EyeSLAM: Real-Time Localization and Mapping of Retinal Vessels during Intraocular Microsurgery

Daniel Braun, Sungwook Yang *Member*, *IEEE*, Cameron N. Riviere, *Senior Member, IEEE*, and Brian C. Becker

***Abstract*—Intraocular microsurgery with robotic assistance aims to boost surgeon performance when operating on delicate sub-millimeter structures of the eye. Fast and accurate mapping and localization of retinal vasculature is critical to increasing the aid that robots can provide during challenging procedures such as photocoagulation and cannulation. We present EyeSLAM, an algorithm that delivers 30 Hz real-time simultaneous localization and mapping of the human retina and vasculature during intraocular surgery. EyeSLAM combines fast vessel detection with fast 2D scan-matching techniques to build and localize a probabilistic map of the vasculature. The algorithm improves upon previous work by using a scan-based localization algorithm for higher convergence rate and increased robustness to quick motion, and more robust vessel detection. We demonstrate that even in the harsh imaging environment of retinal surgery with high magnification, quick shaky motions, textureless retina background, variable lighting, and tool occlusion, EyeSLAM can map 75% of the vessels within two seconds of initialization and localize the retina with RMS error of under five pixels (translation) and one degree (rotation).**

***Index Terms*—Medical imaging, vessel detection, retinal registration, localization and mapping (SLAM), vitreoretinal surgery.**

I. Introduction

Intraocular microsurgery is often regarded as particularly demanding due to the extraordinary precision required to manipulate the small, delicate structures of the retina. The confounding influence of physiological tremor on the surgeon's micromanipulation ability and the challenging nature of the surgical access further increase difficulty [1], [2]. Routine procedures such as membrane peeling require the surgeon to manipulate anatomy less than 10 μm thick [3]–[5], and laser photocoagulation operations benefit from accurate placement of laser burns [6], [7]. Promising new procedures such as vessel cannulation necessitate precise and exacting micromanipulation to inject anticoagulants into veins less than 100 μm in diameter [8], [9].

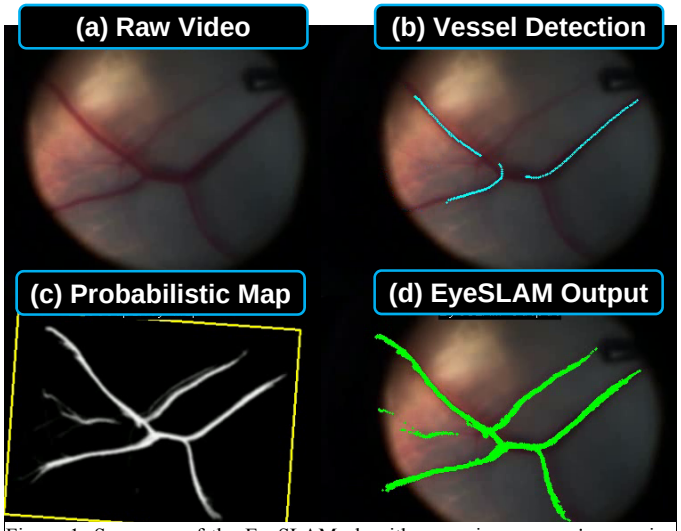


Figure 1: Summary of the EyeSLAM algorithm running on *ex vivo* porcine retina under surgical microscope. (a) Raw input video of retina during surgery via high-magnification microscope is run through (b) fast vessel detection, forming noisy observations which is used to build and localize (c) a 2D occupancy map representing probabilities of vasculature at each point (with yellow box showing the localized camera view), yielding (d) EyeSLAM output of the full vessel map localized to the current frame.

To address micromanipulation challenges in retinal surgical procedures, a variety of assistive robots have been proposed. Master/slave robots developed for eye surgery include the JPL Robot Assisted MicroSurgery (RAMS) system [10], the ocular robot of Ueta et al. [11], and the multi-arm stabilizing micromanipulator of Wei et al. [12]. Retinal surgery with the da Vinci master/slave robot has been investigated [13] and led to the design of a Hexapod micropositioner accessory for the da Vinci end-effector [14]. The Johns Hopkins SteadyHand Eye-Robot [15] shares control with an operator who applies force to the instrument while it is simultaneously held by the robot arm. A unique MEMS pneumatic actuator called the Microhand allows grasping and manipulation of the retina [16]. The Microbots of Dogangil et al. aim to deliver drugs directly to the retinal vasculature via magnetic navigation [17]. A lightweight micromanipulator developed in our lab, Micron, is fully handheld and has been used with vision-based control to aid retinal surgical procedures [18], [19].

While classic robot control can provide general behaviors such as motion scaling, velocity limiting, and force regulation, more specific and intelligent behaviors require knowledge of the anatomy. Vision-based control combines visual

Manuscript received Sept 4th, 2015. This work was supported in part by the National Institutes of Health (grant no. R01 EB000526 and R01 EB007969), the National Science Foundation (Graduate Research Fellowship), and the ARCS Foundation.

Daniel Braun is with the Mechatronics Department, Institut National des Sciences Appliquées of Strasbourg, and with Télécom Physique Strasbourg, 67000 Strasbourg, France.

S. Yang is with the Center for BioMicrosystems, Korea Institute of Science and Technology, Seoul, 02792, South Korea

C. N. Riviere and B. C. Becker and are with the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: camr@ri.cmu.edu).

information of the anatomy with robotic control to enforce tip constraints, or virtual fixtures, which enact task-specific behaviors and provide guidance to the surgeon [20], [21]. In retinal vessel cannulation, knowledge of the vessel location relative to the instrument tip can aid robotic behavior and more effectively help guide the robot during injections into the vessel [22], [23]. During retinal laser photocoagulation, placing burns on retinal vessels should be avoided as this can occlude the vein, possibly causing vitreous hemorrhage [7]. However, most existing methods for vessel detection and retinal registration are not suited to real-time operation, often preferring accuracy over speed for offline use, and do not handle constraints required for intraocular surgery, such as challenging lighting or robustness to tool occlusion.

Becker and Riviere introduced a formulation of simultaneous localization and mapping for retinal vasculature that ran in real-time using fast vessel detection and smoothing noisy observations over time while building and localizing to a map [24]. In this paper, we have extended this work. The new algorithm, called EyeSLAM, offers improved mapping and localization of the retinal vasculature in real time for use in intraocular surgery. It exhibits robustness to variable illumination conditions, high magnification, quick shaky motion, textureless retina, and partial occlusions. Specifically, this paper introduces two new components of the EyeSLAM algorithm: (a) a scan-based localization algorithm to a dynamic map for higher convergence accuracy and increased tolerance of quick motion; and (b) more robust vessel detection with better rejection of spurious detections. The paper also improves upon [24] by including quantitative results showing the effectiveness of EyeSLAM on simulated and recorded surgical video imagery, along with qualitative demonstrations of its usefulness in simulated surgical scenarios.

In Section II, related work in simultaneous localization and mapping (SLAM) and in retinal registration is described. Section III describes our approach of using the fast retinal vessel detection of [25] for feature extraction, an occupancy grid for mapping [26], and fast multi-scale scan matching [27] for localization. In Section IV, we evaluate our approach on videos of paper slides, porcine retina *ex vivo* (see Fig. 1), and human retina *in vivo*, and provide a comparison with our previous work [24].

