



Published in final edited form as:

Proc SPIE Int Soc Opt Eng. 2015 ; 9415: 94150C-. doi:10.1117/12.2082872.

Toward real-time endoscopically-guided robotic navigation based on a 3D virtual surgical field model

Yuanzheng Gong^{*,a}, Danying Hu^b, Blake Hannaford^b, and Eric J. Seibel^a

^aHuman Photonics Lab, Dept. of Mechanical Engineering, Univ. of Washington, Seattle, WA 98195

^bBiorobotics Lab, Dept. of Electrical Engineering, Univ. of Washington, Seattle, WA 98195

Abstract

The challenge is to accurately guide the surgical tool within the three-dimensional (3D) surgical field for robotically-assisted operations such as tumor margin removal from a debulked brain tumor cavity. The proposed technique is 3D image-guided surgical navigation based on matching intraoperative video frames to a 3D virtual model of the surgical field. A small laser-scanning endoscopic camera was attached to a mock minimally-invasive surgical tool that was manipulated toward a region of interest (residual tumor) within a phantom of a debulked brain tumor. Video frames from the endoscope provided features that were matched to the 3D virtual model, which were reconstructed earlier by raster scanning over the surgical field. Camera pose (position and orientation) is recovered by implementing a constrained bundle adjustment algorithm. Navigational error during the approach to fluorescence target (residual tumor) is determined by comparing the calculated camera pose to the measured camera pose using a micro-positioning stage. From these preliminary results, computation efficiency of the algorithm in MATLAB code is near real-time (2.5 sec for each estimation of pose), which can be improved by implementation in C++. Error analysis produced 3-mm distance error and 2.5 degree of orientation error on average. The sources of these errors come from 1) inaccuracy of the 3D virtual model, generated on a calibrated RAVEN robotic platform with stereo tracking; 2) inaccuracy of endoscope intrinsic parameters, such as focal length; and 3) any endoscopic image distortion from scanning irregularities. This work demonstrates feasibility of micro-camera 3D guidance of a robotic surgical tool.

Keywords

Image-guided surgery; 3D surgical field; keyhole surgery; machine vision; robotic navigation; neurosurgery

1. INTRODUCTION

Complete surgical resection of tumor tissue remains one of the most important factors for survival in patients with cancer. Tumor resection in the brain is exceptionally difficult

*Address correspondence to: gong7@u.washington.edu or yzong7@gmail.com..

because leaving residual tumor tissue leads to decreased survival and removing normal healthy brain tissue leads to life-long neurological deficits. Since residual brain tumors can be visibly indistinguishable from healthy brain tissue, invisible during white-light endoscopic imaging, fluorescence is used to provide enhanced contrast. This intraoperative fluorescence guided resection has the potential to improve the extent of resection for the most difficult brain tumors¹.

The scanning fiber endoscope (SFE) is unique in the ability to guide interventions using both wide-field reflectance and fluorescence imaging^{2,3}. The multiple modes of video-rate imaging are at 500-600 lines of resolution at 30 Hz. The small size of the rigid tip (1.6-mm diameter and 9-mm long) and highly flexible and 3-meter long shaft provide key-hole access to surgical fields. By robotically moving the tiny SFE or micro-camera, three-dimensional (3D) virtual reconstructions can be made for minimally-invasive disease surveillance⁴, guiding biopsy², or planning surgical procedures³. By attaching the SFE to a calibrated robotic arm and scanning across a phantom of a cavity from debulked brain tumor with fluorescence targets representing residual tumor tissue, accurate 3D virtual reconstructions of the surgical field can be generated⁵.

Once the fluorescence target is located in the 3D space, the SFE-guided surgical tool tip from the scope/tool set needs to be safely navigated against the target for the residual tumor clean-up operation. The challenges are providing safe, accurate and efficient (space & time) 3D-image guidance that can run in tandem with a surgeon's visual feedback control. The goal is to provide assistance to the surgeon for the tedious and repetitive process of cleaning up residual tumor cells by their fluorescence glow captured by the SFE multimodal imaging which can be automated in the future; an example is the automation of repeated suturing tasks in robotic-assisted surgeries.

