Intraoperative Visualization of Anatomical Targets in Retinal Surgery

Ioana N. Fleming, Sandrine Voros, Balazs Vagvolgyi, Zach Pezzementi, Dr. Jim Handa, Russell Taylor, Gregory D. Hager

Abstract

Certain surgical procedures require a high degree of precise manual control within a very restricted area. Retinal surgeries are part of this group of procedures. During vitreoretinal surgery, the surgeon must visualize, using a microscope, an area spanning a few hundreds of microns in diameter and manually correct the potential pathology using direct contact, free hand techniques. In addition, the surgeon must find an effective compromise between magnification, depth perception, field of view, and clarity of view.

Pre-operative images are used to locate interventional targets, and also to assess and plan the surgical procedure. This paper proposes a method of fusing information contained in pre-operative imagery, such as fundus and OCT images, with intra-operative video to increase accuracy in finding the target areas. We describe methods for maintaining, in real-time, registration with anatomical features and target areas using image processing. This registration allows us to produce information enhanced displays that ensure that the retinal surgeon is always in visual contact with his/her area of interest.

1. INTRODUCTION

Retinal disease remains the leading cause of new blindness in the Western World. The World Health Organization estimates that "over 180 million people are visually disabled world wide and that 40 to 45 million of these are blind" [1]. The most frequent indications for vitreoretinal surgery in the US are typically in the elderly, and include retinal detachment, vitreous hemorrhage, macular pucker, macular hole and diabetic retinopathy. The main challenge linked to microsurgical procedures is to be able to perform surgical gestures on structures with cross sections varying from millimeters down to microns.

Several groups have worked on robotically assisted microsurgery [2,3,4-6]. On the other hand, to our knowledge, only [5] have addressed the problem of the overlay of preoperative images with the intraoperative microsurgical view in the operating room. Indeed, in microsurgery, pre-operative images are used for finding target areas, assessing and planning the entire surgical procedure, but today this is only done qualitatively and offline: the preoperative images are not registered with the intra-operative video data. Development of these features would improve the precision and placement of instruments during delicate maneuvers which would improve outcomes, minimize iatrogenic complications, and shorten surgical time.

For instance, macular hole surgery involves internal limiting membrane (ILM) peeling which consist in incising a few microns of the innermost layer of the neurosensory retina and delaminating it from the rest of the retina. In the incision step, the surgeon needs to delimit accurately the boundaries of the ILM to avoid damaging healthy retinal tissue, which is difficult to perform without quantitative preoperative data. This information is available through the preoperative data and could be included in an intra-operative overlay which would assist the surgeon.

We thus believe that robotic assistance combined with image overlay supported by computational image processing will allow routinely planning and carrying out microsurgical procedures with a better precision and comfort.

In this paper we present a framework for retinal microsurgery which includes two components:
- registration of preoperative diagnostic images with the intra-operative video data,
- tracking anatomical features localized thanks to the registration phase.

A 3D visualization system is used to display this enhanced information to the surgeon during the intervention.

2. BACKGROUND

A key element of our proposed approach is to augment the surgeon's view with additional information, and to acquire and process video from the surgeon's view for the purposes of registration, targeting, and information overlay. The standard setup used today for microsurgery is shown in Fig. 1:

As can be seen in the figure, surgery is typically carried out under direct optical visualization. Currently, the acquisition of video in ophthalmology for documentation purposes is a common practice -- nearly all surgical microscopes provide video acquisition

Fig. 1: Typical intraoperative scene for retinal surgery

capabilities as an option. However, the acquired images are generally markedly inferior to the optical view provided by the surgical microscope. Furthermore, stereoscopic video acquisition, a rarity in any case, is usually not ergonomically acceptable for long-term viewing. As compared to other areas of surgery that rely heavily on video acquisition, most notably minimally invasive surgery viewed by laparoscopic or endoscopic means, the lighting conditions and resolution demands in retinal surgery are extreme. Direct observation of the retina has become possible using a new generation of gradient index (GRIN) lens endoscopes (e.g. the IE 4000, Insight Instruments, Sanford, FL). These instruments enable high-resolution, direct viewing of the retina, but suffer from field of view constraints and lack of stereoscopic viewing capability. They have yet to find routine use in retinal surgery.