II. Related Work

A wealth of published works related to localization and mapping of retinal vessels exists and is summarized in Table I. Most can be grouped into three general categories: vessel detection, retinal registration, and the more general robotic approach of simultaneous localization and mapping (SLAM).

A. Vessel Detection

Vessel detection is the process of extracting vasculature in retinal imagery and often includes calculating the centerlines, width, and orientation of vessels. One set of methods uses local color and intensity information to classify the image on a per-pixel basis [28]–[31]. Another popular approach is to search across the image for vessel-like structures using matched filters at various locations, scales, and orientations [25], [32], [33]. Other algorithms use a bank of Gabor wavelets to do a pixel-wise classification of the image [34]–[36]. However, most focus on offline analysis of low-magnification, wide-area images such as fundus images where accuracy is prioritized over speed. With the exception of speed-focused algorithms such as [25], [31], [36] and other hardware-accelerated methods [37]–[39], most algorithms require more than 1 s to run, which is insufficiently fast for surgical robotic control loops, which often require 10-60 Hz.

One notable exception is the rapid exploratory algorithm of Can et al. [25] that traces the vasculature, yielding a monotonically improving set of partial results suitable for real-time deployment at 30 Hz. Can et al. [25] achieves high-speed vessel detection through a very fast sparse initialization followed by a tracing algorithm. First, a fast search for vessel points along a coarse grid is performed to initialize a set of seed points on vessels. Each seed point, or detected candidate vessel, is then explored in both directions along the vessel with an approximate and discretized matched filter. At each iteration, the best fit for location, orientation, and width of the vessel centerline is estimated through the evaluation of several matched filters. Using orientation estimates to initialize the next iteration, the network of vessels in the image is traced without having to evaluate areas lacking vessels. Because only a small fraction of the total number of pixels in the image is ever processed, most of the vessels can be detected very rapidly. However, the vessel detections of [25] are usually noisier and less complete than slower, more computationally expensive methods.

B. Retinal Image Registration

Numerous approaches to registering, mosaicking, and tracking exist to take a sequence of retinal imagery and calculate relative motion between images. In general, approaches match one or more of several features between images: key points, vasculature landmarks, or vasculature trees. Key point algorithms use image feature descriptors such as SIFT [40] to find and match unique points between retinal images [41]–[44]. Vasculature landmark matching algorithms find distinctive points based on vessel networks, such as vein crossings or bifurcations, and match custom descriptors across multiple images [45]–[47]. Other approaches augment or eschew key points and use the shape of extracted vessels to match vasculature trees [48], [49]. More recent algorithms use hybrid approaches that combine sparse keypoint matching with direct pixel tracking to register video frames together and build a mosaic of the retina in real time [50].

Methods that depend on local key point features [41], [42], [44], [51] often result in poor tracking at high magnification because of the lack of texture on the retina and the non-distinctive nature of individual points on the veins. With optimization, the algorithms of [45], [47] could be run in real-time on modern hardware; however, they only use sparse retinal vessel landmarks, which are relatively few or non-existent at high magnifications. More importantly, they only

TABLE I
COMPARISON OF VARIOUS VESSEL DETECTION, REGISTRATION, AND MAPPING ALGORITHMS FOR THE RETINA

Year	Algorithm	Vessel Detection	Retina Registration	Retina Mosaicking	Vasculature Mapping	Method	Image Size	Reported Time
1989	Chaudhuri et al. [32]	X				Matched Filters	*	50 s
1998	D. E. Becker et al. [45]		X	X		Vasculature Landmarks	640x480	0.9 s
1999	Can et al. [25]	X				Tracing Templates	1024x1024	30 ms
2002	Can et al. [46]		X	X		Vasculature Landmarks	1024x1024	*
2002	Stewart et al. [46]	X	X			Vasculature Landmarks & Tree + ICP	1024x1024	5 s
2003	Chanwimaluang et al. [28]	X				Matched Filters + Local Entropy	605x700	3 min
2003	Mojon et al. [29]	X				Adaptative Local Threshold	*	8 s
2004	Staal et al. [30]	X				Ridge-Based Segmentation	768x584	15 min
2006	Cattin et al. [41]		X	X		Keypoint Features (SURF)	1128x1016	*
2006	Chanwimaluang et al. [49]	X	X			Vasculature Tree + ICP & Correlation	600x900	20 s
2006	Mendonca et al. [63]	X				Morphological Segmentation	565x584	2.5 min
2006	Soares et al. [34]	X				Gabor Wavelets	768x584	3 min
2006	Sofka et al. [33]	X				Multiscale Matched Filters	700x605	10 s
2008	Alonso-Montes et al. [31]	X				Pixel Level Snakes	768x584	0.2 s
2010	Chen et al. [51]		X			Keypoint Features (Harris)	500x500	15 s
2010	Lupascu et al. [35]	X				Gabor Wavelets + AdaBoost	768x584	2 min
2010	Wang et al. [42]		X	X		Keypoint Features (SIFT)	640x480	1 s
2011	Broehan et al. [47]	X	X			Vasculature Tree	720x576	40 ms
2012	Bankhead et al. [36]	X				Wavelets + Spline Fitting	564x584	0.6 sec
2013	B. C. Becker et al. [24]	X	X		X	Vasculature Tree + ICP	402x300	25 ms
2014	Koukounis et al. [39]	X				Matched Filters on FPGA	640x480	35 ms
2014	Zheng et al. [64]	X	X	X		Vasculature Landmarks	*	1.5 s
2014	Richa et al. [50]		X	X		Keypoint Features (SIFT) + SSD	720x1280	15 ms
2015	Chen et al. [65]	X	X			Vasculature Landmarks	*	*
2015	EyeSLAM (our approach)	X	X		X	Vasculature Tree + Scan Matching	400x304	15 ms

A sample of representative algorithms for detection, registration, and mapping of the retina and retinal vessels sorted by publication date. Core capabilities and underlying methods used are noted. Vessel detection is the ability to return where retinal vessels are from a single image of the retina. Registration is the ability to determine the transformation between two images of the retina. Mosaicking is the ability to build up a map of the observed retina, while vasculature mapping is the ability to build up a map of the seen retinal vessels. Times for detection algorithms are for a single image, whereas times for registration algorithms are often for a pair of images. No attempt has been made to normalize for computing advances, so estimates are upper bounds. EyeSLAM fills a niche for real-time performance while achieving retinal vessel detection, registration, and vasculature mapping simultaneously. *No data reported.

perform localization and do not build a map of the vasculature. Also, many of these approaches suffer from interference caused by the instruments, which both occlude existing features in the image and create new, spurious features on the moving shaft. Our approach is an adaptation of the real-time correlative scan-matching method proposed by Olson et al.[27], using vasculature trees to build maps and register motion. This method is not listed in Table I because it is not originally designed for retinal registration, but instead more traditional SLAM applications (see Section II.C). While EyeSLAM may not have as high accuracy as some algorithms listed in Table I, it is unique because it operates in real time and provides both vasculature maps and retinal registration, making it suitable for tight control loops in robotic surgical assistance.

