projective transformation matrix,

$$C = \begin{bmatrix} R & t \\ 0_3^T & 1 \end{bmatrix}, \tag{3.3}$$

is a rigid 4×4 transformation matrix, and a is a scaling factor. The projective transformation matrix P is characterized by the intrinsic camera parameters (in this model just the focal length f) and can be set to a predefined value since in this study we used a digital camera that allows direct focal length adjustment. The rigid transformation matrix C comprises a 3×3 rotation matrix R and a vector $\mathbf{u} = (u_x, u_y, u_z)^T$ that describe the orientation and position of the camera with respect to the 3D scene. R is defined as:

$$R = R_x \cdot R_y \cdot R_z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{bmatrix} \begin{bmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{bmatrix} \begin{bmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{3.4}$$

where $\theta = (\theta_x, \theta_y, \theta_z)$ are the angles of rotation around three Cartesian axes. θ and $\mathbf{u}$ are also called the extrinsic camera parameters.

Using this pinhole camera model, the recorded 2D image is a perspective projection of a 3D scene. However, fusing a 2D image to the corresponding 3D scene is not a trivial task. First, the relationship represented by this 3 × 4 projection transformation matrix T needs to be recovered. Furthermore, projecting the 2D points back to their corresponding 3D locations is an inverse mapping problem, which does not have a unique solution since a point in the 2D image can be related to infinite number of points in the 3D space by the projection transformation T. To solve for

the first problem, we employ a landmark-initialized intensity-based perspective 3D-2D registration algorithm to recover the camera pose. As for the second problem, segmentation is employed to generate a representation of the model so that a unique one to one mapping can be obtained.

To determine the quality of the transformation recovered from the registration procedure, we define some error statistics (TRE, IPE etc.) used in this study. In Figure 3.1, an object centered coordinate system is used to define the objects in the 3D space. P_{gold} represents a point on the object in the object centered coordinate system, P_{reg} represents the same point recovered by a proposed registration transformation in the object centered coordinate system. Target Registration Error (TRE) is defined as the 3D Euclidean distance between P_{gold} and P_{reg}. Image Projection Error (IPE) is the projection distance of the TRE onto the image plane.

3.2.3 Landmark-Based Initial Alignment

Unless the registration has global convergence characteristics, an initial alignment is required to limit the search space within the capture range of the registration algorithm to successfully find the global optimum. To calculate the initial transformation, a landmark-based rigid body registration is employed, where homologous points of anatomical features such as the sulci on the 3D MR brain image and on the 2D photographic image are selected manually. In this initial alignment step, the landmarks in the photograph are used as the targets while the homologues in the 3D image are used as the source. The resulting rigid body transformation $T_{landmark}$ can be viewed as a rough approximation of the true rigid transformation matrix C where the perspective projection degrades to an orthographic projection (i.e., P is the identity matrix).

3.2.4 Intensity-Based Registration Method

The general principle for 3D-2D registration in many applications is that a 2D projection image is first generated from the 3D volume using an estimate for the initial pose, and the 2D projection image obtained at this pose is compared with the 2D image for similarity. This process continues iteratively until the algorithm finds the best match using some numerical optimization procedures.

3.2.4.1 2D Projection Image of Brain via Volume-Rendering

To generate a 2D projection image corresponding to the 3D scene, we first employ a segmentation process [6] commonly known as "skull stripping" to create a rough representation of the brain for visualization purposes. Additionally, this model is used in the final stage as the underlying surface model onto which the photographic image is texture-mapped.

Next, we employ a volume rendering technique based on a GPU-accelerated ray-casting algorithm (described in detail in subsection 3.2.5), which employs the OpenGL graphic library [7] for conventional graphics processing such as setting the viewing environment etc. and OpenGL shading language [8] for GPU programming. Compared to conventional volume rendering approaches that employ texture-mapping, this technique is fast and generates more photo-realistic and artifact-free images. In addition, this approach avoids reconstruction of the cortical surface, which is needed for the conventional surface rendering technique. In the GPU-accelerated volume rendering, a fixed light position is employed and we neglect specular shading effects.

In order to achieve an accurate registration between the 2D projection image generated using volume rendering and the 2D intraoperative photographic image, the anatomical features, which are represented by the sulci and gyri of the cortical foldings of the brain, must be easily recognizable in the 2D projection image. This is achieved through the design of a transfer function, which is used to assign different optical

properties (generally represented by colour and opacity) to each voxel depending on its intensity value in volume rendering. Usually a nonlinear mapping is required to differentiate various tissues using intensity, colour and opacity. This is a difficult task to achieve through manual manipulation of the transfer function. However, there have been some reports of efforts to automate the process, which nevertheless remains a largely unresolved problem due to a non-unique relationship between image voxel values and tissue type. In this work, we employ a simple manual manipulation of the transfer function, in which a user interface widget is used to allow user to adjust the transfer function interactively based on the volume-rendered appearance of the brain.

Finally, to generate a corresponding 2D projection image using volume rendering, a correct viewing perspective that mimics the one used to acquire the photograph must be established. For simplicity, all of the projection calculations we perform using OpenGL are outlined in Appendix A.

3.2.4.2 3D-2D Image Registration

The intensity-based 3D-2D registration employs the Normalized Mutual Information (NMI) introduced in the previous chapter:

$$NMI(I, J) = \frac{H(I) + H(J)}{H(I, J)}, \tag{3.5}$$

as the similarity cost function where

$$H(I) = -\sum_{a} p(a) \log p(a), \tag{3.6}$$

is the Shannon entropy of the images. NMI is a normalized form of the standard mutual information (MI) metric and uses information theory to quantify how well one image is explained by another. Theoretically, the metric is maximized when the images are aligned or when they are maximally dependent. We selected NMI as the similarity measure as it is more suitable for multimodality image registration compared to other metrics such as Normalized Cross Correlation (NCC) as discussed in Chapter 2.