The display of information to the micro-surgeon has become possible using a new generation of high quality image injection systems. These systems overlay or replace the optical view with a computer-generated digital display. Modern image injection systems (e.g. the Leica DI C500) provide a display with resolution of 1024x768 pixels. In some cases (e.g. neurosurgery), these displays can be registered to a tracked reference frame that include the microscope, and can therefore display information registered to the live view. The MAGI project [4,5] provides registered image overlays in a stereo microscope by tracking artificial fiducials and computing a rigid body transformation between the intraoperative stereo images and preoperative datasets. Lievin and Keeve [6] describe a similar sort of stereo augmented reality display. None of these systems has been tested (or appear compatible) with retinal surgery. In [7], methods are described for combining ultrasound data with monocular video streams using advanced head tracking and rendering techniques. While these efforts provide evidence that it is possible to reliably and accurately process video streams from surgery in vivo, they stop short of providing solutions that can respond to deformable anatomy or non-rigid motions, that can perform registration without fiducials, and which integrate to a broader system architecture.

In immediately relevant work, Stewart et al. [8-10] have developed extremely reliable methods for registering retinal fundus images and angiograms. In extensive longitudinal studies, they report sub-pixel registration accuracies, and successful automated registration in more than 99% of cases.

Thus, there are several open questions that must be addressed in order to have a practical image overlay system. First, we must determine if the resolution and dynamic range available with today's digital cameras and display devices is sufficient to allow a surgeon to work without severe loss of visual acuity. Second, while static registration of images has been shown to be possible, it is not clear if this registration can be performed successfully between preoperative and intraoperative imagery, and on an anatomical target that is the subject of manipulation. Finally, during a procedure, the retina is mobile, both due to the rotation of the eye as well as deformation introduced by manipulation. Thus, it is

necessary to not just register images, but to track that registration.

In order to address these questions, we have developed an image overlay system for computer-assisted targeting. With this system, we intend to test whether overlay of pre-operative information has a significant impact on targeting performance. In the remainder of this paper, we describe the development of this system, we present our experimental setup and we include preliminary results.

3. METHODS

Optical coherence tomography (OCT) has recently become a widely used biomedical tissue imaging technique because of its micrometer resolution and cross-sectional imaging capabilities. It is particularly well suited for retinal applications since it provides the desired resolution and millimeter penetration depth. OCT is one of the routine choices today for ophthalmic diagnostic and preoperative planning. OCT provides information about the different layers of the retina and it comes paired with a low quality fundus image known as the *targeting image* containing the location of the cross-sectional planes (Fig. 2).
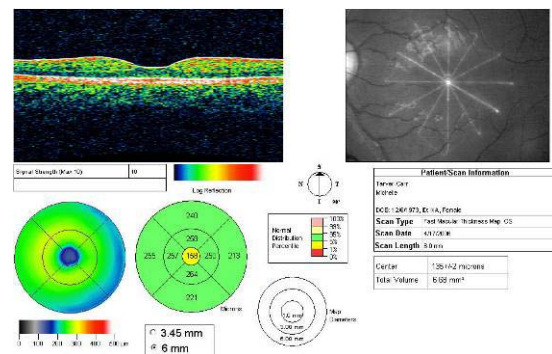

Fig. 2: Optical coherence tomography (OCT) cross-section of the retina (left) and corresponding targeting image (right). The arrow (middle) shows the orientation of the cross-section.

Another complementary diagnostic imaging modality used for preoperative planning of a macular hole surgery procedure is eye fundus imaging, which consists of photographing the interior surface of the eye, including the retina, optic disc, macula, and posterior with an ophthalmoscope. The eye's fundus is the only part of the human body where the microcirculation can be observed directly. The diameter of the blood vessels around the optic disc is about 150 micrometers, and an ophthalmoscope allows observation of blood vessels with diameters as small as 10 micrometers [11] (Fig. 3).