C. Simultaneous Localization and Mapping (SLAM)

The problem this paper addresses is similar to a core problem addressed in robotics: simultaneous localization and mapping, or SLAM. In SLAM, a robot with imprecise, noisy localization (for instance, odometry) explores an unknown environment with local sensors (such as a laser range-finder) with the goal of building a global map and localizing itself relative to this map [52]. Using a probabilistic formulation, SLAM optimizes a joint probability over the map and the localization to simultaneously solve for the true positions of the robot and global environmental features. Early solutions such as the Extended Kalman Filter (EKF) scaled poorly and did not handle ambiguous landmark associations well [52]. Recent particle filter approaches such as FastSLAM are faster and more robust [53]. With the introduction of occupancy grids, which discretize the map and maintain a grid of probabilities representing whether each cell is occupied, SLAM algorithms scale more effectively [26].

Comparing SLAM to our problem, the task of building a temporally consistent map of vasculature and localizing the current observation of vessels to this map exhibits many similarities. However, most implementations of SLAM are tailored to space-carving sensors such as laser range-finders instead of overhead sensors and assume a reasonably good robotic motion model, both of which are poor assumptions in the problem of retinal localization and mapping. Thus, most developed SLAM algorithms are not directly applicable without adaptation for the problem being addressed here. Instead, our approach is to fuse core ideas of SLAM and retinal algorithms together to provide real-time registration and mapping of the retina in a surgical environment.

III. METHODS

Our goal is to design an algorithm that maps and localizes retinal vessels by merging retinal vessel detection with retinal image registration and taking advantage of temporal information as seen in SLAM approaches. A fusion of methods is needed: fast retinal vessel detections algorithms are

noisy, incomplete, and do not handle occlusions [25]; retinal image registration methods that do build vasculature maps are orders of magnitude slower than required for real-time robotic guidance [48], [49]; and SLAM algorithms are not designed or tuned for application in intraocular surgeries. Fig. 2 shows a new algorithm called EyeSLAM derived from [24] that incorporates aspects of [25], [26], [52], [54] to perform >30 Hz vasculature mapping and localization of retinal video using rapidly-detected vessels as features, an occupancy grid for mapping, and fast scan matching for localization to robustly handle noise, tool occlusions, and variable illumination.

A. 2D vs. 3D Models of the Retina

When considering a model of a retina, a 3D sphere seems the most suitable representation for building a map of the inside of the eye. However, a full 3D representation is problematic because 3D estimation in the eye is challenging. Properly calibrating microscope calibration can be difficult [55] and modeling the lens of the eye to achieve intraocular localization is an area of active research [56], especially in conjunction with the nonlinear vitrectomy lenses often used during intraocular surgery. Further complicating matters is the deformation caused by the tool inserted through the sclera (or white of the eye). In practice, most approaches to retinal mapping choose a simpler 2D representation, assuming a roughly planar structure for the retina with an X translation, Y translation, and an in-plane rotation. Scaling can be added, but often has local minima that cause poor tracking [50], so we use the 3DOF representation without scaling. We have found the planar assumption has sufficient power to compactly represent the retina, which is especially true at high magnification where only a small subset of the retina is seen and can be treated as a plane. However, to compensate for small shifting vessel locations caused by 3D rotation, we do add a dynamic aspect to the map (see Section III.D).

B. Problem Definition

Given an series of input video frames $I = [I_0,\ I_1, \dots,\ I_T]$ over a discretized time period $t \in [0,\ 1,\ \dots,\ T]$, the algorithm should output a global map in the form of N vasculature points $M = [M^0, M^1, \dots, M^N]$ and the corresponding camera viewpoint locations $L = [L_0, L_1, \dots, L_T]$ of the input video frames in the map. At time t, we parameterize the i^{th} vein 2D point as $M_t^i = [x_t^i, y_t^i]$. As discussed in Section III.A, we approximate the global map as a planar section of retina. Similar to many other approaches to retinal registration, we assume an affine camera with a viewpoint at time t as a 2D translation and rotation $L_t = [x_t, y_t, \theta_t]$ from the initial position at $t = 0$. As seen from our results, this 3-DOF motion model is sufficient even with a fairly large field of view of the retina. Observations of vessels in the camera at time t are denoted by Z_t.

C. Feature Extraction via Vessel Detection

Finding features for matching at high magnification is difficult. Traditional key point detectors such as SIFT [40] or SURF [57] often fail to find distinctive points on the textureless retina. Likewise, sufficient vasculature landmarks such as crossovers or bifurcations may not be present to function as good features with high magnification. We instead use many points along the entire vein extracted from vessel detection algorithms. These points are more difficult to match individually with a local feature descriptor, so we instead match them as a group based on the shape of the vasculature network. As in [24], we use the highly efficient, but noisy vessel tracing algorithm of Can et al. [25] to detect vessels, which form the anonymous matching points (see Fig. 1b).

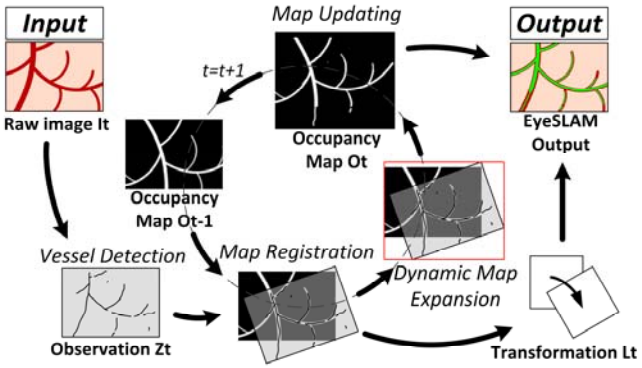
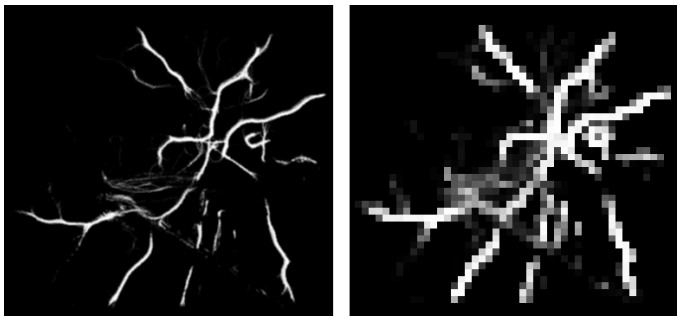


Figure 2: Block diagram showing the steps of EyeSLAM algorithm that maps and localizes retinal vessels during intraocular surgery. Vessels are detected and registered to the map, building an occupancy grid. The map is initialized with the raw vessel detections on the first frame and resized dynamically as needed. The registration allows the map to be transformed back into the current frame, providing full vasculature map along with localization.

To cull spurious detections on vessel-looking structures such as the tip of the instrument or light-pipe, each vessel point must pass a color test and a bloom proximity test. The color test rejects pixels that are too dark or insufficiently red, while the bloom proximity test rejects vessel points that are too close to large white specular blooms in the image. These two simple tests reject many false positives in the detection stage and yield the current observation Z_t as 2D points. The new algorithm EyeSLAM is equipped with a new filter that applies a mask on the dark area circling the visible retina area to mask out fringing effects caused by the microscope. Compared to [24], we have also enhanced the vessel detection method of Can et al. [58] by increasing the allowable change in orientation while tracing, which provides better coverage of the vasculature. Occasional false positive vessels are detected this way, but usually remain as low-probability vessels. Finally, because the vessel detection has coarse discrete filters for fast operation, we apply an averaging smoothing filter across the vessel after tracing, which is another new aspect of the present paper.