To robotically assist the surgeon performing a more automated surgery, accurate 3D navigation of the surgical tool is required, especially for brain tumor margin removal. In this project, the surgical robot is the RAVEN II open-source platform which is designed for the surgeon to provide precise and accurate tool guidance based on external high-quality optical imaging⁶. Without the surgeon in the loop, the RAVEN II currently does not provide accurate closed-loop servo control using the mechanical linkages and sensors within the robot arm, requiring an external measurement system for tracking the tip of the robot arm. To provide 0.5-mm accuracy of the scope/probe set during RAVEN procedures with a phantom of the debulked tumor cavity, an external stereo vision tracking system has been added⁵, but limitations restrict its use to more open surgical procedures by requiring constant environment lighting and direct line-of sight to the tip of the scope/tool set. Thus a new navigational system is required when the scope/tool set enters the debulked surgical cavity for residual tumor removal.

Figure 1 shows the scheme of 3D image-guided surgery of residual tumor clean-up after debulking and making a cavity. The goal is to provide safe, rapid, and accurate navigation from outside the cavity to near contact between the surgical tool and a selected tumor fragment that is fluorescently labeled. The approach of the scope/tool set is illustrated in a series of three steps in Figure 1. The first step is generating an accurate, multimodal

(fluorescence and reflectance) 3D virtual model of the debulked cavity. This step has been accomplished using a phantom and a dual-camera stereo tracking system of the SFE/tool set⁵. As this scope/tool set is raster scanning over the surgical field, the scope/tool set is in the vision field of the stereo tracking system, shown as series numbers 1 to 2 in Fig. 1. An accurate multimodal 3D virtual model of the surgical field can be reconstructed by utilizing structure from motion (SfM) algorithm and the accurate pose information of scope from stereo vision⁵. This reconstructed 3D virtual model of the debulked tumor cavity is considered as pre-known in this study. As the surgical tool enters into the cavity and approaches the region of interest (steps 2 to 3), the ability to track the tool by an external stereo system is lost because the half-closed surgical field blocks the stereo vision. Without the stereo vision as close-loop control signal, the surgical robot cannot be automatically guided to the correct tumor location selected by the surgeon from the 3D virtual model of the debulked tumor cavity. The surgeon could use the video image from the SFE and manually drive the surgical tool to the tissue surface by the RAVEN interface, but the surgeon lacks situational awareness from the limited scope field-of-view and clear depth cues from the non-stereo SFE for rapid and safe navigation. Nonetheless, by registering the 2D live endoscopic image to the pre-known 3D virtual model, the pose of the scope can be recovered in near real-time, as well as the pose of the surgical tool since the spatial relation of tool and scope is fixed. Thus, an optimal surgical path of the tool can be designed to safely, rapidly, and accurately guide the tool to the point of tool-tissue interaction (step 3 in Figure 1). At this final step, the surgeon can guide residual tumor cleanup within the limited field of view of the scope. In the future, this scope-based fluorescence removal can also be made a semi-autonomous robotic surgical procedure using machine vision⁷.

The rest of the paper is structured as follows. In section 2, the conventional feature-based 3D reconstruction algorithm is reviewed and the constrained reconstruction algorithm is proposed and discussed. In section 3, the experiment of 3D image-guided surgery is set up with SFE and surgical tool on micro-positioning stage. The registration between endoscopic image and pre-reconstructed 3D model is performed. Section 4 provides a table with all the details of the experimental results, demonstrating that the proposed algorithm successfully recovers the position and orientation of the endoscope. Conclusion and discussion of the advantages and disadvantages of this 3D image-guided surgery approach are given in Section 5, as well as the future work.

2. METHOD

In our previous work, an accurate multimodal 3D reconstruction of the surgical field (phantom of debulked brain tumor) was generated on RAVEN II surgical robotic system with stereo tracking system⁵. The 3D coordinates of the tumor cells are used to navigate the surgical robot for the following surgical operations. However, without the tracking of stereo vision, such as in the case of minimally invasive keyhole surgery or changing environment lighting conditions, the surgical robot RAVEN was unable to automatically and efficiently approach to the correct 3D location of tumor cells. To continue controlling the surgical tool in a close-loop behavior, a 3D image-guided technique is proposed in this study by registering the 2D endoscopic image with the pre-reconstructed accurate 3D virtual model.