Fig. 3: Eye fundus images

Our targeting system is focused on presenting to the surgeon preoperative targets identified in OCT. This process proceeds by first registering OCT to a pre-operative fundus image, followed by a registration of the fundus image to the microscope view, and finally maintaining the fundus-microscope registration through visual target tracking. This creates a continuously updated registration chain for displaying preoperative targets in the surgical view. We now describe each of these steps in turn.

3.1 OCT to Fundus alignment

We henceforth assume that one or more pathological areas are identified on the OCT scan; these will become the targets of the microsurgical procedure. The span and orientation of each OCT cross-section of the retina is known with respect to the aforementioned *targeting image* (a poor quality fundus image). The first step of our algorithm is an OCT to targeting alignment which consists of adjusting the scale of the OCT cross-section to fit its corresponding representation in the targeting image. The targets on the OCT images can thus be accurately projected on the *targeting image* (Fig. 4 left).
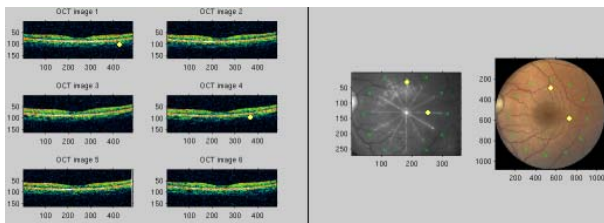

Fig. 4: Retinal map after all OCT sections have been aligned: 6 scanned OCT images on the left. Notice the 2 (two) selected targets on the left in OCT image 1 and OCT image 4. Their projections are overlaid onto the targeting image after alignment (middle right), and also on the fundus image after registration (far right).

The alignment is followed by a registration between the *targeting image* and a high-resolution fundus image. The difficulty of this registration comes from the fact that the targeting image has a poor quality. Thus, we need an algorithm that can reliably align two fundus images while

handling challenges such as low overlap, substantial orientation and scale differences, large illumination variations, and physical changes in the scene. Stewart's Dual Bootstrap ICP is a reliable and accurate retinal registration algorithm, which was tested on 6000 retinal image pairs, reporting successful registration of 99.5% of the pairs containing at least one common landmark, and 100% of the pairs containing at least one common landmark and at least 35% image overlap. The algorithm reports >99% accuracy up to .64 pixel with retinal images [8].

Stewart's feature based approach is to start by extracting and matching keypoints between images. These are used to generate initial similarity transform estimates, each accurate over a small region of the images, called bootstrap region. In each bootstrap region, the algorithm iteratively:

- Refines the transformation estimate using constraints only from within the bootstrap region;
- Expands the bootstrap region; and
- Tests to see if a higher order transformation model can be used (similarity, affine, reduced-quadratic, and quadratic), stopping when the region expands to cover the overlap between images.

The result of the registration between the *targeting image* and the high resolution fundus image is shown in Fig. 4 (right).

3.2 Microscope to Fundus registration

In a second step, a pre-op/intra-op registration must be performed to bring the preoperative information provided by the fundus image (**3.1**) into the operating room. This time, the difficulty is to register a small patch of the retina viewed through the microscope with the high resolution complete fundus image. Stewart's dual bootstrap ICP registration method was used again.

After this process, all the anatomical targets selected by a surgeon preoperatively are known in the initial stereo microscope view (Fig. 5).
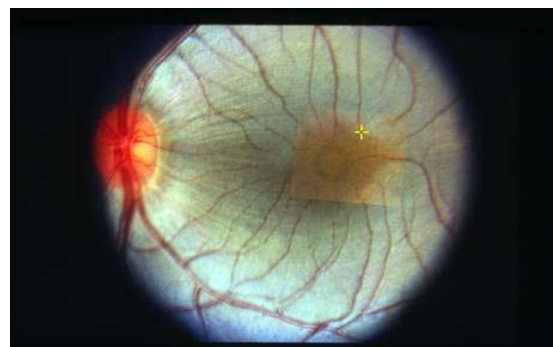

Fig. 5: Registration of a fundus image with a microscopic image makes it possible to determine the position of anatomical targets (yellow cross) in both modalities.