D. Mapping via Occupancy Grids

A global map that holds the current best estimate of all the observed vasculature is maintained using an occupancy grid O_t, which discretizes the map into pixel-sized cells (see Fig. 1(c)). Each pixel in the occupancy grid represents the probability that a vessel occupies that particular spatial location. At each time instant t, the current observations Z_t are transformed to the map with the best estimate of the location



(a) **(b)**

Figure 3: Multi-resolution maps used by the scan-matching registration algorithm. (a) High resolution map. (b) Low resolution map.

L_t^{-1} and used to update the probabilities in the occupancy grid. For each vein point Z_t^i, a defined value is added to the map at the corresponding cell location, which increases the probability that a vessel exists at each detected vein point. The occupancy grid has a maximum allowed value to prevent unbounded evidence from accumulating. To handle a slightly changing environment by 3D rotation, lensing effects, or tool deformation, a decay function decreases the probability of all grid cells, allowing vessels that have not been detected to vanish after some time. While it would be more robust to explicitly model these, instead of using a decay to let the map react to changes, nonlinearities such as non-rigid tool/tissue interaction are difficult to model, especially in real time. Figure 3 shows sample occupancy grid maps that model the probability of vessels at each location with multi-resolution.

Unlike [24], the EyeSLAM algorithm uses a dynamically sized map that can automatically expand to accommodate newly detected vessels that were located outside the previous map boundaries as described on Fig. 2 on the "Dynamic Map Expansion" and "Map Updating" steps. The map update calculates the camera viewpoint on the map to not decay the probability of the grid cells located outside the currently observable field of view. The formulation of the occupancy map reasons about uncertainty over time, smoothing noise while handling occlusions and deformation. Our previous work required finding the centerlines of the vessels as a necessary part of the registration step, which required an expensive image skeletonization process. The new registration method described in Section III.E avoids this time-consuming procedure and operates directly on the occupancy grid map. The final map is now generated using all the vessel points with a high probability value in the occupancy grid map. It is still possible to calculate and build the vessel centerlines locally around the areas of interest if necessary for robotic control, but it is no longer integral to the internal workings of the algorithm, which speeds computations compared to [24].

E. Localization via Scan Matching

To localize eyeball motion (which is mathematically equivalent to localizing camera motion), a 3-DOF planar motion model is chosen. The problem of localization is then to estimate the 2D translation and rotation L_t between the current observations Z_t and the occupancy grid map O_t. The original formulation [24] used Iterative Closest Point (ICP) algorithm for registration (or localization) between a skeletonized version of the occupancy map and the current vessel observations, similar to [59]. While this worked well for smooth motions, it had a tendency to fail with large, jerky motions, which caused divergence and would reset the tracker.

The biggest improvement compared to [24] is the replacing of ICP with the fast correlative scan-matching method proposed by Olson et al. [27]. It is used to scan the whole 3D search window of solutions W_l, parametrized by x_l, y_l and θ_l, and to find the best match B_t between the map O_t and the current observations Z_t transformed to the map with W_l^{-1}. The solution associated with the best match is considered as the best solution for L_t given the probabilistic map. The real challenge is to minimize the processing time while maximizing the quality and the robustness of the solution.

A brute-force method that scans every solution is too slow. As detailed in [27], we adopt a multi-resolution approach

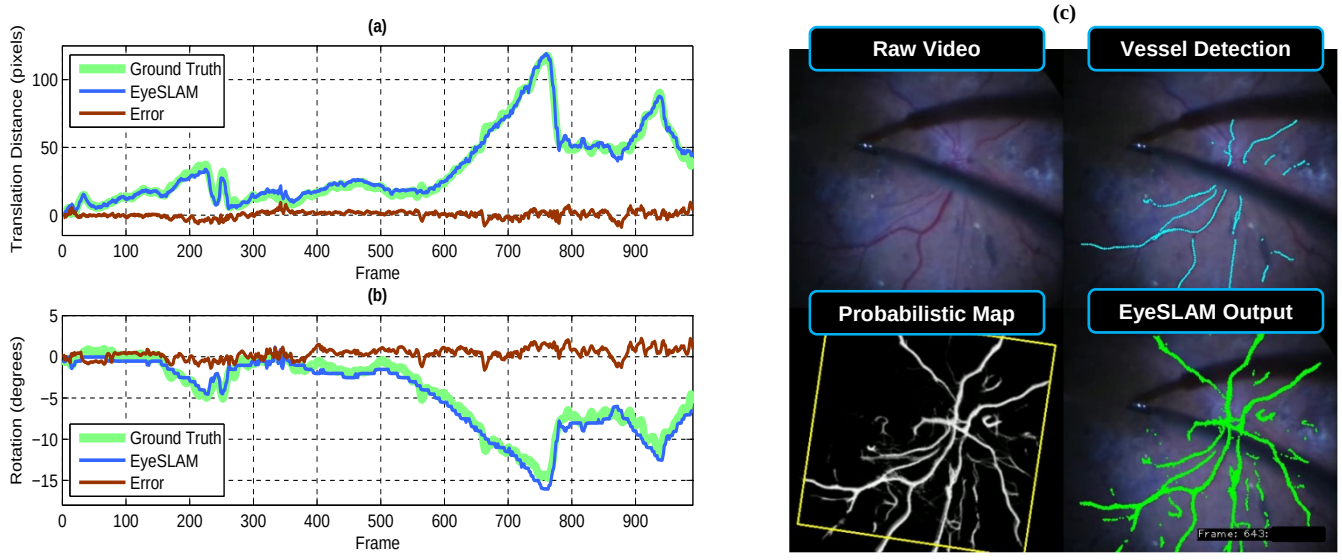


Figure 4: Localization accuracy and error compared to labeled video of human retina *in vivo*. (a) Translation component representing the L2 norm of X and Y (b) Rotation component (c) Sample output of raw video, vessel detection, map, and localized EyeSLAM output from frame 643.

TABLE II
RMS ERROR OF EYESLAM LOCALIZATION

Video Sequence	# Frames	Becker et al. [24] Distance (px)	Becker et al. [24] Theta (deg)	EyeSLAM (ours) Distance (px)	EyeSLAM (ours) Theta (deg)	Overall improvement
Paper Slide (Synthetic)	598	2.8	0.7	1.6	0.4	1.6
Porcine Retina (*ex vivo*)	795	204.2	43.8	7.6	0.9	24.2
Human Retina (*in vivo*)	997	115.1	31.6	4.8	0.8	22.0

Root Mean Squared Error (RMSE) of 2D localization (translation + rotation) for three video sequences. The existing approach and improved EyeSLAM algorithm are compared. The approach of Becker et al. [24] works well for gentle transitions present in the synthetic example, but for more challenging video sequences with jerky motion, it jumps and cannot recover, causing very high overall error. On these videos, the proposed EyeSLAM algorithm is much more robust, with an average RMS error of under 5 pixels in X and Y and under 1° in rotation. Furthermore, this represents a 60% improvement over the earlier algorithm of Becker et al. [24] in ideal conditions and over 20X more accurate localization in sequences with high motion, where EyeSLAM is more able to maintain consistent tracking without jumping.

consisting in scanning the 3-DOF search window (two translations and a rotation) with two different map resolutions (Fig. 3) . A first scan quickly identifies the approximate best solution in the low-resolution map. Afterwards, a second scan is initiated in a restricted search around the initial low-resolution transformation. This is performed on the high-resolution map to more precisely find the best approximation of L_t.