In this study, the scale-invariant feature transform (SIFT) feature detection algorithm is utilized to find the features in the endoscopic image and also the stitched image of the 3D virtual model. Random sample consensus (RANSAC) algorithm is then applied to find the corresponding features between these two images. By providing the accurate 3D virtual model, the 3D coordinates of these matching features can be easily retrieved.

Bundle adjustment is an optimization technique, simultaneously refining the 3D coordinates of all the feature points describing the object, as well as the camera parameters (focal length, center pixel, distortion, position or/and orientation)⁸. Bundle adjustment is often used as the last step in feature-based 3D reconstruction, following the prior steps of feature detection and feature matching. Once we have the 3D coordinates of the feature points and their projections in the 2D endoscopic image, a constrained bundle adjustment algorithm is achieved^{9,10}.

Considering an object with a set of n 3D feature points $\mathbf{Q}_{3 \times n} = \{\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_n\}$ in world coordinate, one image was captured by the SFE at unknown location and orientation, which need to be solved in this study to track the pose of the surgical tool. Let \mathbf{R} represents the rotation matrix and t denotes the translation vector of the current SFE. Thus, the calculated pixel location of its re-projection on the endoscopic image is $p_i (= [x_i, y_i]^T)$:

$$\omega \begin{bmatrix} p_i \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{R} & t \end{bmatrix} \begin{bmatrix} \mathbf{Q}_i \\ 1 \end{bmatrix} \quad (1)$$

\mathbf{K} is the camera calibration matrix including focal lengths, center pixel, first-order and second-order radial distortion coefficients.

Let the position of the observed feature is $\tilde{p}_i (= [\tilde{x}_i, \tilde{y}_i]^T)$, then the re-projection error $r = p - \tilde{p}$ was calculated and rearrange as a column vector, containing the every re-projection error. The conventional bundle adjustment algorithm is given as

$$\text{Objective} \quad \min f(s) = \frac{1}{2} r(s)^T r(s) \quad (2)$$

where s contains many variables, such as 3D coordinates of each point ($3n$), the camera position and orientation (6) and also the identical camera intrinsic parameters (6). In total, the number of variables $N = 3n + 6 + 6$. However, in our case, the bundle adjustment algorithm is constrained by knowing the 3D coordinates of each feature and the intrinsic camera parameters from calibration. Thus, the optimization problem is dramatically simplified, only solving the pose parameters of the camera.

Solving this constrained bundle adjustment can be achieved in real-time. This helps to track the pose of surgical tool in real-time as it approaches the tumor target. The recovered pose information can be used to navigate the robot arm to follow a certain operation path.

3. EXPERIMENT

Prior to experimentation on the RAVEN surgical robot platform, the 3D image-guided surgery technique is performed and tested on a micro-positioning stage as a proof-of-concept study. The micro-positioning stage we used in this experiment is ULTRAlign™ precision integrated crossed-roller bearing linear stage (model 462-xyz-m), on which a high-performance rotation stage is mounted. The high-accurate position and orientation information can be used to analyze the performance of the pose recovery by the proposed algorithm.

The same 3D printed phantom in previous study⁵ was used to mimic the surgical field with realistic surgical features and multiple scattered fluorescence targets, shown in Fig. 2(a). Its accurate multimodal 3D model of surgical field was previously generated on RAVEN surgical robot platform with accurate camera pose information from externally applied stereo vision⁵. This step was achieved as the scope/tool set is raster scanning above the surgical field. The reconstructed 3D model was considered as prior knowledge in this proof-of-concept study.

To acquire the real-time endoscopic image in the surgery procedure, a 1.6-mm multimodal SFE was affixed along with the mock surgical tool, shown in Fig. 2(b). This scope/tool set was then attached on the micro-positioning stage; see Fig. 2(c), which can provide accurate position and orientation information as the ground truth to analyze the feasibility and performance of the algorithm. Since the rotation stage only provides rotation motion along one axis, we restrained the rotation along the other two axes. That means the number of unknown variables in our study is $N=4$, containing 3 position and 1 orientation parameters of the SFE scope tip. This concept will be extended to surgical robot RAVEN II for more application-specific development, such as testing autonomous surgical procedures.