3.3 Tracking and visualization of anatomical targets

The anatomical targets which are now registered with the microscopic view are tracked in stereo using a region-based tracking algorithm [12]. The algorithm is based on direct image matching of selected regions by minimizing the sum-of-squared differences (SSD) of the intensities between a ground truth region (template) and a candidate region. The algorithm uses a parametric motion model which can take into account translation, rotation and scale, as well as deformation. The method can also deal with changes of illumination by using a 'basis' of several templates rather than a single one, which can be particularly attractive in the context of surgery, where strong illumination can cause shadows and specular reflections.

The final microscopic view together with target overlay is displayed to the surgeon in 3D using a polarizing screen. To do so, the left and right images from the microscope are acquired using two video cameras, and rectified to obtain a satisfying 3D visualization.

4. EXPERIMENTAL METHODS AND PRELIMINARY RESULTS
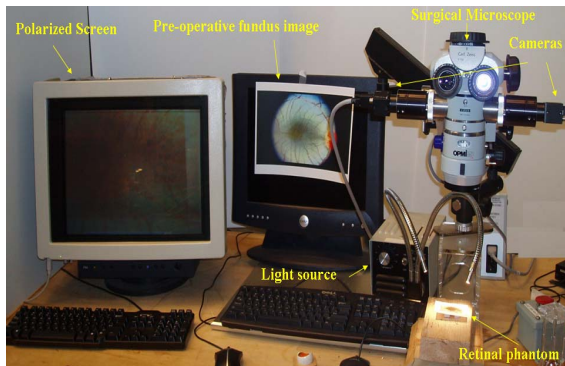
4.1 System Setup

Fig. 6: Experimental system setup.

Our experimental system setup was intended to mimic a vitreoretinal surgery setup (Fig. 6). It included:

- Zeiss stereo surgical microscope (Zeiss OPMI MD; Carl Zeiss Surgical, Inc.) routinely used in today's ophthalmology operating rooms.
- 2 (two) Point Grey Flea®2 cameras were attached to the microscope to grab the left and right frames for image processing. Such cameras are frequently used in the OR for recording purposes.
  - o **Sensor**: Sony 1/3" or 1/2" progressive scan CCDs, Color/BW
  - o **Resolution**: 1024x768
  - o **Frame Rate (Max Res)**: 15 FPS
  - o **Dimensions**: 29x29x30mm
  - o **Interface**: 9-pin IEEE-1394b 800Mb/s interface

- The rectified images were displayed on a 1024x768 resolution polarized screen.
- Registration and tracking algorithms, as well as other image processing routines were performed on a Intel Pentium Xeon Dual Core 2x2.0Ghz computer.

We created a retinal phantom as follows. A small region of an available high-resolution fundus image was printed in high quality on glossy studio photo paper. The print-out was placed under the surgical microscope and designated the intra-operative reference image; the surgical microscope view was also displayed on the stereo monitor. The magnification of the small patch of the retina was chosen so that the vessels' resolution would be similar to the surgeon's view under the microscope during vitreoretinal surgery.

In our preliminary work we restricted target motion to planar rigid body motion (translation and rotation). A single target was tracked using a 80x80 pixel template.

4.2 Objectives

The first step of our experiment consisted in defining an anatomical target (Ta) on the initial microscopic view; we considered this our ground truth. This was followed by multiple attempts to reach the chosen target with a microsurgery tool. Our objectives were to:

1) assess the accuracy of the tracker during a sequence or motions (various translations and rotations),
2) compare the time it takes a human subject to reach a stationary target with or without the information overlay,
3) compare the gesture accuracy (precision in identifying the target) with and without the information overlay.