Incomplete vessel detections at each frame can be noisy, so the final scan-matching estimation of the localization is smoothed using a constant-velocity Kalman filter, yielding the localization L_t. At most 500 vessel points are selected for scan matching (at random) to improve runtime. If too few observations are found, they are discarded and the current localization is kept. Once scan matching completes, the occupancy map is then updated with the newly registered vessel points Z_t to close the loop on the algorithm.

IV. EVALUATION

We have evaluated EyeSLAM both quantitatively and qualitatively on a variety of videos of paper slides, porcine retina *ex vivo*, and human procedures *in vivo*.

A. Video Sequences

For ease of robotic testing in our lab, color video recorded with a surgical microscope is captured at 30 Hz with a resolution of 800x608 at a variety of high magnifications (10-25X). We also tested with human retinas *in vivo* from videos taken of real human eye surgeries. Those videos have variable resolution at different magnification, which are specified if relevant in the results. Fig. 4 shows the proposed algorithm output on a human retina during surgery *in vivo*.

B. Retina Localization Results

Fig. 4 shows translation and rotation transformation with error between EyeSLAM estimates and human-labeled ground truth for a video sequence of human retina during an epiretinal peeling procedure *in vivo*. Quantitatively in Table II, EyeSLAM with scan matching completely outperforms the previous ICP method for experiments on three video sequences. Error is measured relative to a transformation calculated from tracked fiducials for the paper slide or human labeled fiducials as ground truth. As seen in Fig. 4, significant translations in excess of 100 pixels are evident on these video sequences. The global nature of the low-resolution scan-matching step helps prevent jumps due to poor vessel registration. In the planar phantom experiment, where the vessels are well defined, the lighting uniform, and the motion smooth, the two results are close, although EyeSLAM localizes 50% better. On the porcine and human retina, earlier versions of the algorithm have very large RMS errors because of lost tracking and jumps that were not recoverable with the older ICP registration. The new EyeSLAM algorithm provides superior localization, which is important during critical microsurgical operations.

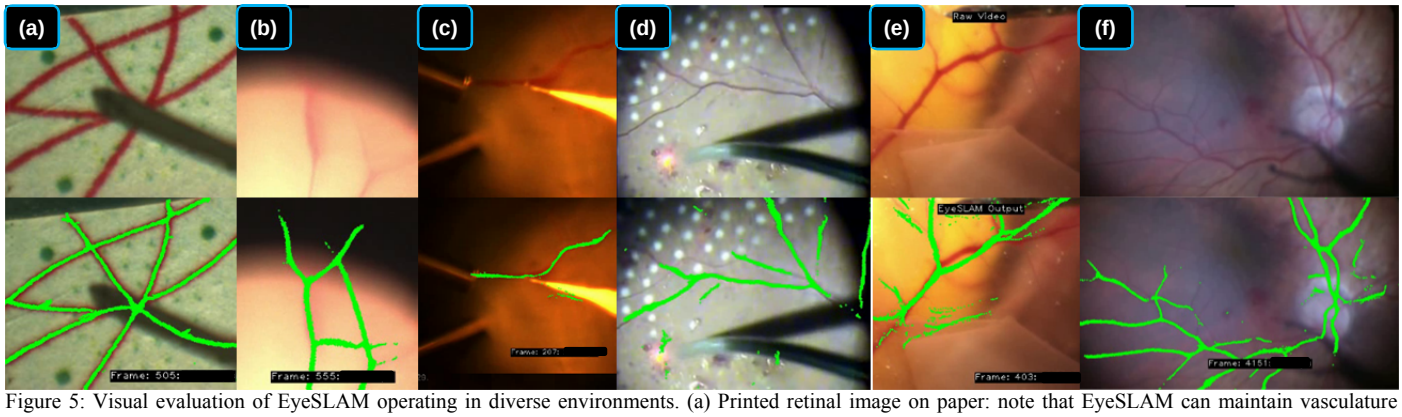


Figure 5: Visual evaluation of EyeSLAM operating in diverse environments. (a) Printed retinal image on paper: note that EyeSLAM can maintain vasculature structure even during tool occlusion. (b) Porcine retina *ex vivo* at high magnification during rapid movement where some of the vessels have moved outside of the view of the microscope; note that EyeSLAM is able to remember where the vessels are, even outside of the microscope FOV. (c) Porcine retina *ex vivo* during retinal vessel cannulation experiment; notice very good performance in a challenging light environment despite a few false positives on the edge of the tool where the color filter is failing. (d) Human retina *in vivo* during panretinal photocoagulation surgery; there are a few false positives on the red laser dot where the color filter is failing. (e) Chick chorioallantoic membrane (CAM) *in vivo*, a model for retinal vessels that is considerably more elastic than the retina; EyeSLAM does not easily correct for deformations induced by the tool, and has some false positive responses on the edge of the membrane. (f) Human retina *in vivo* during epiretinal membrane peeling; note the good mapping and localization.

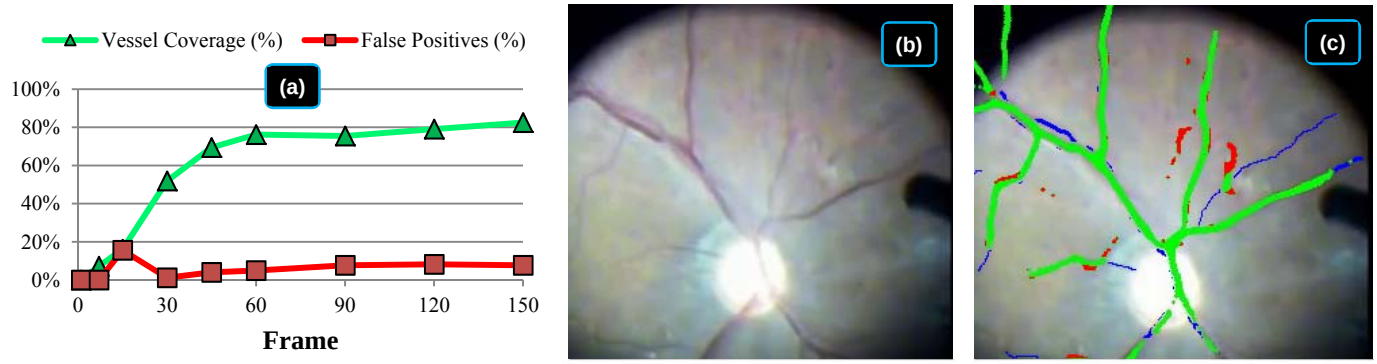


Figure 6: Vasculature map initialization statistics including coverage and false positives for the first five seconds of a video sequence of human retina *in vivo* being prepared to receive photocoagulation treatment. (a) Vessel coverage and false positives of the thresholded probabilistic map evaluated at frame 0, 7, 15, 30, 60, 90, 120, and 150 compared to hand-labeled vasculature at each frame. Coverage is the percentage of hand-labeled vasculature that EyeSLAM correctly built. False positives are the places in the EyeSLAM maps that do not correspond to true vessels. The spike in false positives at frame 15 is due to slight mis-registration during initialization. (b) Raw video of frame 150 after five seconds. (c) Color-coded map at frame 150 with green representing true positives (correctly matched vessels), red representing false positives (spurious vessels), and blue representing false negatives (missed vessels). Best viewed in color.