The estimation of the pose of SFE by registering a 2D image to a 3D virtual model can be achieved in two stages. The first stage is to find the matching features between them. SIFT and RANSAC algorithms were utilized to find the matching points from a video frame and the stitched image of 3D model. The second stage is to refine the pose parameters by these matching points with the constrained bundle adjustment approach that was briefly described in the methodology section since the 3D scene and its 2D projection were known.

Fig. 3(a) shows an example of endoscopic frame with surgical tool blocking part of the image, since the tool was extended along the optical axis of the SFE as the experiment setup. To reduce the effect of the tool signal and computation time, a mask image was generated by edge detection, shown in Fig. 3(b). The white part of the mask represents the useful information that we used to register with the stitched image of the 3D virtual model. The masked SFE reflectance image was shown in Fig. 3(c). Fig. 3(d) shows the masked fluorescence image which is co-registered with the SFE reflectance image. It can be utilized to check if there is any tumor left in the procedure of interoperative surgery.

By registering all the SIFT features between the stitched image and masked endoscopic live video frame, the matching points were detected and lined up with colorful lines, shown in Fig. 4(a). The corresponding 3D coordinates of these features in the 3D space were retrieved

from the pre-reconstructed virtual model, shown in Fig. 4(b). Once the 2D image location and 3D coordinates of these features were obtained, the position and orientation of endoscope can be recovered by Eq. (2). All features are found in the reflectance images only, but when there is fluorescence present in the co-registered fluorescence SFE image, this green color is superimposed over the reflectance color images. Note, the 3D model shows that fluorescence markers were detected and highlighted in the real-size 3D phantom, representing the residual tumor after the debulking procedure, see the green (fluorescence) targets in Fig. 4(b), from which an optimal path of autonomous surgery can be designed.

To test the accuracy and robustness of the recovery of camera position and orientation, five arbitrary poses of the tool/SFE were selected and read from the micro-positioning stage as ground truth. Five reflectance images were captured to register with the pre-reconstructed 3D virtual model. The SIFT and RANSAC algorithms with the same set of parameters were utilized to find the matching points. Taking the image in Fig. 4 as an example, 92 pairs of matching points were found with some field-of-view being blocked by the tool; shown as blue dots in Fig. 4(b), although only 1/10th of matching points were randomly selected for better visualization in Fig. 4(a).

4. RESULTS

The estimation of the five poses was achieved successfully with our constrained bundle adjustment algorithm. All software was implemented in MATLAB, running on a workstation Dell Precision M4700 with 2.7 GHz Intel i7-3740QM CPUs, 20.0 GB memory in a 64-bit Windows operating system. The difference of pose resulted in various computation details, such as number of matching points, computation time, and distance and orientation error to the ground truth, see Table 1.

The difference of number of matching points is caused by the difference of perspective and also the blocked field of view in the endoscopic images. The computation time was relatively stable, always less than 2.7 seconds, and majority of computation time was devoted to finding the matching points between captured video frame and reconstructed 3D model. Experiments showed that the constrained bundle adjustment took only 0.3 sec on average, roughly 10% of the total computation time. The distance between the estimated position and measured one was 3.09 mm on average with 0.043 mm as the standard deviation based these five trials. The orientation error was 2.44 degrees on average with standard deviation of 1.23 degrees. These errors are also consistent with approaching a target in three steps as illustrated in Fig. 1, while more work is required to determine if higher resolution images from a closer endoscope can reduce approach errors while updating the 3D model in near real-time.

5. CONCLUSION AND DISCUSSION

Advantages of providing a 3D image-guided approach to navigate a surgical tool in a minimally-invasive surgery like residual brain tumor clean-up are listed: **1) Optimize the surgery:** Comparing to conventional 2D endoscopic-guided surgery, a 3D virtual model of surgical field provides depth information, which can be used to assist the surgeon to plan the