An optimization procedure:

$$T^* = \arg\max_T E(I, J), \tag{3.7}$$

is employed to find the transformation T^* that maximizes the NMI, where I is the target 2D image, and J is the moving 2D image. The Downhill Simplex optimization technique [**?**] is employed to search for the optimal pose parameters based on the values from the similarity cost function. These pose parameters consist of three translational parameters $\mathbf{u} = (u_x, u_y, u_z)^T$ and three rotational parameters $\theta = (\theta_x, \theta_y, \theta_z)^T$ and are used to construct the rigid body transformation matrix C. The Downhill Simplex algorithm is a derivate-free method that is fast and accurate when the initial solution is close to the optimal solution in the search space. For a 6 degree-of-freedom pose estimation, each parameter is initialized and set with a scaling factor, which is used to define how a parameter is to progress during iterations of optimization. Starting from this initial position, the simplex makes its way through the six dimensional search space until it finds at least a local minimum. All the parameters are considered simultaneously and convergence is reached when the difference between the consecutive evaluations of cost function is within a specified tolerance.

Since the local Downhill Simplex algorithm has a limited capture range, we employ a multi-stage optimization approach similar to the method presented by [9] [10]. First, we employ a search space partition approach, in which the transformation parameter space is partitioned into subspaces, each of which is searched independently. Next, a multi-scale strategy is employed to further increase the robustness of the registration algorithm within each subspace. The 2D projection image and the photographic image are first blurred and then registered to each other. These steps are executed repetitively at different scales, progressing to a fine resolution result in a multi-scale manner. The scales are defined as the width of the full width at half-maximum (FWHM) Gaussian kernel which changes from 8 mm to 2 mm. For Downhill Simplex optimization, the characteristic scale length (also called scaling factor)

for the simplex to proceed is set to 10 mm for translation parameters and 10 degree for rotation parameters. Next, the results from each subspace are compared with the most successful solutions being kept for further searching in the following stage. Finally, a re-optimization is employed to decrease the possibility of the registration being trapped within a local minimum. For the final search and re-optimization, the translation scaling factor and rotation scaling factor are set to 3 mm and 3 degree. During registration, we also apply a mask so that only the exposed brain cortex is used for registration.

3.2.5 Photographic Overlay onto Volumetric Image

To overlay a 2D photographic image onto the volume-rendered 3D brain MR volume, a surface model of the brain is generated as described previously in section 3.2.2. A region of interest (ROI) is manually defined in the photographic image using a closed free-form cardinal spline. Next, a perspective projection is employed using Equation 3.1 to generate the texture coordinates u for each vertex v on the brain surface mesh. Finally, the photographic image is texture mapped onto the brain surface mesh. The ROI is implemented as a mask image, which is used to generate the appropriate transparency (alpha) value of the surface mesh so that only the ROI portion of the photograph is displayed.

In our implementation, we ignore the barrel distortion common in consumer grade cameras since it is relatively small (0.3% which is approximately 5 pixels (0.3 mm) in a 3000 $\times$ 2000 images). For reference, the surgical annotation tag displayed in the Figure 3.7 is approximate 30 pixels wide.

3.2.6 Multimodality Image Fusion and Visualization

The ultimate goal of this study is to provide integrated visualization environment to combine images from different sources, acquired at different times for surgical

planning and guidance. Specifically in the context of epilepsy surgery, preoperative anatomical MR, CT and fMRI images need to be registered together first, so that they can subsequently be fused with the intraoperative photographic image using the proposed method.

To combine multimodal images acquired preoperatively, rigid registration is performed since all the multimodal images are from the same subject. The rigid registration algorithm developed locally employs a NMI metric and multi-resolution strategy. The performance of the registration of multimodal images has been evaluated using the Vanderbilt's Retrospective Image Registration Evaluation (RIRE) project [11] data with accuracy around 1.88 mm.

While image fusion of volumetric data is often employed to combine images from various imaging modalities, the ability to manipulate these data interactively is compromised by the bottleneck between the CPU and the GPU, since every change to the source of fusion image results in a complete download of the updated image from the CPU to the video memory on the GPU. To display the registered multimodal images to the surgeon in an organized and interactive way, a fusion and visualization environment using a GPU-based volume rendering technique was developed. This special volume rendering technique is based on a ray-casting algorithm implemented on the GPU using the OpenGL shading language.

Figure 3.2(a) illustrates the ray-casting procedure. A ray is cast from the viewpoint through a specific pixel position on the screen, intersecting a number of voxels as it traverses the volume, while accumulating the optical luminance from each intersected voxel. This procedure is implemented as follows: First, the viewing direction for each pixel in the viewing window under the texture coordinate system is calculated. Next, an iterative process is performed that checks for every voxel on the ray-casting pathway to acquire the intensity values. These values are then translated to colour and opacity values using both colour and opacity transfer functions, which are implemented as lookup tables in the texture memory. Finally the colour values

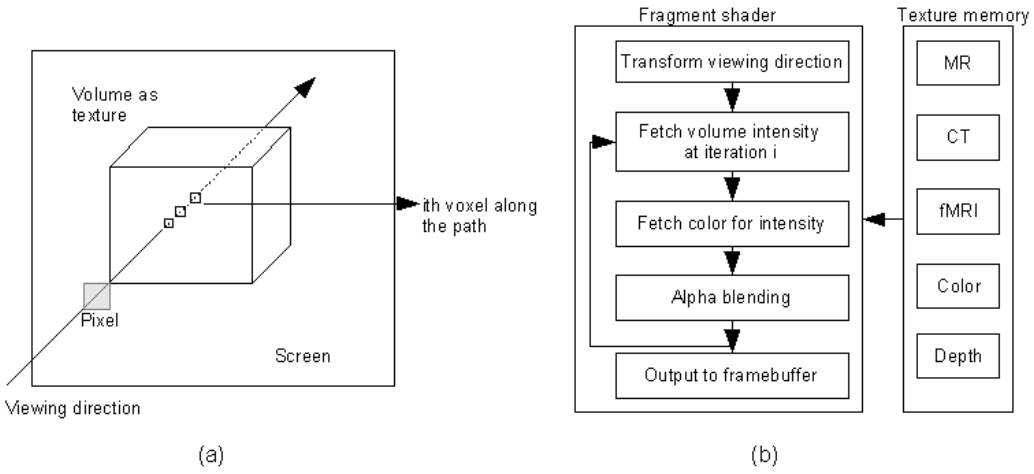

Fig. 3.2: (a) Ray-casting schematic diagram. (b) Ray-casting fragment shader pipeline.

along the ray-casting pathway are alpha-blended into final colour values and output to the display framebuffer for the specific pixel. Figure 3.2(b) shows the fragment shader processing pipeline.

To generate a fused image in the GPU's fragment shader, we employ the following steps. First, the registered CT, MRI and fMRI images are loaded into the video memory as separate 3D textures for volume rendering. Next, we use a simple weighted compositing technique during ray-casting, in which different colour transfer functions are applied to each modality image to get an RGB image. Finally, the fused image is rendered using a specified weight for each modality. The application of the colour transfer function and fusion all happens within the fragment shader so there is no reloading of the image data to the video memory.

3.2.7 Implementation

Our system is based on the AtamaiViewer visualization and navigation platform developed in-house for image-guided procedures. AtamaiViewer is a comprehensive, platform-independent, modular, extensible software platform, which has all the es-

sential features that are needed for common visualization and navigation tasks. It is developed using the Python programming language with the underlying visualization functionality provided by the Visualization Toolkit (VTK) software package [12]. It provides a foundation, which can be extended by user specific modules using plugin techniques.

In our epilepsy surgical planning and guidance system, several modules were developed to facilitate the proposed method. Specific to the display environment, the GPU-based fusion function was implemented using the OpenGL shading language as a generic volume mapper class, derived from the open source VTK software package. The photographic overlay was developed as a module and the perspective texture mapping facility is developed as a class similar to "vtkTextureMapToPlane" class. Moreover, segmentation of the MR brain images and registration of preoperative multimodal images were also implemented as modules in the system respectively.

3.3 Evaluation and Results

3.3.1 Phantom Experiment

To evaluate the intensity-based registration method, we performed a phantom study. First, a 3D CT volume of a standard brain phantom (Kilgore International Inc., Coldwater, MI) was acquired with a GE HiSpeed CT scanner, to which a number of fiducial markers were attached on the left side of the cortical surface of the brain phantom. The brain phantom image matrix is $512 \times 512 \times 320$ with a voxel size of $0.44 \times 0.44 \times 0.64$ mm. The image was re-sampled to a $256 \times 256 \times 256$ matrix size to facilitate rapid volume rendering (Figure 3.3(a)). Next, a surface model of the phantom was created using the marching cube algorithm implemented in VTK (Figure 3.3(b)). Four photographs of this phantom with different poses were then acquired using a consumer grade digital camera (image matrix size of 3000×2000)

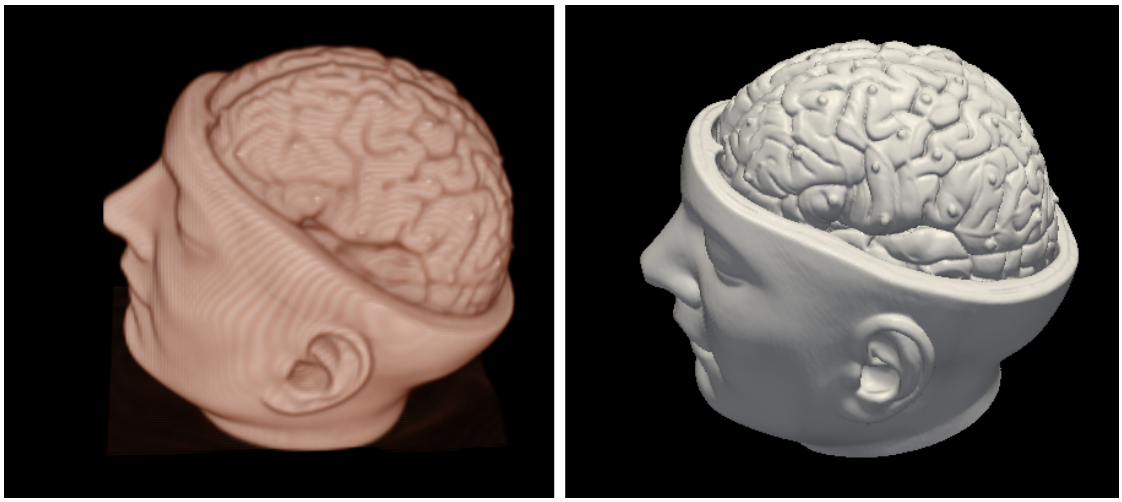

Fig. 3.3: Volume-rendered phantom CT image and surface-rendered model.

(see Figure 3.4) and transferred to the workstation via USB. The focal length was set to 50 mm during photograph acquisition. The photographs were then blurred and re-sampled to 375 $\times$ 250.

The phantom CT image was then imported into the AtamaiViewer system and four landmarks on the volume-rendered 3D CT image and the corresponding landmarks in the photographic images were manually identified as shown in Figure 3.5(a) and Figure 3.5(b) respectively. Landmark registration was then performed and the result was used as the initial registration estimate. Next, we generated ten transformation matrices with a random value chosen within the range of the each parameter with $\Delta_{translation} = 10$ mm and $\Delta_{rotation} = 5^o$ to simulate the initial misalignment and these matrices were used to initialize the matrix C' in Equation 3.5. We then performed the intensity-based 3D-2D registration and volume rendering was displayed within a 375 $\times$ 250 pixel window to match with the re-sampled photographic image. Finally, the fusion was achieved by texture-mapping the photographic image back to the surface model.

To evaluate the quality of the fusion algorithm, 3D registration error measure

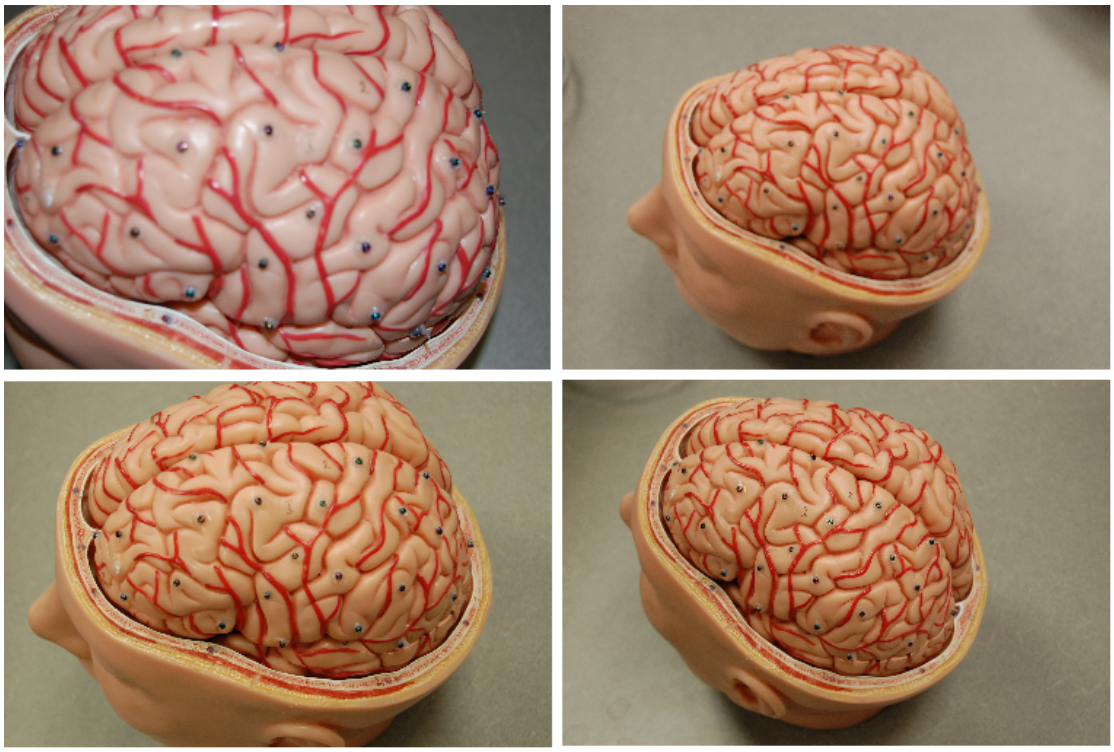

Fig. 3.4: Photographs taken at 4 different poses.

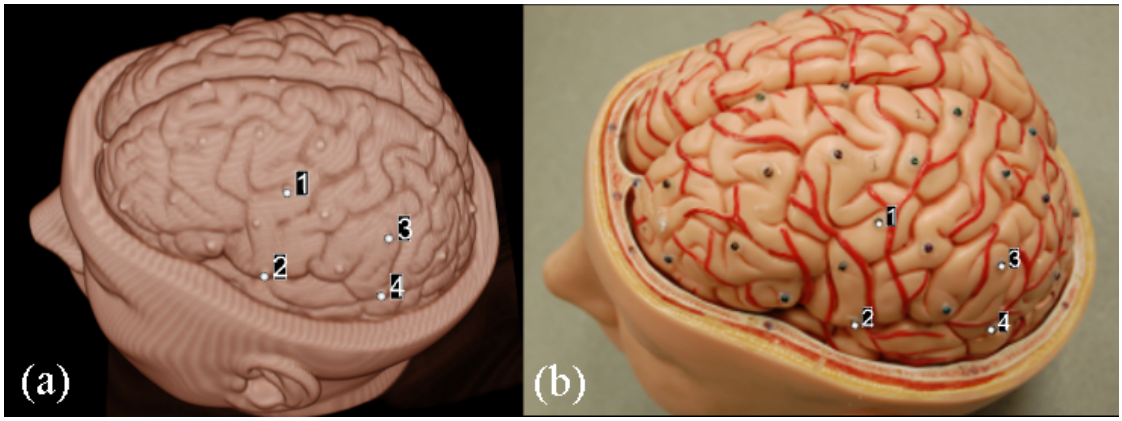

Fig. 3.5: Phantom experiment. (a) Landmarks on the CT image. (b) Landmarks on the photograph.

(MRE) between the homologous fiducial points from the 3D phantom CT image and fused surface model were calculated as:

$$MRE = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(p_s - p_t)^T(p_s - p_t)}, \tag{3.8}$$

where N is the total number of pairs of homologous fiducial landmarks, p_t is the 3D coordinates of the gold standard point i, and p_s is the 3D coordinates of the registered point i. In this study, the 3D coordinates of the homologous fiducial landmarks in the CT volume were identified manually by their centroids. After overlay of the photographic image onto the 3D surface model, the registered 3D coordinates were determined manually as the centroids of fiducials on the fused surface model. The registration results are listed in Table 3.1 for each image. The success rate is the percentage of cases in which the MRE is less than 3 mm. Figure 3.6(a) shows the fusion of part of the photographic image onto the CT volume with 50% transparency. Figure 3.6(b) shows a blown up version of the Figure 3.6(a) at the edge of the overlaid photograph. It can be seen that the gyral curves and the fiducial landmarks match very well between the photographic image and the CT volume.

Table 3.1: Intensity-based registration error in phantom study.

Photograph	3D MRE (mm)	Time (s)	success rate
1	2.48 ± 0.30	20	100%
2	2.24 ± 0.26	21	100%
3	2.87 ± 0.42	20	100%
4	2.15 ± 0.35	20	100%
mean	2.43 ± 0.32	20	100%

3.3.2 Clinical Experiment Using Neuro-Navigation System

To evaluate the fusion method in a more realistic clinical environment, we conducted a clinical *in-vivo* experiment. One of the most important limitations with

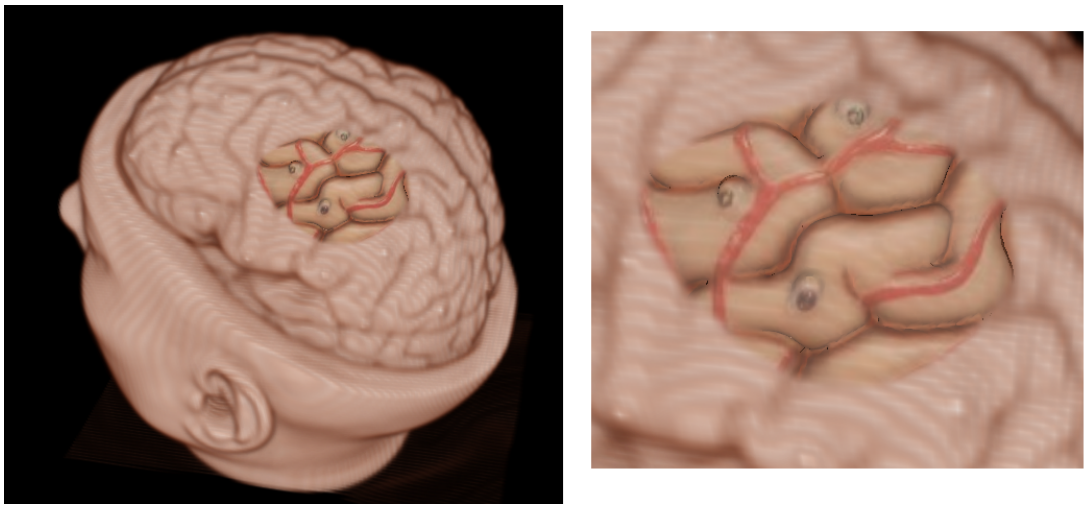

Fig. 3.6: Phantom experiment result. (a) Photographic overlay onto the CT image. (b) Blown-up of region of (a).

clinical evaluation is that the lack of a well defined gold standard, so in this study, a neuro-navigation system was employed to provide such a reference. Nevertheless, the fidelity of the gold standard is compromised during surgery. Firstly, the accuracy of the navigation can influence the final validation results, and brain deformation after craniotomy and dura opening can be up to 1cm [13], which could therefore further degrade this gold standard. For this reason, we also investigated the effect of brain shift on the final results.

In this study, a patient with mesial right temporal lobe epilepsy (TLE) underwent an anterior temporal lobectomy (ATL). Prior to surgery, a preoperative anatomical MR scan was acquired with the fiducial markers affixed to the patient's head skin. Next, the preoperative MR volume was imported into AtamaiViewer, which was employed as a neuro-navigation system in parallel with the regular Stealthstation system. Fiducial landmarks on the patients head were then registered to the landmarks identified on the volume-rendered head image to establish the patient-to-image registration.

After craniotomy and opening of the dura-mater, four anatomical landmark points

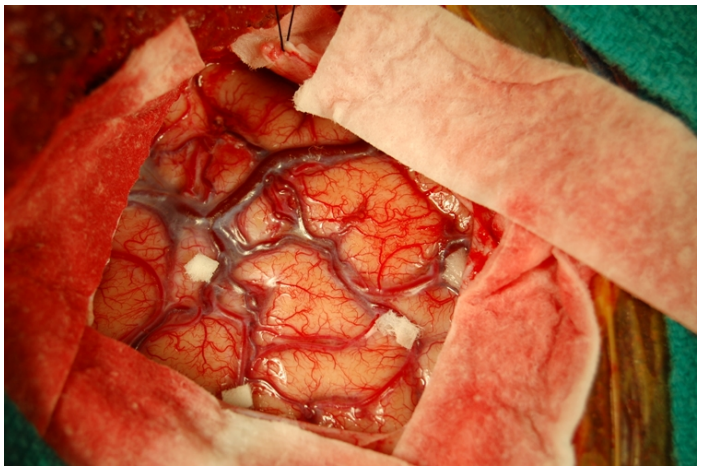

Fig. 3.7: Clinical evaluation: paper tags on the cortical surface.

were identified by the operating surgeon and marked with paper tags on the cortical surface of the patient's brain (Figure 3.7). These landmarks were chosen mainly at the bifurcation of blood vessels.

Next, these landmarks were localized using the pointer tool tracked by the navigation system and their corresponding 3D coordinates under the navigation system were recorded. After 20 minutes, sufficient time for the brain shift to occur, these four landmarks were then measured again using the navigation system. Next, an intraoperative photographic image was acquired to capture the surgical field of view as shown in 3.7. and this image was then imported into the AtamaiViewer. Initial alignment was performed through homologous landmark registration between the image and 3D volume-rendered brain model. Figure 3.8 shows the selection of the four homologous landmarks in the 3D MR volume and 2D image as well as the definition of the ROI. Next, the 3D-2D intensity-based registration was performed to register and fuse the photographic image with the brain model.

Figure 3.9 shows the fusion of the intraoperative photographic image on top of the brain model. The landmarks in the fused photographic image were then manually identified and the coordinates under the 3D navigation system were then compared

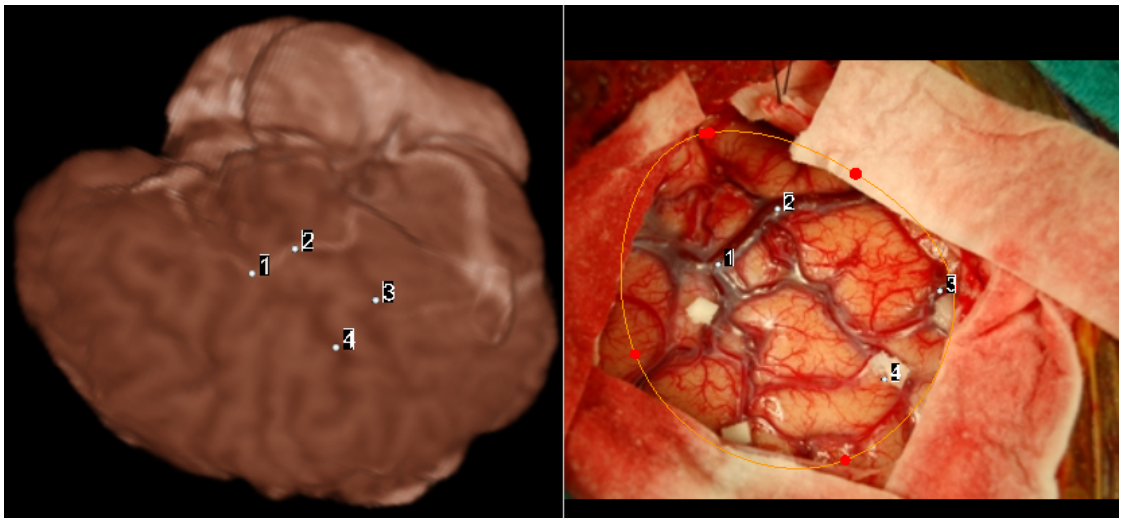

Fig. 3.8: Clinical evaluation: landmarks used to initially align the two images.

to those recorded by the navigation system following brain shift (Table 3.2). Registration errors were reported in three orthogonal directions as a 3D Euclidean distance as well as a 2D in-plane error. 2D in-plane error is defined as the 2D distance in the AP and SI plane. We also listed the brain shift measured using the four landmarks in Table 3.3. In Figure 3.9, the yellow spheres represent the landmarks reconstructed immediately after the opening of the dura mater, the red spheres represent the landmarks reconstructed 20 minutes after the dura opening and the green spheres represent the landmarks reconstructed from the photographic fusion.

Table 3.2: Intensity-based registration error in clinical study.

Point	LR error (mm)	AP error (mm)	SI error (mm)	3D MRE error (mm)	2D in-plane error (mm)
1	4.48	2.44	1.34	5.27	2.78
2	4.15	1.27	0.94	4.44	1.58
3	3.59	3.64	1.96	5.47	4.13
4	2.70	4.35	1.86	5.45	4.73
mean	3.73(0.78)	2.93(1.36)	1.53(0.47)	5.15(0.49)	3.30(1.41)

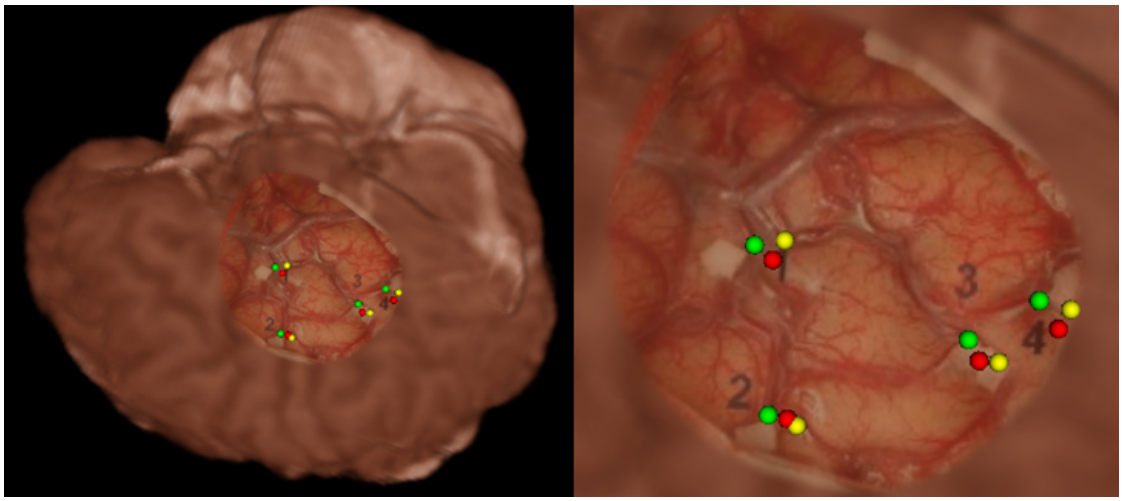

Fig. 3.9: Fusion of the photographic image onto MR brain model and reconstructed landmarks in the clinical study.

Table 3.3: Brain shift measurements in clinical study.

Point	LR error (mm)	AP error (mm)	SI error (mm)	3D MRE error (mm)
1	4.66	0.93	2.27	5.26
2	4.76	1.27	0.42	4.94
3	3.85	2.10	1.18	4.54
4	3.01	0.04	1.93	3.58
mean	4.07(0.81)	1.08(0.85)	1.44(0.82)	4.58(0.73)

3.4 Clinical Applications

The clinical motivation of the fusion algorithm is to facilitate image guidance via both anatomical and functional data during neurosurgery. For epilepsy or tumour resection surgery operated on the left temporal lobe, one of the clinical goals is to minimize the surgical risk of cutting into the eloquent area, especially language areas, which are generally localized in the posterior superior temporal gurus. Electro-cortical stimulation mapping is a clinical standard to elicit the language areas. More recently, preoperative language fMRI has also been shown in some studies to yield good pre-

diction of the language area [14]. Therefore, combining these two techniques could provide better localization of language areas. In addition, this approach could provide enhanced or augmented visualization of the surgical field to the surgeon from the two complementary imaging sources. Here we demonstrate the use of the integrated image and function guidance for epilepsy surgery.

3.4.1 Preoperative Imaging

A patient with left TLE was studied in this experiment. Prior to the surgery, the patient underwent standard anatomical and functional MR imaging. fMRI images were acquired using two language stimulation paradigms (verb generation and sentence completion) to elicit BOLD response in the cortically active regions. In addition, intracranial electroencephalography (EEG) monitoring was employed for seizure localization, whereby subdural electrodes are placed on the surface of the brain to monitor cortical electrical activities. To facilitate the correlation of the position of the electrodes with the neuroanatomy context, a CT scan was performed after the metal EEG electrodes had been placed on the cortex.

3.4.2 Multimodality Image Fusion and Visualization

Skull stripping was employed for the MRI image to extract the brain tissue from the skull and the extracted brain mesh was saved as a cortical surface model data for photographic overlay display. To correlate the position of the electrodes with the neuroanatomy context, the electrodes were segmented from CT images acquired after the electrode implantation step. fMRI images were processed with SPM2 using a general linear model [15] to generate the functional maps that are represented by areas activated by the language tasks. Next, the MRI and fMRI image data were first registered to each other, and then in turn registered to the CT image using AtamaiWarp [16]. The registered CT volume, MR volume and fMRI activation maps

were then displayed using volume rendering, in which different color transfer functions were applied to each modality and the final fused image rendered using a composite blending technique.

3.4.3 Intraoperative Navigation and Fusion

When the studies were performed in the operating room, two navigation systems were employed. One of them, the StealthStation Treon system (Medtronic Navigation Technologies, Inc., Louisville CO 80027, USA), was employed by the neurosurgeons used as the primary guidance system for the procedure. The other was the locally developed AtamaiViewer visualization and navigation system. Each of these systems had its own POLARIS optical tracking camera, but shared the use of a reference tool and a tracked pointer.

Prior to surgery, a preoperative MR scan was acquired with the fiducial markers fixed to the patient's head skin, and the preoperative image data were loaded onto both navigation systems. Fiducial landmarks on the patients' head were then registered to the landmarks identified on the preoperative MR volume to establish the "image-to-patient" registration. After craniotomy and opening of the dura-mater, intraoperative ESM was performed on the patient's left temporal and frontal lobes to elicit the critical language areas and this procedure is described in further details in Chapter 4. Next, a photographic image capturing the stimulated cortical surface was acquired using a digital camera. This image was then transferred to the AtamaiViewer visualization system, and was then overlaid onto the preoperative brain model using the proposed fusion method.

Figure 3.10 shows an example of fusion of the photograph image with the volume rendered anatomical MR volume. Figure 3.10(a) shows the volume rendering of anatomical MR volume fused with functional activation map. Figure 3.10(b) shows the intraoperative photographic image of the cortical surface with tags of language sites on it. Figure 3.10(c) shows the fusion of the photographic image onto the

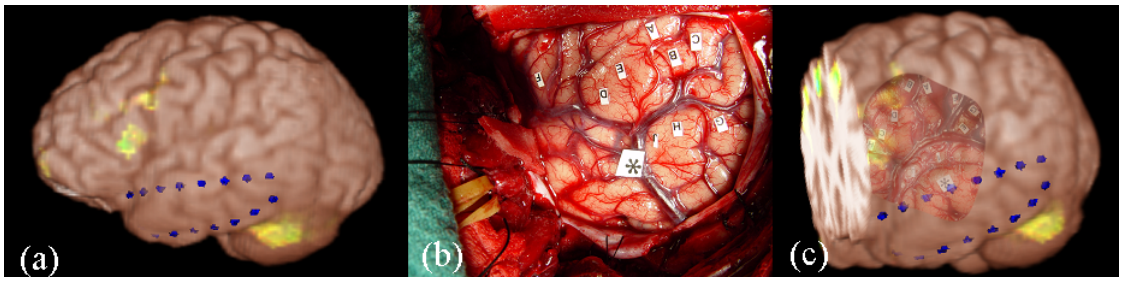

Fig. 3.10: Clinical implementation. (a) Volume rendering of the anatomical MR volume and functional activation map. (b) Intraoperative photograph of left temporal lobe. (c) Overlay of the photograph on top of the volume rendering of the anatomical MR volume and functional activation map.

anatomical MR volume and functional activation map. The green areas represent the regions onto which language activity was mapped. This presentation allows enhanced visualization of the language areas elicited by both the intraoperative ESM and fMRI during neurosurgery.

3.5 Discussion

In this study, we demonstrated the fusion of the intraoperative scene with the preoperative plan. The proposed fusion method provides a means of rapidly capturing the intraoperative environment and overlaying it onto the preoperative model, which facilitates the correlation of the two spaces. Compared to methods proposed in several studies [1] [3], it only requires one untracked single optical image. In addition, it is based on the image intensity generated using a direct volume rendering technique, which in turn does not require accurate landmark localization or feature extraction and surface reconstruction. This approach is both cost effective and user friendly.

The proposed landmark-initialized intensity-based 3D-2D perspective registration algorithm achieved good registration accuracy in the phantom experiment and relatively good accuracy in the clinical experiment. With respect to the registration procedure, landmark initialization brings the initial pose close to the true solution

and helps to reduce the search space substantially. Furthermore, the multi-stage optimization strategy adds to the robustness of the registration. Both approaches proved to be very effective in stabilizing the registration algorithm. The choice of an optimal color transfer function to maximize the features on the brain cortex is also important for this approach. Here we use an empirically selected color transfer function.

One of the limitations of the proposed fusion method is that the relatively larger registration error in the actual clinical situation with 2D in-plane error around 3.3 mm and 3D MRE error around 5.2 mm. This increase in error is likely a result of several factors. First, the quality of clinically acquired MR images does not match with the high quality of phantom brain CT image. This degrades the ability to distinctly render the cortical features, which in turn makes the registration less accurate. Secondly, the cortical surface model of the brain phantom is generated with high accuracy since the brain CT image was acquired without the skull. However, this becomes a much more difficult task for the MR images. Accurate reconstruction of the cortical surface is still an open problem that deserves further investigation on its own.

Although the registration error is relative large in the clinical situations, this is still acceptable for the actual clinical procedures. First, the 3D MRE error is not necessarily a qualified measure for the clinical procedure. When the surgeon examines the photographic overlay, he/she will most likely be visualizing the brain in a direction perpendicular to the cortical surface, making the 2D in-plane registration error more representative measure for the actual procedure. Second, for epilepsy and tumor resection surgery, the clinical standard is to leave a 1cm margin with respect to the eloquent areas, so a 3 mm error remains acceptable for this clinical procedure.

Another concern in the clinical experiment is the problem of brain shift and its effect on the proposed fusion system. The brain shift for this clinical study is not particularly significant (under 5 mm) so the coordinates reported by the neuro-navigation system is assumed to be correct in this case. As shown in the brain shift experiment, the shift mainly took place along the lateral direction with respect to the patient,

which is in line with the direction of gravity. This result agrees with the previous studies [13] on brain shift and demonstrates that less variability along the two in-plane directions. In an actual clinical procedure, the preoperative model does not accurately reflect the intraoperative state. In this case, the fusion algorithm can potentially map the deformed intraoperative cortical image back to the undeformed preoperative model in a sense to rigidly correct for the brain shift. However, this correction is not ideal as the preoperative plan is not updated through this process.

3.6 Conclusion

In this Chapter I have presented a landmark-initialized intensity-based 3D-2D registration method to estimate the camera pose with respect to the 3D patient model and a fusion method to overlay the optical image onto the preoperative volume. The result of the photographic overlay using the phantom image demonstrates good correspondence between the preoperative volumes and intraoperative optical image. A preliminary clinical study showed that the fusion could be achieved with 3D MRE error of approximately 5 mm with 2D in-plane registration error approximately 3 mm. Finally, the proposed fusion method was applied to provide enhanced and augmented visualization for integrated anatomical and functional guidance in several clinical cases with acceptable results.

References

[1] M.J. Clarkson, D. Rueckert, D.L.G. Hill, and D.J. Hawkes. Using photo-consistency to register 2d optical images of the human face to a 3d surface model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1266 – 1280, 2001.

[2] Micheal I. Miga, Tuhin K. Sinha, David M. Cash, Robert L. Galloway, and Robert J. Weil. Cortial surface registration for image-guided neurosurgery using laser-range scanning. *IEEE Transactions on Medical Imaging*, 22(8):973–985, 2003.

[3] Damini Dey, David G. Gobbi, Piotr J. Slomka, Kathleen J.M. Surry, and Terence M. Peters. Automatic fusion of freehand endoscopic brain images to three-dimensional surfaces: Creating stereoscopic panoramas. *IEEE Transactions on Medical Imaging*, 21(1):23 – 30, 2002.

[4] S. Szpala, M. Wierzbicki, G. Guiraudon, and T.M. Peters. Real-time fusion of endoscopic views with dynamic 3-d cardiac images: a phantom study. *IEEE Transactions on Medical Imaging*, 24(9):1207 – 15, 2005.

[5] L. G. Shapiro and G. C. Stockman. *Computer Vision*. Prentice Hall, 2001.

[6] S. M. Smith. Fast robust automated brain extraction. *Human Brain Mapping*, 17:143–155, 2002.

[7] Mason Woo, Jackie Neider, and Tom Davis. *OpenGL Programming Guide*. Addison-Wesley Developers Press, Reading, Mass., 1996 (Second Edition).

[8] J. Kessenich, D. Baldwin, and R. Rost. *The OpenGL Shading Language*. The Khronos Group Inc., 2008.

[9] M. Jenkinson and S. M. Smith. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5:143–156, 2001.

[10] M. Jenkinson, P. R. Bannister, J. M. Brady, and S. M. Smith. Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17:825–841, 2002.

[11]

[12] Will Schroeder, Ken Martin, and Bill Lorensen. *The Visualization Toolkit*. Kitware, Inc., Clifton Park, New York, 2006 (Fourth Edition).

[13] D. L. Hill, C. R. Jr. Maurer, and R. J. Maciunas. Measurement of intraoperative brain surface deformation under a craniotomy. *Neurosurgery*, 43:514–526, 1998.

[14] A. C. Carpentier, K. R. Pugh, M. Westerveld, C. Studholme, O. Skinjar, J. L. Thompson, D. D. Spencer, and R. T. Constable. Functional mri of language processing: Dependence on input modality and temporal lobe epilepsy. *Epilepsia*, 42:1241–1254, 2001.

[15] K. J. Friston, J. T. Ashburner, S. Kiebel, T. E. Nichols, and W. D Penny, editors. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, 2007.

[16] Y. P. Starreveld. *Fast nonlinear registration applied to stereotactic functional neurosurgery*. PhD thesis, the University of Western Ontario, London, ON, Canada, 2002.