C. *Vasculature Mapping Results*

Figure 5 shows a visual, qualitative evaluation of the quality of maps EyeSLAM builds in a variety of simulated and real retinal surgical applications. While there are some incomplete detections and false positives in the map, EyeSLAM is able to largely map and localize in a diverse and challenging set of lighting conditions and environments.

To measure map quality quantitatively, Fig. 6 evaluates vessel coverage and false positives on a typical human video sequence *in vivo*. Calculations are made using hand-labeled vasculature (including very thin vessels) in nine frames in the first five seconds. Initialization is fast, requiring less than 15 frames (0.5 sec) to start building the map and only a few seconds to fully build the map, depending on the image quality and the ability of the vessel detection algorithm. On this sequence, EyeSLAM maps 50% of the hand-labeled vessels within 30 frames (one second) and achieves 75% coverage within 60 frames (two seconds). It does not achieve 100% coverage because it misses some very thin, faint vessels, and the ones located very close together. The false positive rate is under 10% after five seconds. Overall, EyeSLAM converges quickly with good coverage of the well-defined vasculature structures of the retina.

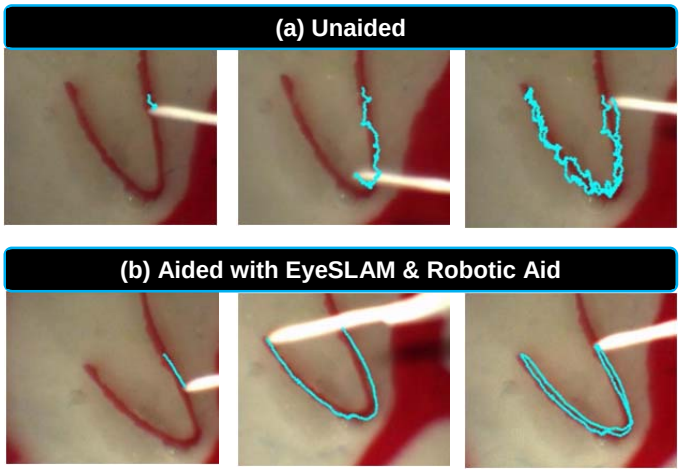


Figure 7: Tracing a retinal vessel in an eyeball phantom. (a) Unaided attempt to trace the vessel and (b) Aided attempt with a robotic micromanipulator enforcing virtual fixtures based off EyeSLAM mapping and localization. The blue line indicates the path of the tip of the instrument registered to each frame of the video using EyeSLAM. Note in both cases, the entire phantom eyeball is moving due to movement of the tool through the sclera port.

D. *Timing Performance*

For speed, images are resized in half, yielding resolutions in the range of 400x304 to 380x360, depending on video sequence source. On an Intel i5-3570K computer, EyeSLAM implemented in C++ runs at 50-100 Hz with a mean runtime of 15 ms on the three videos listed in Table II. This time includes all vessel feature detection, correlative scan-matching localization, and dynamic occupancy-grid mapping running in a single thread. This is about 2X faster than the previous work [24], and is sufficient to run simultaneous EyeSLAM algorithms on stereo microscope views in real-time (>30 Hz).

E. *Surgically-Applied Results*

To test the usefulness of the algorithm in intraocular surgical environments, we simulated the eye and retina in a rubber eyeball phantom filled with saline and fitted with a vitrectomy lens. We asked an operator to perform a vessel tracing task. Fig. 7(a) shows unaided human performance and Fig. 7(b) shows performance aided with the Micron robot enforcing virtual fixtures [60] derived from EyeSLAM. In (a)

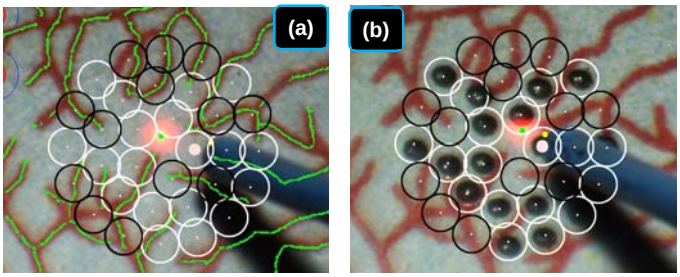


Figure 8: EyeSLAM can be used in the robotic micromanipulator Micron control system to provide accurate targeting information on a paper phantom during a simulated retinal photocoagulation experiment. EyeSLAM provides the necessary localization information to compensate for motion of the retina in real time and register the burn pattern to the pre-operatively specified placement. The EyeSLAM map is used to automatically plan burn patterns that do not overlap or touch the vasculature structure, thus protecting the vessels while applying the needed treatment to the retina. (a) Target placement before laser photocoagulation. Empty black circles represent targets to be avoided within 100 µm of a vessel while white circles show valid targets (b) After the completion of the automated laser photocoagulation.

the tracing is imprecise because of normal physiological hand tremor at sub-millimeter scales whereas in (b) the tracing is smoother because the Micron robot knows vasculature map from EyeSLAM and can enforce virtual fixtures in the control loop to help keep the tip of the instrument on the vessel.

Finally, we demonstrate the efficacy of pairing EyeSLAM localization and mapping with robotic aid in a simulated retinal surgery. Research performed in our lab has used the EyeSLAM algorithm with the Micron manipulator during simulated photocoagulation surgeries performed with vitreoretinal surgeons on synthetic paper slides of retina [61]. In laser photocoagulation, the goal is to accurately place burns on the retina while avoiding the vasculature. EyeSLAM provides the map of the vessels from which the system automatically plans the pattern of laser burns as well as the localization necessary to register that pattern to the retina as it moves during the procedure. The robotic aid aims the laser at the proper location on the retina, compensating for motion measured by EyeSLAM, and activates the laser. The surgeons using the Micron micromanipulator reduced burn placement error by over 50% while providing regular burn sizes with the robotic aid. Fig. 8 shows the automated avoidance of vessels using EyeSLAM during automated laser photocoagulation.

V. Discussion And Future Work

We have presented a new and improved algorithm, EyeSLAM, for retinal mapping and localization that operates in real time at >30 Hz. Designed to handle the dynamic environment of high magnification, variable illumination, and rapid motion inherent in retinal surgery, our approach converges quickly and is robust to occlusion. Fusing ideas from vessel detection, retinal registration, and SLAM, it has proven to be an effective method to temporally smooth vessel detections and build a comprehensive map of the vasculature. We demonstrated EyeSLAM is localization accurate within five pixels in translation and one degree in rotation on representative video sequences and initializes quickly, covering 80% of the vasculature within two seconds. Compared to [24], the new EyeSLAM approach addresses a lot of the earlier shortcomings, especially lag and loss of tracking. EyeSLAM has greatly improved localization, especially being robust to rapid motions common in retinal surgeries, along with additions to make the algorithm faster with more comprehensive vessel detections and an enhanced dynamic map that expands as needed. We have shown EyeSLAM is least 50% better than our earlier formulation with significantly more consistent tracking in more difficult video sequences. We have also show EyeSLAM operates in real time in challenging intraocular environments, providing both mapping and localization of the retinal vasculature to the Micron robot, improving operator performance in synthetic vessel tracing and photocoagulation experiments.

Future improvements should include more robust vessel detection. Another focus is reducing false positives and better modeling occlusion with more sophisticated tool tracking such as [62]. More advanced 3D models could be beneficial and accomplished with stereo vision. More sophisticated closing the loop SLAM algorithms could be studied and applied. Finally, optimization to run in real-time on high-definition video may increase localization accuracy and map quality.

References

[1] F. Holtz and R. F. Spaide, *Medical Retina*. Berlin: Springer, 2007.

[2] S. P. N. Singh and C. N. Riviere, "Physiological tremor amplitude during retinal microsurgery," in *Proc. IEEE Northeast Bioeng. Conf.*, 2002, pp. 171–2.

[3] H. L. Brooks, "Macular hole surgery with and without internal limiting membrane peeling," *Ophthalmology*, vol. 107, no. 10, pp. 1939–1948, 2000.

[4] A. J. Packer, *Manual of Retinal Surgery*. Boston: Butterworth-Heinemann Medical, 2001.

[5] T. A. Rice, "Internal limiting membrane removal in surgery for full-thickness macular holes," *Am. J. Ophthalmol.*, vol. 129, no. 6, pp. 125–146, 1999.

[6] R. N. Frank, "Retinal laser photocoagulation: Benefits and risks," *Vision Res.*, vol. 20, no. 12, pp. 1073–1081, 1980.

[7] D. A. Infeld and J. G. O'Shea, "Diabetic retinopathy," *Postgrad. Med. J.*, vol. 74, no. 869, pp. 129–133, 1998.

[8] J. N. Weiss and L. A. Bynoe, "Injection of tissue plasminogen activator into a branch retinal vein in eyes with central retinal vein occlusion," *Ophthalmology*, vol. 108, no. 12, pp. 2249–57, 2001.

[9] L. A. Bynoe, R. K. Hutchins, H. S. Lazarus, and M. A. Friedberg, "Retinal endovascular surgery for central retinal vein occlusion: Initial experience of four surgeons," *Retina*, vol. 25, no. 5, pp. 625–632, 2005.

[10] H. Das, H. Zak, J. Johnson, J. Crouch, and D. Frambach, "Evaluation of a telerobotic system to assist surgeons in microsurgery," *Comput. Aided Surg.*, vol. 4, no. 1, pp. 15–25, 1999.

[11] T. Ueta, Y. Yamaguchi, Y. Shirakawa, T. Nakano, R. Ideta, Y. Noda, A. Morita, R. Mochizuki, N. Sugita, and M. Mitsuishi, "Robot-assisted vitreoretinal surgery: Development of a prototype and feasibility studies in an animal model," *Ophthalmology*, vol. 116, no. 8, pp. 1538–43, 2009.

[12] W. Wei, R. E. Goldman, H. F. Fine, S. Chang, and N. Simaan, "Performance evaluation for multi-arm manipulation of hollow suspended organs," *IEEE Trans. Robot.*, vol. 25, no. 1, pp. 147–157, 2009.

[13] D. H. Bourla, J. P. Hubschman, M. Culjat, A. Tsirbas, A. Gupta, and S. D. Schwartz, "Feasibility study of intraocular robotic surgery with the da Vinci surgical system," *Retina*, vol. 28, no. 1, pp. 154–158, 2008.

[14] A. P. Mulgaonkar, J. P. Hubschman, J. L. Bourges, B. L. Jordan, C. Cham, J. T. Wilson, T. C. Tsao, and M. O. Culjat, "A prototype surgical manipulator for robotic intraocular micro surgery," *Stud. Heal. Technol. Inform.*, vol. 142, no. 1, pp. 215–7, 2009.

[15] A. Uneri, M. Balicki, J. Handa, P. Gehlbach, R. H. Taylor, and I. Iordachita, "New Steady-Hand Eye Robot with micro-force sensing for vitreoretinal surgery," in *Proc. IEEE Int. Conf. Biomed. Robot. Biomechatron.*, 2010, pp. 814–819.

[16] J. P. Hubschman, J. L. Bourges, W. Choi, A. Mozayan, A. Tsirbas, C. J. Kim, and S. D. Schwartz, "The Microhand: A new concept of micro-forceps for ocular robotic surgery," *Eye*, vol. 24, no. 2, pp. 364–367, 2009.

[17] G. Dogangil, O. Ergeneman, J. J. Abbott, S. Pané, H. Hall, S. Muntwyler, and B. J. Nelson, "Toward targeted retinal drug delivery with wireless magnetic microrobots," in *Proc. IEEE Intl. Conf. Intell. Robot. Syst.*, 2008, pp. 1921–1926.

[18] B. C. Becker, R. A. MacLachlan, L. A. Lobes Jr, G. Hager, and C. Riviere, "Vision-based control of a handheld surgical micromanipulator with virtual fixtures," *IEEE Trans. Robot.*, p. submitted.

[19] S. Yang, R. A. MacLachlan, and C. N. Riviere, "Manipulator design and operation for a six-degree-of-freedom handheld tremor-cancelling microsurgical instrument," *IEEE/ASME Trans. Mechatron.*, vol. 20, no. 2, pp. 761–772, 2015.

[20] L. B. Rosenberg, "Virtual fixtures: Perceptual tools for telerobotic manipulation," in *IEEE Virt. Reality Ann. Int. Symp.*, 1993, pp. 76–82.

[21] A. Bettini, P. Marayong, S. Lang, A. M. Okamura, and G. D. Hager, "Vision-assisted control for manipulation using virtual fixtures," *IEEE Trans. Robot.*, vol. 20, no. 6, pp. 953–966, Dec. 2004.

[22] H. C. Lin, K. Mills, P. Kazanzides, G. D. Hager, P. Marayong, A. M. Okamura, and R. Karam, "Portability and applicability of virtual fixtures

across medical and manufacturing tasks," in *Proc. IEEE Int. Conf. Robot. Autom*, 2006, pp. 225–231.

[23] B. C. Becker, S. Voros, L. A. Lobes Jr., J. T. Handa, G. D. Hager, and C. N. Riviere, "Retinal vessel cannulation with an image-guided handheld robot," in *Proc. Conf. IEEE Eng. Med. Biol. Soc.*, 2010, pp. 5420–3.

[24] B. C. Becker and C. N. Riviere, "Real-time retinal vessel mapping and localization for intraocular surgery," in *IEEE International Conference on Robotics and Automation (ICRA2013).*, 2013, pp. 5360–5365.

[25] A. Can, H. Shen, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms," *IEEE Trans. Inform. Technol. Biomed.*, vol. 3, no. 2, pp. 125–38, 1999.

[26] D. Hahnel, W. Burgard, D. Fox, and S. Thrun, "An efficient FastSLAM algorithm for generating maps of large-scale cyclic environments from raw laser range measurements," in *Proc. IEEE Intl. Conf. Intell. Robot. Syst.*, 2003, vol. 1, pp. 206–211.

[27] E. B. Olson, "Real-Time Correlative Scan Matching," in *IEEE International Conference on Robotics and Automation, 2009 (ICRA'09).*, 2009, pp. 4387–4393.

[28] T. Chanwimaluang, "An efficient blood vessel detection algorithm for retinal images using local entropy thresholding," in *International Symposium on Circuits and Systems*, 2003, vol. 5, pp. 21–24.

[29] D. Mojon, "Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 1, pp. 131–137, Jan. 2003.

[30] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–9, Apr. 2004.

[31] C. Alonso-Montes, D. L. Vilariño, P. Dudek, and M. G. Penedo, "Fast retinal vessel tree extraction: A pixel parallel approach," *Int. J. Circuit Theory Appl.*, vol. 36, no. 5–6, pp. 641–651, Jul. 2008.