surgical procedure, or an autonomous medical robot to generate an optimal surgical/therapy path. **2) Secure the safety of surgery:** Depth information in the 3D image-guided surgery can define a safe working space, so that the damage of the tissue or the endoscope can be avoided as the surgical tool is moving around in the surgical field. **3) Assist the key-hole surgery:** An external stereo optical system can be utilized to track the pose of surgical tool and/or endoscope in open surgery for semi-autonomous tumor ablation⁷, but this is not feasible within enclosed surgical fields, such as key-hole surgeries. The proposed 3D image-guided technique in this paper can recover the position and orientation information of the endoscope (eye on robot arm) in near real-time. **4) Refine the 3D virtual model of surgical field by adding more and updated endoscopic images:** the initial 3D model was generated by raster scanning over the surgical field. New higher-resolution detail can be added to refine the 3D model as the tool is moving towards to the tumor with updates after tissue removal. **5) Adjust the motion of surgical tool:** As the pose of the tool in 3D virtual model frame is known more rapidly in the future, the precise control of the surgical tool can be achieved by taking the recovered pose information as closed-loop feedback.

Recent work in accurately aligning the neurosurgical 3D space intraoperatively after the brain shifts during skull opening comes from the laboratory of Dr. Michael Miga¹¹. Using a phantom of brain surgery, the 3D positional target registration error was recorded for 5 repeated measures using laser range scanning and conoscopic holography. The results of 2.1 mm \pm 0.2 mm error for laser range scanning and 1.9 mm \pm 0.4 mm error for the conoscopic holography compare favorably with our preliminary results. Although both 3D measurement systems are laser-based and provide greater accuracy than our approach, their large size restricts their use to open surgery. In our case, the laser-based micro-camera is even smaller in size than most surgical tools, which allows introduction through less invasive keyholes and registration from within these semi-enclosed cavities. A general goal of surgical robotics is the reduction of size and advancement of keyhole applications¹², which this new navigational approach supports. A future goal is the automation of the tedious task of residual tumor cleanup which is expected to be used in applications where removing every cancer cell (and fewest number of healthy cells) is critically important, such as brain surgeries.

This proof-of-concept study successfully recovered the position and orientation in near real-time using a single, ultrathin, flexible, and multimodal endoscope (attached to a surgical tool that blocks a portion of the visual field) from the 3D virtual model space of simulated surgical field. Future work is improving the speed and accuracy of the real-time SFE video registration to the 3D virtual model using phantoms that allow resection and the RAVEN surgical robot. This work builds toward automating residual tumor cleanup in robotically-assisted keyhole surgeries.

ACKNOWLEDGEMENT

Funding from NIH NIBIB R01 EB016457 “NRI-Small: Advanced biophotonics for image-guided robotic surgery” PI: Dr. Eric Seibel and Dr. Blake Hannaford, (co-I). The authors appreciate Richard Johnston and David Melville at Human Photonics Lab, as well as Di Zhang at Rapid Manufacturing Lab, University of Washington for their technical support.