5. RESULTS

To assess the precision of the algorithm we tracked one target over time and recorded N=9 images at different positions and orientations of the reference image (Fig. 7 a,b). We registered both the initial frame containing the ground truth Ta and the recorded frames with the high resolution fundus image (3543x2236 pixels) (Fig. 7 c). Once both the ground truth and the tracked target Tb were projected onto the same image we computed the distance between their locations; we considered this measurement the error of the tracking algorithm (in pixels). As calculated on a 3543x2236 resolution image, the tracker error was found to be of mean 3.86 pixels (std. dev. 2.25 pixels). Assuming the diameter of the eye varies between 23.5mm and 25mm we can roughly estimate that the error committed by the tracking algorithm would range between 0.04 and 0.044 mm (std. dev. 0.023 - 0.025 mm). It is worth noting that this error represents a sum of both our tracking algorithm and the registration method employed.
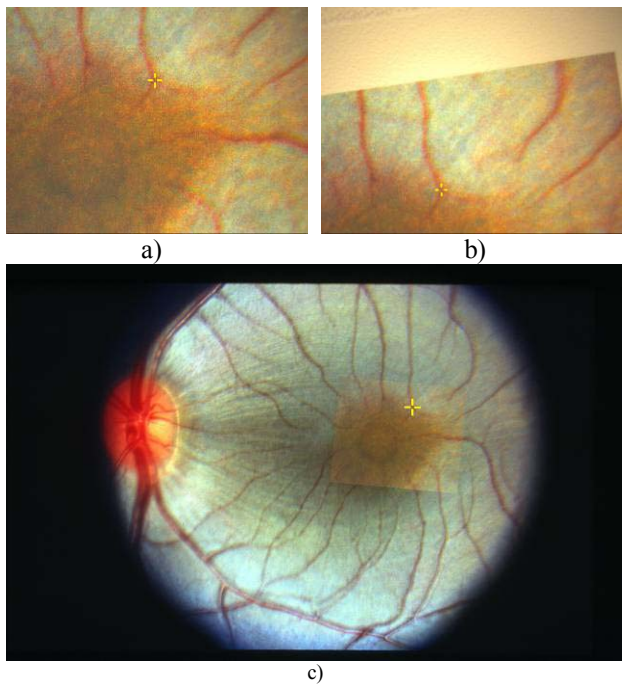
a)   b)

c)

Fig. 7: a) initial position of the target; b) target tracking, after successive translations and rotations; c) registration of image b) with the fundus image.

To compare the gesture accuracy and the targeting time with and without the information enhanced system, the authors performed six targeting trials using the overlay system and six trials without it. We collected one microscope view frame at the beginning and one at the end of each trial (Fig. 8). The targeting time was on average in 8.59 s without the overlay (std. dev. 4.80 s), compared to 8.26 s with the overlay (std. dev. 2.13 s).

All the saved images were registered back with the same high resolution fundus image. The average precision in identifying the target was measured in pixels on the 3543x2236 resolution image. It was found to be of 50.83 pixels (std. dev. 54.57 pixels) without the overlay system and 8.41 pixels (std. dev. 5.1 pixels) with the overlay system, which roughly translates into a range between 0.527mm and 0.583mm without the overlay system compared to 0.087mm and 0.096mm with the overlay system.

While testing our tracking algorithm we came across differences between the chosen retinal phantom and a real microscopic view during retinal surgery. The print out was of high quality but because it was just a printed image of the fundus it lacked the details the surgical microscope reveals. The presence of different textures and minute details is what influenced our choice of a region based tracking method in the first place. The lack of these features can influence the processing speed of the algorithm. To better estimate the performance (frames per second FPS) of the tracking algorithm under conditions closer to vitreoretinal surgery we applied it to a biological model. Leng et al. [13] reported that the chorioallantoic membrane (CAM) of a 12 days old chicken embryo is a good test model system for retinal

surgery since at around this age the secondary veins on the surface of the CAM are of a size that closely matches the size of the veins on the human retina. A previously collected video recording of a CAM vein cannulation procedure (under IRB approval, [14]) was used to asses the tracking algorithm. Under these conditions we noticed a tracking processing speed of 31-33 FPS.
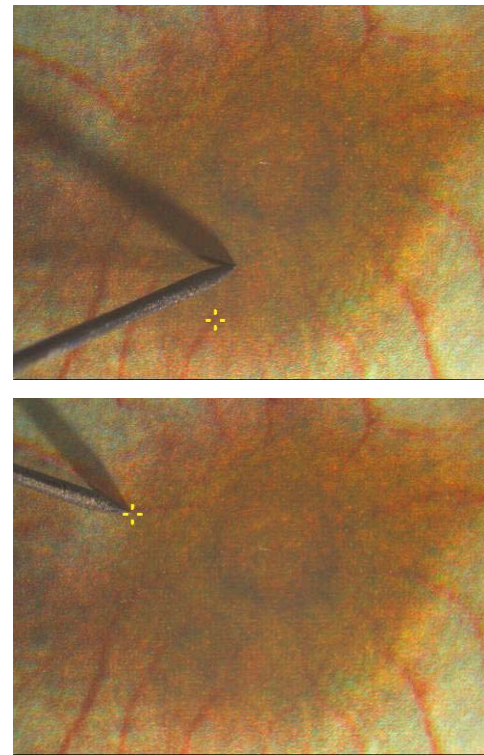

Fig. 8: Microscope view of the surgical instrument reaching target (yellow cross) without overlay (up) and with overlay (down). The difference in precision is easily noticeable.

6. DISCUSSION

The experiments presented in this paper are preliminary but show that the use of an overlay system displaying preoperative data during a microsurgical intervention could significantly improve the vitreoretinal surgical experience. Our framework has the potential to improve the positioning accuracy of surgical instruments. We did not observe a significant decrease in targeting time while using the information overlay given the fact that we had a very small sample size; the authors were the only subjects testing the framework. A potentially significant difference in the targeting time was however observed between the first and last attempt which could be due to a learning curve. In order to obtain more statistically significant results, we are in the process of filling for IRB approval to perform a more extensive usability study. The study will include several trained ophthalmologists and it intends to explore the benefits, drawbacks and learning curve of an information overlay enhanced surgical procedure in a statistically robust way.

For this preliminary proof of concept, the phantom used to model the retina was a printed fundus image. This simplistic model differs from the real clinical situation since it does not take into account the deformation that might occur in reality. We were careful, however to choose registration and tracking algorithms that claim to be robust to the variation of illumination and deformation, which are situations that often occur during surgery. One alternative which we could consider in our future experiments is the use of a biological model (CAM) or an animal model; historically rabbits.

Stewart's registration algorithm which was used both for the pre-op OCT to fundus alignment and pre-op/intra-op fundus to microscope registration can be time consuming depending on machine speed and image size. It must be noted however that this is acceptable since the registrations are either done pre-operatively or only once at the beginning of the intervention. Furthermore, by registering OCT images with fundus images, lesions that were only visible on the OCT cross sections can now be localized first on fundus images and then on the surgeon's microscope view.

The algorithm used during our preliminary investigations supports only single target tracking. As a result, any occlusion of the target would affect the performance of the tracker. We are currently adapting our algorithm to detect occlusions and also to support tracking multiple targets; if one of the targets is occluded, its position could easily be computed using the other targets' transformations.

Concerning the visualization method, we made use of a polarized screen rather than a head mounted display (HMD). The choice was made taking in consideration the low resolution of our available HMD. Using a HMD will be more ergonomic in the OR, allowing the surgeon to see the 3D images without changing position; no head tilt would be necessary. The adaptation of our overlay system to a HMD is very straightforward and we plan on using a higher quality HMD for our future experiments.

7. CONCLUSIONS AND FUTURE WORK

We have developed a new framework for retinal microsurgery which makes use of an information enhanced visual system for computer-assisted target tracking. Although the results are very preliminary, they show that our approach of using an overlay system can improve the gesture accuracy and reduce the operating time. We are working on a more extensive study to further assess performance augmentation at a statistically significant level.

Although in this paper we focused only on presenting the overlay system, extensive long-term work is also being done to improve the feature set of our microsurgical workstation. In addition to the retina surface tracking presented in this paper, this work includes stereo tracking of the surgical instrument and velocity-based virtual fixtures using tool and target localization. This will allow for a precise planning and monitoring of regions that are to be reached or avoided during eye surgery. Work is also being done on making the system OR-compatible by introducing sterilizable and/or disposable components to the Steady Hand Robot.

REFERENCES

[1] ---------, "World Health Organization News Release", [www.who.int/mediacentre/news/releases], 2004.

[2] R.Kumar, G.D.Hager, A.Barnes, P.Jensen, L.L.Whitcomb, and R.H.Taylor.An augmentation system for fine manipulation. In proceeding of medical image computing and computer assisted intervention. *Lecture Notes in Computer Science*,1935 Springer-Verlag: 956–965, 2000

[3] R.H.Taylor, P.Jensen, L.L.Whitcomb, A.Barnes, R.Kumar, D.Stoianovici, P.Gupta, Z.Wang, E.deJuan, and L.Kavoussi. Steady-hand robotic system for microsurgical augmentation. *The International Journal of Robotics Research*, 18(12):1201–1210, 1999.

[4] P. J. Edwards, A. P. King, J. Maurer, C.R., D. A. de Cunha, D. J. Hawkes, D. L. G. Hill, R. P. Gaston, M. R. Fenlon, A. Juczyzck, A. J. Strong, and C. L. Chandler, "Design and Evaluation of a system for microscope-assisted guided interventions (MAGI)," *IEEE Trans Medical Imaging*,19(11), pp. 1082-1093, 2000.

[5] A. P. King, P. J. Edwards, J. Maurer, C.R., D. A. de Cunha, R. P. Gaston, M. Clarkson, D. L. G. Hill, D. J. Hawkes, M. R. Fenlon, A. J. Strong, T. C. S. Cox, and M. J. Gleeson, "Stereo augmented reality in the surgical microscope," *Presence: Teleoperators and Virtual Environments*,9(4), 2000.

[6] M. Lievin and E. Keeve, "Stereoscopic augmented reality system for computer assisted surgery," *Proceedings Computer Assisted Radiology and Surgery*, pp. 27-30, 2001

[7] M. Bajura, H. Fuchs, and R. Ohbuchi, "Merging virtual objects with the real world: seeing ultrasound imagery within the patient," in *Computer-Integrated Surgery*, R. H. Taylor, S. Lavallee, G. Burdea, and R. Mosges, Eds. Cambridge, Mass.: MIT Press, 1996, pp. 245-254.

[8] C. V. Stewart, C.-L. Tsai, and B. Roysam, "The Dual Bootstrap Iterative Closest Point Algorithm with Application to Retinal Image Registration," *IEEE Trans on Medical Imaging*,22(11), pp. 1379-1394, 2003.

[9] H. Shen, C. V. Stewart, B. Roysam, G. Lin, and H. L. Tanenbaum, "Frame-rate spatial referencing based on invariant indexing and alignment with application to on-line retinal image registration," *IEEE Trans. on Pattern Analysis and Machine Intelligence*,25(3), pp. 370-384, 2003.

[10] C. V. S. A. Can, B. Roysam, and H.L. Tanenbaum, "A feature-based algorithm for joint, linear estimation of high-order image-to-mosaic transformations: mosaicing the curved human retina," *IEEE Transactions on Pattern Analysis and Machine Intelligence*,24, pp. 412-419, 2002.

[11] P. J. Saine, M. E. Tyler, Ophthalmic Photography: Retinal Photography, Angiography, and Electronic Imaging, 2nd Edition, Butterworth-Heinemann Medical; ISBN: 0750673729

[12] G. Hager and P. Belhumeur, "Efficient Region Tracking with Parametric Models of Geometry and Illumination," *IEEE Trans Pattern Analysis and Machine Intelligence*,20(10), pp. 1125-1139, 1998.

[13] T. Leng, J.M. Miller, K.V. Bilbao, D.V. Palanker, P. Huie, and M.S. Blumenkranz, "The chick chorioallantoic membrane as a model tissue for surgical retinal research and simulation", *Retina*, 24(3):427–434, June 2004.

[14] B. Mitchell, J. Koo, I. Iordachita, P. Kazanzides, A. Kapoor, J. Handa, G. Hager, R. Taylor, "Development and application of a new Steady-Hand manipulator for retinal surgery," *IEEE Conference on Robotics and Automation* , (623-629), 2007