[32] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters.," *IEEE Trans. Med. Imag.*, vol. 8, no. 3, pp. 263–9, Jan. 1989.

[33] M. Sofka and C. V. Stewart, "Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures," *IEEE Trans. Med. Imag.*, vol. 25, no. 12, pp. 1531–1546, Dec. 2006.

[34] J. V. B. Soares, J. J. G. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree, "Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification," *IEEE Trans. Med. Imag.*, vol. 25, no. 9, pp. 1214–1222, Sep. 2006.

[35] C. A. Lupascu, D. Tegolo, and E. Trucco, "FABC: Retinal vessel segmentation using AdaBoost," *IEEE Trans. Inform. Technol. Biomed.*, vol. 14, no. 5, pp. 1267–74, Sep. 2010.

[36] P. Bankhead, C. N. Scholfield, J. G. McGeown, and T. M. Curtis, "Fast retinal vessel detection and measurement using wavelets and edge location refinement," *PLoS One*, vol. 7, no. 3, Jan. 2012.

[37] R. Perfetti, E. Ricci, D. Casali, and G. Costantini, "Cellular neural networks with virtual template expansion for retinal vessel segmentation," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 54, no. 2, pp. 141–145, Feb. 2007.

[38] A. Nieto, V. M. Brea, and D. L. Vilarino, "FPGA-accelerated retinal vessel-tree extraction," in *International Conference on Field Programmable Logic and Applications*, 2009, pp. 485–488.

[39] D. Koukounis, C. Ttofis, A. Papadopoulos, and T. Theocharides, "A high performance hardware architecture for portable, low-power retinal vessel segmentation," *Integr. VLSI J.*, vol. 47, no. 3, pp. 377–386, 2014.

[40] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[41] P. Cattin, H. Bay, L. Van Gool, and G. Székely, "Retina mosaicing using local features," in *Medical Image Computing and Computer-Assisted Intervention*, 2006, vol. 4191, pp. 185–192.

[42] Y. Wang, J. Shen, W. Liao, and L. Zhou, "Automatic fundus images mosaic based on SIFT feature," *J. Image Graph.*, vol. 6, no. 4, pp. 533–537, 2011.

[43] J. Chen, R. Smith, J. Tian, and A. F. Laine, "A novel registration method for retinal images based on local features.," in *Proc. Conf. IEEE Eng. Med. Biol. Soc.*, 2008, vol. 2008, pp. 2242–5.

[44] B. C. Becker, S. Yang, R. A. MacLachlan, and C. N. Riviere, "Towards vision-based control of a handheld micromanipulator for retinal cannulation in an eyeball phantom," in *Proc. IEEE Int. Conf. Biomed. Robot. Biomechatron.*, 2012, pp. 44–49.

[45] D. E. Becker, A. Can, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Image processing algorithms for retinal montage synthesis, mapping, and real-time location determination," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 1, pp. 105–118, 1998.

[46] A. Can, C. V. Stewart, B. Roysam, and H. L. Tanenbaum, "A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3, pp. 347–364, Mar. 2002.

[47] A. M. Broehan, T. Rudolph, C. A. Amstutz, and J. H. Kowal, "Real-time multimodal retinal image registration for a computer-assisted laser photocoagulation system.," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 10, pp. 2816–24, Oct. 2011.

[48] C. V Stewart, C.-L. Tsai, and B. Roysam, "The dual-bootstrap iterative closest point algorithm with application to retinal image registration.," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1379–94, Nov. 2003.

[49] T. Chanwimaluang, G. Fan, and S. R. Fransen, "Hybrid retinal image registration," *IEEE Trans. Inform. Technol. Biomed.*, vol. 10, no. 1, pp. 129–142, Jan. 2006.

[50] R. Richa, R. Linhares, E. Comunello, A. von Wangenheim, J.-Y. Schnitzler, B. Wassmer, C. Guillemot, G. Thuret, P. Gain, G. Hager, and others, "Fundus Image Mosaicking for Information Augmentation in Computer-Assisted Slit-Lamp Imaging," *Med. Imaging, IEEE Trans.*, vol. 33, no. 6, pp. 1304–1312, 2014.

[51] J. Chen, J. Tian, N. Lee, J. Zheng, R. T. Smith, and A. F. Laine, "A partial intensity invariant feature descriptor for multimodal retinal image registration.," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1707–18, Jul. 2010.

[52] S. Thrun, "Simultaneous Localization and Mapping," in *Robotics and Cognitive Approaches to Spatial Mapping*, vol. 38, M. Jefferies and W.-K. Yeap, Eds. Springer Berlin / Heidelberg, 2008, pp. 13–41.

[53] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit, "FastSLAM: A factored solution to the simultaneous localization and mapping problem," in *Proceedings of the National Conference on Artificial Intelligence*, 2002, pp. 593–598.

[54] E. B. Olson, "Real-Time Correlative Scan Matching," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2009, pp. 4387–4393.

[55] Y. Zhou and B. J. Nelson, "Calibration of a parametric model of an optical microscope," *Opt. Eng.*, vol. 38, no. 12, pp. 1989–1995, 1999.

[56] C. Bergeles, K. Shamaei, J. J. Abbott, and B. J. Nelson, "Single-camera focus-based localization of intraocular devices.," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 8, pp. 2064–74, Aug. 2010.

[57] H. Bay, A. Ess, T. Tuytelaars, and L. van Gool, "SURF: Speeded Up Robust Features," *Comp. Vis. Img. Understand.*, vol. 110, no. 3, pp. 346–59, 2008.

[58] A. Can, H. Shen, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms," *IEEE Trans. Inform. Technol. Biomed.*, vol. 3, no. 2, pp. 125–138, 1999.

[59] C. V Stewart, C.-L. Tsai, and B. Roysam, "The dual-bootstrap iterative closest point algorithm with application to retinal image registration.," *IEEE Trans. Med. Imaging*, vol. 22, no. 11, pp. 1379–94, Nov. 2003.

[60] B. C. Becker, R. A. MacLachlan, G. D. Hager, and C. N. Riviere, "Handheld micromanipulation with vision-based virtual fixtures," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2011, pp. 4127–4132.

[61] S. Yang, L. A. Lobes Jr, J. N. Martel, and C. N. Riviere, "Handheld Automated Microsurgical Instrumentation for Intraocular Laser Surgery," *Lasers Surg. Med.*, p. to be published, 2015.

[62] R. Sznitman, R. Richa, R. H. Taylor, B. Jedynak, and G. D. Hager, "Unified detection and tracking of instruments during retinal microsurgery," *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 35, no. 5, pp. 1263–1273, 2013.

[63] A. M. Mendonca and A. Campilho, "Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction," *IEEE Trans. Med. Imag.*, vol. 25, no. 9, pp. 1200–1213, Sep. 2006.

[64] Y. Zheng, E. Daniel, A. A. Hunter, R. Xiao, J. Gao, H. Li, M. G. Maguire, D. H. Brainard, and J. C. Gee, "Landmark matching based retinal image alignment by enforcing sparsity in correspondence matrix," *Med. Image Anal.*, vol. 18, no. 6, pp. 903–913, 2014.

[65] L. Chen, X. Huang, and J. Tian, "Retinal image registration using topological vascular tree segmentation and bifurcation structures," *Biomed. Signal Process. Control*, vol. 16, pp. 22–31, 2015.