REFERENCES

- [1]. Liu JC, Meza D, Sanai N. Trends in fluorescence image-guided surgery for gliomas. *Neurosurgery*. 2014; 75(1):61–71. [PubMed: 24618801]
- [2]. Yang C, Hou VW, Girard EJ, Nelson LY, Seibel EJ. Target-to-background enhancement in multispectral endoscopy with real-time background autofluorescence mitigation for quantitative molecular imaging. *Journal of Biomedical Optics*. 2014; 19(7):076014.
- [3]. Gong Y, Soper TD, Hou VW, Hu D, Hannaford B, Seibel EJ. Mapping surgical fields by moving a laser-scanning multimodal scope attached to a robot arm. *Image-Guided Procedures, Robotic Interventions, and Modeling, SPIE Medical Imaging, Proc. SPIE 9036*. 2014:90362S-1–8.
- [4]. Burkhardt MR, Soper TD, Yoon WJ, Seibel EJ. Controlling the trajectory of a scanning fiber endoscope for automatic bladder surveillance. *Mechatronics, IEEE/ASME Transactions on*. 2014; 19(1):366–373.
- [5]. Gong Y, Hu D, Hannaford B, Seibel EJ. Accurate 3D virtual reconstruction of surgical field using calibrated trajectories of an image-guided medical robot. *Journal Medical Imaging*. 2014; 1(3): 035002–035002.
- [6]. Hannaford B, Rosen J, Friedman DW, King H, Roan P, Cheng L, Glozman D, Ma J, Kosari SN, White L. Raven-II: an open platform for surgical robotics research. *Biomedical Engineering, IEEE Transactions on*. 2013; 60(4):954–959.
- [7]. Hu, D.; Gong, Y.; Hannaford, B.; Seibel, EJ. Semi-autonomous simulated brain Tumor ablation with Raven II surgical robot using behavior tree. *IEEE International Conference on Robotics and Automation (ICRA)*; Seattle, WA. May 26th – 30th, 2015; (in press)
- [8]. Triggs, B.; McLauchlan, PF.; Hartley, RI.; Fitzgibbon, AW. In *Vision algorithms: theory and practice*. Berlin Heidelberg; Springer: 2000. Bundle adjustment—a modern synthesis; p. 298-372.
- [9]. Wong KH, Chang MMY. 3D model reconstruction by constrained bundle adjustment. *ICPR IEEE*. 2004; 3:902–905.
- [10]. Gong Y, Meng D, Seibel EJ. Bound constrained bundle adjustment for reliable 3D reconstruction. *Optics Express*. (under review).
- [11]. Simpson AL, Sun K, Pheiffer TS, Rucker DC, Sills AK, Thompson RC, Miga MI. Evaluation of conoscopic holography for estimating tumor resection cavities in model-based image-guided neurosurgery. *IEEE Transactions on Biomedical Engineering*. 2014; 61(6):1833–1843. [PubMed: 24845293]
- [12]. Marcus H, Nadi D, Darzi A, Yang GZ. Surgical robotics through a keyhole: from today’s translational barriers to tomorrow’s “disappearing” robots. *IEEE Transactions on Biomedical Engineering*. 2013; 60(3):674–681. [PubMed: 23372075]

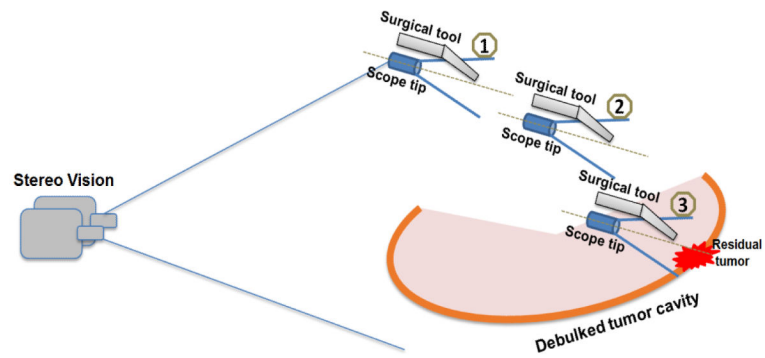


Figure 1. The scheme of image-guided surgery based on a 3D virtual model. The position and orientation of the scope/tool set can be recovered in near real-time as the tool approaches to the target, by registering 2D endoscopic video frames to the pre-reconstructed 3D virtual surgical field model. The optimal path of surgical tool can be designed and re-adjusted towards the residual tumor, showing with series of steps 1 to 3.

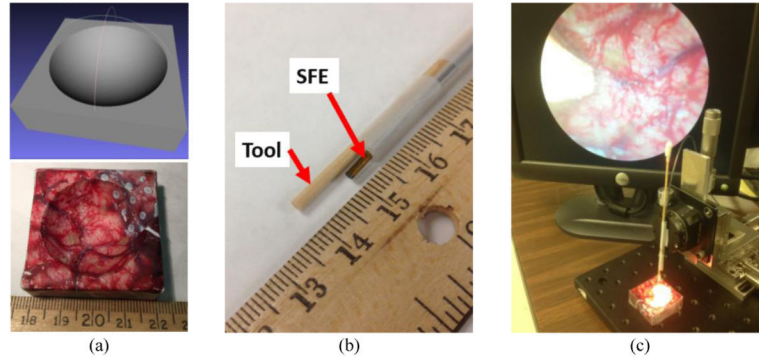


Figure 2. The experimental setup with a micro-positioning stage. (a) A phantom of known geometry was CAD-designed (up) and 3D-printed (bottom) to represent the debulked brain tumor cavity, and its accurate 3D virtual model was reconstructed as a pre-knowledge. (b) A multimodal SFE of 1.6mm outer diameter was affixed along a surgical tool. The spatial relation between tool and forward-viewing endoscope is fixed as shown. (c) The tool with SFE was attached on a micro-positioning stage, which can provide high-accurate pose data. The mock surgical tool tip is shown in the endoscopic video.

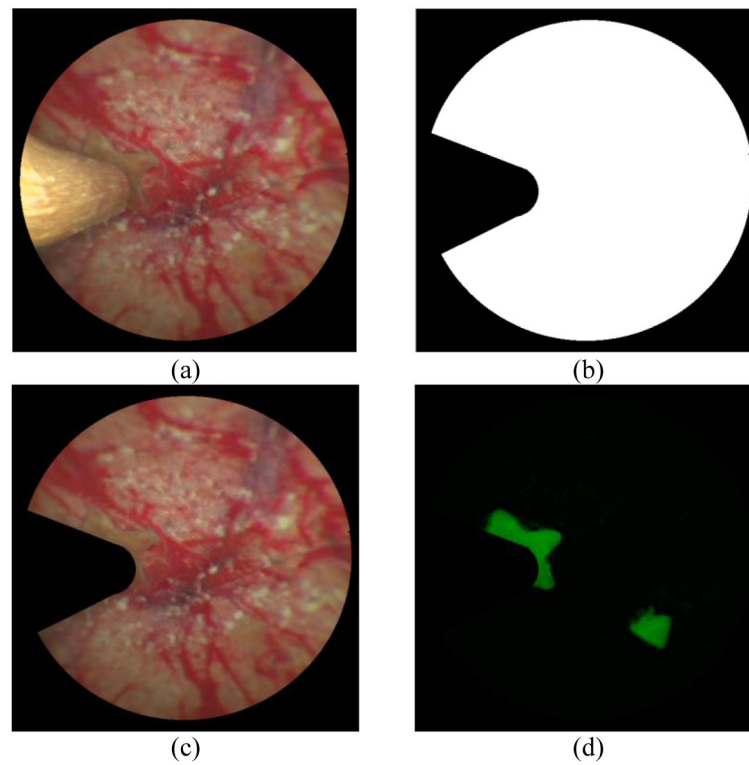


Figure 3.

Image processing procedure for an endoscopic frame with surgical tool. (a) An endoscopic frame with the surgical tool blocking part of the reflectance image. (b) The mask image of the endoscopic frame. (c) The endoscopic image only containing the color reflectance from the phantom. (d) The endoscopic image only containing the fluorescence signal from the phantom.

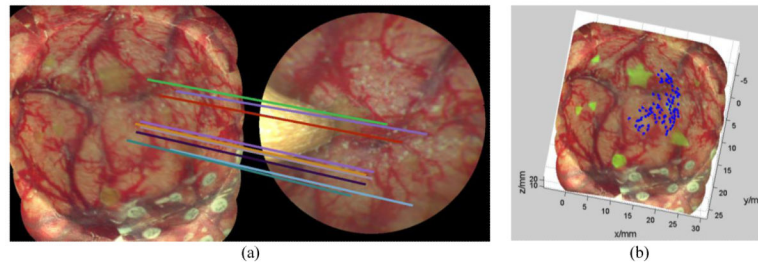


Figure 4. The procedure of finding the 3D coordinates of matching features. (a) Matching features between 3D reconstructed virtual model of phantom (left) and a single SFE video frame (right), with feature alignment shown by the multiple colored lines. (b) The 3D virtual model of surgical field with matching features to SFE video frame labeled in blue. The 3D coordinates of the features can be retrieved from the virtual model.

Table 1

The experiments results of five different poses of SFE-image-guided navigation to 3D virtual model.

	Number of matching points	Computation time	Distance Error	Orientation Error
Pose 1	54	2.31 sec	3.12 mm	1.47 degree
Pose 2	48	2.45 sec	3.08 mm	1.52 degree
Pose 3	92	2.67 sec	3.11 mm	2.64 degree
Pose 4	68	2.66 sec	3.06 mm	4.47 degree
Pose 5	84	2.27 sec	3.00 mm	2.10 degree
Average	(NA)	(NA)	3.09 mm	2.44 degree
Std.	(NA)	(NA)	0.043 mm	1.23 degree

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript