

Vision-Based Control of a Handheld Micromanipulator for Robot-Assisted Retinal Surgery

THESIS DOCUMENT

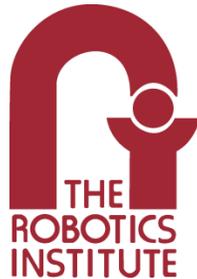
BRIAN C. BECKER

briancbecker@cmu.edu

CMU-RI-TR-12-31

THESIS COMMITTEE

Cameron N. Riviere (Chair)
George A. Kantor
George D. Stetten
Gregory D. Hager (Johns Hopkins University)



The Robotics Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania, US, 15213

*Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Robotics*

Copyright © September 2012 Brian C. Becker. All rights reserved.

ABSTRACT

Surgeons increasingly need to perform complex operations on extremely small anatomy. Many existing and promising new surgeries are effective, but difficult or impossible to perform because humans lack the extraordinary control required at sub-millimeter scales. Using micromanipulators, surgeons gain higher positioning accuracy and additional dexterity as the instrument removes tremor and scales hand motions. While these aids are advantageous, they do not actively consider the goals or intentions of the operator and thus cannot provide context-specific behaviors, such as motion scaling around anatomical targets, prevention of unwanted contact with pre-defined tissue areas, compensation for moving anatomy, and other helpful task-dependent actions.

This thesis explores the fusion of visual information with micromanipulator control and enforces task-specific behaviors that respond in synergy with the surgeon's intentions and motions throughout surgical procedures. By exploiting real-time microscope view observations, a-priori knowledge of surgical operations, and pre-operative data prepared before the surgery, we hypothesize that micromanipulators can employ individualized and targeted aids to further help the surgeon. Specifically, we propose a vision-based control framework of virtual fixtures for handheld micromanipulator robots that naturally incorporates tremor suppression and motion scaling. We develop real-time vision systems to track the surgeon and anatomy and design fast, new algorithms for analysis of the retina. Virtual fixtures constructed from visually tracked anatomy allows for complex task-specific behaviors that monitor the surgeon's actions and react appropriately to cooperatively accomplish the procedure.

Particular focus is given to vitreoretinal surgery as a good choice for vision-based control because several new and promising surgical techniques in the eye depend on fine manipulations of tiny and delicate retinal structures. Experiments with Micron, the fully handheld micromanipulator developed in our lab, show that vision-based virtual fixtures significantly increase pointing precision by reducing positioning error during synthetic, but medically relevant hold-still and tracing tasks. To evaluate the proposed framework in realistic environments, we consider three demanding retinal procedures: membrane peeling, laser photocoagulation, and vessel cannulation. Preclinical trials on artificial phantoms, ex vivo, and in vivo animal models demonstrate that vision-based control of a micromanipulator significantly improves surgeon performance ($p < 0.05$).

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support and assistance of many people. First, a special thanks to my advisor Cameron Riviere for his many years of mentorship and adding the phrase “deadlines are forcing functions” to my vocabulary. Without him, this research would never have been possible. Likewise, my thesis committee members George Kantor, George Stetten, and Greg Hager have been invaluable with their insights. A huge thank you goes to Dr. Louis Lobes for providing the surgical knowledge and skill to the project and spending all-too-many hours patiently performing tediously repetitive and buggy surgical experiments week after week, year after year.

Second, I'd like to thank my family for their unwavering support. My parents, who always encouraged me even if they often did not understand why I was always at campus no matter what time of day or night they called, nor why the PhD program took so long, were a bedrock I could always depend on. And I could always count on my sister to cheer me up, offer insightful life advice (what little life the PhD program left me with!), and take me shopping so even if I lived like a hobo, I would at least not be dressed like one :-)

Third, I offer a huge thank you to my friends and colleagues at CMU who kept me sane: Umashankar Nagarajan, Michael Furlong, Brina Goyette, Prasanna Velagapudi, Nathan Wood, Joydeep Biswas, Heather Jones, Alvaro Collet Romea, Heather Justice, Pyy Matikainen, Santosh Divvala, Gaarrrt Gallagher, Portia Taylor, Nathan Brooks, Ed Hsiao, and Siddharth Sannan. I will always treasure the 6:15 pm supper train, random hangouts in the robolounge, racquetball and tennis sessions to work off steam, 10 pm coffee time, working late on crazy side projects, 2 am meteor shower trips (and finding glowing bioluminescent mushrooms), all-nighters while finishing homework and making videos, Saturday Pamela's trips, and 4 am Eat 'n' Park excursions, which often turned into grand science experiments.

Finally, I offer a big hearty cheer to God who created the world with such depth and beauty that a PhD is barely sufficient to begin understanding a tiny fraction of the wondrous universe we live in.

TABLE OF CONTENTS

Problem Domain.....	16
1 Introduction.....	17
1.1 Taxonomy of Micromanipulators.....	17
1.2 Micron: Handheld Micromanipulator.....	19
1.2.1 Handpiece.....	19
1.2.2 Position Measurement.....	20
1.2.3 Tremor Suppression.....	20
1.2.4 Microscope & Cameras.....	21
1.3 Problem Statement.....	21
1.4 Proposed Approach.....	22
1.5 Thesis Statement.....	23
2 Vitreoretinal Microsurgery.....	24
2.1 Anatomy of the Eye.....	24
2.2 Pars Plana Vitrectomy.....	25
2.3 Laser Photocoagulation.....	26
2.3.1 Treatment Procedure.....	27
2.3.2 Behavioral Aids.....	27
2.4 Vessel Cannulation.....	27
2.4.1 Treatment Procedure.....	28
2.4.2 Behavioral Aids.....	28
2.5 Membrane Peeling.....	28
2.5.1 Treatment Procedure.....	29
2.5.2 Behavioral Aids.....	29
2.5.3 Arteriovenous Sheathotomy.....	30
2.5.4 Treatment Procedure.....	30
2.5.5 Behavioral Aids.....	30
2.6 Retinal Models.....	31
2.6.1 Paper, Rubber, and Sorbothane Slides (In Vitro).....	31
2.6.2 Table Tennis Ball (In Vitro).....	32
2.6.3 JHU Eye Phantom (In Vitro).....	32
2.6.4 Open-Sky Porcine Retina (Ex Vivo).....	33
2.6.5 Incubated Chicken Eggs (In Vivo).....	34
2.6.6 Porcine Retina in Johns Hopkins Eye Phantom (Ex Vivo).....	34
2.7 Discussion.....	34
Theoretical Frameworks.....	36
3 Virtual Fixture Framework for Visual Control.....	37
3.1 Background.....	38
3.1.1 Control System and Notation.....	39
3.1.2 Feedforward Tremor Compensation.....	39
3.1.3 Vision System.....	41

3.1.4	System Setup	41
3.2	Virtual Fixture Framework	41
3.2.1	Point Virtual Fixture	41
3.2.2	Higher-Order Subspace Virtual Fixtures.....	42
3.2.3	Generalizable Spline Virtual Fixtures.....	42
3.2.4	Tremor Suppression.....	43
3.2.5	Motion Scaling.....	44
3.2.6	Generalized Control Law	45
3.2.7	Combining Virtual Fixtures	45
3.2.8	Visual Cues to Maintain Eye-Hand Coordination.....	45
3.3	Evaluation.....	46
3.3.1	Experimental Procedure.....	46
3.3.2	Generating Fixtures from Vision	47
3.3.3	Results.....	49
3.4	Discussion.....	50
4	Modeling Tremor for Feedforward Control.....	52
4.1	Tremor Modeling with Autoregressive Modeling.....	52
4.1.1	Equipment	53
4.1.2	Experimental Procedure.....	53
4.1.3	Modeling Techniques	53
4.1.4	Results with AR and ARMA Models.....	55
4.1.5	Results with Nonlinear and Multivariate Models.....	55
4.1.6	Generalized ARMA Model.....	55
4.1.7	Discussion	56
4.2	State Estimation with Kalman Filters	57
4.2.1	Kalman Filter Algorithm	57
4.2.2	State Estimation with Kalman Filter.....	58
4.2.3	Results.....	60
4.3	Tremor Rejection with Feedforward Control.....	60
4.3.1	Background	60
4.3.2	Methods	61
4.3.3	Experimental Protocol.....	62
4.3.4	Results.....	62
4.3.5	Discussion	64
5	Real-Time Retinal Image Analysis	65
5.1	Tip Tracking	65
5.2	Adaptive Camera Calibration to ASAP	66
5.3	3D Reconstruction with Dense Stereo	66
5.4	Hardware Video Encoding.....	67
5.5	EyeSLAM: Localizing and Mapping the Retina	67
5.5.1	Retinal Detection.....	68
5.5.2	Retinal Image Registration	69
5.5.3	Simultaneous Localization And mapping (SLAM).....	70

5.5.4	Introduction of EyeSLAM and Problem Definition	70
5.5.5	Feature Extraction.....	71
5.5.6	Mapping.....	72
5.5.7	Localization	73
5.5.8	Setup & Timing Performance.....	73
5.5.9	Evaluation: Convergence, Intermediate Results, and Comparisons	74
5.5.10	Evaluation of EyeSLAM In Vein Tracing Tasks	76
5.5.11	Discussion	76
Surgical Applications.....		78
6	Membrane Peeling.....	79
6.1	System Setup.....	80
6.2	Tremor Suppression	80
6.3	Multi-Part Fixtures	80
6.3.1	Motion Scaling Virtual Fixture.....	81
6.3.2	Hard Stop Virtual Fixture	81
6.3.3	Virtual Fixture Control	81
6.4	Displacement Measurements	82
6.4.1	Dense Stereo Reconstruction.....	82
6.4.2	Surface Estimation	82
6.5	Velocity Limiting.....	83
6.6	Visual Cues	83
6.7	Experiment and Results	83
6.7.1	Clamped Experiments.....	83
6.7.2	Handheld Peeling Experiments.....	84
6.8	Discussion.....	88
7	Laser Photocoagulation.....	89
7.1	System Setup.....	89
7.2	Tracking of Burn Locations and Tip Positions	89
7.3	Coordinate System Calibration	90
7.4	Initial Control Mechanism.....	91
7.5	Surface Reconstruction	91
7.6	Control System	92
7.7	Experimental Procedure	95
7.8	Effects of Visualization and Tool Ergonomics	96
7.9	Experimental Protocol	97
7.10	Results.....	97
7.11	Discussion.....	98
8	Vessel Cannulation	101
8.1	System Setup.....	101
8.2	Vision System	101
8.3	Tremor Compensation and Motion Scaling	102
8.4	Control System	103

8.5	Experiments	104
8.6	Results.....	105
8.7	Chicken Egg Membrane	105
8.8	Learning Curves.....	106
8.9	Discussion.....	106
Discussion.....		107
9	Conclusions.....	108
9.1	Summary	108
9.2	Contributions	108
9.2.1	Virtual Fixture Framework For Handheld Micromanipulation.....	108
9.2.2	Model-Predictive, Feed-Forward Control for Tremor Rejection	109
9.2.3	Localization And Mapping of the Retina (EyeSLAM).....	110
9.2.4	Retinal Eye Phantom Model with Porcine Retina	110
9.2.5	Extensive Evaluation of Micron with Retinal Procedures	111
9.3	Validation of Thesis Statement	111
10	Future Directions	113
10.1	New 6-DOF Micron Micromanipulator	113
10.2	More Realistic Evaluation	114
10.3	New Areas of Surgical Applicability.....	114
10.4	Expanded Virtual Fixtures Framework.....	114
10.5	More Robust Vision Systems	115
10.6	More Sophisticated Hardware	115
10.7	Integrating Force Sensing.....	116
10.8	Conclusion.....	116
11	Works Cited	117

LIST OF FIGURES

Fig. 1. Handheld instruments contain actuators that move the tip relative to the handle.....	18
Fig. 2. Micron system with electronics connecting the PSD trackers and the actuated handheld tool.	19
Fig. 3. Micron actuation with piezoelectric bender motors mounted in a 3-pointed star configuration. Tip movement is caused by flexing the bender motors.....	19
Fig. 4. (a) Micron three-link parallel piezoelectric manipulator. Each leg of the manipulator has two actuators mechanically series-connected in a folded configuration, generating approximately 600 μm of motion. Each actuator assembly is rigidly fixed to the base plate and is connected to the star-shaped output plate by a polypropylene flexure. (b) Manipulator workspace. Lateral (XY) tip motion is larger than axial (Z) motion due to the tool lever arm. This volume defines the largest disturbance that can be cancelled and limits the achievable motion scaling. (c) Micron 6DOF sensing of the tip pose and handle pose with LEDs tracked by Position Sensitive Detectors (PSDs).	20
Fig. 5. Surgeon using Micron under a surgical optical microscope in an eyeball phantom with cameras capturing and displaying the live, processed video to computer screens.....	21
Fig. 6. Diagram of the human eye (credit to chabacano, used under Creative Commons Attribution-Share Alike 3.0 license). (1) posterior vestibule (2) ora serrata (3) ciliary muscle (4) ciliary zonules (5) canal of Schlemm (6) pupil (7) anterior chamber (8) cornea (9) iris (10) lens cortex (11) lens nucleus (12) ciliary process (13) conjunctiva (14) inferior oblique muscle (15) inferior rectus muscle (16) medial rectus muscle (17) retinal arteries and veins (18) optic disc (19) dura mater (20) central retinal artery (21) central retinal vein (22) optic nerve (23) vorticoose vein (24) bulbar sheath (25) macula (26) fovea (27) sclera (28) choroid (29) superior rectus muscle (30) retina.....	24
Fig. 7. Example of the “open sky” approach to removing vitreous, where a “toilet shape” U cut is made. A sponge is used to draw the vitreous, which is subsequently cut away. Photo credit to [60].	25
Fig. 8. Pars plana vitrectomy with three ports through the sclera for infusion, light, and an active tool. Photo credit to [60].	25
Fig. 9. Motivation for this work derives from retinal conditions and their effect on eyesight. Laser photocoagulation treats retinal ailments such as diabetic retinopathy and macular degeneration, which affect 4.1 million and 1.8 million Americans, respectively. Photo credit to NIH.	26
Fig 10. Placement of laser burns around a retinal break. Similar ring patterns are used around macular holes and other retinal work. Photo credit [67].	27
Fig. 11. Internal Limiting Membrane (ILM) peeling around a macular hole. Photo credit [67].	29
Fig. 12. Demonstration of performing an arteriovenous sheathotomy by separating the sheath that holds the vein and artery together. Photo credit [181].	30
Fig. 13. Planar targets with macroscopic and microscopic views. (a) Printed paper slides for laser photocoagulation (b) Laser etched rubber slides for general micromanipulation tasks (c) Plastic wrap on sorbothane slides for membrane peeling tasks	31

Fig. 14. Table tennis eye model from the bottom half of a ball with a hair painted red to mimic a retinal vessel. (a) Macroscopic view (b) Microscopic view with the green-painted tip of Micron.....32

Fig. 15. Rubber JHU eyeball phantom. (a) Painted inside illuminated with vitrectomy fiber optic light-pipe (b) Micropipette attached to Micron inserted through the sclerotomy port on the right. 32

Fig. 16. (a) Preparation of pig eyes (b) excised porcine retina with retinal vessel lying flat in an open-sky environment.33

Fig. 17. Ex-vivo porcine retina in the modified Johns Hopkins University eye phantom. (a) Phantom filled with silicone oil and equipped with a standard 27 gauge light-pipe vitrectomy port on the left. On the right is a thin rubber film glued over a larger hole so the shaft of the instrument has more flex, similar to the sclera in surgery. (b) Standard two-hand grip of the light-pipe and Micron, as viewed from the front (c) Two-handed grip viewed from back right (d) Two-handed grip view from back left (e) Low magnification view of retina in eye phantom viewed through the vitrectomy lens and silicone with white-painted micropipette on right. (f) High magnification view of porcine retina in the eye phantom.35

Fig. 18. System setup with (a) Micron, (b) ASAP position sensors, (c) surgical microscope, (d) image injection system, (e) stereo cameras, and (f) phantom half eyeball..... 37

Fig. 19. Example of handheld micromanipulation with position-based virtual fixtures, which drives the tip position PT to a goal position PG on the virtual fixture V . The goal position is calculated by the orthogonal projection MO of the null position PN , which exactly reflects the hand motion. Note: figure not to scale.39

Fig. 20. Block diagram of the control and vision system. The vision system registers the cameras to ASAP and generates virtual fixture geometry from the anatomy. The control system includes feedback and feedforward paths for more precise control of the instrument tip. The virtual fixture controller uses the filtered hand motion to enforce the subspace constraints of the fixtures.40

Fig. 21. Virtual fixtures constraining the tip to a subspace with increasing degrees of freedom: (a) point, (b) curve, (c) surface, and (d) volume.42

Fig. 22. Effect of motion scaling parameter λ on the tip position. (a) Hard virtual fixture ($\lambda = 0$). (b) Soft virtual fixture with $\lambda = 1/2$; control is shared 50% between virtual fixture and hand motion. (c) Fixtures off ($\lambda = 1$).43

Fig. 23. Micron enforcing a point hard virtual fixture with visual cues: green represents the goal, blue represents the null position. (a) Nearly coincident circles are desirable, indicating Micron is near the center of the range of motion. (b) The position and size of the blue circle indicates drift. Here, the operator has drifted left, down, and upwards in the X, Y, and Z directions, respectively. If Micron was turned off, the tip would snap to the blue circle..... 45

Fig. 24. (a) Laser etched target onto rubber surface (b) Generating a 3D circle virtual fixture from the tracked target (c) White-painted, tapered tip of Micron.....46

Fig. 25. (a) Half eyeball phantom with a dyed hair tapped to yellow paper affixed to the inside of a ping pong ball with green-painted tip of Micron visible. (b) 3D reconstruction and 3D spline

representation of the vein, with dark blue representing low Z positions and dark orange representing high Z positions. Although the vein lies on the surface, it is visualized with red.....47

Fig. 26. Mean RMS error across seven trials of each combination of task and scenario, with standard deviation error bars. Hard fixtures significantly reduce error compared to unaided and shelving filter scenarios ($p < 0.05$). (top) 3D mean RMS error (bottom) 2D mean RMS error48

Fig. 27. Hold Still results (a) unaided (b) aided with shelving filter (c) aided with soft fixtures (d) aided with hard fixtures.49

Fig. 28. Vein Tracing results (a) unaided (b) aided with shelving filter (c) aided with soft fixtures (d) aided with hard fixtures.49

Fig. 29. Circle Tracing results (a) unaided (b) aided with shelving filter (c) aided with soft fixtures (d) aided with hard fixtures.....50

Fig. 30. Move and Hold results (a) unaided (b) aided with shelving filter (c) aided with soft fixtures (d) aided with hard fixtures. Deviations from the virtual fixtures in (c) and (d) result from saturation of the actuators, caused by tremor or drift in excess of the range of motion of Micron....50

Fig. 31. (a) ASAP measuring and displaying instrument position on LCD screen. (b) Instrument tip with LEDs being inserted through face mask eye. The two notches on opposite sides function as simulated sclerotomies to brace the tool.52

Fig. 32. AR and ARMA results for X (top), Y (middle), and Z (bottom) tremor components for a typical subject under Scenario 1 (instrument braced against side of eye). For this sample, tremor amplitude is 45.4, 62.9, and 5.4 μm RMS in x, y, and z, respectively.....54

Fig. 33. AR and ARMA results for X (top), Y (middle), and Z (bottom) tremor components under Scenario 2 (instrument not braced against side of eye). For this sample, tremor amplitude is 46.4, 31.9, and 24.1 μm RMS in x, y, and z, respectively.....55

Fig. 34. AR and ARMA results for X (top), Y (middle), and Z (bottom) tremor components for a typical subject under Scenario 3 (instrument braced, but no visual feedback). For this sample, tremor amplitude is 254.0, 166.1, and 15.2 μm RMS in x, y, and z, respectively.56

Fig. 35. Box and whisker plot of the ARMA coefficients (the error term coefficients yield a similar plot with smaller coefficients).56

Fig. 36. Mean percentage improvement for AR and ARMA in comparison with last-value prediction.56

Fig. 37. Kalman filter input is the 6DOF position of the Micron handle (i.e. hand position) and the output is the 3DOF position, velocity, and acceleration along with the 3DOF angular position and velocity.....57

Fig. 38. State estimation of hand motion by Kalman Filter compared to raw measurements.59

Fig. 39. Handle velocity ph' caused by tremor causes error in the tip position pt' over the latency period.60

Fig. 40. (a) Tip position with Kalman filter estimated velocity (b) Velocity-predicted error by integrating velocity over latency time T compared to tip position error. (c) Inverted velocity-predicted error is used as feedforward control.....	62
Fig. 41. Comparison of the feedback vs. feedback+feedforward controller for tremor suppression ((a) top: transverse, (b) bottom: axial). The controller is swapped out dynamically at $t = 1$ s. Although tremor suppression is not completely effective in either case, the addition of the feedforward component is clearly beneficial.	63
Fig. 42. Power spectrum of the two controllers under handheld conditions: the baseline controller and the feedforward controller. For comparison to a tremor-free case, the baseline feedback controller operating in a clamped rubber vise is shown also.	63
Fig. 43. Frequency response of the two controllers under handheld conditions: the baseline feedback controller and the feedback+feedforward controller. Notice the additional attenuation of error at low frequencies in the 5-20 Hz tremor range. This frequency response is for the transverse axis of the manipulator; the lateral response is similar.	64
Fig. 44. White-painted tip of Micron being tracked with the hollow black circle representing the centroid of the shaft, and the black line originating from the centroid indicates the direction and length of the tip.	65
Fig. 45. Multiple view camera geometry, showing the projective relationship $M1$ and $M2$ that map the 3D point P to image coordinates $p1$ and $p2$ viewed by two cameras. Note this is for the general case; the cameras views in the Micron setup are parallel.	66
Fig. 46. 3D Quadratic parametric modeling of the phantom eyeball. (a) One of the raw stereo frames (b) Estimated depth map (c) 3D quadratic fit estimated with RANSAC (d) Error between the estimated 3D quadratic and the actual depth map (white indicates no error, darker is higher error).	67
Fig. 47. The proposed mapping and localization algorithm for retinal vasculature running in real-time on recorded <i>in vivo</i> human retina during a retinal peeling with blue dye. Video source: http://youtu.be/CTnavOgDsXA	68
Fig. 48. Block diagram detailing the steps in our proposed algorithm EyeSLAM that maps and localizes retinal vessels during intraocular surgery. Retinal vessels are detected in each frame, localized to a skeletonized map of the occupancy grid with iterative closest point (ICP), and the map probabilities are updated.....	71
Fig. 49. Snapshots of the feature extraction, mapping, and localization of the EyeSLAM algorithm at various times during 21 s video of porcine retina of an attempted vessel cannulation.	72
Fig. 50. Output of the EyeSLAM algorithm at various points during a 48 s video of the Johns Hopkins eye phantom.	73
Fig. 51. Comparison of our approach (30 Hz) to the vessel detection algorithms of Can et al. [131] (80 Hz), Sofka et al. [138] (0.3 Hz), and Bankhead et al. [141] (8 Hz) on a 20 s video of human retina during an <i>in vivo</i> membrane peel. Video source: http://youtu.be/_naooJFuxPI	74

Fig. 52. Example of a small, sudden jerk at time J and the resulting behavior. The time required for re-convergence is about 150 ms.75

Fig. 53. Tracing a retinal vessel in an eyeball phantom unaided (left) and with the aid a robotic micromanipulator enforcing virtual fixtures based off of our vessel mapping and localization algorithm. Note in both cases, the entire eyeball is moving.76

Fig. 54. Comparison of EyeSLAM to the vessel detection algorithms of Can et al. [131], Sofka et al. [138], and Bankhead et al. [141] on a 18 s video of porcine retina in the Johns Hopkins eye phantom.77

Fig. 55. The membrane peeling procedure shown with plastic wrap on a rubber slide using the white-painted nitinol tip: (a) Approach (b-c) Engagement (d-e) Lifting (f-g) Delamination (h) Peeled membrane.....79

Fig. 56. Micron setup for membrane peeling. (a) Micron (b) ASAP optical sensors (c) Surgical microscope. (d) Red rubber slide mounted on the load cell with plastic wrap for membrane peeling (e) Stereo cameras. (f) Image injection system overlays visual cues in the microscope eyepieces....80

Fig. 57. Illustration of varying levels of position-based virtual fixtures: (a) Virtual fixture (b) Tip position (c) Null position. As the surgeon moves the null position below the virtual fixture boundary, the tip position is influenced by the activated behavior. Without Micron, the tip position follows the null position. With a hard stop, the tip position follows the virtual fixture boundary closest to the input null position. Motion scaling compromises by sharing control between the fixture and human.....81

Fig. 58. (a) Left image of the stereo pair (b) Corresponding disparity map. Note the disparity map has been exaggerated to illustrate the overall reconstruction; the maximum height difference in the surface is only a couple hundred microns. Figure best viewed in color.....82

Fig. 59. Displacement above the surface and force measurements during clamped experiments with three different behaviors: (a) Tremor compensation with shelving filter, (b) Simple low-pass tremor compensation + motion scaling below 50 μm , and (c) Low-pass tremor compensation + motion scaling below 50 μm + a hard stop at -50 μm . In all cases, Micron was manually lowered and raised 600 μm using a micropositioner over a 10-15 s period.84

Fig 60. Micron peeling example with sorbothane slide. (a) Stereo disparity map with black square showing RANSAC plane fit area (b) Force graph (c) Tip displacement. Note (a), (b), and (c) are not visible to the surgeon during the procedure.....84

Fig 61. Surgeon results for membrane peeling with maximum upwards and downwards force per trial, averaged over 10 randomized unaided and aided runs. (a) Rubber slide (b) Sorbothane slide. *Significant results ($p < 0.05$).....85

Fig. 62. Displacement and force results from a typical novice membrane peeling: (a) Unaided trial (b) Aided trial. Shaded areas indicate which behaviors are active during the experiment. Note that since Micron is off in (a), the tip position is the null position. The force-limiting effect of the hard stop is marked with a dashed line in the aided force graph (b). Figure best viewed in color.....86

Fig 63. Displacement and force results from a typical aided membrane peel performed on sorbothane by a surgeon. (a) Unaided (b) Aided87

Fig. 64. The setup for laser photocoagulation with Micron, an active handheld micromanipulator. Micron (a) is shown held in the hand of the operator. The position of Micron is measured optically in 3D by the two ASAP PSDs (b) and stereoscopically in 2D via two cameras (c) mounted to the microscope (d). Real-time video with overlaid information is displayed on the 3D computer monitor (e). The laser (f) connects to Micron (a) via fiber optic (g). Targets are placed under the hollowed out eye of the face model (h).89

Fig. 65. Micron laser setup. The Micron handle attaches to the laser via a fiber optic cable. Cameras track the finder beam and target, while ASAP optically tracks the tip and pose of Micron. The goal location is determined by intersecting the ray connecting the pivot point and the target with the neutral plane. The piezoelectric actuators then deflect the tip from the neutral position to the goal position and fire the laser to burn the target..... 92

Fig. 66. Graphical depiction of the entire process of the proposed laser photocoagulation system. .93

Fig. 67. Block diagram of system, showing data and execution flow. The preoperative procedure is executed first, followed by the initial calibration. Following these two steps, the vision and control systems are then executed and run simultaneously for the duration of the operation. 94

Fig. 68. (Left) Paper slide size in comparison to a US penny. (Right) Results from paper slide trials. Green solid dots indicated preoperatively specified burn locations, and red hollow circles indicate the centroid of the burn. (a) Top left: 0.5 Hz unaided. (b) Bottom left: corresponding aided trial from 0.5 Hz trial. (c) Top right: 1.0 Hz unaided. (d) Bottom right: corresponding aided from 1.0 Hz trial. 95

Fig. 69. Results from porcine retina trials. Green solid dots indicated preoperatively specified burn locations, and red hollow circles indicate the centroid of the burn. (a) Top left: 2.0 Hz unaided. (b) Bottom left: corresponding aided trial from 2.0 Hz trial. (c) Top right: 2.0 Hz unaided with errors depicted. (d) Bottom right: corresponding aided trial with errors depicted. 97

Fig. 70. Mean error vs. duration for all experiments (except the Aided case where the surgeon was asked to perform the operation slowly and deliberately) shows that the surgeon can be either slow and accurate or fast but with lots of error. When the procedure is performed Aided with Micron, the results are both fast and accurate. 98

Fig. 71. Mean error bar chart with error bars representing the standard deviation for all three trials with laser repeat set to 0.5, 1.0, 2.0 Hz from left to right. Each scenario consisted of four unaided trials and four aided trials interleaved. 99

Fig. 72. Mean duration bar chart with error bars representing the standard deviation for all three trials with laser repeat set to 0.5, 1.0, 2.0 Hz from left to right. Each set consisted of four unaided trials and four aided trials interleaved. The apparent discrepancy in the 0.5 Hz case is because the surgeon was instructed to take his time during the aided trials. 99

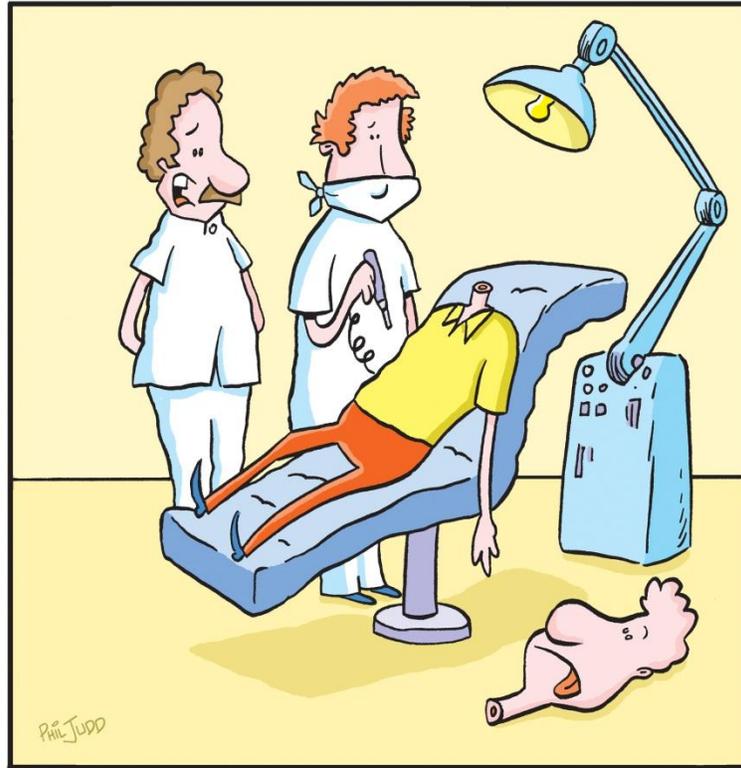
Fig. 73. Micron handheld micromanipulator with glass pipette attached with tubing to a 20 mL syringe..... 101

Fig. 74. Cannulation setup with Micron micromanipulator, microscope, stereo cameras, and PSD optical sensors.....	101
Fig. 75. Stereo cameras track 3D positions of green and blue fiducials on the micropipette to estimate the tip position. The vision system also detects vessels in the image to calculate the center of the vessel nearest the micropipette tip.....	102
Fig. 76. Example traces of the micropipette tip during cannulation on porcine retina overlaid onto microscope images, showing tremor during the procedure without Micron (Unaided) and with Micron (Aided).....	105
Fig. 77. Example traces of the micropipette tip during cannulation on chicken eggs overlaid onto microscope images, showing tremor during the procedure without Micron (Unaided) and with Micron (Aided).....	105
Fig. 78. Learning curve of cannulation success for 40-60 μm vessels over four sessions spaced over a 3 month period.....	106
Fig. 79. Surgeon's pointing accuracy increases with virtual fixtures in circle and vein tracing tasks.	109
Fig. 80. Feedforward control uses tremor models to anticipate, predict, and correct error caused by latencies in actuation for improved tremor rejection.....	109
Fig. 81. EyeSLAM performing real-time localization and mapping of the retinal vasculature in a difficult video. Notice the high magnification, very sparse vasculature tree, noisy features, and occlusions present.....	110
Fig. 82. Retinal eye model with porcine retina in a modified Johns Hopkins eye phantom.....	110
Fig. 83. Three retinal procedures reported: (a) membrane peeling (b) laser phogocoagulation (c) vessel cannulation.....	111
Fig. 84. New 6-DOF Micron manipulator prototype using a Stewart-platform to achieve a larger range of motion with a remote center of motion.....	113
Fig. 85. Example of using the virtual fixture framework in simulated stapes footplate surgery.	114
Fig. 86. Orientation virtual fixtures constrain not the tip position, but where the tip is pointing. An example of this is laser targeting.	115

LIST OF TABLES

Table 1: Comparison of microsurgery manipulation aids [27]	18
Table 2: Various problems occurring in the retina and corresponding solutions that are used by vitreoretinal surgeons.....	26
Table 3: Max error in positioning tasks. Mean of the 3D max error across all seven trials of each combination of task and scenario. Significantly reduced results compared to the hard fixtures case are bolded ($p < 0.05$). Because vein tracing experiments used lower magnification and larger motions were necessary (10 mm), saturation was encountered more frequently than in other tasks.	49
Table 4: Error of controllers with mean RMS performance represented as percent improvement over the baseline feedback-only controller. Best attainable tremor suppression should be approximated by the case where Micron is clamped in a vise (i.e. no tremor).....	62
Table 5: Popular vessel detection algorithms and their running times listed by year.....	69
Table 6: Existing retinal registration or mosaicking approaches along with the method of registration and time required to process a pair of images. *No times listed in the paper.....	69
Table 7: Reduction in overall mean error of laser placements between unaided and aided trials.....	98
Table 8: Cannulation results for 40-60 μm diameter vessels.....	105

PROBLEM DOMAIN



"Still getting the hang of
laser surgery?"

Microsurgery is hard, even for experienced surgeons. As surgeons attempt to manipulate small anatomy in the sub-mm range, human control of surgical instruments becomes degraded [1]. In particular, retinal surgeries are especially challenging because of the extremely small and delicate nature of retinal anatomy. Retinal vasculature is often less than 100 μm and membranes such as the Internal Limiting Membrane (ILM) in the eye are only 5-10 μm thick [2]. New surgical procedures in the eye depend on safely manipulating such small structures [3-10]. For retinal operations such as laser photocoagulation, vessel cannulation, and membrane peeling, physiological tremor becomes a serious impediment. Peak-to-peak tremor of vitreoretinal surgeons can exceed 100 μm [11], creating potentially dangerous unintentional movements and loss of dexterity which can lead to complications or the inability to successfully complete surgical procedures.

In response to these demanding operations, robotic surgery aims to combine the skill of surgeons with the precision of machines [12-16]. Numerous robotic solutions for eye surgery have been introduced [17-26], but this thesis focuses on Micron [27], the fully handheld micromanipulator developed at Carnegie Mellon University. Specifically, the problem of interest is how to fuse visual information gathered from cameras mounted to the microscope to design intelligent control behaviors for Micron that unobtrusively aid surgeons during vitreoretinal procedures.

1 INTRODUCTION

A popular solution to the problem of manipulation at small scales is to use robotic aids, such as micromanipulators [28]. Micromanipulators provide the surgeon with instrumentation to more precisely control micron-level movements. Usually, micromanipulators feature a number of useful behaviors, including tremor suppression, motion scaling, or limits on movement speed or operating volume. Tremor suppression acts as a filter to reduce involuntary, physiological human movements that usually occur around 8-20 Hz [29], while allowing lower-frequency voluntary motion. Motion scaling increases movement precision by reducing all movements made by the operator by a fixed amount; thus large hand motions become identical, but smaller tip motions. For safety, often the micromanipulator enforces a maximum velocity on the tip of the instrument and can prevent the surgeon from straying too far away from the center of the workspace. These features are generally either always active (i.e. tremor suppression) or activated manually by the operator. As evidenced by the widespread use of surgical robots such as the da Vinci by Intuitive Surgical and the published work on micromanipulators [18], [19], [30–32], these tools represent an effective way to perform delicate micromanipulation tasks.

1.1 TAXONOMY OF MICROMANIPULATORS

For the purposes of this thesis, we define a taxonomy of micromanipulators consisting of three general classes: master/slave, cooperative, and active handheld. We describe the defining attributes of each class with examples. Since several micromanipulators exhibit aspects of more than one class, the boundaries should serve as general guidelines rather than strict divisions.

Master/slave micromanipulators separate the human input into one subsystem and the instrument actuation into another (often remotely located) subsystem. Surgical Intuitive’s da Vinci [30] is perhaps the most popular master/slave surgical robot and has been outfitted with a prototype manipulator for intraocular microsurgery [22]. The micromanipulators designed by Hunter et al. [33], Jensen et al. [34] and Ueta et al. [19] are more examples of a master-slave robots designed specifically for retinal microsurgery. Because there is no mechanical linkage between the input and output, the available workspace is large and control is usually quite versatile as arbitrary tremor suppression filters and behaviors can modify the signals on the drive-by-wire connection. However, the additional pathways from input to control introduce extra latency and remove physical feedback, such as the ability to sense forces applied to the tissue. Because of the equipment required for the two separate subsystems and the complexity of synchronizing them, master/slave configurations are often more expensive than other micromanipulators.

The class of cooperative micromanipulators involves robotic arms and linkages that the human can grasp to jointly share control between the human and the robot [35–37]. Johns Hopkins University’s Steady-Hand Eye-Robot [23], [31] is an example of a robot arm for microsurgery that is operated cooperatively by the human and robot. The human guides the micromanipulator while the robot acts as a mechanical low-pass filter to smooth movement. Hard constraints on movement are particularly easy to enforce with non-backdrivable linkages. Because of the inherent stiffness in the mechanical linkages, movement is often sluggish and pressures exerted on tissue by the tip are not felt by the operator, although recent force-sensing instrumentation at the tip adds force feedback

[24], [38]. More precise movements are usually achieved with virtual fixtures by commanding scaled velocities based on the forces applied to the robot handle [39]. A number of virtual fixtures frameworks have been developed for the SteadyHand [39–42] that use forces applied to the cooperatively controlled handle, providing even more fine-grained control to the operator.

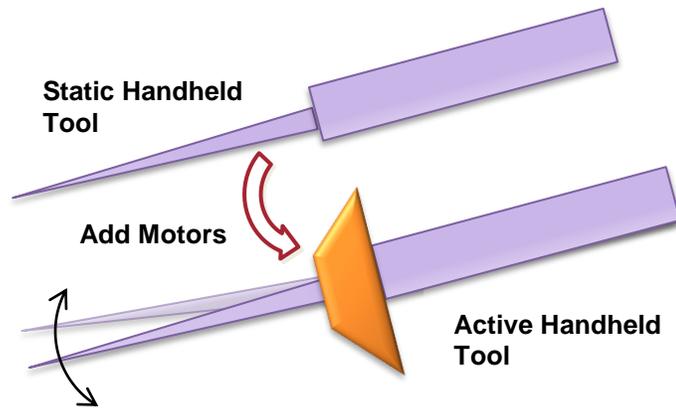


Fig. 1. Handheld instruments contain actuators that move the tip relative to the handle.

Stetten et al. introduces a cooperative tool that senses forces at the instrument tip during microsurgery and amplifies, transmits, and applies them to the hand, allowing the operator to feel magnified forces during micromanipulation [37]. Unlike other cooperative micromanipulators, no robot arm is used and thus the tool can be considered to belong to both the cooperative and handheld classes. However, no control of the tip is possible; instead the device relies on the magnified forces to guide surgeon to more deftly manipulate tissue without applying excessive force. Thus, we make the distinction between handheld tools such as [37] and active handheld tools, which actively actuate and control the instrument tip.

Finally, the micromanipulator class of most interest to the focus of this thesis is the active handheld micromanipulator proposed by Riviere et al. [43], [44]. In a handheld micromanipulator, the entire instrument, including actuation, is contained in the shaft of the instrument, and control of the instrument tip happens by offsetting the tip relative to the current handheld position (see Fig. 1). The Micron robot [27], [32] developed in our lab at Carnegie Mellon University is the first of its

	Unaided	Master/slave	Cooperative	Micron
Motion scaling:	No	Yes	No	Yes
Workspace intrusion:	No	Slave arm & master	Arm	Active tool, sensor sightlines
Force feedback:	1:1	Research area	Yes (superimposed on damping)	1:1
Set and forget hold:	No	Yes	Yes	No
Features:	Current practice	Could combine all of the above features, telemedicine	Inexpensive position-output actuators and simple control	Hand-held operation improves user acceptance and safety, mechanical simplicity
Challenges / costs:	Tremor limits accuracy & repeatability	Unproven force feedback performance / greatest mechanical and control complexity (high cost)	Dexterity fundamentally limited by force → rate user interface and low control bandwidth	Manipulator size and range, high bandwidth control / measurement subsystem cost

Table 1: Comparison of microsurgery manipulation aids [27]

kind, a fully handheld micromanipulator not much larger than a typical un-actuated surgical tool. Behaviors are accomplished by moving the tip relative to the hand motion via actuators between the handle and the tip. For instance, high-speed control performs tremor suppression by rapidly moving the instrument opposite to hand tremor. While remaining very lightweight and flexible, handheld micromanipulators have a limited range of motion and require high rate sensing and control to actively compensate for hand motion in addition to enforcing behaviors. However, there are a number of key advantages to this design

including familiarity, reliability, and inexpensiveness. Handheld designs are familiar to surgeons, who can leverage their extensive experience with handheld instruments operated under traditional surgical-grade microscopes to avoid lengthy training processes. If the instrument stops working properly, a surgeon can simply turn it off, and complete the surgery as usual with the inert handheld instrument. This increases safety as the procedure can always be performed no differently than a typical, unaided surgery with regular tools. Likewise in case of device malfunction, a surgeon can more readily move the tool's small form factor away from any tissue. Finally, handheld micromanipulators are often comparatively inexpensive. All proposed research will be performed with Micron, a tool with demonstrated benefits including tremor compensation, motion-scaling, and other aids shown in experimental studies with surgeon subjects [27], [45], [46]. Table 1 shows a comparison of these three approaches across multiple factors [27].

1.2 MICRON: HANDHELD MICROMANIPULATOR

Handheld micromanipulators present a challenging engineering problem, requiring high control bandwidth, low error position measurements, and sufficient range-of-motion to complete the desired micromanipulation tasks. To date, Micron [27] and its descendent design ITrem [47], are the only handheld micromanipulators designs with active control of the tip that have been built and reported. We briefly describe the Micron system, shown in Fig. 2.

1.2.1 HANDPIECE

Micron, pictured in Fig. 2, Fig. 3, and Fig. 4, is a fully handheld micromanipulator that self-

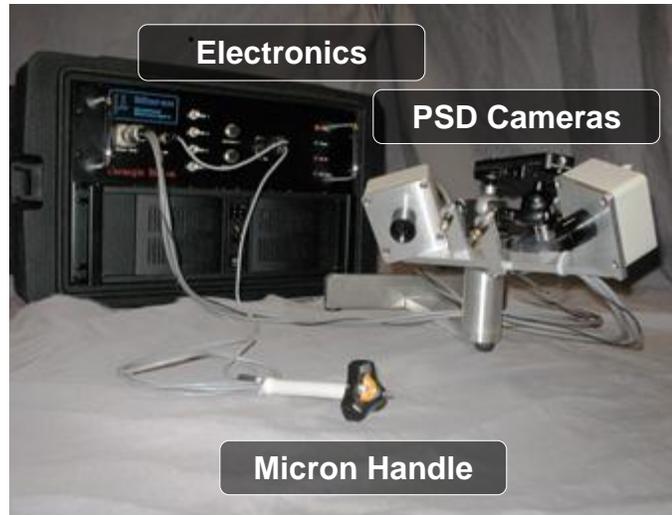


Fig. 2. Micron system with electronics connecting the PSD trackers and the actuated handheld tool.

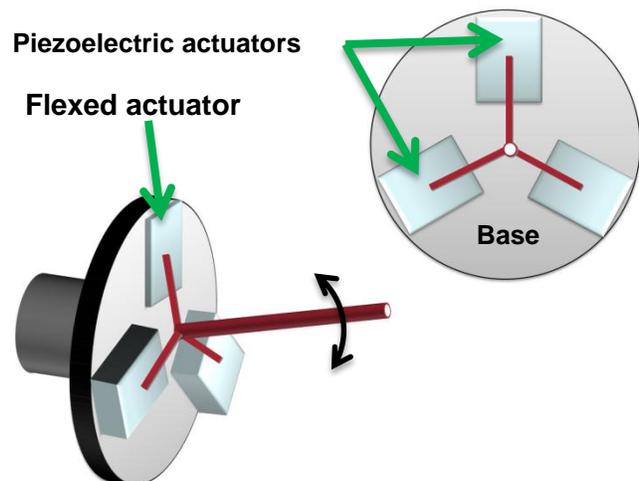


Fig. 3. Micron actuation with piezoelectric bender motors mounted in a 3-pointed star configuration. Tip movement is caused by flexing the bender motors.

actuates via Thunder® TH-10R piezoelectric actuators (Face International Corp., Norfolk, Virginia, USA) located between the handle grip and the tip of the instrument. Arranged in a 3-pointed star configuration, the piezoelectric actuators can flex in and out individually to give Micron a 3DOF range of movement that is approximately 0.6 mm axially and 1-2 mm transversely, depending on the tool mounted to Micron’s base. Fig. 3 shows the arrangement of motors and a diagram of actuator operation. Actuators are driven with charge control at a 2 kHz bandwidth in a closed loop feedback control system that uses the position measurement system described in the Section 1.2.2. Control wires, laser fiber optics, and other surgical instrumentation fit down the hollow center of the shaft and to the tip of the instrument via flexible linkages.

1.2.2 POSITION MEASUREMENT

Low-latency, high-bandwidth positioning information is obtained from custom optical tracking hardware named ASAP [48]. Shown in Fig. 2, two position sensitive detectors (PSDs) at 60 ° track LEDs mounted to Micron within a 4 cm³ workspace at a rate of 2 kHz, with measurement accuracies of <10 μm RMS at the tip of the instrument. The 3D positions of three LEDs on the actuated shaft holding the end-effector are reconstructed and used to calculate the 6DOF position tracking of Micron’s tip with Horn’s equation [49]. An additional LED attached to the handle provides hand movement information. See Fig. 4(c) for LED arrangement.

1.2.3 TREMOR SUPPRESSION

Tremor compensation can be achieved by inserting filters between the hand motion

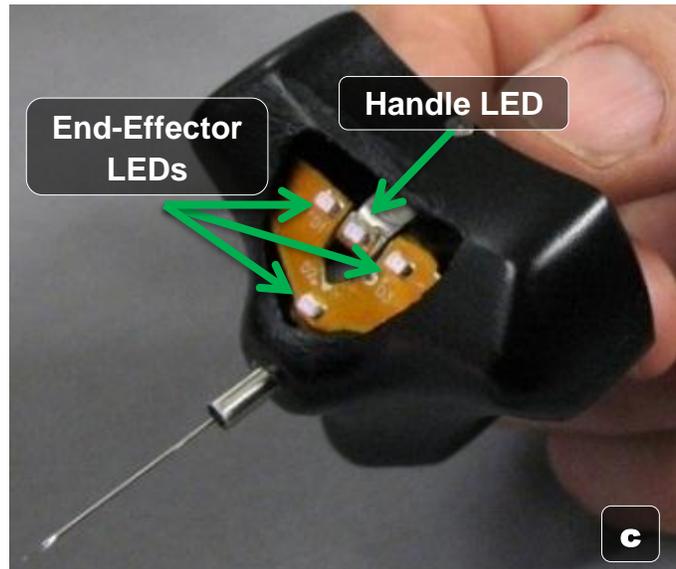
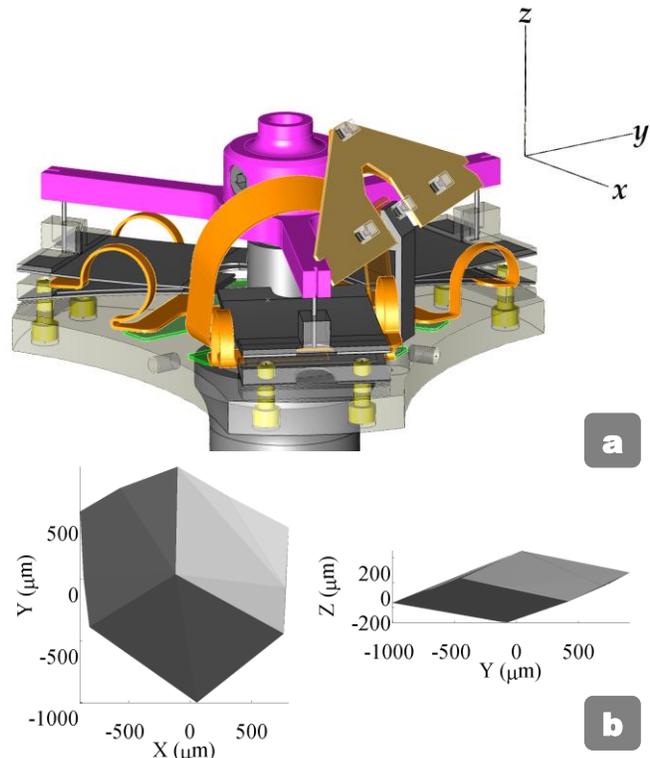


Fig. 4. (a) Micron three-link parallel piezoelectric manipulator. Each leg of the manipulator has two actuators mechanically series-connected in a folded configuration, generating approximately 600 μm of motion. Each actuator assembly is rigidly fixed to the base plate and is connected to the star-shaped output plate by a polypropylene flexure. (b) Manipulator workspace. Lateral (XY) tip motion is larger than axial (Z) motion due to the tool lever arm. This volume defines the largest disturbance that can be cancelled and limits the achievable motion scaling. (c) Micron 6DOF sensing of the tip pose and handle pose with LEDs tracked by Position Sensitive Detectors (PSDs).

and the drive mechanism to reduce the higher-frequency tremor while retaining the operator's gross lower-frequency movements. A 2nd order, low-pass Butterworth filter based on the handle motion is used to implement tremor reduction. Motion scaling, often used in conjunction with tremor suppression, scales hand movements so the tip only moves a set fraction of the hand movement. For example, a scale factor of $\frac{1}{2}$ would transform hand movement of 2 mm to 1-mm movement at the robotic tip. Recent developments in tremor compensation include a low-pass shelving filter [27], which acts as a hybrid tremor suppression filter that affords relative motion scaling and reports an additional 30-50% reduction in tremor compared to the low-pass only filter.

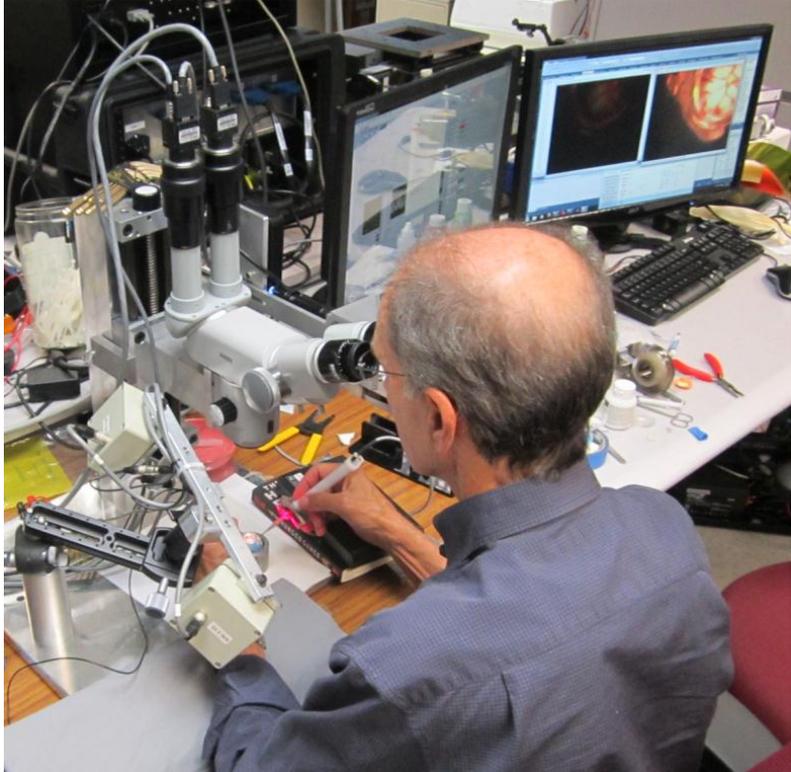


Fig. 5. Surgeon using Micron under a surgical optical microscope in an eyeball phantom with cameras capturing and displaying the live, processed video to computer screens.

1.2.4 MICROSCOPE & CAMERAS

The operator uses a Zeiss® OPMI® 1 microscope to view the workspace under 25x magnification. Two Flea2® cameras (Point Grey Research, Richmond, B.C., Canada) are mounted to the eyepiece. They can acquire custom-defined resolutions at variable rates over a 2x3 to 12x18 mm area depending on microscope magnification; nominally, they can capture 1024x768 @ 30 fps in a raw color Bayer format. The real-time video is displayed to the operator on a 22" 3D computer display (Trimon ZM-M220W, Zalman Tech Co., Ltd., Seoul, Korea). The computer display enables the system to easily overlay extra information on the real-time video, such as target locations or depth cues, to aid the operator. Alternatively, an image injection system based on a picoprojector can display custom images, video, and visual cues directly into the right microscope eye-piece [50]. Fig. 5 shows the surgeon using Micron in a typical operational setup.

1.3 PROBLEM STATEMENT

Unfortunately, most existing techniques utilized in robotic micromanipulators such as tremor suppression and motion scaling lack any sort of understanding of the surgeon's intentions, plans, or goals. Used as black-boxes, each behavior operates blindly upon activation in the same manner regardless of what the surgeon needs or the situation requires. It is the surgeon's sole responsibility to turn behaviors on and off as well as tweak parameters to keep the behaviors at maximum

helpfulness during the entire procedure. For instance, a robotic system may have a motion-scaling knob defining the scaling parameter that the operator must manually change when nearing the targeted tissue or when transitioning between steps in a procedure. In fact, it is possible that a behavior designed for aid may become a hindrance if ill-tuned or accidentally applied inopportunistly. The fundamental assumption underlying this thesis is that further aids to surgeons necessitate at least basic domain knowledge of surgical procedures coupled with real-time monitoring and at least rudimentary understanding of a surgeon's actions and goals.

Much like an auto-pilot system eliminates the need for constant course corrections and tweaking many different knobs by understanding the aviator's flight plan and the current situation via instrumentation, our goal is to provide task-specific automation behaviors by understanding the surgeon's *operation plan* and the current state of the procedure *via visual feedback*. However, just as an auto-pilot does not eliminate the need for trained and experienced pilots, neither should a robotic surgical system supplant surgeons. Instead, it will only provide them with tools that reduce tedious repetition, automate mundane procedures, and improve precision so surgeons are free to exploit their extensive judgment and experience without having to worry about low-level performance.

In using visual information, a significant body of research is dedicated to related algorithms such as tracking surgical instruments [51], understanding surgical actions [52], [53], and integrating and evaluating such technologies in robotic systems [31], [42], [45]. While all are necessary for a comprehensive robotic system that is able to synergistically aid surgeons during operations, one particularly promising avenue of research is virtual fixtures. Virtual fixtures enforce constraints on the behavior of the instrument by software controlled hardware. The motivation behind virtual fixtures is that if the exact behavior the operator wants to follow is known, inaccuracies can be eliminated by introducing an artificial fixture that constrains the user to desired behavior. For instance, if you know an artist wants to draw a straight line or a particular shape, a ruler or stencil improves performance by more precisely controlling the drawing instrument than the unaided operator could otherwise manage. For domains such as drafting where the work consists of many simple primitives, extensive use of fixtures is beneficial. Some surgical procedures lend themselves well to compositions of primitive movements, a fact exploited by researchers to use virtual fixtures to show performance improvements in specific surgeries [39], [42], [54–56].

1.4 PROPOSED APPROACH

This thesis focuses on combining virtual fixtures, tremor suppression, motion scaling, and other behaviors into a unified vision-based control framework that targets the unique capabilities and challenges of handheld micromanipulators. While keeping the theoretical aspects of our work applicable in the general surgical mindset, this thesis focuses on applications to retinal surgery, often considered one of the most demanding areas of microsurgery because of the extremely small structures that must be routinely manipulated [57]. Using domain knowledge coupled with real-time vision analysis of the surgeon's actions during the procedure, the proposed theoretical framework and practical system implementation provides task-specific behavioral aids to the surgeon. Model based control and 3D visual sensing enables the application of the proposed framework to targeted retinal surgeries.

1.5 THESIS STATEMENT

We propose three hypotheses which this thesis will investigate:

1. Micromanipulation tasks can be decomposed into steps, during which applying tip position constraints (virtual fixtures) with a handheld tool increase operator precision and safety.
2. Real-time visual analysis of anatomy and the surgeon's movements provides the context that allows the controller to select and customize the virtual fixtures.
3. Vision-based virtual fixtures aid surgeons during microsurgical operations.

Summarizing these hypotheses yields our thesis statement:

Fusing real-time visual information into the control system of a fully handheld robot enables a rich set of task-specific behaviors known as virtual fixtures that can be applied to increase surgeon precision and performance during micromanipulation, such as vitreoretinal procedures.

2 VITRORETINAL MICROSURGERY

The theoretical aspects of handheld micromanipulation presented in this thesis are kept generalizable to many different types of micromanipulation, ranging from microsurgery to industrial processes to biomedical experimentation. However, it is neither practical nor advantageous to design and conduct experimentation covering such a wide variety of manipulation tasks. Instead, this thesis focuses on microsurgery, and in particular retinal microsurgery. Although many areas of surgery are challenging from a manipulation viewpoint, vitreoretinal surgery is considered to be one of the most demanding fields of microsurgery [58]. Thus we propose to use retinal surgery as the practical experimental test bed for the more general theoretical frameworks developed in this thesis.

This section highlights some of the common procedures in the retina and a few promising techniques developed recently. The purpose of this chapter is two-fold. First, it serves to motivate the continued development of handheld micromanipulators (see Fig. 9) and drive the design requirements for the framework of beneficial micromanipulator behaviors developed in Chapter 3. Second, it provides important background information for the validation of the thesis work. We begin by overviewing the structure of the human eye and describing the general group of retinal surgeries known as pars plana vitrectomies. We then overview some of the specific surgeries that are performed in conjunction with a vitrectomy.

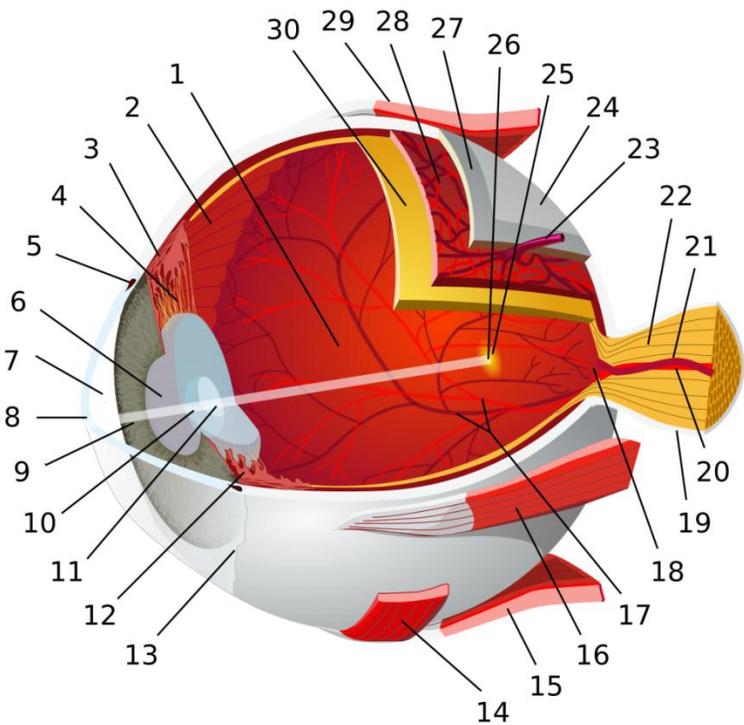


Fig. 6. Diagram of the human eye (credit to chabacano, used under Creative Commons Attribution-Share Alike 3.0 license). (1) posterior vestibule (2) ora serrata (3) ciliary muscle (4) ciliary zonules (5) canal of Schlemm (6) pupil (7) anterior chamber (8) cornea (9) iris (10) lens cortex (11) lens nucleus (12) ciliary process (13) conjunctiva (14) inferior oblique muscle (15) inferior rectus muscle (16) medial rectus muscle (17) retinal arteries and veins (18) optic disc (19) dura mater (20) central retinal artery (21) central retinal vein (22) optic nerve (23) vorticosse vein (24) bulbar sheath (25) macula (26) fovea (27) sclera (28) choroid (29) superior rectus muscle (30) retina.

2.1 ANATOMY OF THE EYE

The human eye is a complex organ with many subsystems, as depicted in Fig. 6. The entire eyeball is held inside the sclera (27), a strong, thin hollow sphere that protects the eye. Light enters the clear cornea (8) through the pupil (6) in the iris (9) and is focused by the lens (11). It passes through a clear, jelly-like substance known as vitreous humor (also referred to as the vitreous body or just

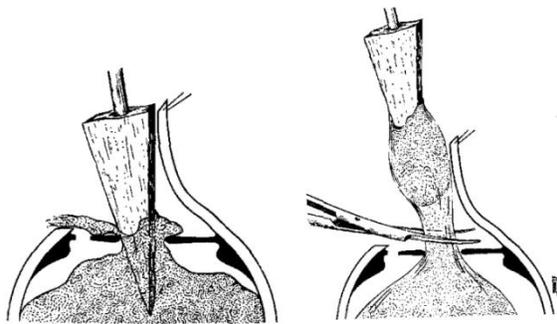
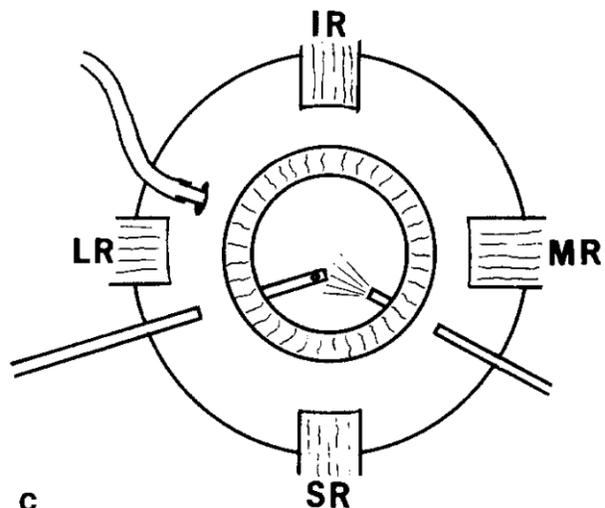


Fig. 7. Example of the “open sky” approach to removing vitreous, where a “toilet shape” U cut is made. A sponge is used to draw the vitreous, which is subsequently cut away. Photo credit to [60].

vitreous) that keeps the eye inflated. After passing through the vitreous, light hits specialized sensor cells in the retina, the rods and cones, (30), which transmits the light information to the brain via the optic nerve (20). The retina is approximately 130 μm to 550 μm thick and is nourished by various vasculature (17, 20, 21). The retina is separated from the vitreous by the internal limiting membrane. Many problems occurring in the retina require surgical intervention, as detailed in subsequent sections.

2.2 PARS PLANA VITRECTOMY

Surgical access to the retina is difficult because of its placement behind the sclera and vitreous. In the late 1960s, Kasner developed an “open sky” vitrectomy technique for cataract surgery [59]. His solution was to cut through the sclera and remove the vitreous with a procedure known as a vitrectomy. The term vitrectomy literally means to excise the vitreous. During a vitrectomy, a “U” shape incision is cut into the sclera, the cornea temporarily removed, and the vitreous excised with sponge and scissors (see Fig. 7). Machemer [60] is credited with popularizing the general technique known as pars plana vitrectomy, which provides much less intrusive access to the retina. He extended Kasner’s technique by replacing the “open sky” approach of making a large U-shaped incision in the sclera with several small ports in the pars plana section of the sclera through which instruments could be inserted. Furthermore, he developed a vitreous cutter that could be inserted through the port, which cut and removed the vitreous via a blade sliding across an opening connected to suction pumps. Generally, a pars plana vitrectomy is performed as shown in Fig. 8, with three ports: a light-pipe for illumination, an active tool (such as vitreous cutter, forceps, etc.), and an infusion line to maintain proper ocular pressure. Throughout the 70s, the pars plana vitrectomy procedure was used to treat previously incurable problems, including non-clearing vitreous hemorrhages, vitreous traction causing retinal detachment, and giant retinal tears. Today, one may view the pars plana vitrectomy (usually referred to simply as a vitrectomy) as a general procedure, which may have many different goals and sub-procedures. Common retinal problems and associated sub-procedures performed during treatment are listed in Table 2.



C Fig. 8. Pars plana vitrectomy with three ports through the sclera for infusion, light, and an active tool. Photo credit to [60].

Problem/Solution	Laser Photocoagulation	Vessel Cannulation	Membrane Peeling	Arteriovenous Sheathotomy
Diabetic Retinopathy	X			
Retinal Detachment	X		X	
Epiretinal Membrane			X	
Macular Hole			X	
Vessel Occlusion	X	X		X
Foreign Body Removal	X			
Macular Degeneration	X			

Table 2: Various problems occurring in the retina and corresponding solutions that are used by vitreoretinal surgeons.

We choose this subset of procedures because they are known to be particularly challenging (Fig. 9). In the following sections, we discuss each vitrectomy treatment option briefly and describe micromanipulator behaviors that could be applicable to the procedure.

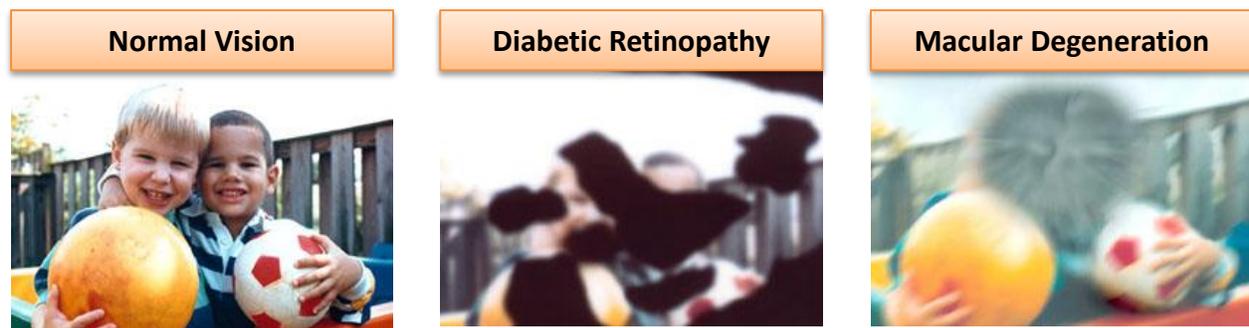


Fig. 9. Motivation for this work derives from retinal conditions and their effect on eyesight. Laser photocoagulation treats retinal ailments such as diabetic retinopathy and macular degeneration, which affect 4.1 million and 1.8 million Americans, respectively. Photo credit to NIH.

2.3 LASER PHOTOCOAGULATION

Laser photocoagulation of the retina is a common adjunct in pars plana vitrectomy surgery for common diseases such as diabetic retinopathy [9], retinal detachment [6], macular edema [4], [8], branch vein occlusion [7], and intraocular foreign body removal [3]. After the anatomical goals of clearing the media and successfully re-opposing the retina to the pigment epithelium, three common laser procedures are used. Panretinal photocoagulation lowers the production of vascular endothelial growth factor (VEGF) by the ischemic retinal cells and decreases the potential for regrowth of neovascular tissue. Peripheral panretinal photocoagulation in retinal detachment seals retinal breaks and stimulates fibrous metaplasia which prevents retinal breaks and retinotomies from leading to re-detachment of the retina. Focal laser patterns are used to surround retinal breaks iatrogenically created to drain subretinal fluid, to treat accidental retinal breaks occurring when scar tissue is being peeled, and to surround traumatic retinal breaks prior to removing intraocular foreign bodies.

Accuracy in laser photocoagulation is important for optimal clinical results as inadvertent photocoagulation of a retinal vein can cause occlusion of the vein, possibly leading to vitreous hemorrhage [61]. Laser applications frequently are applied within one millimeter of the foveola, requiring careful avoidance of unintended targets such as the optic nerve to avoid permanent central vision loss [62]. Furthermore, accidental coagulation of the macular venule or arteriole can cause foveal ischaemia or intraretinal fibrosis [63]. For best results in previously photocoagulated retinas, previous burn locations are also to be avoided [64]. Research efforts to improve positioning of burns have yielded automated approaches [65], but system ergonomics and complexity have prevented clinical adaptation [66].

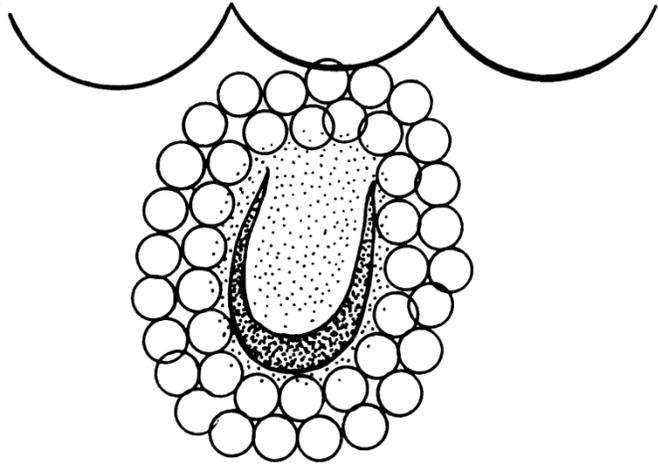


Fig 10. Placement of laser burns around a retinal break. Similar ring patterns are used around macular holes and other retinal work. Photo credit [67].

2.3.1 TREATMENT PROCEDURE

Laser photocoagulation is used in a wide variety of treatments, but can be grouped into two typical scenarios. First, to seal retinal breaks, two or three rings of laser burns as seen in Fig 10 are applied around the surgery site [67]. Second, in pan-retinal photocoagulation, a grid of hundreds to thousands of laser burns is applied to the retina. Burn sizes are typically around 200-400 μm in diameter with laser pulse durations of 10-100 ms. Laser pulses are generally applied at a constant rate of 0.5 – 2 Hz when the surgeon depresses the foot pedal.

2.3.2 BEHAVIORAL AIDS

Because laser photocoagulation is such a demanding, yet repetitive procedure, with an active handheld micromanipulator, several behaviors can be designed to increase speed and accuracy:

- **Select Target:** Select the best target based on closest targets and hand motion.
- **Maintain Orientation:** Keep the tip (and thus the laser) pointed at a targeted burn location during firing of the laser to increase accuracy of burns.
- **Fire Laser:** Automatically fire laser upon target acquisition.
- **Avoid Zones:** Prevent the laser from burning sensitive anatomy such as vessels.
- **Maintain Distance:** Keep the tip of the fiber optic within a specified range of the retina to keep burn energy in the proper range.

2.4 VESSEL CANNULATION

New microsurgical procedures in the eye are yielding encouraging results in the treatment of diseases such as retinal vein occlusion (RVO). RVO occurs when a clot obstructs blood flow in a central or branch vein of the eye (CRVO and BRVO, respectively). As the second most common retinal vascular disease, RVO affects an estimated 16.4 million adults worldwide [68] and has no

proven effective treatment [69]. A promising experimental surgical procedure is retinal endovascular surgery (REVS), which involves cannulating the vein and directly injecting clot-dissolving plasminogen activator (t-PA) to remove obstructions [5]. In a test of 28 eyes with CRVO, 54% exhibited improved visual acuity within 3 months [70], although follow up studies have reported mixed results [5], [71].

In vivo retinal vessel cannulations were first reported with cats and rabbits using a micromanipulator to fix the micropipette and damp hand movements [72]. Although subsequent research has demonstrated successful cannulations in a variety of animal and human models, there are difficulties involved in manipulating such tiny vessels. It has been noted that larger vessels are easier to puncture than smaller vessels more distal to the optic nerve [73]. In experiments with human cadaver eye models, Tang and Han recorded that only 10 out of 18 attempted cannulations were successful and commented that manipulating vessels with currently available instrumentation is generally traumatic [74]. Because retinal vessels are so small (less than 200 μm), it is common to cannulate larger, nearby vessels, which is thought to lessen efficacy [75]. Overall, the procedure is considered difficult to perform [76] primarily because of a surgeon's limited dexterity while attempting to pierce and then keep the cannula in the vessel [75].

2.4.1 TREATMENT PROCEDURE

Before cannulation, intraocular pressure is lowered to about 5 mm Hg, allowing vessels to dilate; following vessel puncture, the pressure should be increased to lessen bleeding [75]. Once the vessel has been punctured, 3.4 ml of rt-PA with a concentration of 200 $\mu\text{g}/\text{ml}$ is injected into the cannulated vessel towards the optic nerve head [58]. The infusion process takes several minutes, during which the cannula should remain in the vessel.

2.4.2 BEHAVIORAL AIDS

In vessel cannulation, precision of movement is of primary importance to increase success rates and reduce collateral damage to surrounding retina. Such precision can be accomplished via:

- **Motion Scaling:** To increase precision with $<100 \mu\text{m}$ vessels, motion around the vessel can be scaled. Anisotropic scaling in the direction of the vein allows unimpeded thrusting puncture movements.
- **Steady Approach:** Since tremor amplitude can be larger than the width of the vessel, removal of tremor is important (indeed, [72], [75], [77] use a micromanipulator).
- **Maintain Position:** During infusion, the cannula should be kept as motionless as possible to prevent trauma to the vessel or accidental dislodging of the cannula.
- **Prevention Zones:** Throughout the procedure, keep the tip of the instrument from contacting the retina to prevent tears or tissue trauma.

2.5 MEMBRANE PEELING

Membrane peeling is a common retinal surgery to remove epiretinal or internal limiting membranes (ILM) [67]. Epiretinal membranes are thin fibrous layers that grow on the retina over time or in conjunction with other retinal diseases. They often distort the retinal surface (a condition known as "macular puckers"), making straight lines appear wavy. The ILM is a layer less than 5 μm

thick that naturally separates the vitreous from the retina. Macular holes occur when the vitreous pulls away and tears the retina, a condition that is common as the eye ages [67]. Kelly and Wendell first demonstrated using vitrectomy procedures to relieve pressure from the vitreous on the retina near macular holes [78]. To facilitate closure of the macular hole, some researchers propose removal of the ILM around the hole to alleviate tangential pressure from the surrounding retina [79]. During the treatment of macular holes, a common condition where the vitreous pulls and tears the retina, removal of the ILM around the hole is advocated to alleviate tangential pressure from the surrounding retina [2]. Removing these membranes is necessary to restore visual acuity and involves a technique known as membrane peeling, a technically challenging procedure that sometimes requires repeated attempts over several minutes [67]. Unsuccessful peeling can result in poor visual outcome [80] and one study showed that up to 50% of patients exhibited inadvertent injury in the nerve fiber layer after peeling [81].

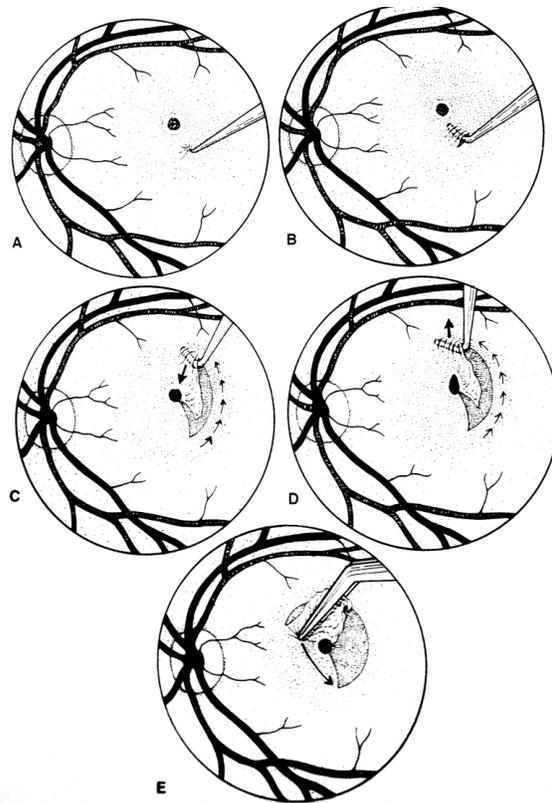


Fig. 11. Internal Limiting Membrane (ILM) peeling around a macular hole. Photo credit [67].

2.5.1 TREATMENT PROCEDURE

Removing the epiretinal membrane or ILM requires a technique known as membrane peeling (shown in Fig. 11). A microvitreoretinal (MVR) or other blade is used to puncture the membrane. A pick or ILM elevator is then inserted just under the membrane to lift it up. This is very challenging work and sometimes takes minutes' worth of attempts [67]. Once the membrane has been lifted with the stripping motion, the pick is advanced. The surgeon tunnels between the membrane and the retina, periodically lifting the membrane off the retina. Alternatively, forceps can be used to remove the peeled membrane. One study showed that up to 50% of patients exhibited inadvertent injury and defects in the nerve fiber layer after performing ILM [81].

2.5.2 BEHAVIORAL AIDS

Because membranes are so thin and are firmly adhered to the retina, behaviors to aid the procedure include limitations on forces that could rip, tear, or damage the retina:

- **Downward Motion Limits:** Preventing fast or extraneous motion in the downward direction to avoid excessive force or accidental trauma to the retina.
- **Motion Scaling:** Increasing precision in the Z motion could reduce the number of tries it takes surgeons to insert the pick between the ILM and the retina.

- **Follow Curved Path:** Once the pick has been inserted, help with the fine tunneling motions (advance in a curve, lift up slightly) could also reduce accidental injury to the retina.
- **Velocity Limiting:** Capping tip velocity during delamination could reduce high upward forces and prevent damage to the retina.

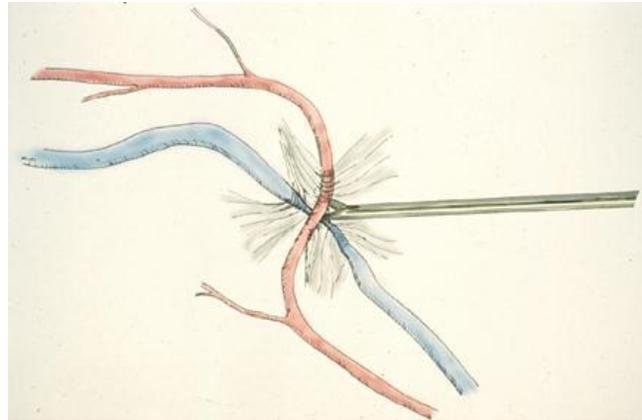


Fig. 12. Demonstration of performing an arteriovenous sheathotomy by separating the sheath that holds the vein and artery together. Photo credit [181].

2.5.3 ARTERIOVENOUS SHEATHOTOMY

One hypothesis in vessel occlusion is the adventitial sheath that binds arteries and veins together at arteriovenous crossings restricts blood flow, causing turbulence, which may lead to thrombosis. This is supported by the fact that the arteriole crosses over the vein between 70-99% of the time when BRVO occurs at arteriovenous crossings [82], [83]. It is thought that by removing the sheath, the vein will decompress, allowing for vessel recannulation. Osterloh and Charles first succeeded in showing surgical decompression via arteriovenous sheathotomy [84]; as a result, researchers have corroborated a variety of similar procedures [85], [86]. However, other experiments have reported failure to surgically separate the vessels because of the difficulty of the operation [87–89].

2.5.4 TREATMENT PROCEDURE

As depicted in Fig. 12, an arteriovenous sheathotomy procedure aims to separate the artery from the underlying retinal vein by severing the sheath that connects the two. A bent microvitreoretinal (MVR) blade or scissors are used to gently cut the sheath and pry the vein and artery apart from each other. Complications include nerve fiber layer defects, hemorrhage, and retinal detachment. Interesting, some research has indicated that simply freeing the vessels from the retina might be enough to restore blood flow [89].

2.5.5 BEHAVIORAL AIDS

Performing a arteriovenous sheathotomy required delicacy to cut and tease the vessels apart while not damaging the vessels; behaviors to aid the procedure could include:

- **Tremor Compensation:** Because of the extremely delicate nature of the operation and the many reported failures and hemorrhages, reducing tremor is essential.
- **Avoidance Zones:** Setting limits on tip movements too far into the retina could prevent tearing or scratching the retina.
- **Motion Scaling:** Reducing movements around the critical junction of the vein and artery would increase surgeon precision in manipulation.
- **Velocity Limiting:** Preventing fast tip movements could help avoid accidental hemorrhage of the blood vessels.

2.6 RETINAL MODELS

Because of the difficulty in testing on real, live human eyes at this preliminary stage in Micron's development, alternative models must be substituted. Such models should be easy to obtain and run experiments on, but as realistic as possible. Depending on the particular task being evaluated, different aspects of realism may be required. For general micromanipulation tasks that validate the theoretical aspects of our work, patterns etched on rubber slides may be sufficient; for cannulation tasks, a realistic vasculature that mimics retinal veins is most important; however, for photocoagulation procedures, the retina itself needs to be most realistic to allow laser burns (i.e. actual retina where burns are evident or an easily burned substitute such as paper slides). Throughout this thesis, we make use of several different simulated eye environments for the evaluation of various procedures, which are discussed in the following subsections.

2.6.1 PAPER, RUBBER, AND SORBOTHANE SLIDES (IN VITRO)

For many tasks, simple planar environments as shown in Fig. 13 are sufficiently good models. For general micromanipulation tasks, rubber slides are durable, mimic tissue flexibility, and can be laser etched with very fine patterns to easily perform tasks such as point, hold, and tracing.

Similarly, printed paper slides can easily be tailored to show any pattern, albeit with fairly low resolution at high magnification unless a very high DPI printer is used. Paper slides are particularly useful for laser photocoagulation tasks as the laser burns are very evident. It is important to ensure that the paper is not entirely white, otherwise the laser may reflect off without leaving a burn.

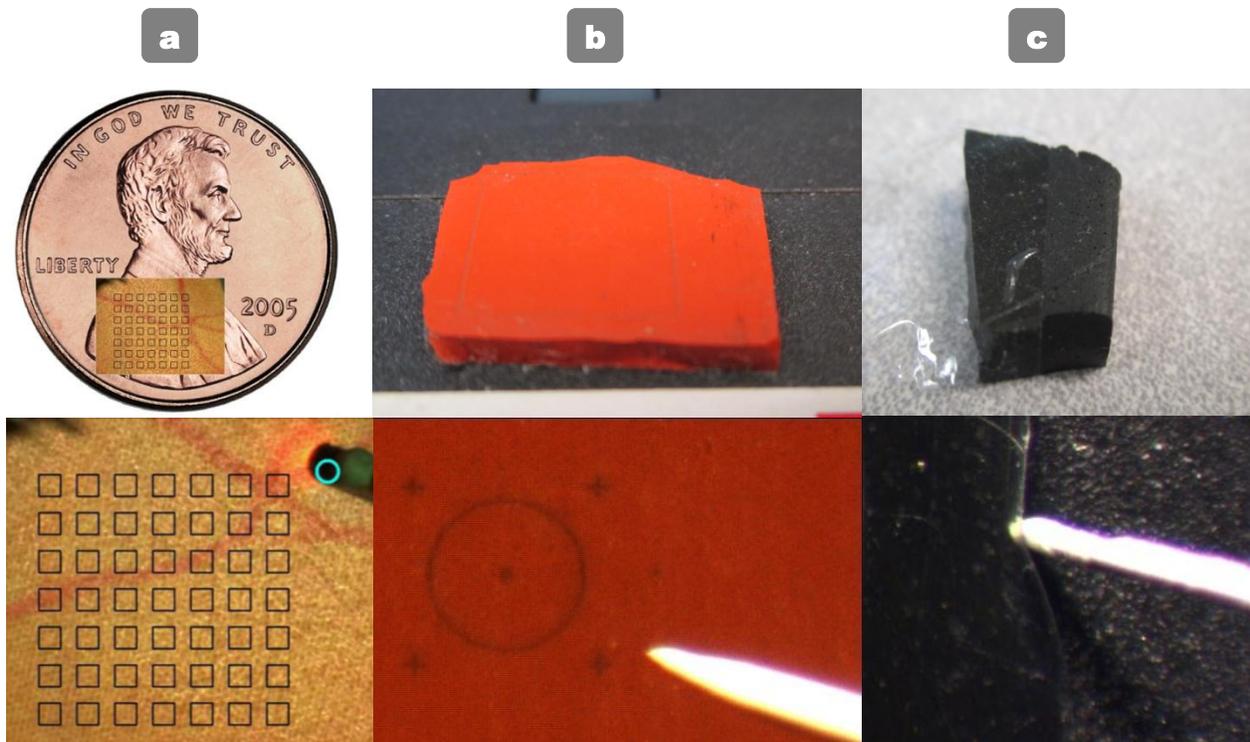


Fig. 13. Planar targets with macroscopic and microscopic views. (a) Printed paper slides for laser photocoagulation (b) Laser etched rubber slides for general micromanipulation tasks (c) Plastic wrap on sorbothane slides for membrane peeling tasks.

For planar environments where it is important for the tip of Micron to interact with the surface in a similar way to retina, sorbothane is a good choice, as it is a viscoelastic urethane polymer designed to have tissue-like deformation properties [90], [91]. In tests, we have found that a sorbothane slide with plastic wrap stretched across the top is a good model for retinal membrane peeling and exhibits the appropriate amount of force (5-15 mN) during peeling tasks with surgeons.

2.6.2 TABLE TENNIS BALL (IN VITRO)

Planar environments are not well suited for tasks requiring a shape that more closely mimics the eye shape. For these tasks where shape is important, we use the bottom half of a table tennis ball (Fig. 14). Approximately the same size as a human eye, a table tennis ball eye phantom is constructed by affixing textured paper to the bottom of the ball and taping a red-painted hair to mimic retinal vessels. This model, in particular, allows evaluation of dense stereo algorithms and vessel tracking algorithms in a very clean environment where ground truth is more easily assessed.

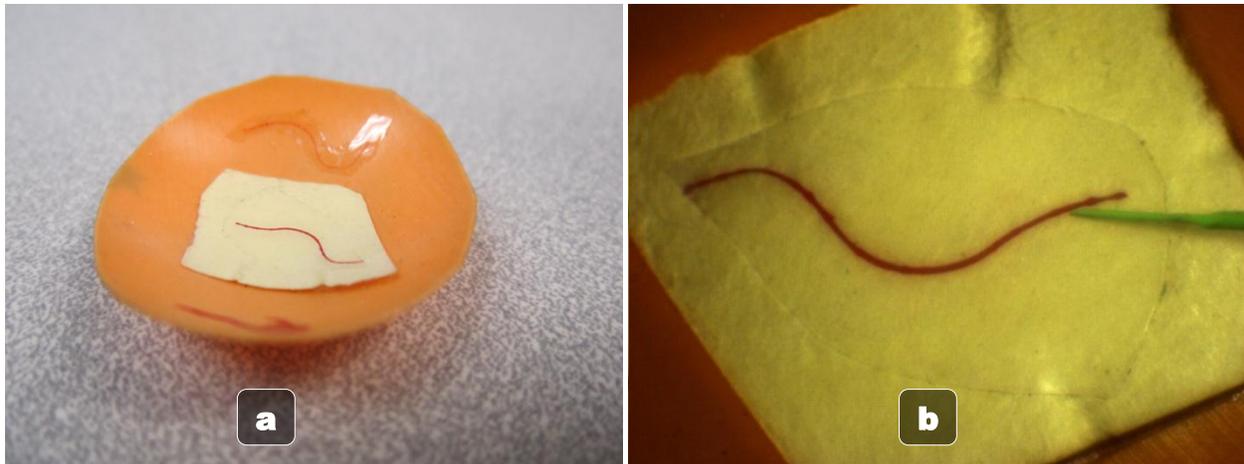


Fig. 14. Table tennis eye model from the bottom half of a ball with a hair painted red to mimic a retinal vessel. (a) Macroscopic view (b) Microscopic view with the green-painted tip of Micron.

2.6.3 JHU EYE PHANTOM (IN VITRO)

As shown in Fig. 15, the Johns Hopkins University eyeball phantom is a hollow, rubber sphere constructed to be the approximate size of a human eyeball. The inside of the phantom is lined with opaque white rubber and painted red and yellow to mimic the retinal structures such as the optic nerve and vasculature. Since painting very fine veins is difficult, we glued strands of red-painted human hair with a diameter of 80-120 μm to mimic very fine vessels found in the retina. The top of the phantom is open for direct line of sight with the microscope, and can be fitted with a vitrectomy lens to simulate lensing effects.

For illumination, we use a standard 0.5 mm (25 gauge) vitrectomy light pipe, which is inserted through the upper left side of the eye with a 0.5 mm trocar. The Micron

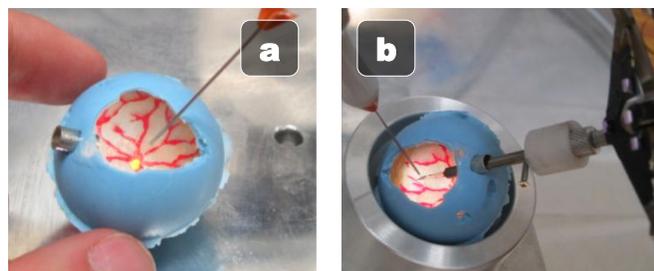


Fig. 15. Rubber JHU eyeball phantom. (a) Painted inside illuminated with vitrectomy fiber optic light-pipe (b) Micropipette attached to Micron inserted through the sclerotomy port on the right.

instrument is inserted through the upper right side of the eye with a larger 4.2 mm port to accommodate the attached micropipette and protective sheath. To simulate the eye rotation in the socket that typically occurs during vitreoretinal surgery, we place the eyeball phantom in a spherical metal cup lubricated with KY jelly. This construction allows the eye to rotate easily and freely during micromanipulation, as would be the case during a pars plana vitrectomy.

2.6.4 OPEN-SKY PORCINE RETINA (*EX VIVO*)

For more realistic experiments where actual vasculature or retina is needed to complete the procedure, such as laser photocoagulation or vessel cannulations, we find porcine retina is most convenient (Fig. 16). We order them from Hatfield Quality Meats over the phone, and whole eyes are overnighed in a bag on ice several days later. Because the lens of the pig clouds after death, we excise the retina and vasculature from the eye and place it on paper towel or felt slides for the “open-sky” operations. “Open-sky” procedures are done in open air on a flat surface without any of the sclerotomy or illumination constraints typically found in retinal surgery.

Because the vitreous adheres to the retina more firmly in pigs than in humans, the process of getting an intact portion of porcine retina out of the eye is tricky. The procedure that seems to work best is to make an incision just below the sclera, and cut a circumference across the eye, through the lens and through the optic nerve, splitting the eye into left and right pieces. The vitreous is kept in each half of the eye. The bottom of each half of the eye is cut away, around the optic nerve, using tweezers to keep the retina from slipping away from the pigment epithelium. Using high-grade surgical scissors, the gelatinous vitreous is excised and the retina piece with sclera still attached is trimmed to yield a 1-2 cm² piece of retina that contains retinal vessels.

The entire procedure is high-risk, and even after practice, it may take several eyes before sufficiently good quality retina pieces with intact vasculature are excised. The chunks of retina can be prepared before the surgery and moistened with an eye-dropper for several hours. Note that because the model is *ex vivo*, retinal vessels have collapsed and are thus harder to cannulate; in fact, during cannulation, the blood in the vessel is usually pushed away simply by pressing the micropipette against the vessel.



Fig. 16. (a) Preparation of pig eyes (b) excised porcine retina with retinal vessel lying flat in an open-sky environment.

2.6.5 INCUBATED CHICKEN EGGS (IN VIVO)

The chicken egg membrane has been suggested as a good model for retinal veins [92]. This procedure requires fertilized eggs (during the summer, they can generally be found fairly cheaply on Ebay) and an egg incubator. The goal is to let the chicken embryos grow for about 10-14 days and then use the vessels and membranes that grow just on the inside of the shell. After “candling” the egg to determine the top, the egg is carefully cracked open, and then the exposed membrane is carefully peeled to reveal the vasculature. The advantage to this model is that it is an *in vivo* model and vessels bleed catastrophically if cannulations are performed incorrectly. However, because the vessels are suspended in the egg fluid, they are significantly more elastic and stretchy than retinal vessels. This makes them a poor model as it not unusual to push several millimeters before cannulation is achieved, which can exceed Micron’s saturation limits and carry the micropipette and vein outside the camera’s workspace. Another downside is that eggs are difficult to obtain during the winter since the cold will likely kill the embryos, even if packages are carefully insulated.

2.6.6 PORCINE RETINA IN JOHNS HOPKINS EYE PHANTOM (EX VIVO)

While open-sky models are good to test cannulation without additional complications, they are not very representative of the constraints on motion, angular approach, remote center of motion, illumination, and other situations encountered in real-life vitreoretinal surgery. The most realistic model we have designed is based on the Johns Hopkins eye phantom model with porcine retina placed at the bottom. A vitreotomy lens is fit to the top and the entire eyeball filled with fluid. While saline is a natural choice for fluid, it causes the retina to float away. The retina can be pinned down to a rubber slide before being placed in in the eye phantom, but this is tedious and not always effective. We have found a much better solution is to fill the eye with oil, preferably clear oil such as silicone. If oil is used, the retina can be either dropped into the eye phantom and oil poured directly on top to fill up the eye or, alternatively, the eyeball can be filled with oil and the retina piece be lowered in with tweezers. Because the wall of the rubber eye phantom is so thick, it is much more rigid than the actual sclera of the eye. To more accurately mimic the flexibility experienced at the sclerotomy, we find that cutting a fairly large hole in the side of the eye phantom and gluing a thin piece of rubber glove over the hole provides a more stretchy and realistic sclerotomy entrance. See Fig. 17 for depictions of the eye phantom modifications, usage, and microscope view of the resulting retina. In the opinion of surgeons working with our project that this model is closest to a real surgical environment.

2.7 DISCUSSION

We have described the accepted pars plana vitrectomy procedure and analyzed four vitreoretinal surgeries, their treatment procedures, and how a handheld micromanipulator could be beneficial during the operation. While Micron is applicable to a wide range of micromanipulation tasks, we chose these surgeries because of the precision required to perform the operation, making them ideal candidates for robotic aid. Because of the prevalence of laser photocoagulation and the skill required by cannulation and membrane peeling procedures, we focus on these three surgical techniques for the experimental sections of this thesis. Furthermore, we have investigated, compared, and designed various eye models we can use to evaluate Micron while performing these procedures.

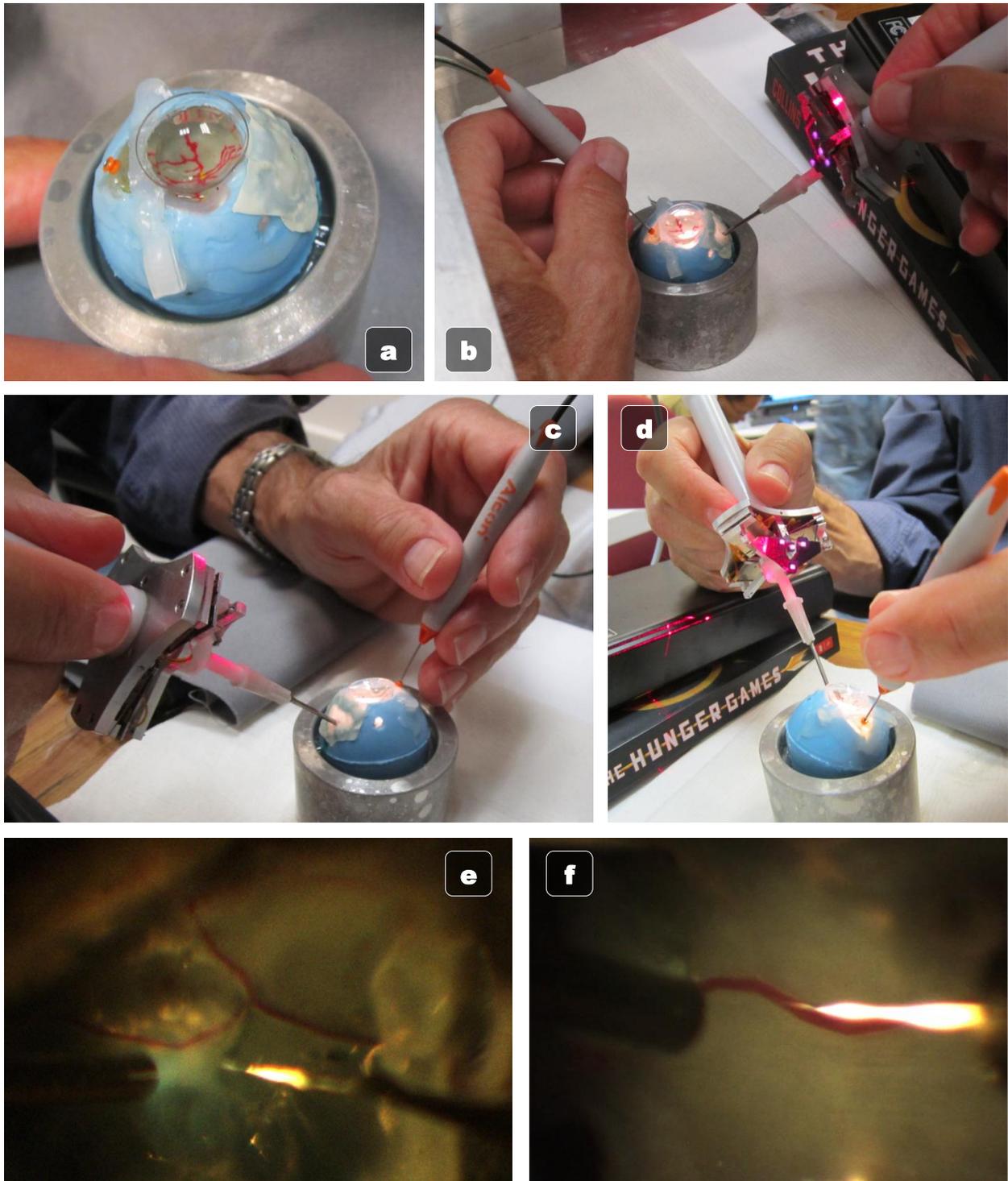
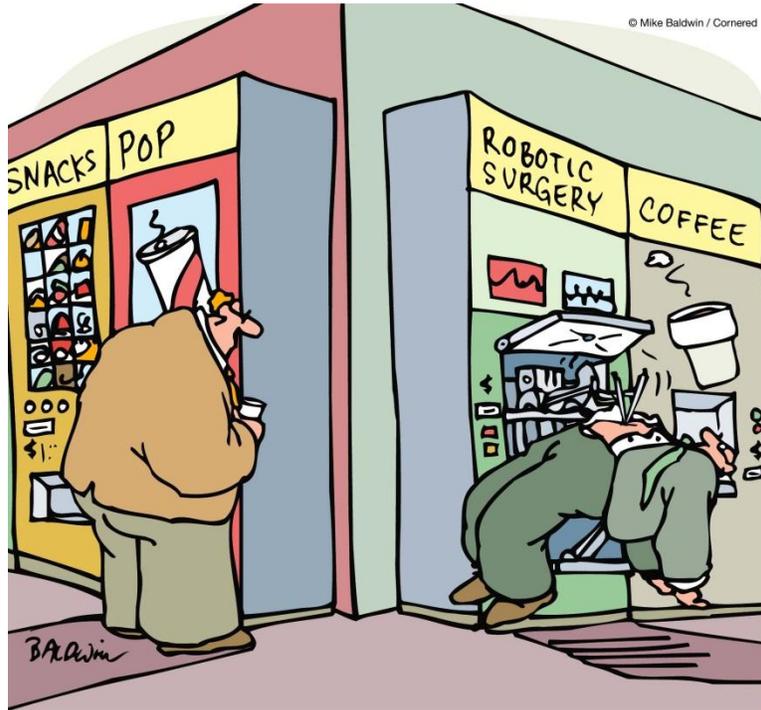


Fig. 17. Ex-vivo porcine retina in the modified Johns Hopkins University eye phantom. (a) Phantom filled with silicone oil and equipped with a standard 27 gauge light-pipe vitrectomy port on the left. On the right is a thin rubber film glued over a larger hole so the shaft of the instrument has more flex, similar to the sclera in surgery. (b) Standard two-hand grip of the light-pipe and Micron, as viewed from the front (c) Two-handed grip viewed from back right (d) Two-handed grip view from back left (e) Low magnification view of retina in eye phantom viewed through the vitrectomy lens and silicone with white-painted micropipette on right. (f) High magnification view of porcine retina in the eye phantom.

THEORETICAL FRAMEWORKS



Given the capabilities and limitations of a handheld micromanipulator such as Micron and the requirements of retinal surgical procedures, the question this thesis seeks to answer is: what is the best framework for designing fast, effective, and intelligent behaviors that are helpful, yet unobtrusive to a vitreoretinal surgeon? What type of control strategies, visual analysis, and system designs need to be developed? To answer these questions, we propose theoretical frameworks for integrating control systems with computer vision. Specifically, we develop a framework for visual control that unifies tremor compensation, motion scaling, and virtual fixtures. The proposed virtual fixtures framework uses position-based hand movements rather than input forces and is targeted for the unique characteristics of handheld micromanipulation. The framework is validated with a number of synthetic, but medically relevant tasks.

In addition to proposing a virtual fixtures framework for Micron, we also address and optimize the low-level control system and vision system. On the control system side, we improve performance by developing tremor models and using them to develop a feed-forward control path that anticipates and corrects for future error based on tremorous hand movement. On the vision side, we focus on engineering a real-time visual analysis system that performs tip tracking, adaptive registration of the cameras with ASAP, stereo vision, and 3D stereo vision. For analyzing retinal structures, we evaluate vessel detection algorithms and propose a new algorithm, EyeSLAM, for real-time simultaneous localization and mapping of the retina and vasculature based on fast vessel detection algorithms, occupancy grids, and iterative closest point (ICP) algorithms.

3 VIRTUAL FIXTURE FRAMEWORK FOR VISUAL CONTROL

In biology and microsurgery, proper manipulation of extremely small anatomical features often requires precision and dexterity that exceeds the capability of an unaided operator. Physiological tremor, or high-frequency involuntary hand movements with amplitudes of over 100 μm [11], is a large contributing factor to the difficulty of micromanipulation. Furthermore, lower frequency drift in gross hand positioning over time, caused by poor depth perception through the microscope and the relatively low bandwidth of the eye-hand feedback loop, reduces pointing accuracy [27]. Various robotic platforms such as the Robot Assisted MicroSurgery (RAMS) system [18], the Johns Hopkins SteadyHand [23], [31], and the Eye-Robot [19] have been proposed to deal with these problems in retinal micromanipulation. In general, robotic control aids for micromanipulation can be grouped into three categories: tremor compensation, motion scaling, and virtual fixtures.

Tremor compensation is a key component of many surgical robots [31], [32], the goal usually being to suppress frequency bands dominated by tremor in order to eliminate unwanted motion while preserving the operator's intended movements. Cooperative robots such as SteadyHand accomplish this mechanically by utilizing the inherent stiffness of the robot arm to damp high-frequency movement. In robots where the actuation is independent of the hand movement, tremor compensation can be achieved by inserting filters between the hand motion and the drive mechanism. In either case, reduction in the higher-frequency tremor while retaining the operator's gross lower-frequency movements is the end goal.

Motion scaling, often used in conjunction with tremor suppression, scales movements so the tip only moves a fraction of the hand movement. For example, a scale factor of $\frac{1}{2}$ would transform hand movement of 2 mm to a 1-mm movement at the robotic tip. This scaling behavior is a key feature in master/slave robots [93], [94], but is more difficult in cooperative robots, which are controlled by force at the hand/robot interface, rather than hand motion [23]. Recent developments in tremor compensation for handheld devices describe a dual-stage low-pass shelving filter [27], which is a hybrid tremor suppression filter that affords what can be viewed as relative motion scaling. An additional 30-50% reduction in tremor compared to a simple low-pass filter is reported.

Virtual fixtures, in contrast to tremor compensation or motion scaling behaviors that operate in the general case, instead aim

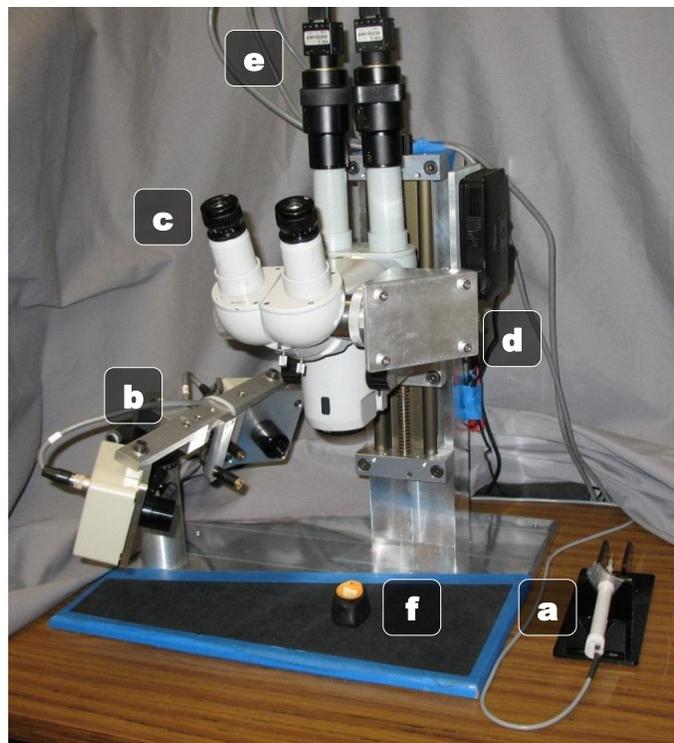


Fig. 18. System setup with (a) Micron, (b) ASAP position sensors, (c) surgical microscope, (d) image injection system, (e) stereo cameras, and (f) phantom half eyeball.

to improve specific motions or tasks [39]. Similar to a ruler that aids in the specific instance of drawing straight lines, virtual fixtures can be thought of as software guides that constrain the robot motion in specific ways. Proposed originally by Rosenberg [95] as a method to overlay abstract sensory information onto a force-reflecting master workspace, virtual fixtures were developed to address latency in teleoperation tasks. They were also found to reduce cognitive load and increase performance. Successful efforts have been reported in using virtual fixtures with the SteadyHand robot for surgically-relevant tasks [42], [55].

Our work is with active handheld micromanipulators, best exemplified by the Micron device depicted in Fig. 2. Micron incorporates motors between the handle and the tip, effectively creating an active handheld tool in which the end-effector can actuate independently of the hand motion within a limited range of motion (1x1x0.5 mm). With this capability, active tremor suppression [32], relative motion scaling [27], and simple virtual fixtures [46] have been individually investigated. We propose a more general position-based virtual fixture framework for handheld micromanipulators that integrates both tremor compensation and motion scaling. Fig. 18 shows the system setup.

3.1 BACKGROUND

Virtual fixtures and the concept of robotics constraints have gained popularity with the practical application of guiding the tip of robotic end-effectors during surgical manipulation tasks. Funda et al. [96] presented an optimization technique for commanding multi-linkage robotic arms via joystick, while keeping joints within their limits, maintaining orientation of the end-effector, and imposing task-space constraints on position. Kapoor et al. [41] extended Funda's work in constrained optimization to jointly handle simultaneous primitives such as maintaining a direction, rotating around a line, and staying above a plane, with a weighted optimization. Davies et al. [97] demonstrated the ACROBOT (Active Constraint Robot) with virtual fixtures to prevent the surgical cutter from entering pre-defined avoidance zones. Everett et al. [94] showed decreased operation time and operator fatigue in telemanipulating objects viewed from a camera by dynamically scaling the motion based on the relative position and orientation of the robotic end-effector.

In search of more general frameworks, Kumar et al. [98] decomposes surgical tasks into basic system states where the surgeon initiates primitive actions with a foot pedal to execute subroutines to position, orient, approach, insert, etc. Bettini et al. [39] introduce a comprehensive virtual fixtures framework that translates forces applied to a shared manipulator arm into end-effector velocities after applying motion constraints, demonstrating a number of useful behaviors such as trajectory following, positioning, and volumetric restrictions. Hard virtual fixtures disregard any forces in non-preferred directions, in order to prevent deviations from the fixtures, while soft virtual fixtures allow a mixture between human commands and the active virtual fixture. Such virtual fixtures and robotic pose constraints can be generated from cameras observing the surgical workspace, leading to research in using virtual fixtures to enable the end-effector to track surfaces [54], [99], follow trajectories in tight spaces [42], and aid in simulated surgical procedures [42], [56], [100].

A significant amount of research has been accomplished in applying virtual fixtures to various robotic platforms and micromanipulation problems [39], [41], [97]. In most formulations of virtual fixtures, the user manipulates a robot arm that is attached to the instrument directly or remotely via teleoperation. Forces on the robot arm are transformed to velocity commands at the instrument tip. These velocities are shaped first in software via

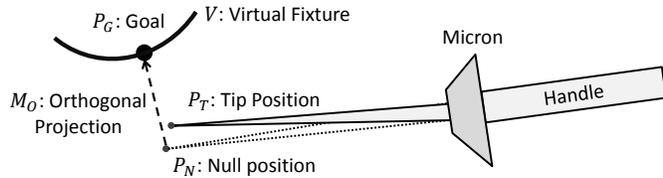


Fig. 19. Example of handheld micromanipulation with position-based virtual fixtures, which drives the tip position P_T to a goal position P_G on the virtual fixture V . The goal position is calculated by the orthogonal projection M_O of the null position P_N , which exactly reflects the hand motion. Note: figure not to scale.

constraints as specified by the active virtual fixture. If the robot is non-backdriveable (such as the Johns Hopkins SteadyHand), strict adherence to the virtual fixture can be enforced by ensuring that velocity components are zero in directions that move the tip away from the fixture. Alternatively, the approach of Stetten et al. [37] provides amplified force feedback to the operator.

Unlike most existing virtual-fixture-enabled robots, Micron is not manipulated by the operator through the application of forces to a robot arm or through joy-stick control. It is a fully handheld device that purely senses position and handle movements; thus, the input to the virtual fixtures must be hand motion. This fundamental difference, the use of hand position instead of force or joystick as the control input, necessitates the development of a unique formulation of virtual fixtures specifically designed for this class of handheld micromanipulators.

3.1.1 CONTROL SYSTEM AND NOTATION

As depicted in Fig. 19, the tip of Micron in 3D space is defined as $P_T \in \mathbb{R}^3$. As Micron actuates, the tip moves independently of the handle, so it is important to define a measure of how much the tip has moved. We define the null position as the 3D tip position $P_N \in \mathbb{R}^3$ under the assumption that Micron is off; i.e., P_N exactly reflects the hand motion. One can think of P_N as being mechanically tied to the handle; thus $P_T = P_N$ unless Micron actuates the tip. Fig. 3 and Fig. 19 graphically depict the difference between the null position P_N and the actual tip position P_T . With high-bandwidth, low-latency tracking via ASAP, Micron performs position-based control using closed-loop feedback with approximated linear inverse kinematics to bring the tip position P_T coincident with the goal P_G . The ASAP controller $C(s)$ in Fig. 20 uses proportional, integral, and derivative (PID) feedback.

3.1.2 FEEDFORWARD TREMOR COMPENSATION

When a 3D goal P_G is specified in handheld operation, a typical tremor disturbance in the hand motion P_N is rejected 90% by a 2nd-order low-pass Butterworth filter $F(s)$ with a corner frequency of 2 Hz. However, the remaining 10% of tremor represents 10-20 μm deviations of the tip position from the goal, making it the largest contributing factor in the error reported for hard virtual fixtures [101]. Analysis revealed that a 3-ms latency in the Micron manipulator $G(s)$ is largely responsible for tip positioning error [102]. We denote T as the latency between when the ASAP controller $C(s)$ sends a command to the Micron manipulator $G(s)$ and when the effect is seen in the tip position. During this time, tremor at the handle is moving the null position with a velocity \dot{P}_N ,

and to a first-order approximation, the latency in actuation introduces error P_Δ at the tip position that accumulates over time T :

$$P_\Delta = \int_0^T \dot{P}_N(t) dt \quad (1)$$

To address this latency, a feedforward control path was designed to anticipate, estimate, and compensate for the error P_Δ before it is seen by the feedback loop. As shown in Fig. 20, a constant-acceleration Kalman filter $K(s)$ is run at 2 kHz to optimally estimate the position, velocity, and acceleration states of the tip and handle. The velocity of the handle \dot{P}_N , which represents tremor velocity, is used by the tremor prediction module $P(s)$ to estimate tip error P_Δ caused by the delay T in actuation $G(s)$. Although $P(s)$ could predict $\dot{P}_N(t)$ using a nonlinear algorithm to such as the double adaptive bandlimited multiple Fourier linear combiner [103], we found a simple prediction of current velocity at time t works well, e.g., $\dot{P}_N(t+k) = \dot{P}_N(t) \forall k \in [0, T]$. Once $P(s)$ integrates the velocity over latency period T , the predicted error P_Δ is used as a feedforward term to adjust the goal position P_G to form the latency-compensated goal position P_G^* . Combined with the existing feedback control strategy, an additional 50% reduction in error is seen, for a total tremor rejection

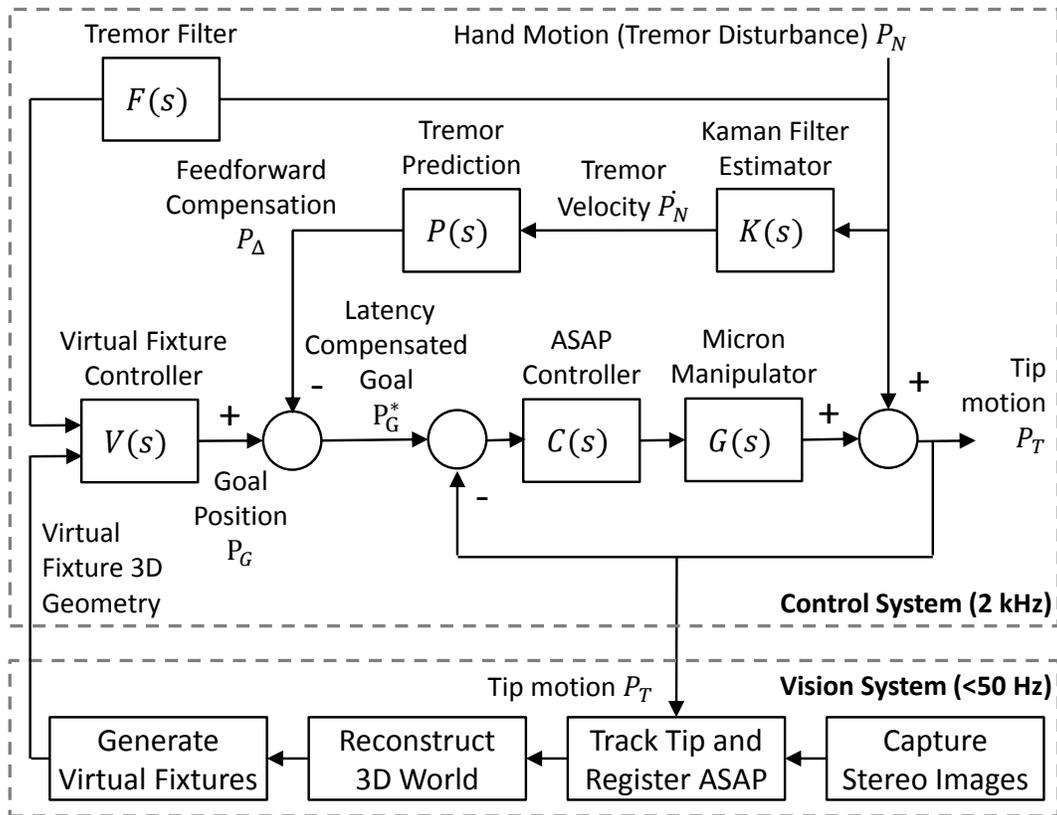


Fig. 20. Block diagram of the control and vision system. The vision system registers the cameras to ASAP and generates virtual fixture geometry from the anatomy. The control system includes feedback and feedforward paths for more precise control of the instrument tip. The virtual fixture controller uses the filtered hand motion to enforce the subspace constraints of the fixtures.

rate of 95%. See Chapter 4 for more details on tremor modeling and feedforward control.

3.1.3 VISION SYSTEM

Vision-based control of the manipulator is necessary because the anatomy of interest is localized via intraoperative imagery. Stereo cameras attached to the microscope view the same workspace as the operator and capture video at up to 50 Hz with resolution up to 1024x768. The cameras track both the instrument tip and the anatomical targets. However, traditional image-based visual servoing techniques [104] operating at the low camera rates cannot provide the high bandwidth and low latencies needed for active control of handheld micromanipulators. Instead, the control loop must use the much faster 2 kHz ASAP positioning system. Thus, virtual fixtures are generated using stereo images, transferred to the ASAP coordinate frame, and the resulting geometry is passed to the virtual fixture controller $V(s)$ that runs at 2 kHz in the Micron control system. The registration between the ASAP measurement system and the stereo cameras is maintained using a recursive least squares calibration procedure described in Section 5.2.

3.1.4 SYSTEM SETUP

The operator uses Micron under a surgical microscope (Zeiss OPMI 1). A 27 gauge hypodermic needle (418 μm outer diameter) is attached to Micron. A 127 μm nitinol wire is inserted into the needle to serve as a very fine instrument tip. The cameras track both the tip of the instrument and anatomical targets. The entire setup can be seen in Fig. 18.

3.2 VIRTUAL FIXTURE FRAMEWORK

Represented as a subspace defined in 3D Euclidean space by the stereo vision cameras, the virtual fixture must constrain the tip position of Micron, using only the hand motion of the operator to guide and position the instrument tip on the virtual fixture. In the case of hard virtual fixtures, the tip of Micron should always lie on the subspace representing the fixture; in the case of soft virtual fixtures, the error between the hand motion and the virtual fixture should be scaled. In both cases, tremor compensation should smooth the tip movement.

3.2.1 POINT VIRTUAL FIXTURE

We begin the formulation of virtual fixtures for Micron by considering the simplest fixture possible: fixing Micron's tip to a single goal point $P_G \in \mathbb{R}^3$ in space. While the point virtual fixture is active, the control system should enforce the relationship $P_T = P_G$, regardless of where or how the operator moves the handle (within Micron's range of motion). We define the simple control law:

$$P_T = P_G \quad (2)$$

We use approximated linear inverse kinematics with a PID feedback + feedforward controller and a notch filter at the manipulator resonance to drive the tip to the goal position. Controller latency is 3 ms and settling time is at most 50 ms if the goal point is within Micron's current range of motion. Once the point virtual fixture is activated, Micron uses a 1 kHz controller to drive and maintain the tip on the goal. In response to shifting anatomy, moving the virtual fixture is possible by setting a new goal point P_G . To avoid high-frequency oscillations, large changes in P_G should be smoothed with either filtration (lowpass, Kalman, etc.) or trajectory planning.

3.2.2 HIGHER-ORDER SUBSPACE VIRTUAL FIXTURES

Higher-order subspaces can be built on top of the point virtual fixture to obtain more interesting behaviors. Each additional level adds a degree of freedom to the tip motion. For instance, the point virtual fixture restricts all motion. However, a line virtual fixture allows the tip to freely travel along a line while restricting motion orthogonal to the line. Likewise, we can define the hierarchy of virtual fixtures as seen Fig. 21: point, curve, surface, and volume. To reconnect the need for these virtual fixtures back to medical applicability, it is worth considering possible a medical relevance for each virtual fixture:

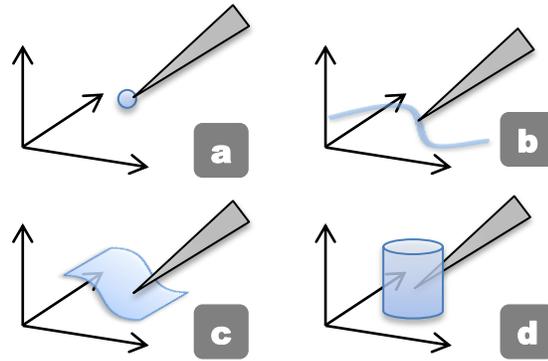


Fig. 21. Virtual fixtures constraining the tip to a subspace with increasing degrees of freedom: (a) point, (b) curve, (c) surface, and (d) volume.

- **Point** (0 DOF): Steadying cannula during injection
- **Curve** (1 DOF): Following path for laser ablation, guiding suture/needle along a blood vessel
- **Surface** (2 DOF): Maintaining a constant standoff distance, navigating in narrow crevices
- **Volume** (3 DOF): Restricting the tip volumetrically to prevent tissue contact outside of “safe” areas

Although not covered in this thesis, such virtual fixture formulations extend to higher degrees of freedom to include orientation virtual fixtures. Each virtual fixture in the hierarchy can degenerate to a special case of a lower DOF fixture. For instance, a line virtual fixture can be implemented with a point virtual fixture that moves along the line.

In fact, we assert that all virtual fixtures can be implemented easily with just the point virtual fixture. The key to implementing higher-order fixtures with a point fixture is selecting the correct point on the higher-order virtual fixture. Algebraically, any point on a fixture will satisfy the constraint imposed by the fixture. Geometrically, it is most intuitive to select a point on the virtual fixture as close as possible to the operator’s intended position. Thus for an arbitrary virtual fixture V , the nearest goal position P_G is the orthogonal projection of the null position P_N onto the virtual fixture V . We define this orthogonal projection by the mapping $\mathbf{M}_O: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ that selects the goal point P_G as the closest point from the null position P_N to the virtual fixture V :

$$P_G = \mathbf{M}_O(V, P_N). \quad (3)$$

For simple geometric structures (lines, planes, circles, cylinders, etc.), analytic solutions for \mathbf{M}_O exist. For more complex shapes, numerical solutions or iterative approximations to \mathbf{M}_O are available [105].

3.2.3 GENERALIZABLE SPLINE VIRTUAL FIXTURES

When applying the position-based virtual fixtures framework to surgical environments, simple geometric shapes will not suffice for representation. A more general fixture representation is

necessary, due to the complex nature of anatomical structures. In many fields, such as graphics, meshes are used to approximate arbitrarily complex objects. While such an approach has merits and in fact explicitly encodes surface normal information into the mesh, very fine meshes are needed to represent anatomy which is non-planar. Furthermore, the surface normals necessary for the orthogonal projection \mathbf{M}_O are discontinuous at the vertices of the mesh, adding an extra level of complexity to the mapping. Because of the unwieldy nature of meshes and their large storage requirement for fine approximations to curved anatomy, meshes were not considered for the virtual fixture framework.

A more suitable representation is one of piecewise curves, or splines, which have the beneficial property of representing continuous curves and surfaces with low dimensionality. Additionally, the orthogonal mapping \mathbf{M}_O can be calculated numerically [106]. Fitting splines can be done directly in image space [107], or after a point-cloud representation of the anatomy of interest has been extracted. Because virtual fixtures must be available in the 3D ASAP coordinate system, the latter approach is taken. Although a variety of spline representations may be used, we use basis-splines (B-splines) to describe general anatomy because of their versatile yet simple representation, with widespread support for fitting and evaluation. Surfaces such as the retina can be represented by a B-spline surface. Encoding a notion of “side” to the surface allows for volumetric fixtures, such as forbidding the instrument to pass through the B-spline surface virtual fixture representing the retina.

3.2.4 TREMOR SUPPRESSION

In point virtual fixtures, all degrees of freedom are proscribed by the fixture, eliminating all unwanted tremor. However, for higher-order virtual fixtures, tremor parallel to the subspace is not affected by the orthogonal projection \mathbf{M}_O . In fact, tremor is evident in all 3 DOF for volume virtual fixtures when the tip is fully inside the volume. To reduce tremor in these higher DOF fixtures, one may use a tremor suppression filter $\mathbf{F}_T^n: \mathbb{R}^n \rightarrow \mathbb{R}^n$ $n \in \{1,2,3\}$. For instance, we can insert the tremor suppression filter between the 3D null position P_N and the orthogonal projection \mathbf{M}_O :

$$P_G = \mathbf{M}_O(V, \mathbf{F}_T^3(P_N)) \quad (4)$$

Alternatively, the tip position can be re-parameterized on the subspace representing the virtual fixture with a lower degree of freedom. In this case, we redefine the mapping

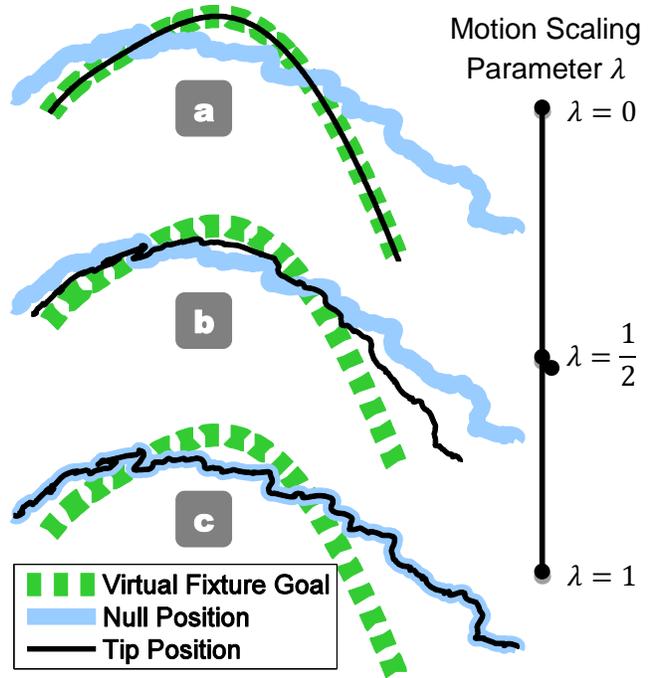


Fig. 22. Effect of motion scaling parameter λ on the tip position. (a) Hard virtual fixture ($\lambda = 0$). (b) Soft virtual fixture with $\lambda = 1/2$; control is shared 50% between virtual fixture and hand motion. (c) Fixtures off ($\lambda = 1$).

$M_O^n: \mathbb{R}^3 \rightarrow \mathbb{R}^n, n \in \{1,2\}$ as transforming P_N in 3D space to a n D re-parameterized goal position on the virtual fixture via orthogonal projection. For instance, P_G on a curve can be re-parameterized by arc length and filtered with a one-dimensional tremor suppression filter F_T^1 . For an n DOF virtual fixture, we can calculate the goal point as:

$$P_G = F_T^n(M_O^n(V, P_N)) \quad (5)$$

For experiments in this paper, we use (5) for circle tracing, to prevent any nonlinearity in the orthogonal projection from affecting the tremor suppression, and (4) for all other fixtures. It is also important to note that noise in virtual fixture placement can inject high-frequency movement similar to tremor. Therefore, if virtual fixtures are being generated by video, high-frequency components should be eliminated. We accomplish this with simple low-pass filtration, since the anatomy does not move rapidly.

3.2.5 MOTION SCALING

So far we have described hard fixtures where the tip position cannot deviate from the constraint imposed by the virtual fixture. However, it is also possible to derive soft virtual fixtures that share control between the virtual fixture and the operator. If hard virtual fixtures are analogous to a stiff metal ruler, then soft virtual fixtures can be likened to a compliant foam ruler that aids in drawing lines but can be partially overridden by the operator. An additional parameter $\lambda \in [0,1]$ defines how much the operator can override the virtual fixture. In our formulation, λ represents the proportion of the hand motion P_N vs. the goal point P_G that Micron uses to actuate the tip:

$$P_T = (1 - \lambda)P_G + \lambda P_N \quad (6)$$

In essence, λ functions as a weighted average of the goal and null position. $\lambda = 0$ corresponds to a hard virtual fixture where the operator can only control where on the virtual fixture the tip goes, and $\lambda = 1$ disables virtual fixtures entirely and leaves the operator in complete control (Fig. 5).

However, for values of λ between 0 and 1, (6) can be directly manipulated into a motion scaling paradigm:

$$P_T = P_G + \lambda(P_N - P_G) \quad (7)$$

$$e = P_N - P_G \quad (8)$$

$$P_T = P_G + \lambda e \quad (9)$$

By introducing the error e as the difference between the null position P_N controlled by the operator and the goal point P_G calculated by the virtual fixture, it is clear the parameter λ directly represents the motion scaling factor. For example, if $\lambda = 1/2$, then all hand motions that deviate from the virtual fixture will be scaled by one half. Thus, we can see that sharing control between the virtual fixture and the operator yields motion-scaling behavior (see Fig. 22b for an example).

3.2.6 GENERALIZED CONTROL LAW

To summarize the control law that incorporates virtual fixtures, tremor suppression, and motion scaling into a handheld micromanipulator framework based on position control, we assume for the sake of simplicity that tremor suppression on the null position is sufficient and re-parameterization is not necessary. Assuming all operations are in 3D, the generalized control law then becomes:

$$P_G = \mathbf{M}_O(V, \mathbf{F}_T(P_N)) \quad (10)$$

$$e = \mathbf{F}_T(P_N) - P_G \quad (11)$$

$$P_T = P_G + \lambda e \quad (12)$$

First, (10) selects the goal point on the virtual fixture closest to the tremor suppressed null position. Eq. (11) then calculates the error between what the virtual fixture and where the operator is currently pointing. Finally, (12) drives the tip to either the virtual fixture or, if λ is non-zero, scales the error to achieve motion scaling about the fixture.

3.2.7 COMBINING VIRTUAL FIXTURES

Virtual fixtures can be easily combined into compound virtual fixtures by dynamically selecting which fixture is active. For example, we demonstrate a box virtual fixture built from four plane segments. One difficulty encountered when combining virtual fixtures occurs when the mapping \mathbf{M}_O becomes degenerate. For instance, when two lines or planes form an angle, the null position P_N might be positioned equidistant between both virtual fixtures. Noise in sensor measurement might subsequently introduce tip position oscillations as multiple virtual fixtures compete for control of the tip. One solution to prevent the tip from oscillating between virtual fixtures is to disable all fixtures except the current one until the tip nears the edge, so that only when the tip reaches the boundary between fixtures can Micron transition from one virtual fixture to another. This approach was used for our multi-fixture experiments.

3.2.8 VISUAL CUES TO MAINTAIN EYE-HAND COORDINATION

A final consideration in our formulation of position-based virtual fixtures is one of practical implementation. The input to virtual fixtures is the null position, which is measured by the ASAP sensors but is otherwise unseen to the operator. Thus, only the tip position can be used by the operator to guide the tool. Hard virtual fixtures fix the tip to the virtual fixture regardless of any hand movement, removing all error and effectively disrupting the eye-hand coordination feedback loop.

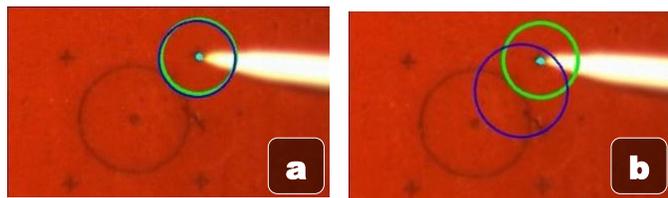


Fig. 23. Micron enforcing a point hard virtual fixture with visual cues: green represents the goal, blue represents the null position. (a) Nearly coincident circles are desirable, indicating Micron is near the center of the range of motion. (b) The position and size of the blue circle indicates drift. Here, the operator has drifted left, down, and upwards in the X, Y, and Z directions, respectively. If Micron was turned off, the tip would snap to the blue circle.

With the eye-hand feedback loop broken, the operator will unknowingly drift away from the virtual fixture until the virtual fixture is no longer within Micron's range of motion, at which point active

control becomes impossible. Thus it is imperative to provide the operator with a surrogate sense of error to restore eye-hand coordination.

In order to prevent this unbounded drifting behavior, we display visual cues that indicate the 3D location of the unseen null position. As shown in Fig. 23, we choose visual cues in the form of two circles: a green one to show the goal location and a blue one to show the null position, which reflects the actual hand movements. The distance between the circle centers represents the error Micron is currently eliminating. Z error is displayed by varying the radius of the null position circle, e.g. a growing radius represents upwards drift. The operator is instructed to keep the two circles roughly coincident to prevent drifting and saturation of the motors. In all tests, the operator used these visual cues to maintain positioning (the learning curve was found to be minimal).

3.3 EVALUATION

Several experiments were performed with Micron to validate the proposed virtual fixture framework. Two scenarios were investigated: simple, synthetic tasks on a rubber slide (similar to [27]), and a more complex vein-tracing task in an eyeball phantom. As shown in Fig. 24, the rubber slide had a laser-etched target consisting of four crosses arranged as the corners of a 600 μm square, with a circle of 500 μm diameter in the middle. To simulate a more realistic task in a more relevant surgical environment, an eye phantom with a curved vein was constructed on the bottom half of a 42-mm diameter sphere (see Fig. 25). The curved “vein” was a red painted hair taped on a yellow paper background that was firmly glued to the inside of a table-tennis ball. Because of the differences in scale, 29X magnification was used for the rubber slide experiments, and 11X magnification was used for the eyeball phantom.

3.3.1 EXPERIMENTAL PROCEDURE

Four sets of experiments were run for each task, the first three being synthetic tasks on the rubber slide similar to [27] and the last mimicking a more realistic surgical task of cannulating a blood vessel:

- **Hold Still:** Hold the tip of the instrument 500 μm above a fiducial on a plane for 30 s. A point fixture located above the top right cross was used.
- **Circle Tracing:** Trace a circle with a 500 μm offset from rubber surface two times. A 3D circle fixture derived from the tracked rubber target was used.
- **Move/Hold:** Move between the four corners of a square sequentially, pausing at each corner. Four plane segments oriented vertically and connecting each of the four corners were used, forming a compound box fixture. Because of the vertical orientation of the planes, the Z-movement is unconstrained, allowing the operator to choose the vertical

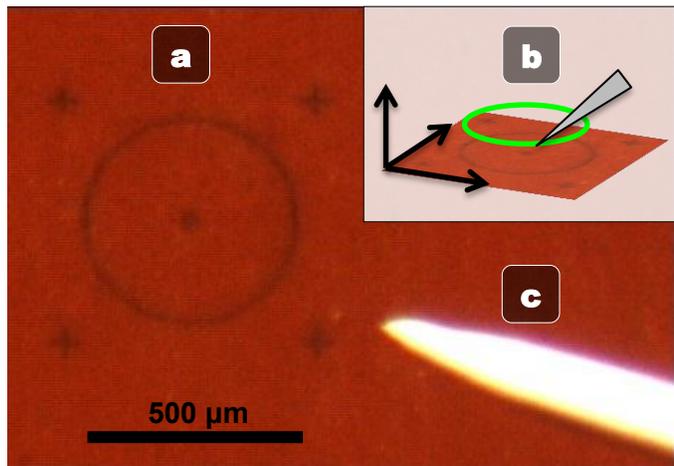


Fig. 24. (a) Laser etched target onto rubber surface (b) Generating a 3D circle virtual fixture from the tracked target (c) White-painted, tapered tip of Micron.

distance from the surface.

- **Vein Tracing:** Follow a 10 mm segment of a curved vein on the inside of a sphere, maintaining a 500 μm vertical distance from the spherical surface. The vein was represented with a B-spline in 3D space.

Experiments were performed in random order to alleviate ordering effects. Visual cues were displayed on a 3D monitor for the synthetic tasks and in the microscope via an image injection system for the vein tracing task. Four different scenarios were tested:

- **Unaided:** Micron was turned off.
- **Aided with Shelving Filter:** Micron was turned on with the tremor suppressing shelving filter from [27].
- **Aided with Soft Virtual Fixtures:** Micron was turned on with virtual fixtures and with the motion scaling factor $\lambda = 1/5$, so errors were reduced by 5X.
- **Aided with Hard Virtual Fixtures:** Micron was turned on with virtual fixtures but no motion scaling ($\lambda = 0$).

3.3.2 GENERATING FIXTURES FROM VISION

The discussion of virtual fixtures has assumed the availability of each fixture. In reality, virtual fixtures are generated from anatomical structures perceived by the

computer vision system. In addition to generating fixtures, the vision system is also responsible for maintaining the adaptive registration by tracking the tip position. For ease of tracking, the Micron nitinol tip is painted with acrylic paint. Simple image processing techniques including thresholding and blob tracking allow the vision system to accurately locate the instrument tip. The OpenCV library [108] is used to generate virtual fixtures from visual information. While the tracking and virtual fixture formulation presented here could handle moving fixtures, this chapter does not

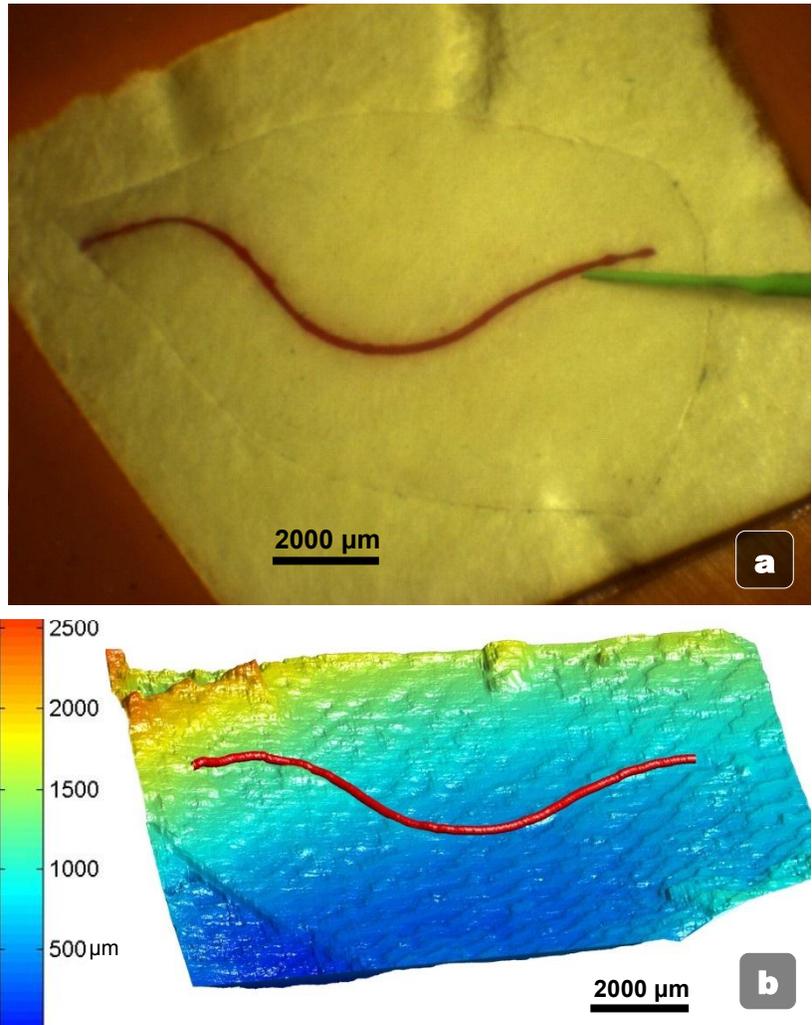


Fig. 25. (a) Half eyeball phantom with a dyed hair tapped to yellow paper affixed to the inside of a ping pong ball with green-painted tip of Micron visible. (b) 3D reconstruction and 3D spline representation of the vein, with dark blue representing low Z positions and dark orange representing high Z positions. Although the vein lies on the surface, it is visualized with red.

explicitly explore this aspect. Section 5.5 discusses how to explicitly deal with moving retinal structures in the virtual fixtures framework.

Rubber Slide Fixtures

Stereo cameras attached to the microscope operate at 54 Hz with a custom resolution of 504x324. Hierarchical template matching is used to locate the target. The template matching method is enhanced to handle occlusions of the target caused by the tip of the instrument and its shadow, yielding very fast and robust tracking. Once the target is located in both images, the center of the target can be backprojected into 3D using the camera registration. Individual virtual fixtures (point, circle, box) are derived from the 3D location of the target and the relative distances of the components. All tracking and virtual fixture generation happens at 50 Hz.

Eye-Phantom Vein Fixtures

Although the same tip tracker is used to localize the instrument and maintain adaptive registration, the vision system for the eyeball phantom was expanded to handle the more complex geometry of the experiment. Higher-resolution 1024x768

video is analyzed at 30 Hz. Lifting the planar assumption and moving into an environment with curved surfaces necessitates a good 3D understanding of the environment. We use stereo vision to densely reconstruct the surface of the eyeball phantom retina using the fast Semi-Global Block Matching (SGBM) algorithm [109]. Although a quadratic or spherical fit could be attempted, we leave the disparity map in a raw pixel-wise representation. The vein is localized with color trackers, filtered with blob analysis, and skeletonized with morphological and distance transforms. The vein is then fit to a B-spline in the XY image space and XZ direction (based on the disparity map). Corresponding points on the spline in the left and right images are backprojected to form a full 3D representation of the vein on the surface (see Fig. 25).

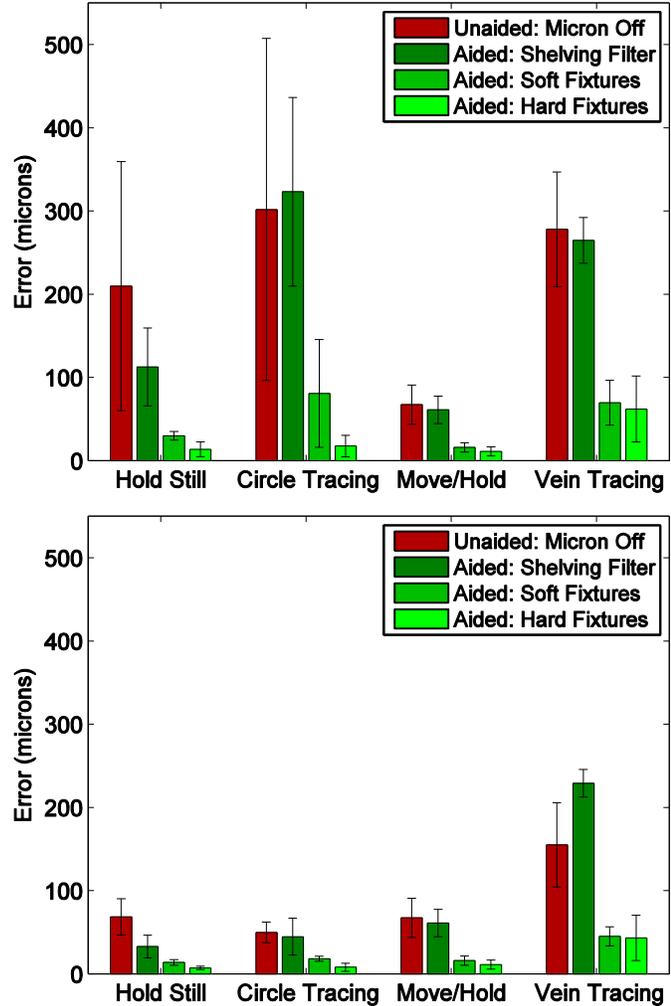


Fig. 26. Mean RMS error across seven trials of each combination of task and scenario, with standard deviation error bars. Hard fixtures significantly reduce error compared to unaided and shelving filter scenarios ($p < 0.05$). (top) 3D mean RMS error (bottom) 2D mean RMS error

The anatomy of greatest interest is often directly underneath of the tip of the instrument and occluded from view. Because the tip occludes both the vein when fitting the spline and the surface when calculating the disparity map, the initial unoccluded spline representation is stored. When occlusion of the vein is detected, Iterative Closest Point (ICP) [110] is used to register the occluded vein with the initial vein and fill in occluded parts. The GNU Scientific Library (GSL) library [111] is used to then fit the splines. Occlusion of the disparity map is handled by linearly interpolating over occluded patches. Full 1024x768 stereo depth maps and 3D spline representations are updated at 2 Hz and provided to the controller which runs at 2 kHz.

3.3.3 RESULTS

Each set of synthetic experiments on the rubber slide was performed by three surgeons at least once, for a total of seven sets. The vein tracing experiment in the eyeball phantom was performed by a single vitreoretinal surgeon for a total of seven sets. All error was measured as the Euclidean distance between the tip position sensed by the ASAP optical trackers and the closest point on the virtual fixture as generated by the registered stereo cameras. Error was recorded at 2 kHz.

Fig. 26 displays 2D and 3D mean RMS error with standard deviation across all three trials for each combination of task and scenario. It is important to note that the Move/Hold scenario has the same 3D and 2D error because the vertical plane virtual fixtures do not restrict Z-movement. Additionally, the Vein Tracing task has a

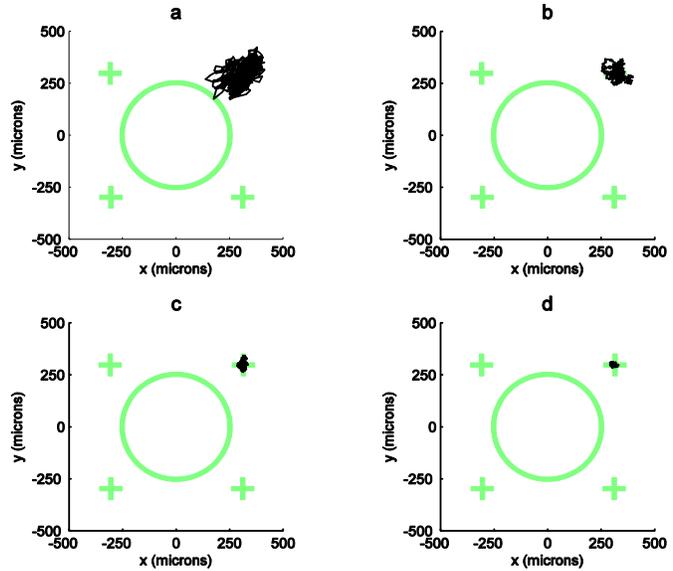


Fig. 27. Hold Still results (a) unaided (b) aided with shelving filter (c) aided with soft fixtures (d) aided with hard fixtures.

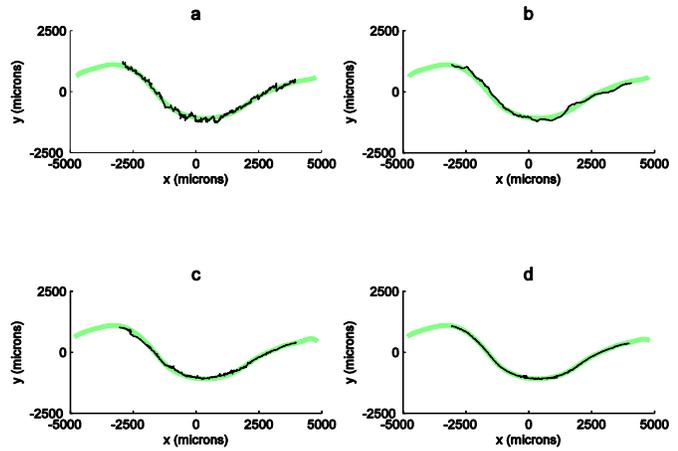


Fig. 28. Vein Tracing results (a) unaided (b) aided with shelving filter (c) aided with soft fixtures (d) aided with hard fixtures.

Task	Unaided Micron Off (μm)	Aided		
		Shelving Filter (μm)	Soft Fixtures (μm)	Hard Fixtures (μm)
Hold Still	480±214	261±99	99±40	87±41
Circle Tracing	710±500	711±262	302±260	105±72
Move/Hold	326±501	305±167	132±96	126±43
Vein Tracing	701±177	492±121	415±226	383±177

Table 3: Max error in positioning tasks. Mean of the 3D max error across all seven trials of each combination of task and scenario. Significantly reduced results compared to the hard fixtures case are bolded ($p < 0.05$). Because vein tracing experiments used lower magnification and larger motions were necessary (10 mm), saturation was encountered more frequently than in other tasks.

higher incident of saturation and thus more error, which is partially because of the lower magnification and the greater tip traversal required (almost 10 mm, compared to the 0.5-mm circle). In fact, 98% of the error during the Vein Tracing task with hard virtual fixtures occurs while the manipulator is in saturation. Maximum 3D error for each set of trials is presented in Table 3.

In all cases, mean and maximum error for hard virtual fixtures is significantly less than in the unaided case for both 2D and 3D measurements ($p < 0.05$) as calculated by a two-tailed t-test. Furthermore, mean RMSE error of hard virtual fixtures for both 2D and 3D is significantly less than the state-of-the-art shelving filter from [27] ($p < 0.05$). Fig. 27, Fig. 29, Fig. 30, and Fig. 28 show one trial of each task and scenario with a trace of the 2D tip position overlaid in black on the target represented by thick light green lines.

3.4 DISCUSSION

In this chapter, we have presented a derivation of virtual fixtures that depends not on forces applied to a robot arm or remote controlled joystick, but on motion of the handle of a handheld instrument. This new position-based virtual fixture formulation is necessary for the class of handheld micromanipulators such as Micron that use handle positions as input to the control system. Virtual fixtures are generated in real time from stereo cameras attached to the microscope, providing task-dependent behaviors to the operator. Visual cues are displayed to the operator to maintain eye-hand coordination and help guide the operator during the procedure.

Hard virtual fixtures totally constrain the tip, enforcing constraints that should never be violated, e.g. forbidden areas or snap-to behaviors. Implementation of soft virtual fixtures under our formulation directly results in intuitive motion scaling. In all cases, tremor suppression filters

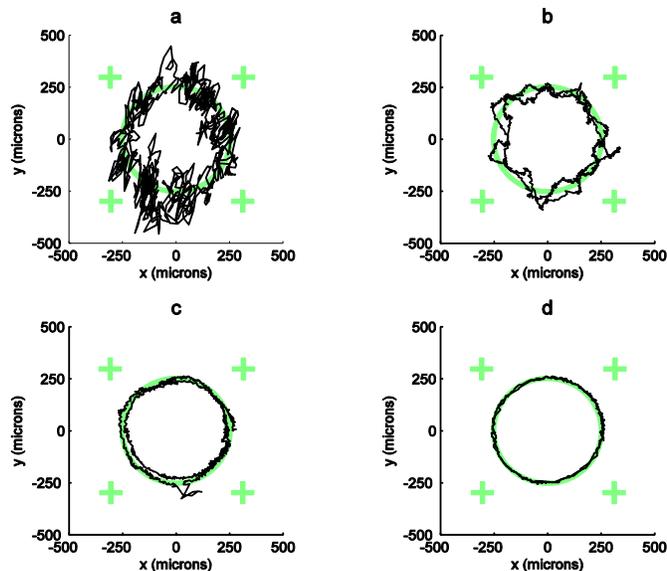


Fig. 29. Circle Tracing results (a) unaided (b) aided with shelving filter (c) aided with soft fixtures (d) aided with hard fixtures.

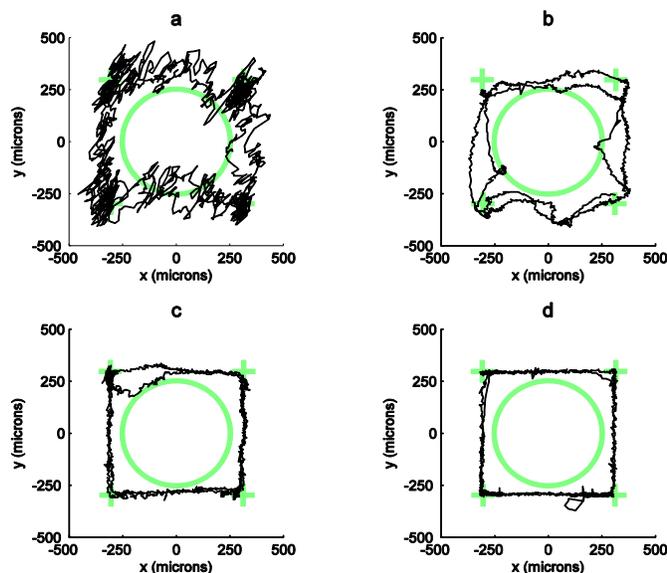


Fig. 30. Move and Hold results (a) unaided (b) aided with shelving filter (c) aided with soft fixtures (d) aided with hard fixtures. Deviations from the virtual fixtures in (c) and (d) result from saturation of the actuators, caused by tremor or drift in excess of the range of motion of Micron.

compensate for unwanted high frequency motion. Furthermore, the virtual fixture framework easily adapts to general parameterizations such as splines to model complex anatomy. Using Micron as a test platform, virtual fixtures have been validated with medically relevant artificial tests such as vein tracing. Compared to unaided performance and state-of-the-art tremor compensation [27], virtual fixtures significantly reduce tip positioning error ($p < 0.05$).

4 MODELING TREMOR FOR FEEDFORWARD CONTROL

A handheld micromanipulator such as Micron has a number of advantages. First, it is small and lightweight, making it easy-to-use and inexpensive. Second, handheld instruments are intimately familiar to surgeons, so Micron can leverage surgeons' experience and skills with little training. Third, small handheld instruments offer greater safety because the surgeon can more easily override or remove the instrument in cases of malfunction. Finally, if the equipment stops working, the surgeon can simply switch Micron off and use it as a normal handheld instrument.

However, handheld micromanipulators pose additional challenges over purely mechanical damping or master/slave configurations. Because the handle and tip are mechanically coupled, the actuator between them must operate at very high control frequencies with very low latencies. If the actuator responsible for moving the tip relative to the handle cannot react fast enough to counter hand motion, tremor compensation and other micromanipulation tasks become degraded. Since real-world systems exhibit some latency, a pure feedback control system without sufficient bandwidth or too much latency will result in imperfect compensation of tremor, as evidenced in reported residual errors of 10-60 μm during hold-still tasks [101]. This error must be viewed in the context of retinal surgery, where membranes in the eye are only tens of microns thick, and tearing them can damage eyesight.

This chapter focuses on modeling tremor, estimating the state of Micron, and developing feedforward controllers for improved performance. We begin by developing autoregressive models of tremor and analyzing trends in tremor with respect to bracing, visual feedback, and inter-person variation. A model of the tremor can be used for prediction during operation, and can also serve as a substitute for the human user in simulations of the control system. We then investigate using a Kalman filter as an observer to optimally estimate seen and unseen variables of tremor such as position and velocity to model tremorous motions. Finally, we develop a feedforward controller to anticipate tremor and compensate for actuator latency.

4.1 TREMOR MODELING WITH AUTOREGRESSIVE MODELING

Previous research in modeling tremor has employed autoregressive techniques, but these generally have not involved visual feedback and not with magnification [112], [113]. Furthermore, most studies have simplified the spectrum by bandlimiting techniques such as prefiltering [113], analyzing the acceleration signal rather than displacement [112], or windowing for spectral

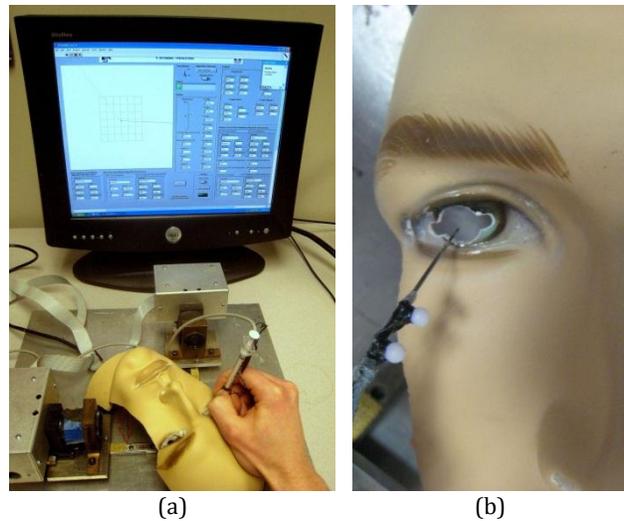


Fig. 31. (a) ASAP measuring and displaying instrument position on LCD screen. (b) Instrument tip with LEDs being inserted through face mask eye. The two notches on opposite sides function as simulated sclerotomies to brace the tool.

analysis [114]. In contrast, we wish to model of the entire spectrum of movement (except for any constant bias, linear trend, or high-frequency sensor noise) during pointing or station-keeping tasks in a simulated vitreoretinal surgical environment, with magnified visual feedback, using standard autoregressive (AR), AR moving average (ARMA), multivariate AR, and nonlinear-augmented AR techniques.

4.1.1 EQUIPMENT

To accurately measure the Micron instrument tip, an optical system called ASAP (Apparatus to Sense Accuracy of Position) is used to provide the real-time position and orientation of the instrument. Two LEDs mounted to the instrument inside diffuse spheres pulse square waves, one at 2 kHz and the other at 3 kHz. The two signals are sensed by orthogonal position-sensitive detectors (PSD's) and demodulated to measure the tip position with an error of less than 10 μm RMS [48].

4.1.2 EXPERIMENTAL PROCEDURE

Tremor data was acquired from five non-medical personnel subjects. The setup, shown in Fig. 52, shows a subject inserting the instrument through the hollowed out eye of the face mask. The instrument tip position and orientation is recorded by ASAP and displayed on the LCD computer monitor at 25X magnification. This "virtual Micron" setup mimics retinal surgery, a practical application of the Micron system.

The study involved three test conditions:

1. The first scenario asked the subject to hold the instrument vertically through the eye. The subject braced the instrument in a notch at the side of the eye and maintains the X, Y, and Z position of the tool tip at a set point on the screen (pointing task). This approximated a typical retinal surgery where the surgeon braces the tool against the sclera (the white of the eye).
2. Secondly, subjects were required to perform the same test as above, except without the bracing (i.e. the instrument shaft is not in contact with the face).
3. The last scenario asked the subjects to close their eyes and maintain the instrument as still as possible while bracing it against the notch in the eye.

In all three cases, the hand rested on the forehead of the face for stabilization. Each subject completed the three scenarios, rested briefly, and then completed the three scenarios again in reverse order. This second dataset enabled cross-validation of the models. All tremor was recorded at a sampling rate of 1 kHz for a duration of 60 s, with 45 s extracted from the middle of each recording to avoid ramp-up and ramp-down effects. The means and linear trends in the data were removed for normalization purposes. To reduce sensor noise, the signal was decimated to 40 Hz.

4.1.3 MODELING TECHNIQUES

Autoregressive techniques simplify a dataset to a mathematical model dependent only on a small number of parameters [114]. These parameters are estimated from the data and can be used to recreate the signal or predict future values of the signal. Autoregressive techniques model a signal as a linear combination of past inputs and outputs. A standard autoregressive (AR) model of order p

represents the value of signal x at position n by a weighted average a of the previous signal values and some error \mathcal{E}_n :

$$x_n = \sum_{k=1}^p a_k x_{n-k} + \mathcal{E}_n \quad (13)$$

Multivariate AR models exploit relationships that exist between multiple signals [115]. For Micron, three tremor signals corresponding to the X, Y, and Z axis are available and may provide additional information when modeled together. In this case, X_n is a three element vector of the X, Y, and Z tremor components and the averaging coefficients are 3x3 matrices A such that:

$$X_n = \sum_{k=1}^p A_k X_{n-k} + \mathcal{E}_n \quad (14)$$

An autoregressive moving average (ARMA) considers not only previous signal values, but a weighted average b of some additional input [116]. For a single signal, the input is taken to be the error terms \mathcal{E}_n of previous predictions:

$$x_n = \sum_{k=1}^p a_k x_{n-k} + \sum_{k=1}^q b_k \mathcal{E}_{n-k} + \mathcal{E}_n \quad (15)$$

These autoregressive techniques can only model linear systems. Although there is some evidence that tremor is a linear process [117], a complete characterization of tremor has not been achieved. Thus, autoregressive techniques augmented with nonlinear regression via wavelet networks [118] were briefly investigated to model any nonlinear processes that might be appearing in the tremor or visual feedback loop.

For each subject and scenario, all four model types were estimated for orders one to forty on the first dataset. Except for the multivariate autoregression, each X, Y, and Z component of the tremor was treated as a separate signal and modeled by itself. To validate how well the constructed model represents the particular subject and scenario being analyzed, it was applied to the second dataset recorded under identical conditions. The evaluation criterion for the model was the root mean squared error (RMSE) magnitude of the difference between the true recorded signal and the predicted signal. A simple last-value predictor, $\tilde{x}_n = x_{n-1}$, was also calculated as a baseline

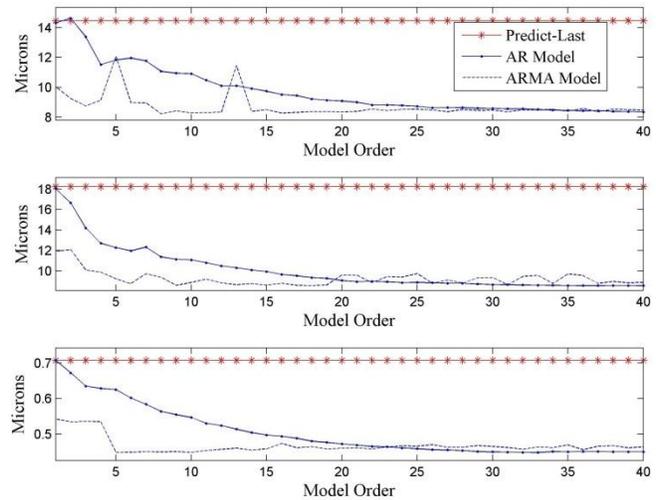


Fig. 32. AR and ARMA results for X (top), Y (middle), and Z (bottom) tremor components for a typical subject under Scenario 1 (instrument braced against side of eye). For this sample, tremor amplitude is 45.4, 62.9, and 5.4 μm RMS in x, y, and z, respectively.

for comparison.

4.1.4 RESULTS WITH AR AND ARMA MODELS

The mean RMS tremor amplitude is 73.7 ± 53.9 , 100.6 ± 100.2 , and 115.5 ± 89.8 μm for Scenarios 1-3 with 3.4 ± 3.0 μm sensor noise. Each model is applied to predict tremor. Fig. 32, Fig. 33, and Fig. 34 show the result of applying AR and ARMA models of increasing order. The data in each figure are derived from one person across scenarios and tremor signal components. The RMSE magnitudes achieved by predictions of AR and ARMA models 1 to 40 are displayed and compared to the simple baseline last-value predictor. Order 40 predicts the next value from one second of past signal.

As evidenced by the graphs, both AR and ARMA models perform better than the baseline predictor. While for high orders AR and ARMA models converge, the ARMA model tends to converge much more quickly, by sixth order or lower, which is in agreement with [113], [114]. The results shown are from a typical subject; for some subjects the RMSE trends slightly upwards for higher orders of AR, probably an artifact of over-fitting.

Numerically, the last-value predictor achieves a mean reduction of 81.5% in RMSE magnitude over all the subjects and scenarios. Using this as a baseline predictor, Fig. 36 shows that AR and ARMA models improve the prediction by roughly 20-50%, depending on the model order used.

4.1.5 RESULTS WITH NONLINEAR AND MULTIVARIATE MODELS

The nonlinear model used was a wavelet network [118] that augmented the standard autoregressive model. One interesting result of adding the wavelet nonlinear regression is the tendency of the nonlinear regression to overfit the data. For the dataset used to construct the model, the nonlinear approach improved the prediction. However, when tested on the cross-validation dataset, the results showed very slight improvement. For a fifth-order model, the nonlinear regression reduced the RMSE by 0.1%.

In contrast, the multivariate standard autoregressive model led to a 0.2% increase in RMSE on the cross-validation set for a second-order model and oftentimes yielded an unstable fit at higher model orders. In many instances, for orders larger than 15, the multivariate models no longer outperform last-value prediction. While an improvement can be achieved on the training dataset, performance was markedly degraded on the cross-validation set.

4.1.6 GENERALIZED ARMA MODEL

While a model per person per scenario is indeed useful, a unified model that generalizes across both people and scenarios would be more advantageous as it alleviates

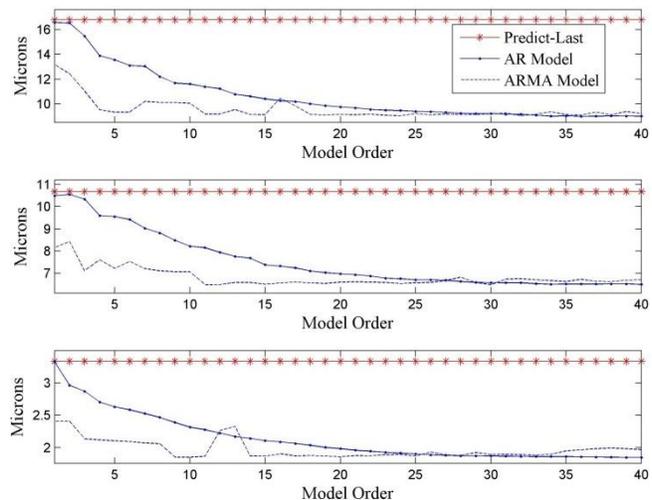


Fig. 33. AR and ARMA results for X (top), Y (middle), and Z (bottom) tremor components under Scenario 2 (instrument not braced against side of eye). For this sample, tremor amplitude is 46.4, 31.9, and 24.1 μm RMS in x, y, and z, respectively.

the need for a pre-usage calibration step. Because it has proven superior in prediction than an AR model and is less complex than a nonlinear model, only the ARMA model was investigated. First, the variances of the model parameters were examined. Fig. 35 shows a box and whisker plot of the first six coefficients for all ARMA models. This shows the median values for the coefficients, in addition to the lower and upper quartiles, and the furthest data points within a factor of 1.5 of the interquartile range. Outliers are represented as crosses.

The ARMA error coefficients are similar for different trials. By taking the mean of these coefficients, an averaged ARMA model can be

constructed that represents all subjects and all scenarios. Applying the averaged model to each cross-validation run shows a slight drop in overall accuracy from 48.0% to 47.5% improvement over the baseline. A reduction in accuracy of approximately half a percent may be a preferable alternative to a time-consuming calibration routine that learns the ARMA coefficients for each new user. The predicted RMSE for Scenarios 1-3 is 8.9 ± 9.3 , 9.2 ± 9.2 , and $9.8 \pm 8.9 \mu\text{m}$.

4.1.7 DISCUSSION

In this preliminary study of modeling and predicting tremor using autoregressive techniques with five subjects under three different scenarios, a sixth-order ARMA model predicted tremor 48.0% better, on average, than a baseline last-value predictor. A single ARMA model derived from an average of all ARMA models shows that the problem can be generalized across all three scenarios and subjects with only a 0.55% drop in accuracy. With this single ARMA model, the tremor was predicted with a mean of $8.2 \pm 5.9 \mu\text{m}$ RMSE using raw hand movement signals with a mean of $96.6 \pm 84.5 \mu\text{m}$ RMS amplitude. We also investigated other autoregressive models. The standard AR model performs nearly as well as ARMA, but only at much higher model orders. The multivariate AR model did not generalize well from the data used to estimate the model to the cross-validation

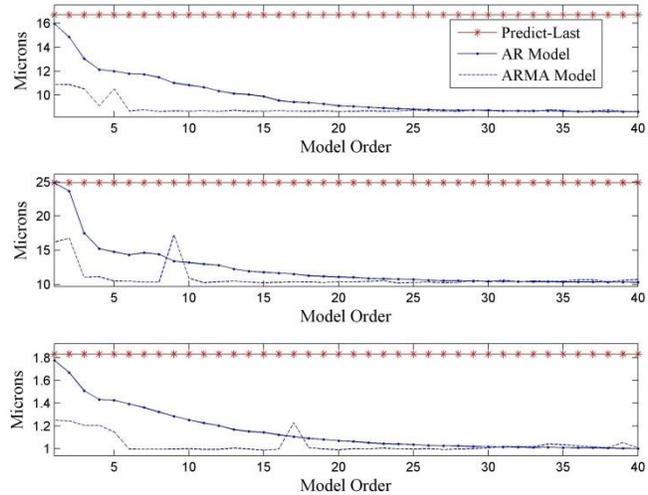


Fig. 34. AR and ARMA results for X (top), Y (middle), and Z (bottom) tremor components for a typical subject under Scenario 3 (instrument braced, but no visual feedback). For this sample, tremor amplitude is 254.0, 166.1, and 15.2 μm RMS in x, y, and z, respectively.

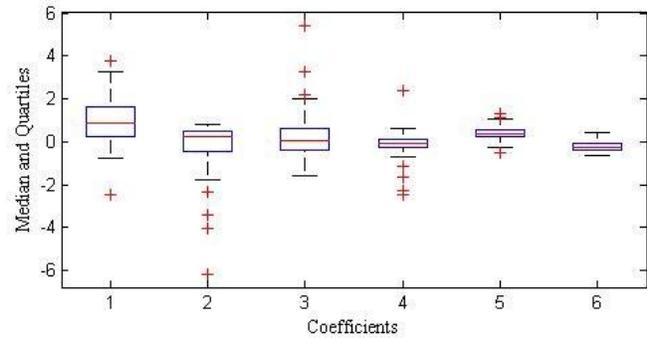


Fig. 35. Box and whisker plot of the ARMA coefficients (the error term coefficients yield a similar plot with smaller coefficients).

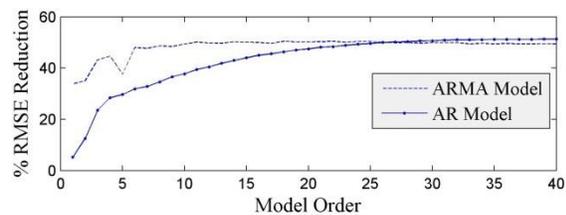


Fig. 36. Mean percentage improvement for AR and ARMA in comparison with last-value prediction.

data. AR models augmented with nonlinear wavelet networks marginally increased prediction performance, but likely not enough to warrant the additional complexity.

4.2 STATE ESTIMATION WITH KALMAN FILTERS

While autoregressive modeling does an adequate job at modeling tremor when the signal is decimated to 40 Hz to reduce the noise, AR and ARMA models are less suited to handle full-rate signal rates of 2 kHz with high noise. Alternative models of tremor such as the weighted Fourier linear combiner WFLC [103] have demonstrated good results, but are best suited to modeling tremor acceleration, which is more or less zero-meaned. Since Micron uses an absolute positioning measurement system, the WFLC would need to be modified to handle the integrative effect of acceleration over time (i.e. also estimate an offset bias as well). Instead, to more accurately model tremor at high bandwidth with absolute positioning, we turn to a modeling technique that optimally estimates both measured and unseen states from noisy estimates and knowledge of the underlying physical process: the Kalman filter. The goal is to model the 6 DOF pose position and velocity of the hand and the tremor movements as shown in Fig. 37.

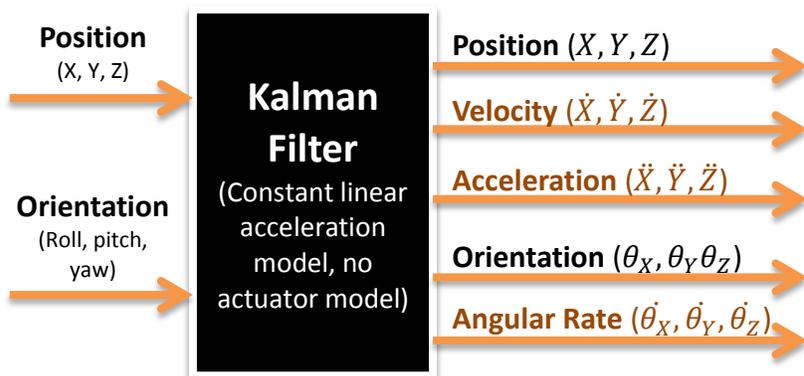


Fig. 37. Kalman filter input is the 6DOF position of the Micron handle (i.e. hand position) and the output is the 3DOF position, velocity, and acceleration along with the 3DOF angular position and velocity.

of the hand and the tremor movements as shown in Fig. 37.

4.2.1 KALMAN FILTER ALGORITHM

Optimal estimation of the current state of a system is a wide field [119], but one of the most popular state estimators is the Kalman filter [120], [121]. A main assumption of the Kalman filter is that observations, or sensor readings, z of a system follow some dynamics. The current state of the system $x \in \mathbb{R}^n$ is the n minimum number of variables necessary to predict future behavior of the system. If the system is linear, states evolve according to the dynamic model A of the system and with the inputs u to the system (i.e., actuation). More formally, we can represent the state transitions discretely at time step k as:

$$x_k = A_k x_{k-1} + B_k u_k \quad (16)$$

$$y_k = C_k x_k \quad (17)$$

where B_k and C_k describe the impact of inputs u on the state x and how observations y are made from the current state x , respectively. Assuming Gaussian noise on the observations z and the dynamic model A characterized by covariances R and Q , respectively, the Kalman filter optimally estimates the system's state x at each time step:

$$\hat{x}_k = A_k x_{k-1} + B_k u_k \quad (18)$$

$$\hat{P}_k = A_k P_{k-1} A^T + Q_k \quad (19)$$

$$\hat{y}_k = z_k - H_k \hat{x}_k \quad (20)$$

$$S_k = H_k \hat{P}_k + R_k \quad (21)$$

$$K_k = \hat{P}_k H_k^T S_k^{-1} \quad (22)$$

$$x_k = \hat{x}_k + K_k y_k \quad (23)$$

$$P_k = (I - K_k H_k) \hat{P}_k \quad (24)$$

where (18-19) are known as the predict step, which use the dynamic model and inputs to the system to predict the state \hat{x} at the next time step. Equations (20-24) are known as the update steps, where the residual between the predicted state \hat{x} and the measured state z is used to calculate the Kalman gain K , which optimally mixes between the measurements and the dynamic model to yield the best estimate of the current state x . P is the error covariance that represents the accuracy of the state estimate x . The Kalman filter has been widely successful in a number of different applications.

4.2.2 STATE ESTIMATION WITH KALMAN FILTER

Full 6-DOF pose information for the tip and handle is calculated from the LEDs via triangulation and the application of the closed-form Horn calculation [49] on the recovered 3D positions of the LEDs. Because the raw 2 kHz pose data is noisy and numerically calculating velocity or acceleration data would be extremely noisy, we use a Kalman filter for complete state estimation of the pose of Micron's tip and handle position. We choose a linear Kalman filter similar to [122] since the sample rate is sufficiently high to well approximate the nonlinearities in the system. Furthermore, recent work has shown that nonlinear versions such as the Extended or Unscented Kalman filters have difficulties modeling tremor with nonlinear harmonic models [123]. Because the hand motion contains low-frequency voluntary movement (<1 Hz) combined with involuntary tremor motions (10-20 Hz), we use a constant acceleration motion represented by the state transition:

$$\begin{bmatrix} p \\ \dot{p} \\ \ddot{p} \end{bmatrix}_k = \begin{bmatrix} 1 & dt & \frac{1}{2}dt^2 \\ 0 & 1 & dt \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p \\ \dot{p} \\ \ddot{p} \end{bmatrix}_{k-1} \quad (25)$$

with $P = [p, \dot{p}, \ddot{p}]^T$ representing the position, velocity, and acceleration of the system.

To represent orientation, we investigated quaternions but used Euler angles instead for the following reasons. First, Micron's orientation only changes slightly during operations, usually by only 10-20°, so gimbal lock is not an issue. Second, Euler angles' representation with three variables instead of quaternion's overrepresented four variables leads to easier calculations, as issues of heteroscedasticity and re-normalization do not arise. More importantly, since the Kalman filter must be run at 2 kHz in a real-time operating system, using three variable representations for orientation results in faster calculations, especially in the slow inverse calculation of (7). Third,

quantitative analysis revealed that in typical usage scenarios, the difference between estimations using Euler angle representations and those using quaternions is negligible. We use the constant angular velocity model for the orientation:

$$\begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix}_k = \begin{bmatrix} 1 & dt \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix}_{k-1} \quad (26)$$

with $\Theta = [\theta, \dot{\theta}]^T$ representing the orientation angle and angular velocity in the world frame. Denoting the state transitions of position and orientation as A^p and A^o and the full 15 state vector $x = [P_X, P_Y, P_Z, \Theta_X, \Theta_Y, \Theta_Z]$, we can build the full 6-DOF state transition in block-diagonal form:

$$\begin{bmatrix} P_X \\ P_Y \\ P_Z \\ \Theta_X \\ \Theta_Y \\ \Theta_Z \end{bmatrix}_k = \begin{bmatrix} A_X^p & \mathbf{0}_3 & \mathbf{0}_3 & & & \\ \mathbf{0}_3 & A_Y^p & \mathbf{0}_3 & & & \\ \mathbf{0}_3 & \mathbf{0}_3 & A_Z^p & & & \\ & \mathbf{0}_{6 \times 9} & & A_X^o & \mathbf{0}_2 & \mathbf{0}_2 \\ & & & \mathbf{0}_2 & A_Y^o & \mathbf{0}_2 \\ & & & \mathbf{0}_2 & \mathbf{0}_2 & A_Z^o \end{bmatrix} \begin{bmatrix} P_X \\ P_Y \\ P_Z \\ \Theta_X \\ \Theta_Y \\ \Theta_Z \end{bmatrix}_{k-1} \quad (27)$$

Measurements at each time-step are represented as $z = [p_x, p_y, p_z, \theta_x, \theta_y, \theta_z]$. Measurement and process covariances are set to $R = [10^3, 10^3, 10^2, 10^3, 10^3, 10^2] * I_{6 \times 6}$ and $Q = 10^{-1} * I_{15 \times 15}$, respectively. Because the tip and handle are only loosely coupled, separate Kalman filters are used for state estimation of the tip and handle to reduce computational requirements. For the handle, no hand actuation input is available, so control input $u_k = 0$. For the tip Kalman filter, the forward kinematics J and model B of the actuators are calculated once from a ~ 30 s calibration procedure. Because the actuators are in the body frame while all other state variables are in the world frame, the actuator input u_k is first rotated into the world frame by the current orientation R_k estimated from the state $[\theta_{X,k}, \theta_{Y,k}, \theta_{Z,k}]^T$:

$$\hat{x}_k = Ax_{k-1} + BR_k u_k \quad (28)$$

giving us a prediction model that takes into account both the internal dynamics of the system and any actuation. By integrating system dynamics, actuator models, and measurements, the Kalman filter estimates of the position, velocity, acceleration, orientation, and angular velocity of both the tip and handle of Micron.

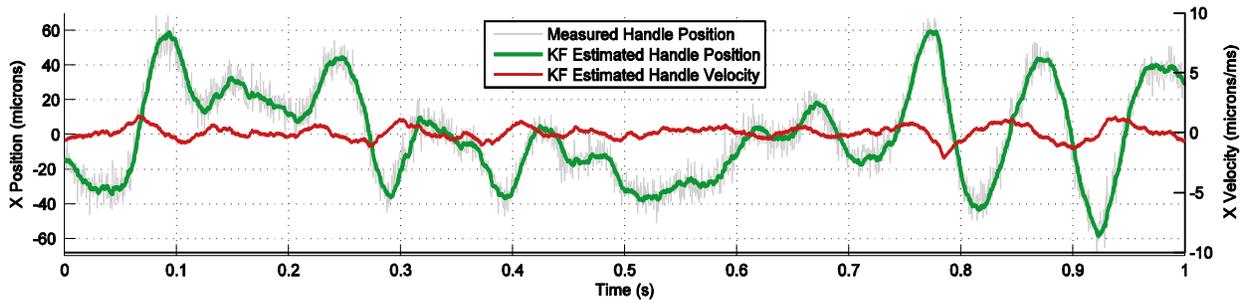


Fig. 38. State estimation of hand motion by Kalman Filter compared to raw measurements.

4.2.3 RESULTS

Fig. 38 shows the Kalman filter state estimation of the handle position and velocity for a one second time slice of the X axis. The Kalman filter is able to accurately and smoothly estimate not only position, but higher order derivatives while ignoring significant amounts of sensor noise. The Y and Z axis graphs are similar to the figures shown for the X axis.

4.3 TREMOR REJECTION WITH FEEDFORWARD CONTROL

Given an accurate enough model of hand motion representing the tremor, it should be possible to use that model to predict tremor motion into the future and correct for latencies in actuation. For instance, Micron's piezoelectric actuators have a 3 ms delay in actuation, e.g. from the time a command is commanded, 6 cycles pass before motion is sensed at the output (2 kHz sampling rate) [27]. In this section, we use a model of tremor to predict tremor motion into the future and feed the predictions forward into the control system to anticipate and correct errors before they occur.

4.3.1 BACKGROUND

In a handheld micromanipulator as depicted in Fig. 39, we assume there is some set point $p^s \in \mathbb{R}^3$ that is selected as the goal position for the tip of the micromanipulator. Once the set point has been generated by a higher-level control system, the low-level control system attempts to track the set point with the tip of instrument. Disturbances to the tip arise from a variety of sources including vibrations, resonances, or contact with tissue. However, the largest disturbance in a micromanipulator system is generally hand motion, which includes both voluntary movement and involuntary tremor. Typically, tremor is treated as the disturbance, but voluntary motion can be thought of as a disturbance too, such as the case when the set point is being generated from virtual fixtures or motion scaling. For the purposes of this paper, we consider all hand motion that does not coincide with the set point motion to be a disturbance to the control system.

In feedback systems, disturbances are handled as they cause the output to drift away from the set point; this error is then fed back into the control system to bring the output back to the set point. One popular feedback control approach is PID (Proportional, Integral, and Derivative) gains on the error. However, feedback control can only react to disturbances after they have so affected the system state that the deviation from the goal is noticeable. This error is then exacerbated by latencies in actuation as it may take several cycles before the actuators respond to eliminate the error.

If the disturbance is predictable or correlated with some other predictable variable (e.g. gravity, friction), feedforward control can couple the set point directly to the control variable. By modeling the disturbance and its effect on the system, control effort can be applied before error occurs. To compensate for latencies in the actuation, short-term future predictions from the model can be used to drive the actuators in anticipation of how the disturbance will affect the system. When suppressing tremor in a handheld micromanipulator, the hand motion that does not correspond to the set-

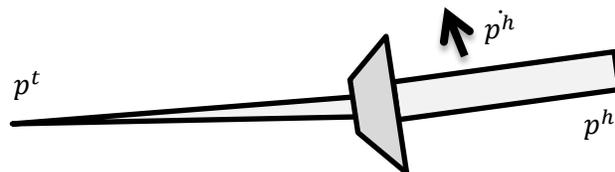


Fig. 39. Handle velocity p^h caused by tremor causes error in the tip position p^t over the latency period.

point motion is the disturbance. Such feedforward rejection of tremor requires good estimation of both observed and unobserved motion states of the system.

By using a Kalman filter for state estimation of both the handle and tip, we propose a feedforward control strategy that anticipates hand motion to more effectively suppress tremor in handheld micromanipulators.

4.3.2 METHODS

During handheld operation, the tip p^t may deviate from the set-point goal p^s by some error:

$$e = p^s - p^t \quad (29)$$

Inverse kinematics J^{-1} translate measured error e into corrective actuation with a PID controller, generating feedback control u_{FB} :

$$u_{FB} = K_P J^{-1} e + K_I \int J^{-1} e + K_D \frac{d}{dt} J^{-1} e \quad (30)$$

where K_P , K_I , and K_D are the proportional, integral, and derivative gains. In a zero latency environment, a properly tuned PID controller will bring the tip to the set-point. However, let us introduce latency in actuation, denoted by T seconds. To a first order approximation, the tip p^t will move by the integral of the handle velocity p^h over the latency period T before the velocity can be accounted for by the Micron accuation. Unless the hand velocity p^h corresponds exactly to the set-point velocity p^s , the error in latency will be:

$$e_L = \int_0^T p^h - p^s dt \quad (31)$$

Assuming the set-point velocity is negligible ($p^h \gg p^s$), our error e_L reduces to the integral of p^h over the latency of the system T and feedforward control u_{FF} can be integrated into the system with a PID controller similar to (14):

$$u_{FF} = K_P J^{-1} e_L + K_I \int J^{-1} e_L + K_D \frac{d}{dt} e_L \quad (32)$$

Both feedback and feedforward inputs are applied at each time step:

$$u = u_{FB} + u_{FF} \quad (33)$$

One analogy might be that of a ship steering system. To keep the ship on trajectory, minor course corrections will be made as necessary; this is the feedback system working to correct small deviations in course. However, if a large disturbance occurred, such as a strong wind, the slow actuation of the ship may result in the feedback system taking a long time to restore proper course bearing. On the other hand, if it is known that there will soon be wind blowing east, the ship can begin steering westerly to be ready for the future disturbance. Likewise, to more robustly maintain

a set point, hand movements estimated by the Kalman filter can be anticipated before they occur instead of after accumulating into significant error.

4.3.3 EXPERIMENTAL PROTOCOL

We present results of the Kalman filter state estimation of Micron and its effect on tremor suppression. Using the state-estimation to anticipate tremor, we use implement and test feedforward control in hold-still tasks with Micron, our handheld micromanipulator. Tests were performed with Micron under a board-approved protocol to evaluate the Kalman filter state estimation and the tremor suppression effects of the feedback+feedforward controller. All experiments were performed with a single individual familiar with Micron but without surgical experience. The task of interest is hold-still, in which the operator attempts to hold the tip motionless at a fixed 3D set point. Visual cues are presented to the operator on a 3D monitor to help the operator avoid drifting too far from the set point, which would saturate the actuators. To alleviate operator bias, controllers are executed at random in 5 s intervals without notifying the operator.

Ground truth for the tip position is constructed by low-passing the tip position with a 100 Hz bi-directional, zero-lag 2nd order Butterworth filter. Because manipulator resonances begin around 100 Hz, this serves as a reasonable ground truth. Furthermore, the zero-lag low-pass filter is a more unbiased estimator of the ground truth than the Kalman filter because the system feeds the output of the Kalman filter back into the control system. Combined with the actuator model, this could introduce biases into the estimation. Each controller is evaluated by calculating RMS error between the ground truth tip position and the set point goal.

4.3.4 RESULTS

Fig. 40 shows the performance of the Kalman filter and how well the velocity output of the Kalman filter predicts tip error. When

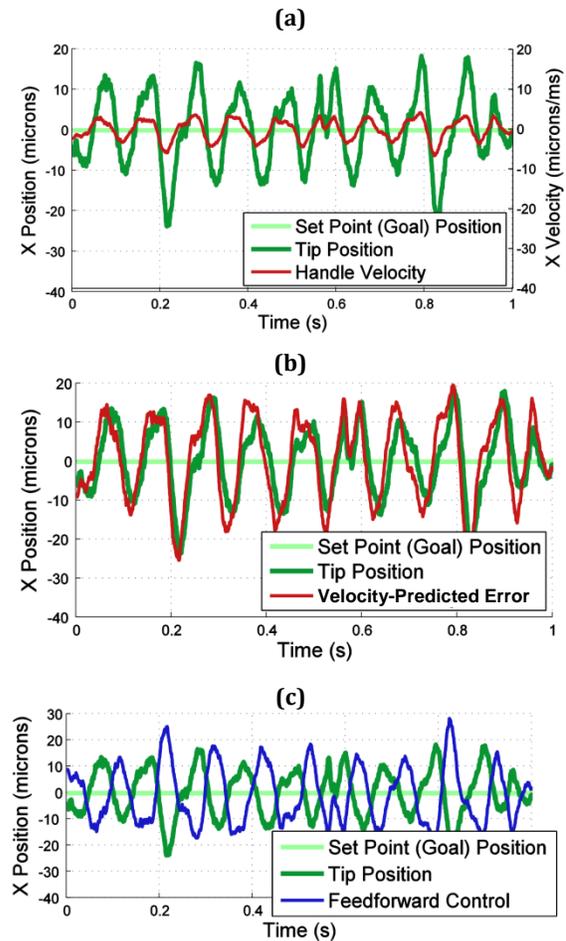


Fig. 40. (a) Tip position with Kalman filter estimated velocity (b) Velocity-predicted error by integrating velocity over latency time T compared to tip position error. (c) Inverted velocity-predicted error is used as feedforward control.

Method	RMS Error (μm)	Improvement Over Baseline (%)
Baseline Feedback	7.9	-
Clamped (No Tremor)	1.4	-
Feedback+Feedforward	3.4	56.7

Table 4: Error of controllers with mean RMS performance represented as percent improvement over the baseline feedback-only controller. Best attainable tremor suppression should be approximated by the case where Micron is clamped in a vise (i.e. no tremor).

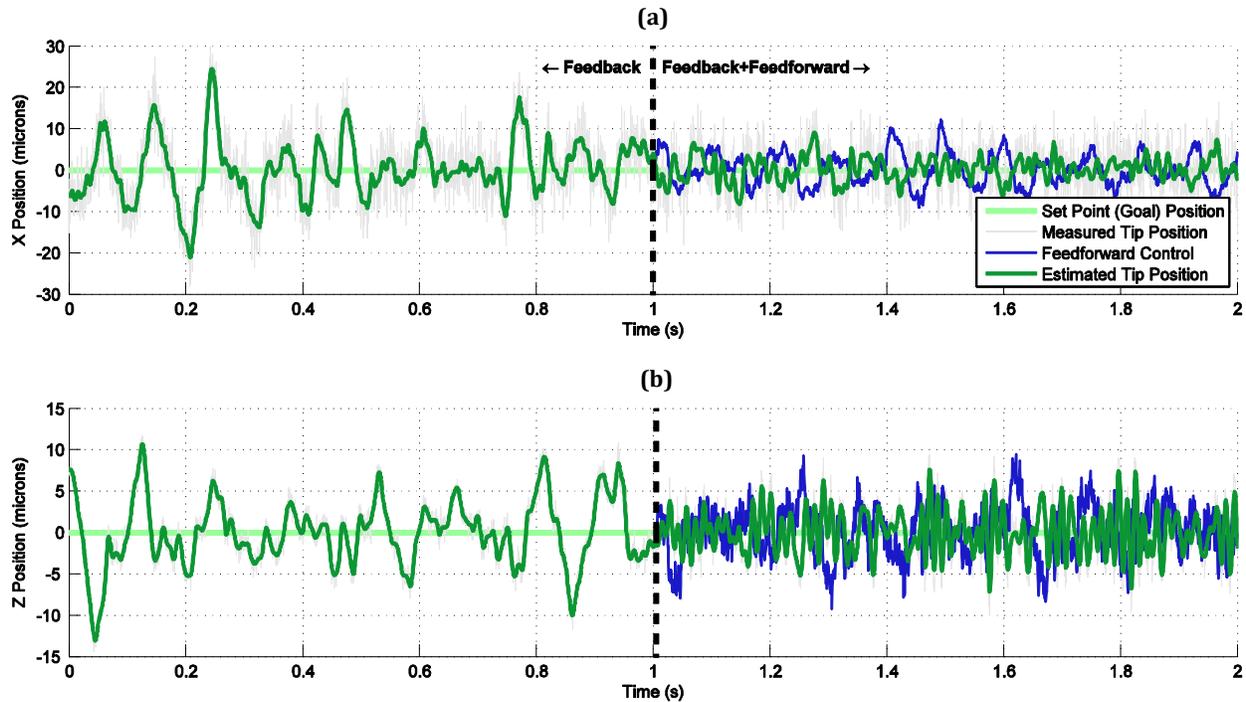


Fig. 41. Comparison of the feedback vs. feedback+feedforward controller for tremor suppression ((a) top: transverse, (b) bottom: axial). The controller is swapped out dynamically at = 1 s. Although tremor suppression is not completely effective in either case, the addition of the feedforward component is clearly beneficial.

feedforward control with Kalman filter estimation is added to the feedback controller, more tremor is suppressed. Fig. 41 demonstrates the additional suppression by showing the activation of the feedback+feedforward controller during a hold still task. Because of imperfect estimation, actuation latencies, and other nonlinearities, perfect cancellation is not achieved; however, tremor-caused disturbances of 10-20 μm magnitudes can generally be reduced to 5-10 μm with the feedforward approach. Table 4 lists the feedback+feedforward controller as reducing the RMS error by 56.7% as compared to only using the feedback controller. Fig. 42 shows a power spectrum of the error for each controller. The feedback+feedforward controller performs better at low frequencies, and peaks somewhat around 60 Hz. However, this is not a concern, as very little energy exists at this frequency. As a best case bound for handheld tremor suppression, we present the results of Micron maintaining a set point in the absence of any hand motion and tremor; this is achieved by clamping the handle with rubber in a vise.

The closed-loop frequency response of the new feedback+feedforward controller compared to the baseline controller is shown in Fig. 43. A sine wave chirp signal sweeping 1 Hz to 1 kHz is injected into the pose of the

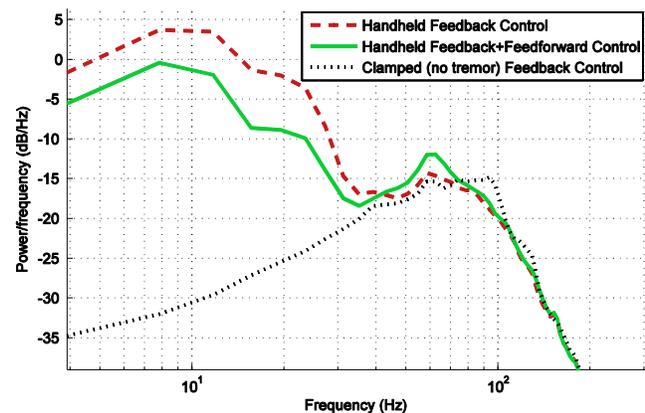


Fig. 42. Power spectrum of the two controllers under handheld conditions: the baseline controller and the feedforward controller. For comparison to a tremor-free case, the baseline feedback controller operating in a clamped rubber vise is shown also.

instrument during handheld operation. The tip position response to the injected stimulus is recorded over many runs to average out handheld motion. Notice the feedback+feedforward controller does significantly better than the the feedback-only controller in the lower frequencies where tremor is most strongly exhibited. At 10 Hz, where peak tremor power occurs, the attenuation in the response is -7 dB. By the Bode Integral Theorem, we expect to see an increase in the response at higher frequencies corresponding to the decrease at lower frequencies, which we do. However, as evidenced in the power spectrum in Figure 5, there is significantly less energy in the system at these higher frequencies so overall the system demonstrates a net gain in tremor compensation.

4.3.5 DISCUSSION

We have demonstrated that applying a Kalman filter for state estimation of hand motion in a feedforward controller can produce superior tremor suppression. A Kalman filter for both the tip and the handle of Micron is run at 2 kHz to fully estimate the position, velocity, acceleration, orientation, and angular velocities. By modeling the effect of physiological tremor on the process plant, feedforward control uses estimated handle velocity to anticipate and reject extraneous hand motions. Experiments show a reduction in RMS error of more than half, which corresponds to approximately -7 dB rejection ratio. This significantly improves accuracy around tiny anatomy such as retinal vessels and reduces trauma to surrounding tissue caused by tremor.

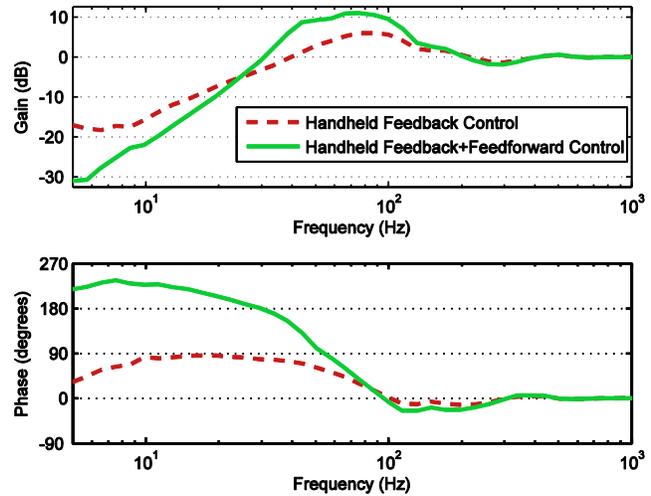


Fig. 43. Frequency response of the two controllers under handheld conditions: the baseline feedback controller and the feedback+feedforward controller. Notice the additional attenuation of error at low frequencies in the 5-20 Hz tremor range. This frequency response is for the transverse axis of the manipulator; the lateral response is similar.

5 REAL-TIME RETINAL IMAGE ANALYSIS

In order to provide assistance through the virtual fixture framework that controls the behavior of Micron, visual information about the tip position and anatomy must be analyzed from the stereo cameras mounted to the microscope. This visual analysis must be completed in real-time because, unlike a static or master/slave robot that remains stationary in-between planning operations, Micron must rapidly compensate for hand movement and tremor. In this chapter, we explore, select, and develop computer vision techniques suitable for high-speed operation and integration into the control system. The Open Computer Vision (OpenCV) library [108] is extensively used to make the most of multi-threaded CPU and GPU-accelerated algorithms.

5.1 TIP TRACKING

In order to provide useful behaviors of the Micron tip relative to some anatomy, the tip of the instrument must be tracked. Although a variety of sophisticated surgical instrument trackers are available such as [124], we find that by applying colored fiducials to the tip, simpler methods can be used with significantly less computational complexity. Depending on the environment, background color, and light configuration, different fiducials are used. For many applications, painting the entire shaft an infrequently seen color is an effective solution. If visual calculation of orientation is required, two colored fiducials can be applied along the shaft of the instrument (one at the tip and one further up).

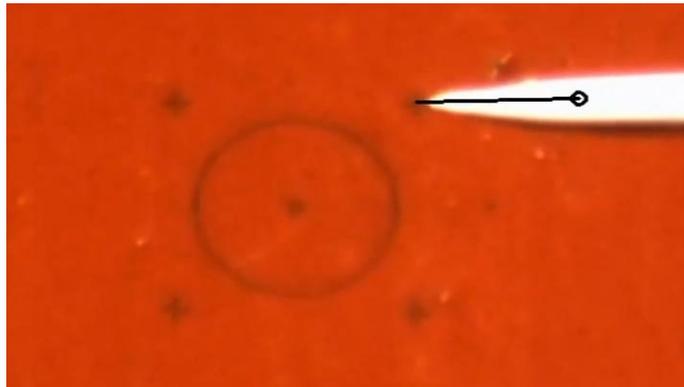


Fig. 44. White-painted tip of Micron being tracked with the hollow black circle representing the centroid of the shaft, and the black line originating from the centroid indicates the direction and length of the tip.

Thresholding or fast, trainable color detection algorithms [125] are used to create noisy binary masks of the tip, which are subsequently filtered by a combination of median filters and morphological operations. The resulting clean binary masks are inputted to a fast blob segmentation algorithm [126]. Segmented blobs are analyzed and the largest blob is selected subject to constraints such as minimum size and color intensity. In the case of a single, localized fiducial, the centroid of the blob can be directly used as the tip position. However, if the entire shaft is a colored fiducial, Hu image moments [127] are calculated to determine the axis of the instrument shaft, and a linear scan originating at the centroid of the blob along the major axis of the tip finds the tip position. Fig. 39 demonstrates tracking the entire shaft to find the tip position.

Tip tracking runs at 30 Hz at full resolution for both the left and right images. Separate threads are spawned for each camera and tip tracking is run simultaneously. By factoring in camera shutter time, duration of transfer to CPU, and the time required for analysis of the image, each tip position in the left and right image can be synchronized to ASAP-measured tip positions. By connecting the strobe output that triggers upon the beginning of image integration to the analog input of ASAP, precise synchronization is possible.

5.2 ADAPTIVE CAMERA CALIBRATION TO ASAP

The ASAP measurement system provides very fine positioning information of the instrument and allows for general tremor compensation via filtering in the frequency domain. Control of Micron operates in the coordinate system defined by the ASAP measurement system. However, it has no knowledge of what an operator sees in the microscope. Thus, vision-based control of the manipulator is necessary because the anatomy of interest is localized via intraoperative imagery. Traditional image-based visual servoing techniques [104], [128] operating at the camera rates of 30 Hz cannot provide the high bandwidth and low latencies

needed for active control of handheld micromanipulators. The solution is to provide information from the camera frame in the ASAP coordinate system so the control loop can use the much faster 2 kHz ASAP positioning system. Thus, virtual fixtures generated from vision must be transferred from the stereo images to the ASAP coordinate frame, requiring calibration and registration of the cameras.

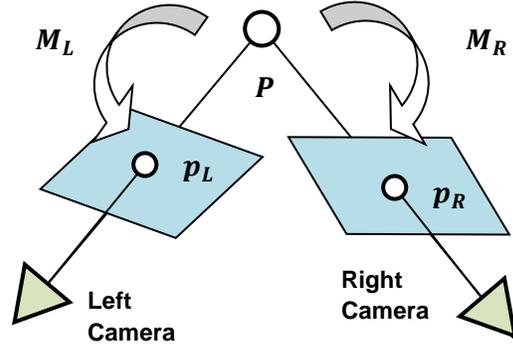


Fig. 45. Multiple view camera geometry, showing the projective relationship M_1 and M_2 that map the 3D point P to image coordinates p_1 and p_2 viewed by two cameras. Note this is for the general case; the cameras views in the Micron setup are parallel.

Fig. 45 shows the multi-view geometry between the cameras and their corresponding images. Calibration is performed by collecting roughly 30 s of the 2D tip positions $p \in [x, y, 1]^T$ in the image with the corresponding 3D tip positions $P \in [X, Y, Z, 1]^T$ as measured by ASAP [48]. The projective camera matrix $M \in \mathbb{R}^{3 \times 4}$ that satisfies the homogenous relationship $p \equiv MP$ for each camera is estimated using the Direct Linear Transform [129]. After the initial 30-s calibration procedure, an adaptive registration dynamically adjusts the calibration with each new camera frame to account for any nonlinearities and microscope focusing:

$$M \leftarrow M + \eta(p - MP)P^T \quad (34)$$

where η is a small gain (e.g. $\eta = 10^{-6}$). Thus, an up-to-date, invertible mapping always exists between the cameras and ASAP coordinate frames. This allows anatomy localized and tracked in the cameras to be reconstructed in the 3D ASAP coordinate system, so that virtual fixtures can be generated and enforced at 2 kHz.

5.3 3D RECONSTRUCTION WITH DENSE STEREO

For many operations, it is important to know the location of the retina in 3D to either enforce depth limits or display relative depths to the operator via visual cues. While multi-view geometry is simple enough to back-project sparse points such as the tip position into 3D ASAP coordinates, a different strategy is needed for dense 3D information. We have discovered that simple block-matching stereo algorithms produce reasonably consistent, accurate, and fast depth maps if enough care is taken to setup and optimize the process. We start out by estimating a homography from the left camera to the right camera using features found on a highly textured planar surface. By warping

the left camera view to the right camera view with the estimated homography, we explicitly force the resulting depth map to have an approximately zero mean disparity. Because block-matching only checks in one direction, we explicitly offset one image by half of the maximum level of disparity. We can then use a fairly small 32 maximum levels of disparity, which increases speed. To account for the low texture of the retina, we resize the image to 320x240 and use a large block size of 51x51 pixels. For faster than realtime computation, we use the GPU accelerated block stereo matching algorithm from the OpenCV library [108].

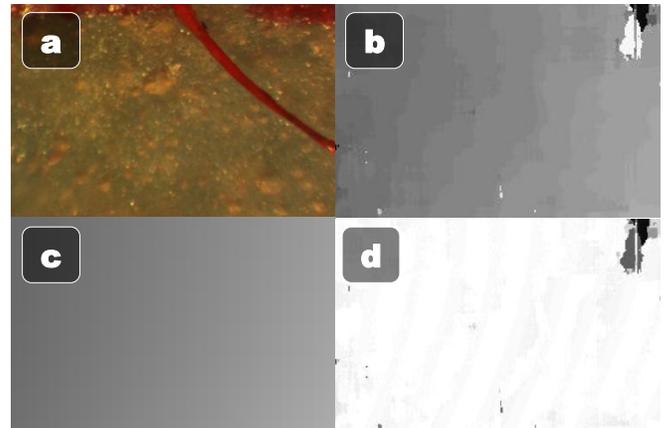


Fig. 46. 3D Quadratic parametric modeling of the phantom eyeball. (a) One of the raw stereo frames (b) Estimated depth map (c) 3D quadratic fit estimated with RANSAC (d) Error between the estimated 3D quadratic and the actual depth map (white indicates no error, darker is higher error).

As seen in Fig. 46, the resulting depth map is still fairly noisy and subject to patches of missing or incorrect depth data. However, since the retina is a smooth section of a roughly spherical eye, we reduce the full depth map to a parametric model by estimating the best fit 3D quadratic to the depth map using the RANSAC algorithm [130]. The stereo depth map calculation and 3D surface modeling can be performed at 30 Hz thanks to a GPU-accelerated implementation of stereo block matching.

5.4 HARDWARE VIDEO ENCODING

When writing computer vision applications that process fairly high-resolution imagery in real-time, significant amounts of CPU are used and lots of video data is generated. In fact, generally, the challenge with most vision systems is to choose algorithms that do not consume all available CPU. With an already-taxed CPU, storing processed video is difficult. Two stereo cameras running at 1024 x 768 @ 30 Hz generate 135 M/s in full RGB video, in addition to any supplementary data generated by the system. This is a lot of data to process, but worse yet: what do you do with the data once you've processed it? One of the most important aspects of any experiment is recording data for post-processing analysis. Saving the full 135 M/s is difficult with traditional harddrives and consumes disk space very rapidly, while encoding the video consumes a significant amount of CPU that is already being used by the rest of the vision system. One solution is to use a lightweight, lossless encoder such as HuffYUV that reduces the size of the video by 4-5X without a heavy CPU load. However, newer Intel CPUs starting with the code-name Sandy Bridge line (Intel i7-2600K) have introduced an on-die video processing pipeline named Quick Sync that encodes and decodes H.264 video. With dedicated ASIC hardware built-into the GPU portion of the CPU, high-definition video can be encoded in excess of 100 fps with almost no traditional CPU usage. Thus, high-quality, full-resolution video can be saved with a state-of-the-art encoding for later analysis and presentation with only a negligible impact on system performance.

5.5 EYESLAM: LOCALIZING AND MAPPING THE RETINA

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 [42], [46]. During retinal laser photocoagulation, placing burns on retinal vessels should be avoided as

this can occlude the vein, possibly causing vitreous hemorrhage [61]. However, existing methods for vessel detection or retinal registration are not suited to real-time operation, preferring accuracy over speed for offline use, and do not handle constraints required for intraocular surgery, such as robustness to tool occlusion.

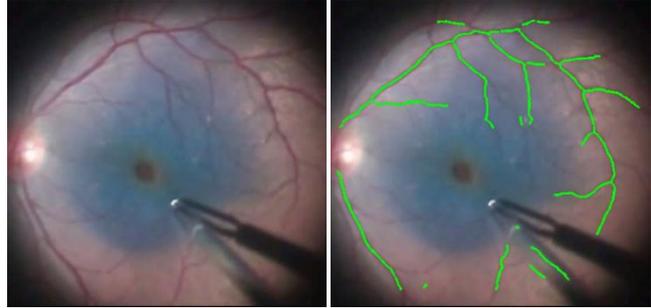


Fig. 47. The proposed mapping and localization algorithm for retinal vasculature running in real-time on recorded *in vivo* human retina during a retinal peeling with blue die. Video source: <http://youtu.be/CTnavOgDsXA>

In this section, we propose an approach to map and localize the vasculature of the retina in real-time that is robust to tool occlusions

and variable illumination conditions, for use in intraocular surgery. We investigate a wealth of published work related to localization and mapping of retinal vessels exists and can be grouped into three general categories: vessel detection, retinal registration, and the more general robotic approach of simultaneous localization and mapping (SLAM). Our approach of using the fast retinal vessel detection of [131] for feature extraction, an occupancy grid for mapping [132], and iterative closest point (ICP) [110] for localization is described. We evaluate our approach on videos of an *in vitro* eyeball phantom, *ex vivo* porcine retina, and *in vivo* human retina (see Fig. 47).

5.5.1 RETINAL DETECTION

Vessel detection is the process of extracting vasculature in retinal imagery and often includes calculating the center lines, width, and orientation of vessels. One set of methods uses local color and intensity information to classify the image on a per-pixel basis [133–136]. Another popular approach is to search across the image for vessel-like structures using matched filters at various locations, scales, and orientations [131], [137], [138]. Other algorithms use a bank of Gabor wavelets to do a pixel-wise classification of the image [139–141]. 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 [131], [136], [141] and other hardware-accelerated methods [142], [143], most algorithms require more than 1 s to run, which is insufficiently fast to benefit robotic control loops. Table 5 shows a sampling of the more recent and popular approaches and the approximate times required to run each algorithm.

Algorithm	Year	Time
Chaudhuri et al. [137]	1989	50 s
Can et al. [131]	1999	0.03 s
Chanwimaluang et al. [133]	2003	3 min
Jiang et al. [134]	2003	8 s
Staal et al. [135]	2004	15 min
Soares et al. [139]	2006	3 min
Sofka et al. [138]	2006	30 s
Mendonca et al. [144]	2006	2.5 min
Alonso-Montes et al. [136]	2008	0.9 s

Lupascu et al. [140]	2010	2 min
Bankhead et al. [141]	2012	0.2 s

Table 5: Popular vessel detection algorithms and their running times listed by year.

One notable exception is the rapid exploratory algorithm of Can et al. [131] that traces the vasculature, yielding a monotonically improving set of partial results suitable for real-time deployment at 30 Hz. Can et al. [131] 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 center line 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 [131] are noisier and less complete than other, more computationally expensive methods.

5.5.2 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 [145] to find and match unique points between retinal images [146–149]. Vasculature landmark matching algorithms find distinctive points based on vessel networks, such as vein crossings or bifurcations, and match custom descriptors across multiple images [150–152]. Other approaches augment or eschew key points and use the shape of extracted vessels to match vasculature trees [153], [154]. As seen in Table 6, popular approaches are usually slow.

Algorithm	Method	Year	Time
Becker et al. [150]	Vasculature Landmarks	1998	0.2 s
Can et al. [151]	Vasculature Landmarks	2002	*
Stewart et al. [153]	Vasculature Tree + Landmarks	2002	3 s
Chanwimaluang et al. [154]	Vasculature Tree	2006	20 s
Cattin et al. [146]	Keypoint Features	2006	*
Wang et al. [147]	Keypoint Features	2010	2 s
Chen et al. [155]	Keypoint Features	2010	15 s
Broehan et al. [156]	Vasculature Landmarks	2011	0.04 s
Becker et al. [149]	Keypoint Features	2012	0.03 s

Table 6: Existing retinal registration or mosaicking approaches along with the method of registration and time required to process a pair of images. *No times listed in the paper.

Methods that depend on local key point features [146], [147], [149], [155] 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 [150], [152] 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 perform localization and do not build a map of the vasculature. Also, all 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 most similar to Stewart et al. [151] which uses a robust, dual-boot ICP algorithm to register vasculature landmarks (bifurcations and crossovers) and trees (centerlines of the vessels). Stewart et al. achieves very accurate results, but our algorithm is 100X faster and handles occlusions, dynamic lighting conditions, and interference with the instrument tip. Additionally, our algorithm also constructs a retinal vasculature map that is temporally consistent.

5.5.3 SIMULTANEOUS LOCALIZATION AND MAPPING (SLAM)

The problem this work 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 [157]. 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 [157]. Recent particle filter approaches such as FastSLAM are faster and more robust [158]. 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 [132].

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 over-head sensors and assume a reasonably good robotic motion model, both of which are poor assumptions in the problem of retinal localization and mapping.

5.5.4 INTRODUCTION OF EYESLAM AND PROBLEM DEFINITION

Our goal is to design an algorithm that maps and localizes retinal vessels by merging retinal vessel detection and retinal image registration and taking advantage of temporal information as seen in SLAM approaches. A fusion of methods is needed: fast retinal vessel detection algorithms are noisy, incomplete, and do not handle occlusions [131]; retinal image registration methods that do build vasculature maps are orders of magnitude slower than required for real-time robotic guidance [153], [154]; and SLAM algorithms are not designed or tuned for application in intraocular surgeries. Fig. 48 shows a new algorithm called EyeSLAM that incorporates aspects of [131], [151], [157] to perform 30 Hz vasculature mapping and localization of retinal video using rapidly-detected vessels as features, an occupancy grid for mapping, and iterative closest point (ICP) for localization to robustly handle noise, tool occlusions, and variable illumination.

We define our problem as follows: 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} vasculature point as a 2D location $M_t^i = [x_t^i, y_t^i]$. Because typical retinal surgeries have high magnification, we approximate the global map as a plane and can then define the location of the camera viewpoint at time t as a 2D translation and rotation $L_t = [x_t, y_t, \theta_t]$ from the initial position at $t = 0$. Observations of vessels in the camera at time t are denoted by Z_t . The remainder of this section describes the EyeSLAM algorithm that takes these video inputs and returns these mapping and localization outputs.

5.5.5 FEATURE EXTRACTION

Finding features for matching at high magnification is difficult. Traditional key point detectors such as SIFT [145] or SURF [159] fail to find distinctive points on the textureless retina. Likewise, high magnification and a small field of view means vasculature landmarks such as crossovers or bifurcations may not be present in sufficient numbers to function as good features. We instead use many anonymous feature points extracted from vessel detection algorithms. These features cannot be matched individually with a local feature descriptor, but can instead be matched as a group based on the shape of the vasculature network of vessels. We use the highly-efficient, but noisy vessel tracing algorithm of Can et al. [131] to detect vessels (see Fig. 49).

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 bloom proximity test. The color test rejects pixels

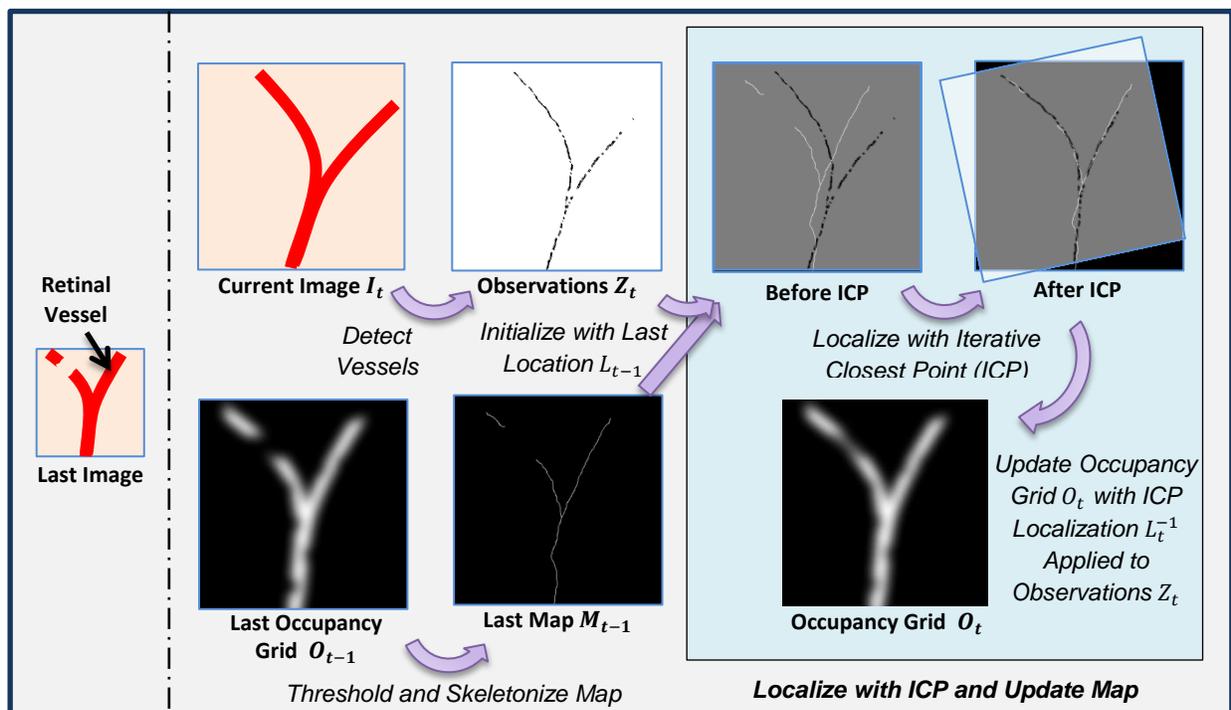


Fig. 48. Block diagram detailing the steps in our proposed algorithm EyeSLAM that maps and localizes retinal vessels during intraocular surgery. Retinal vessels are detected in each frame, localized to a skeletonized map of the occupancy grid with iterative closest point (ICP), and the map probabilities are updated.

that are too dark or insufficiently red, while the bloom proximity test rejects vessels 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.

5.5.6 MAPPING

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. 49). Each pixel in the occupancy grid represents the probability that a vessel occupies that particular spatial location. At each time instance t , the current observations Z_t are transformed to the map with the best estimate of the location L_t^{-1} and used to update the probabilities in the occupancy map by adding a Gaussian around each detected vessel point Z_t^i , as each observation increases the probability that a vessel exists at the detected point. The occupancy grid has a maximum value to prevent unbounded evidence from accumulating. A global decay function decreases the probability of all grid cells, allowing vessels that have not been detected to vanish after some time.

The formulation of the occupancy map reasons about uncertainty over time, smoothing out noise

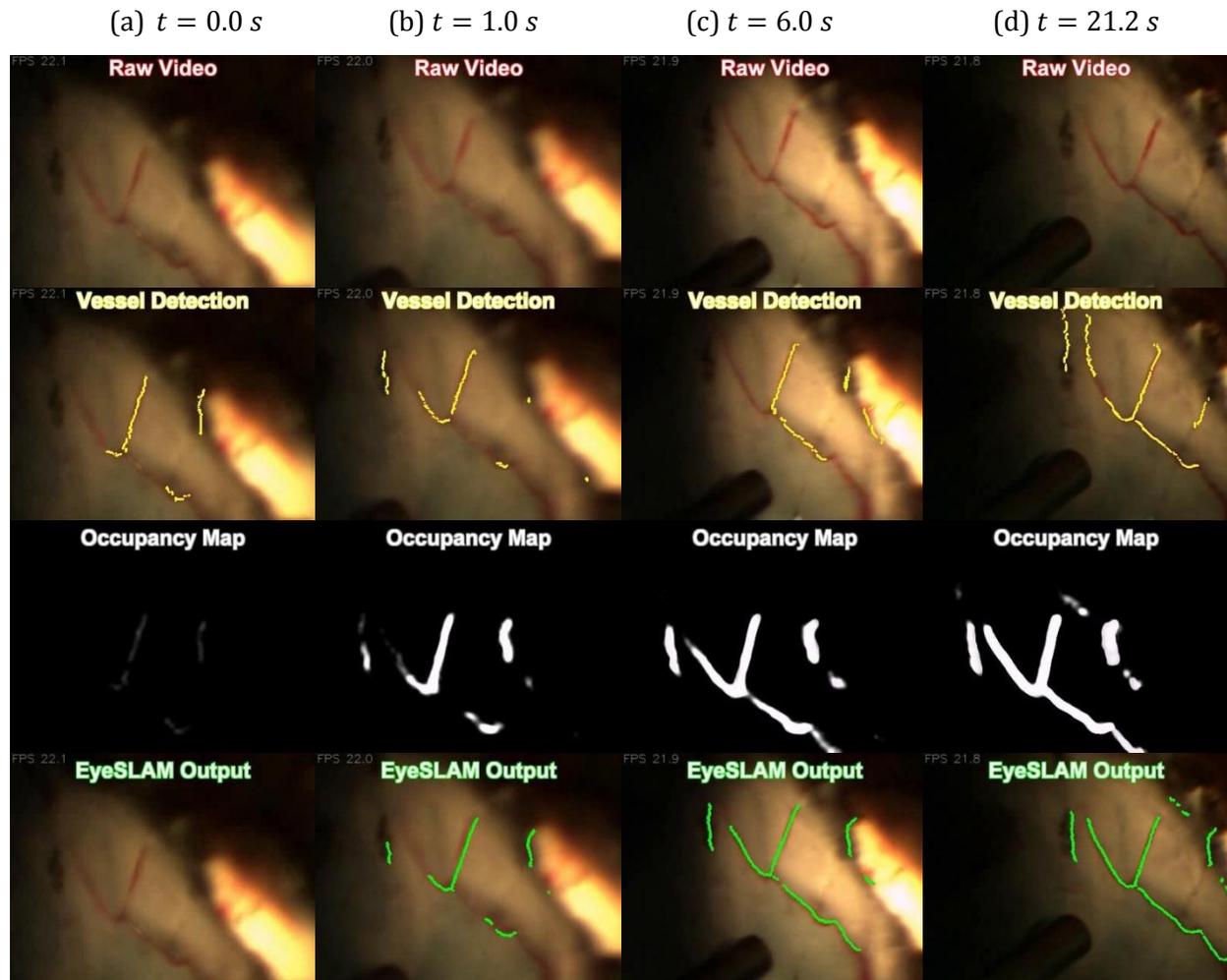


Fig. 49. Snapshots of the feature extraction, mapping, and localization of the EyeSLAM algorithm at various times during 21 s video of porcine retina of an attempted vessel cannulation.

and handling temporary occlusions. A final map containing the centerlines of the most probable vessels is constructed by skeletonizing the occupancy grid, which is approximated by thresholding, computing the distance transform, calculating the Laplacian, and thresholding again. This yields a map of 2D points M_t of the most confident vessels in the occupancy grid (see Fig. 49).

5.5.7 LOCALIZATION

To localize eye-ball motion (which is mathematically identical 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 map M_t , both of which are represented by an un-ordered, anonymous set of 2D points. Iterative closest point (ICP) is used to find point correspondences and calculate the transformation. To guarantee real-time performance, Z_t and M_t are randomly sampled to have a maximum of 500 points. To prevent spurious detections from causing large mismatches and adversely affecting the solution, candidate correspondences are only used to estimate the transform if their distance is under some threshold. Horn’s quaternion-based method [49] is used to estimate the rigid transform instead of an affine or similarity transform because scale and shear are negligible. Incomplete vessel detections at each frame are noisy, so the final estimation of the localization is smoothed using a constant-velocity Kalman filter, yielding the localization L_t . The occupancy map is then updated with the newly registered vessel points Z_t to close the loop on the algorithm.

5.5.8 SETUP & TIMING PERFORMANCE

We have evaluated the proposed algorithm on a variety of videos of recorded in vitro eye phantom, ex vivo porcine retina, and in vivo human procedures. For ease of robotic testing in our lab, color video recorded of an eyeball phantom or porcine retina is captured at 30 Hz with a resolution of 800x608 at a variety of high magnifications (10-25X). Each frame is converted to grayscale by

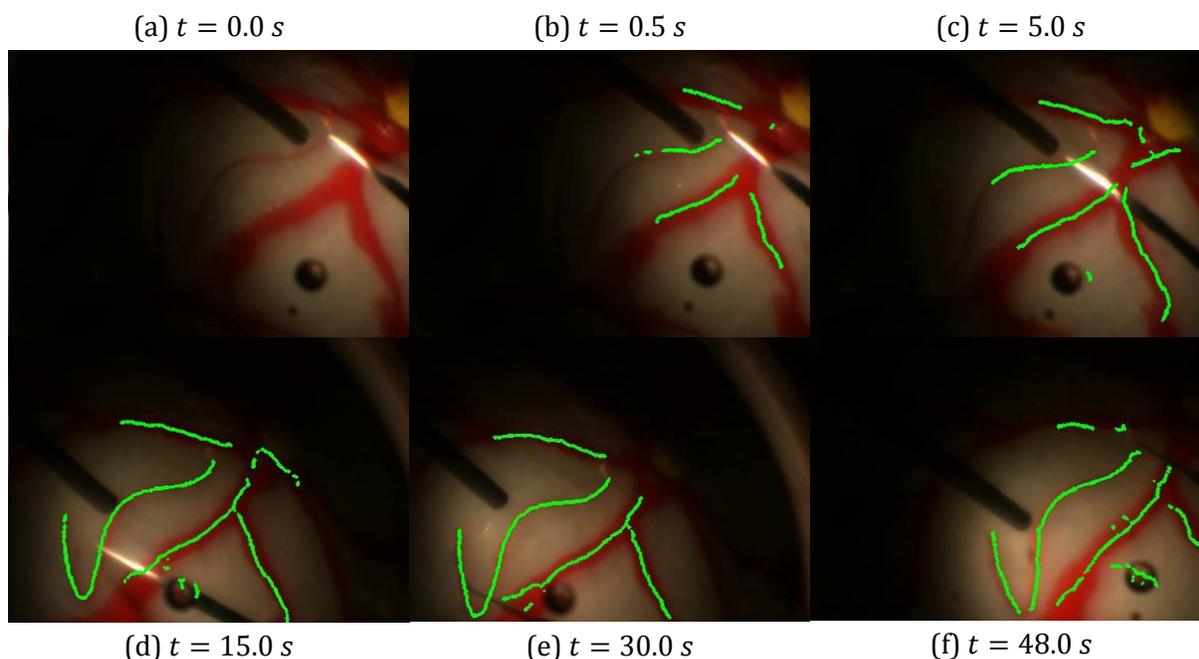


Fig. 50. Output of the EyeSLAM algorithm at various points during a 48 s video of the Johns Hopkins eye phantom.

selecting the green channel, a common practice in many vessel detection algorithms. For efficiency, the image is scaled to half-size. On a fast, modern computer (Intel i7-2600K), our algorithm implemented in C++ with OpenCV [108] runs at 30-40 Hz, including all vessel feature detection, occupancy grid mapping, and ICP localization run in a single thread. Fig. 47 shows the proposed algorithm output on an *in vivo* human procedure.

5.5.9 EVALUATION: CONVERGENCE, INTERMEDIATE RESULTS, AND COMPARISONS

Initialization is fast, requiring less than a second to start building the map and only a few seconds to build a full map. Fig. 50 shows the output of the algorithm running on a 48 s clip of a surgeon tracing a vein in an eyeball phantom. Within half a second, the map for visible areas is initialized. Because the light-pipe only illuminates portions of the eye at a time, the map is built as new vessels become visible. The proposed algorithm is able to handle the occlusion of the vessels by the instrument shaft and light-pipe. Notice the algorithm has been able to accurately map and correctly localize even though the entire view has moved, rotated, and been occluded by the tool under variable illumination.

Fig. 49 shows the intermediate steps of the algorithm on *ex vivo* porcine retina in an eyeball phantom filled with saline. The view is through a vitrectomy lens and illumination is provided by the surgeon solely through a light-pipe. Incomplete vessel detections are merged over time into the occupancy grid to form a full and accurate map after a few seconds. Fig. 49(d) shows that the localization is still maintained even after movement of the eyeball by the surgeon.

Finally, to demonstrate why current vessel tracing algorithms are insufficient for real-time surgical environments, Fig. 54 compares EyeSLAM to three existing vessel tracing algorithms on video of porcine retina taken in the saline-filled eyeball phantom with a vitrectomy lens. It only takes 1.0 s for EyeSLAM to converge, and it provides a more complete and robust mapping and localization than any of the other methods. In particular, notice in Fig. 54c, EyeSLAM correctly ignores the micropipette and specularities and is more robust to temporary illumination changes. Fig. 51 also compares vessel tracing algorithms on video of human retina *in vivo* taken in during a membrane peel. EyeSLAM provides a more



Fig. 51. Comparison of our approach (30 Hz) to the vessel detection algorithms of Can et al. [131] (80 Hz), Sofka et al. [138] (0.3 Hz), and Bankhead et al. [141] (8 Hz) on a 20 s video of human retina during an *in vivo* membrane peel. Video source: http://youtu.be/_naooJFuxPI

complete output than any of the other methods. In particular, notice from Fig. 51(a) to (b), our algorithm has learned to ignore the tool shaft over time, exhibiting more robustness to occlusion and illumination. Overall, the temporal fusion of the proposed approach increases coverage and consistency.

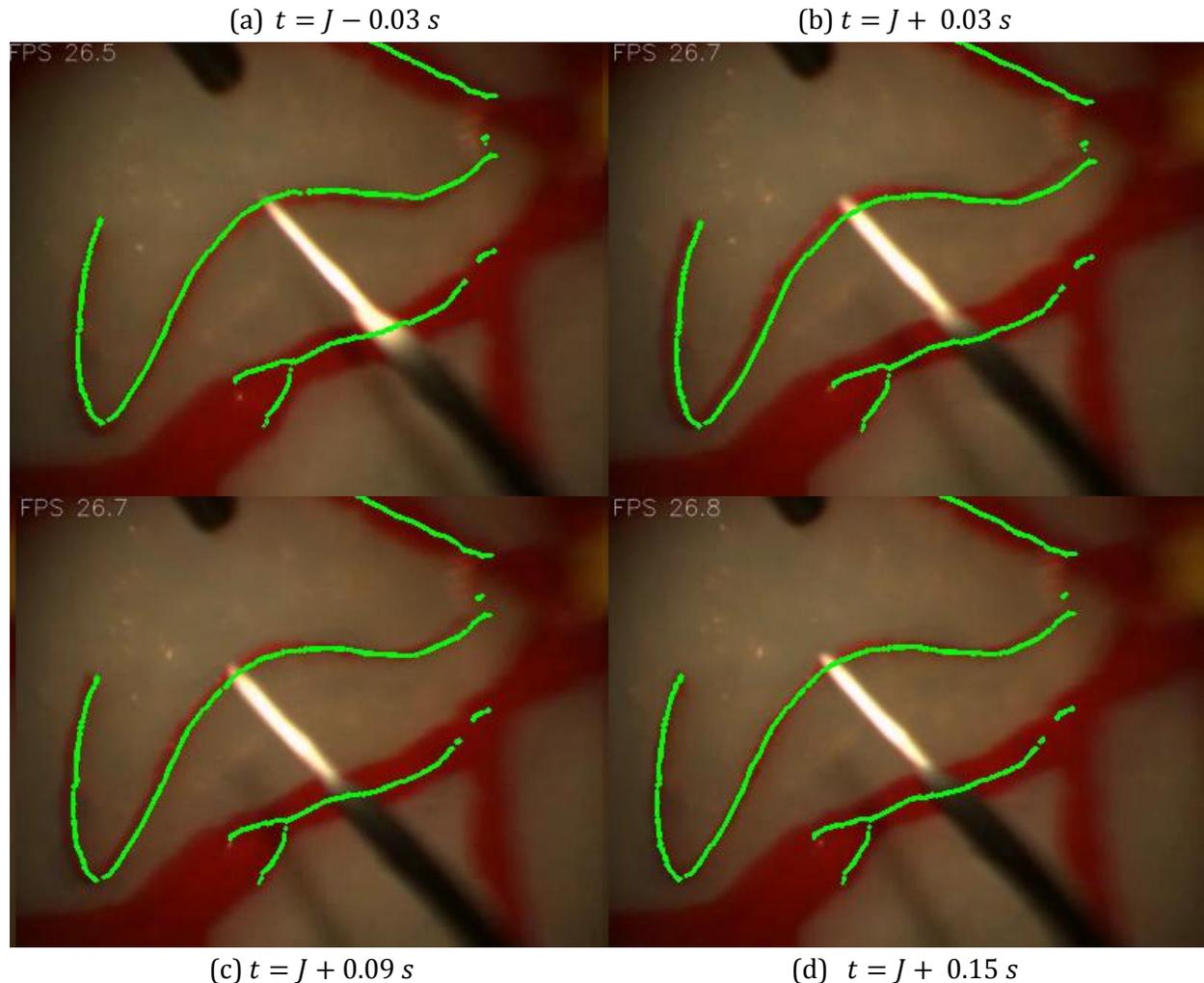


Fig. 52. Example of a small, sudden jerk at time J and the resulting behavior. The time required for re-convergence is about 150 ms.

Because of the noise in vessel detection algorithms, the Kalman filter can be tuned to decrease jitter at the expense of introducing a low-pass, laggy feel to the localization. Because rapid movement is not expected and smoothness of localization is usually desired, we can eliminate most jitter by placing a high uncertainty on the measurements inputted into the Kalman filter. However, if a fast jerk is experienced, the map may skip and jump instead of track smoothly, causing existing vessels to build up in a new location of the map while the old locations decay out. In this case, failure is not catastrophic and will re-converge; to speed up convergence, EyeSLAM can be reset and the map will re-initialize in a few seconds. Fig. 52 demonstrates a mild failure case of EyeSLAM to correctly deal with a small, rapid jump in the eye; it takes 150 ms (5 frames) for the localization to fully converge on the jump.

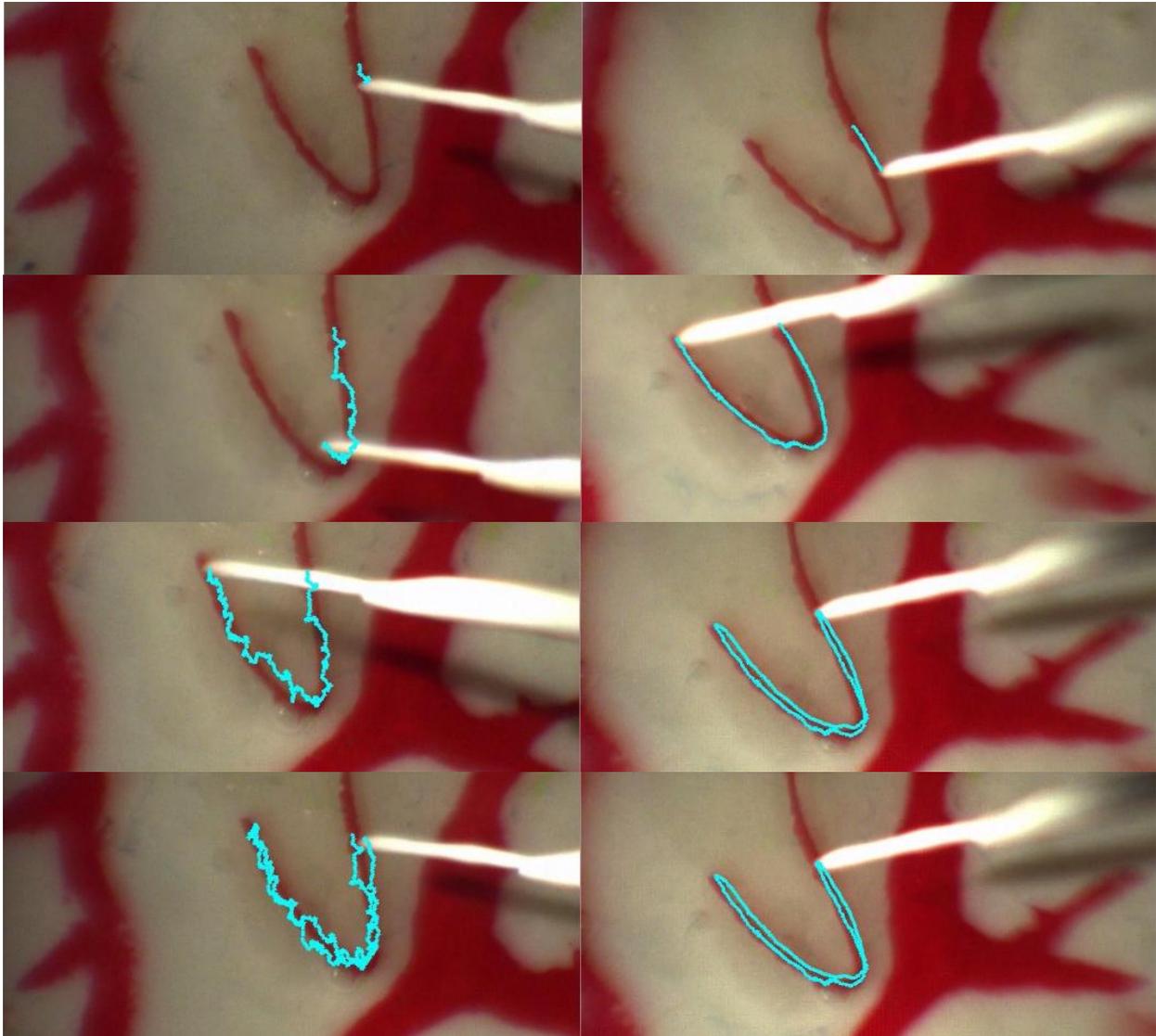


Fig. 53. Tracing a retinal vessel in an eyeball phantom unaided (left) and with the aid a robotic micromanipulator enforcing virtual fixtures based off of our vessel mapping and localization algorithm. Note in both cases, the entire eyeball is moving.

5.5.10 EVALUATION OF EYESLAM IN VEIN TRACING TASKS

To test the usefulness of the algorithm in intraocular surgical environments, we simulated vitreoretinal surgery in a rubber eyeball phantom filled with saline and fitted with a vitrectomy lens. A light-pipe provided illumination and the Micron robotic micromanipulator was used to provide vein tracing assistance from the output of the EyeSLAM algorithm. Fig. 53 shows unaided without robotic aid and unaided with Micron enforcing virtual fixtures using EyeSLAM. Even with significant movement of the eye phantom, EyeSLAM is able to track the vein in real-time and Micron is able to enforce a much smoother tracing of the vessel.

5.5.11 DISCUSSION

We have presented a new algorithm for retinal mapping and localization that operates in real time at 30 Hz. Designed to handle the dynamic environment of high-magnification, variable illumination

retinal surgery, our approach converges quickly and is robust to occlusion. In comparisons on retinal video, it has proven to be an effective method to temporally smooth vessel detections and build a comprehensive map of the vasculature. Shortcomings to the algorithm include some lag when smoothing jitter with the Kalman filter and loss of tracking in the case of large, sudden movements. Future improvements should include more accurate vessel detection algorithms and implementing robust ICP versions that do not fail when only a small portion of the vessels are detected. More effective handling of uncertainty during localization, similar to SLAM approaches, would also be beneficial.

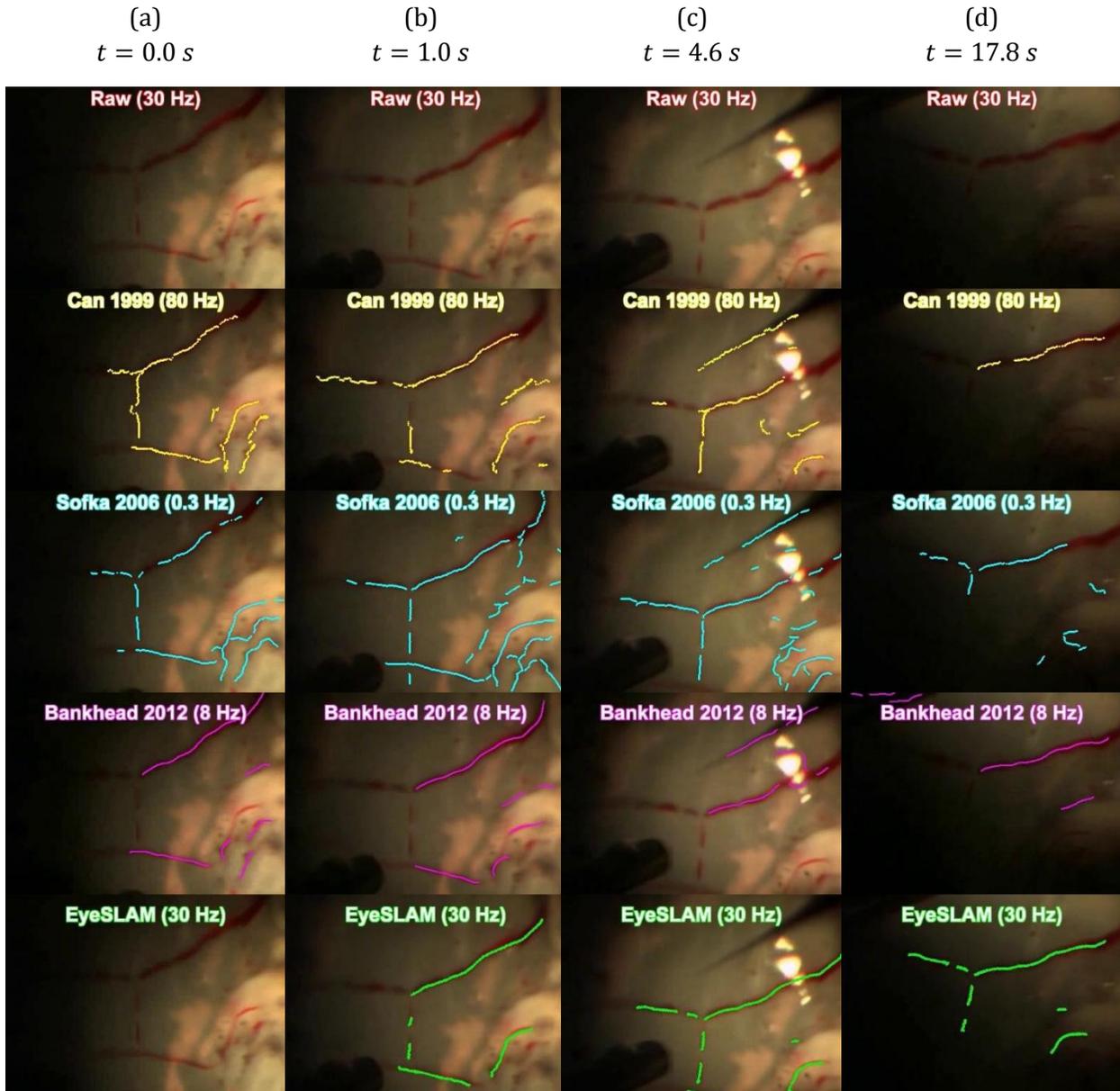


Fig. 54. Comparison of EyeSLAM to the vessel detection algorithms of Can et al. [131], Sofka et al. [138], and Bankhead et al. [141] on a 18 s video of porcine retina in the Johns Hopkins eye phantom.

SURGICAL APPLICATIONS



**"No, Meester Bond...
I expect it to cure your astigmatism"**

In the theoretical frameworks section of this thesis, we proposed position-based virtual fixtures for handheld micromanipulators, optimized the low-level control system of Micron with feedforward tremor predictions, and developed vision systems for analysis of the surgeon and the retina to generate fixtures for individual controllers. Having verified each theoretical component individually in synthetic tests, we now work to integrate and apply the developed theory in practical scenarios. We holistically evaluate the overall vision-based control of Micron in more realistic preclinical trials of retinal procedures; specifically, we target three operations in vitrectomy surgery:

- **Membrane peeling** in a synthetic phantom of plastic wrap on a rubber or sorbothane slides.
- **Laser photocoagulation** on planar paper slides and flat sections of excised porcine retina in open-sky environment
- **Vessel cannulation** in excised porcine retina in a realistic eye-model that incorporates lens effects, sclerotomy constraints, and inner eyeball fluid dynamics.

6 MEMBRANE PEELING

Membrane peeling is a common retinal procedure that involves three steps: engaging, lifting, and delaminating. During the engaging stage, the surgeon slips a bent blade or hook just under the membrane. The membrane is then lifted slowly to prevent tears. Often this involves slight up and down movements while advancing the tool to tunnel under the membrane. Once an edge of the membrane has been partially lifted, the surgeon carefully delaminates, or peels, the

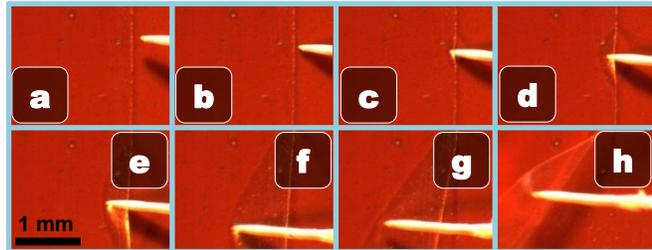


Fig. 55. The membrane peeling procedure shown with plastic wrap on a rubber slide using the white-painted nitinol tip: (a) Approach (b-c) Engagement (d-e) Lifting (f-g) Delamination (h) Peeled membrane.

membrane from the retina. See Fig. 55 for an example of peeling on a phantom. High forces and thus high velocities should be avoided to reduce trauma to the tissue. As general guidelines, Gupta et al. measured forces during vitreoretinal procedures to be under 7.5 mN [160] while Balicki et al. report typical tip velocities during membrane peels to be 100–500 $\mu\text{m/s}$ [24].

Various micromanipulators and robotic aids exist to help surgeons during vitreoretinal microsurgery, such as the Japan Eye Robot [19], Carnegie Mellon’s University’s Micron [27], and Johns Hopkins University’s SteadyHand [23]. Designed specifically for retinal surgery, SteadyHand is a cooperatively controlled robot in which control is shared between the surgeon and a stiff, non-backdrivable robot arm. Recently, Iordachita et al. [38] developed a high-resolution, tiny force sensor designed to fit on the end-effector of a vitreoretinal tool and demonstrated the ability to monitor forces during membrane peeling of a chicken embryo. Using force feedback from [38] and scaling techniques, Balicki et al. demonstrated membrane peeling in a phantom with the SteadyHand, successfully limiting forces applied to the membrane [24]. However, the focus was on the delaminating step and did not include engaging the membrane; furthermore, large forces of more than 2 N were needed to control the robot arm, which caused fatigue and decreased precision. In a survey on tactile feedback, 63 vitreoretinal surgeons ranked visual feedback as much more important than tactile feedback during surgery [161]. In fact, over half of surveyed surgeons indicated that tactile feedback during ILM peeling was not important (30.2%) or not very important (22.2%), suggesting force-feedback may not be needed for good performance. However, it is important to note that robotic approaches for using or augmenting force feedback [24], [35], [37], [38] may be beneficial to surgeons as well.

We develop microsurgical aids for the surgeon during membrane peeling operations using visual feedback rather than force feedback. For instance, optical sensors enable tremor suppression to filter out unintentional and involuntary hand motion, thereby avoiding fast, jerky movements that might tear the membrane or retina. We design vision-based virtual fixtures to provide finer control near the surface using motion scaling. A hard stop virtual fixture just under the surface prevents the application of excessive downward force. Likewise, velocity limiting during the membrane delamination helps maintain a smooth, even force. We report force results in a phantom involving peeling plastic wrap from a rubber and sorbothane slide.

6.1 SYSTEM SETUP

Micron is designed to aid microsurgical tasks under high magnification. In this experiment, a Zeiss OPMI 1 surgical microscope with 29X magnification is used. A 27 gauge hypodermic needle is attached to Micron and equipped with a bent 5 mil nitinol wire to simulate a microvitreal pick. The nitinol wire is tapered and painted white to facilitate tracking by cameras (see Fig. 5). Two PointGrey Flea2 cameras attached to the microscope beamsplitters record the same view the surgeon sees. Capturing 1024x768 resolution images at 30 Hz with a pixel size of 3.5 μm , the cameras are used to reconstruct the surface and instrument tip locations. To measure force, all tests are performed on top of a load cell, which measures Z-force at 2 kHz and has a resolution of 0.3 mN RMS. Fig. 56 shows the system setup.

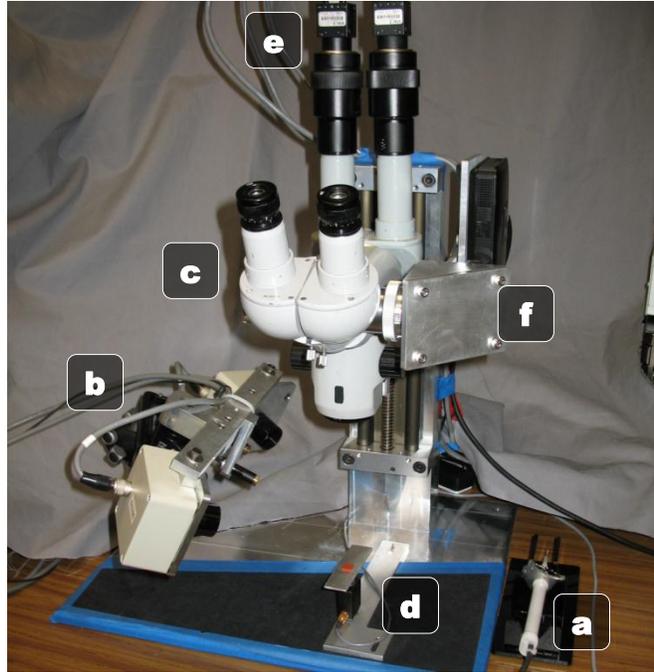


Fig. 56. Micron setup for membrane peeling. (a) Micron (b) ASAP optical sensors (c) Surgical microscope. (d) Red rubber slide mounted on the load cell with plastic wrap for membrane peeling (e) Stereo cameras. (f) Image injection system overlays visual cues in the microscope eyepieces.

6.2 TREMOR SUPPRESSION

Because tremor, the involuntary physiological motion of the human hand, is on the order of 100 μm [11], it is a chief source of difficulty when manipulating at sub-mm scales. Tremor suppression aims at reducing tremor components in hand motion, while allowing voluntary motions. We use two forms of tremor suppression, depending on the stage of the procedure. During the engaging and lifting steps, a first order low-pass Butterworth filter with a corner frequency of 1.5 Hz improves manipulation precision while leaving most of the manipulator range available for the more important depth-limiting virtual fixtures. During the delaminating phase, a velocity limit of 500 $\mu\text{m}/\text{s}$ is enforced to prevent the tip from moving too rapidly and exerting too much upwards force on the membrane.

6.3 MULTI-PART FIXTURES

While tremor suppression aims to improve micromanipulation accuracy by reducing tremor, this will not aid cases where the surgeon presses too deeply into the retina while trying to engage or lift the membrane. Assuming displacement to force proportionality (i.e. $F = -kx$), regulating the Z-depth excursions into the retina is a more viable solution for limiting retinal forces. This key force proportionality observation allows indirect control of force with only vision measurements. We propose a multi-part virtual fixture control system: motion scaling near the retina and a hard stop at a maximum allowable depth.

6.3.1 MOTION SCALING VIRTUAL FIXTURE

During membrane engagement, it is desirable to scale all motion down by a fixed factor to increase manipulation precision. For instance, with a scaling factor of two, a 50 μm hand motion will yield a 25 μm motion at the tip of the instrument. Higher scaling factors afford more precision, but use more of the manipulator range; thus for a good balance, we choose a scaling factor of three. Likewise, to conserve manipulator range and ensure unimpeded motion parallel to the membrane, only motions normal to the surface are scaled. Motion scaling is activated just above the membrane to benefit the engaging and lifting steps.

6.3.2 HARD STOP VIRTUAL FIXTURE

In an effort to prevent excessive downward force, the tip can be stopped altogether from moving too deeply into the surface. This “hard stop” virtual fixture imposes a limit at a pre-defined particular Z-depth. After characterizing the Young’s modulus of typical retinal tissue, the depth limit can be chosen based on the $F = -kx$ relationship for a force that is excessive. The hard stop can be maintained as long as the manipulator range is not exceeded. Because the apparent stop to downward motion breaks the eye-hand coordination loop, it is important to notify the surgeon that the hard stop has been reached. This can be accomplished audibly as in [24] or via some other cue.

6.3.3 VIRTUAL FIXTURE CONTROL

Since the measurements available to Micron are the 6 DOF poses of the tip and handle, all control of Micron is position-based. For instance, tremor suppression works by calculating a filtered goal position from the null (handle) position and performing position control to bring the instrument tip coincident with the calculated goal. Likewise, our formulation of virtual fixtures uses the input null position from the handle motion to drive the tip position in a way that is compatible with the intentions of the user, yet still complies with the virtual fixture. For instance, a hard stop virtual fixture defined by a surface can be enforced by projecting the null position onto the surface, thus finding a goal position that satisfies the virtual fixture while simultaneously being as close to the desired, human-indicated position as possible. Motion scaling around a surface is accomplished by projecting the null position to the surface, then linearly interpolating between the virtual fixture position on the surface and the null position indicated by the handle. Fig. 57 depicts the relationship between fixtures strategies. For details on the implementation of position-based virtual fixtures with Micron, see Chapter 3.

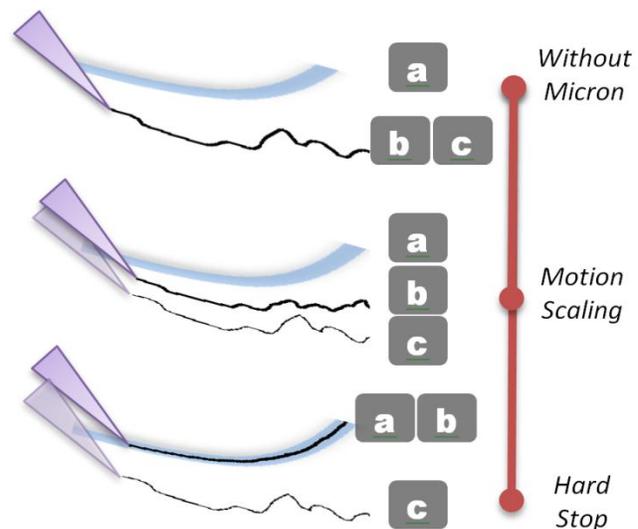


Fig. 57. Illustration of varying levels of position-based virtual fixtures: (a) Virtual fixture (b) Tip position (c) Null position. As the surgeon moves the null position below the virtual fixture boundary, the tip position is influenced by the activated behavior. Without Micron, the tip position follows the null position. With a hard stop, the tip position follows the virtual fixture boundary closest to the input null position. Motion scaling compromises by sharing control between the fixture and human.

6.4 DISPLACEMENT MEASUREMENTS

Central to the virtual fixture control of limiting depth is the distance measurement from the tip of the instrument to the surface. Without good estimation of this displacement, the multi-part virtual fixture control will be less effective at providing helpful aids and thus limiting forces. To sense this displacement, we use dense stereo reconstruction and locally planar surface estimation.

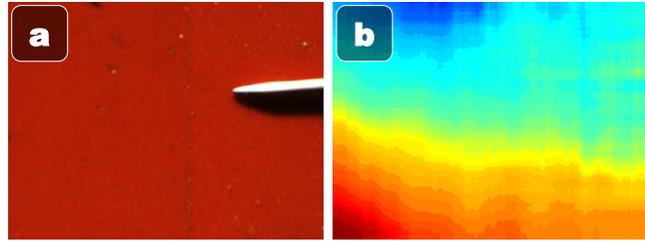


Fig. 58. (a) Left image of the stereo pair (b) Corresponding disparity map. Note the disparity map has been exaggerated to illustrate the overall reconstruction; the maximum height difference in the surface is only a couple hundred microns. Figure best viewed in color.

6.4.1 DENSE STEREO RECONSTRUCTION

Although a variety of sensing modalities exist to measure displacement from a surface, we find an inexpensive, yet effective method is dense stereo reconstruction using the stereo images from the cameras attached to the surgical microscope. While any sufficient discussion of dense stereo is too long for this paper, a review of stereo vision techniques may be found in [162]. For efficiency, we resize the camera images from 1024x768 to 640x480. We then rectify the images with a pre-computed homography obtained from an automatic but offline matching of scale invariant feature transform (SIFT) key points between the left and right images [145].

The Open Computer Vision (OpenCV) library [108] is used for processing images and performing block stereo matching. Because the surface is known to be relatively smooth and continuous with limited local variations, a large block-size is used (experimentally tuned to 151x151 pixels). This also has the advantage of better rejecting camera noise/blur, the tip of the instrument, and local deflections of the surface caused by contact with the tip. Full dense reconstruction with 16 levels of disparity (and 4 fractional bits of precision) runs at 10 Hz. Fig. 58 shows an example of the disparity depth map.

6.4.2 SURFACE ESTIMATION

Although displacement measurements could be made directly from the calculated disparity at the tip location, stereo matching noise, low frame rates, and occlusion by the tip are problematic. To remedy these problems, we first transform the calculated disparity into 3D world coordinates and then locally approximate the surface. To transform the disparity into the 3D world space used by the optical tracking, the stereo cameras are registered to the optical trackers with a 30 s calibration procedure. This initial registration is re-used for each run; nonlinearities and changes in microscope zoom are corrected with an adaptive least-squares calibration (see Section 5.2).

Locally approximating the surface immediately below the tip of the instrument has three advantages: less noise, less lag, and robustness to occlusion. We assume a locally planar surface representation, although higher fidelity models such as quadratic or non-parametric (mesh/B-spline) fits could also be employed. For additional robustness to outliers, noise, and tip occlusion, outlier rejection is performed using 20 iterations of RANSAC; the least-squares surface fit is then calculated from the resulting inliers. With a robust local surface representation transformed into 3D

world space, displacements can be measured at 2 kHz with each new tip position update from the high-bandwidth optical trackers, resulting in high bandwidth control. The precision of the surface estimation is 2.0 μm RMS. Low-pass filters smooth noise in plane estimation. The combined dense stereo reconstruction and surface estimation code runs at 10 Hz on a quad-core Intel Core i7-2600k.

6.5 VELOCITY LIMITING

While the motion scaling and hard stop virtual fixtures aid during the engagement and lifting steps, a different mechanism is needed for the delaminating step when the tip is relatively far away from the surface. During delamination, very low tip speeds are desirable to avoid ripping the membrane. Assuming a velocity to force proportionality, we place limits on the velocity of the tip to avoid excessive upwards force. During delamination, velocities are capped at 100 $\mu\text{m}/\text{s}$. Combined with the shelving filter [27] that acts as relative motion scaling, very smooth, low velocity peeling movements are possible.

6.6 VISUAL CUES

Because Micron has a limited range of motion, providing feedback to the user during operation about how much of the range of motion is being used is beneficial. In fact, when the eye-hand coordination feedback loop is broken, e.g. during velocity limiting or the operation of virtual fixtures, it is necessary to provide some indication of the error so the operator does not drift away from the goal unbounded. For example, when the hard stop is active, the tip is prevented from dipping below a preset Z-depth. However, from the perspective of the user, the tip stops for no apparent reason, leading the user to assume that more downward motion is needed. To prevent this, cueing is required to appropriately signal the operator. Our system displays a visual cue in the microscope eyepiece that turns red when the hard stop is reached, indicating that the tip should be lifted to avoid damaging the retina.

6.7 EXPERIMENT AND RESULTS

All experiments are performed on a phantom for repeatability (Fig 60). The phantom is a red rubber or black sorbothane slide mounted to a load cell that measures vertical force during the entire operation. To simulate the membrane, Glad® ClingWrap (10-12 μm thick) is laid on top of the rubber and smoothed down using a Q-Tip to form a tight seal. The peeling procedure is then to engage the edge of the plastic wrap, lift, and pull it away from the rubber slide (see Fig. 55). Primary evaluation and criteria for success are force measurements.

6.7.1 CLAMPED EXPERIMENTS

Several experiments were conducted with Micron clamped in a vise and attached to a manual micropositioner to verify the efficacy of various subsystems of the proposed system. First, to test the quality of force measurements, we applied a constant force to the load cell and calculated a precision of 0.3 mN RMS. To measure the combined precision of the dense stereo reconstruction, surface estimation, actuation, and force sensing, the tip of Micron was deflected 50 μm below the estimated surface. The RMS precision of the force under this scenario was measured to be 0.5 mN, indicating good system performance in a static environment.

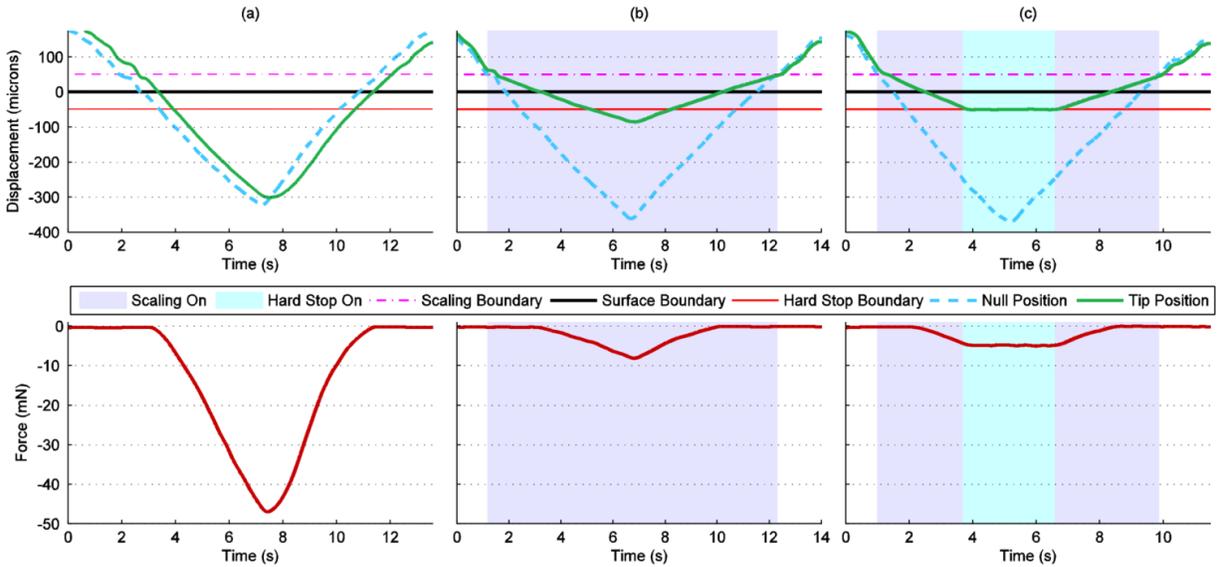


Fig. 59. Displacement above the surface and force measurements during clamped experiments with three different behaviors: (a) Tremor compensation with shelving filter, (b) Simple low-pass tremor compensation + motion scaling below $50\ \mu\text{m}$, and (c) Low-pass tremor compensation + motion scaling below $50\ \mu\text{m}$ + a hard stop at $-50\ \mu\text{m}$. In all cases, Micron was manually lowered and raised $600\ \mu\text{m}$ using a micropositioner over a 10-15 s period.

To demonstrate the behavior of the depth virtual fixtures during the engaging and lifting step, we simulate a down and up maneuver by lowering the micropositioner holding Micron $600\ \mu\text{m}$. The multi-stage virtual fixtures consist of motion scaling that is activated automatically when the tip drops below $50\ \mu\text{m}$ and a hard stop virtual fixture that prevents the tip from moving below $-50\ \mu\text{m}$ (which corresponds to about 3-5 mN of force with our phantom). The tip behavior and force were evaluated under three conditions: shelving tremor suppression, low-pass tremor suppression with motion scaling, and the combination of suppression, scaling, and hard stop. From Fig. 59, we observe that both the motion scaling and hard stop virtual fixtures produce the desired behavior.

6.7.2 HANDHELD

PEELING EXPERIMENTS

Finally to validate the entire system during handheld operation, 20 trials of membrane

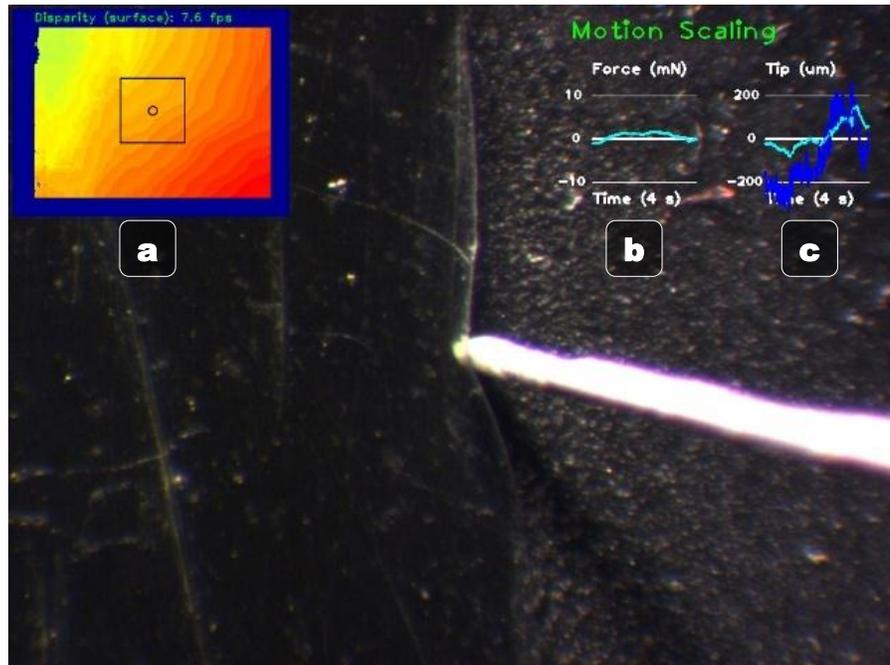


Fig 60. Micron peeling example with sorbothane slide. (a) Stereo disparity map with black square showing RANSAC plane fit area (b) Force graph (c) Tip displacement. Note (a), (b), and (c) are not visible to the surgeon during the procedure.

peelings were performed in the phantom under the surgical microscope by an experienced vitreoretinal surgeon. Two scenarios were tested: unaided without Micron and aided with the described system. As before, motion scaling is activated 50 μm above the surface and a virtual fixture hard stop 50 μm below the surface is enforced. The transition between depth-limiting virtual fixtures to tremor suppression with velocity limiting occurs when the tip is 50 μm or more above the surface.

Throughout the experiment, force data was gathered at 2 kHz and passed through a bidirectional 50 Hz low-pass filter to remove noise. Qualitatively, Fig. 62 shows unaided and aided displacement and force during a typical membrane peel operation on rubber (performed by a novice) while Fig 63 shows unaided and aided displacement and force during a typical membrane peel operation on sorbothane (performed by surgeon). Notice the hard-stop behavior works well to limit downward force.

Quantitatively, the maximum upwards and downwards forces averaged over the 10 runs are shown in Fig 61. The virtual fixtures significantly reduce the downward force in the rubber slide experiments ($p < 0.05$) and significantly reduce both upwards and downwards maximum force in the sorbothane experiments ($p < 0.05$). The sorbothane slide more closely approximates the biomechanics of tissue, and, according to the surgeon performing the task, is more representative of the actual membrane peeling task than the rubber slides. Our approach of using a vision-based hard stop is particularly effective with a 77-83% reduction in maximum downward force.

The mean completion time for the membrane peel on sorbothane by the surgeon was 38.7 ± 11.7 and 35.6 ± 3.9 aided with no statistical difference between the two.

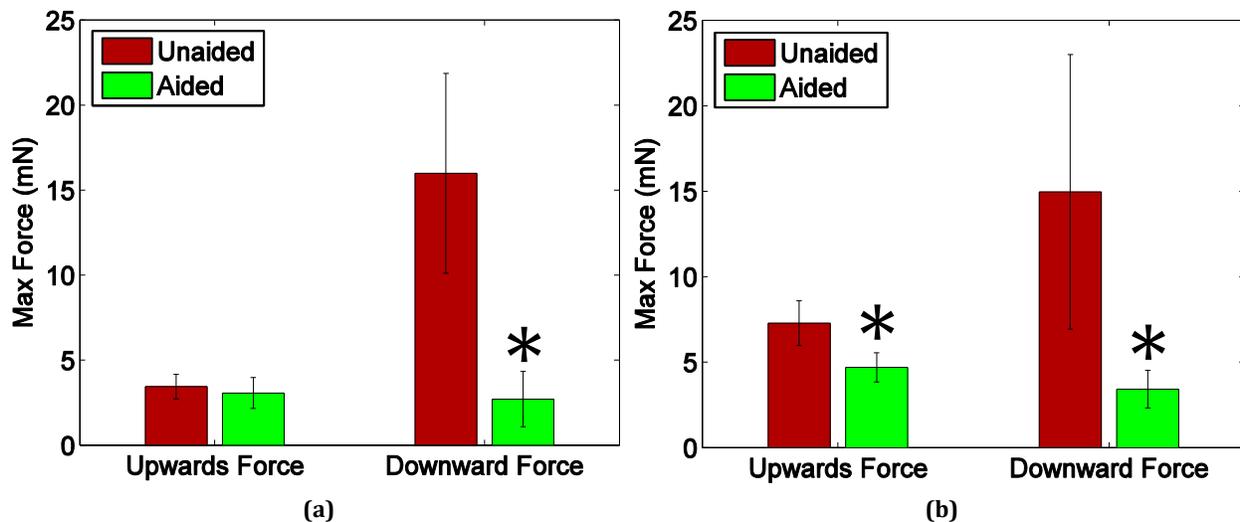
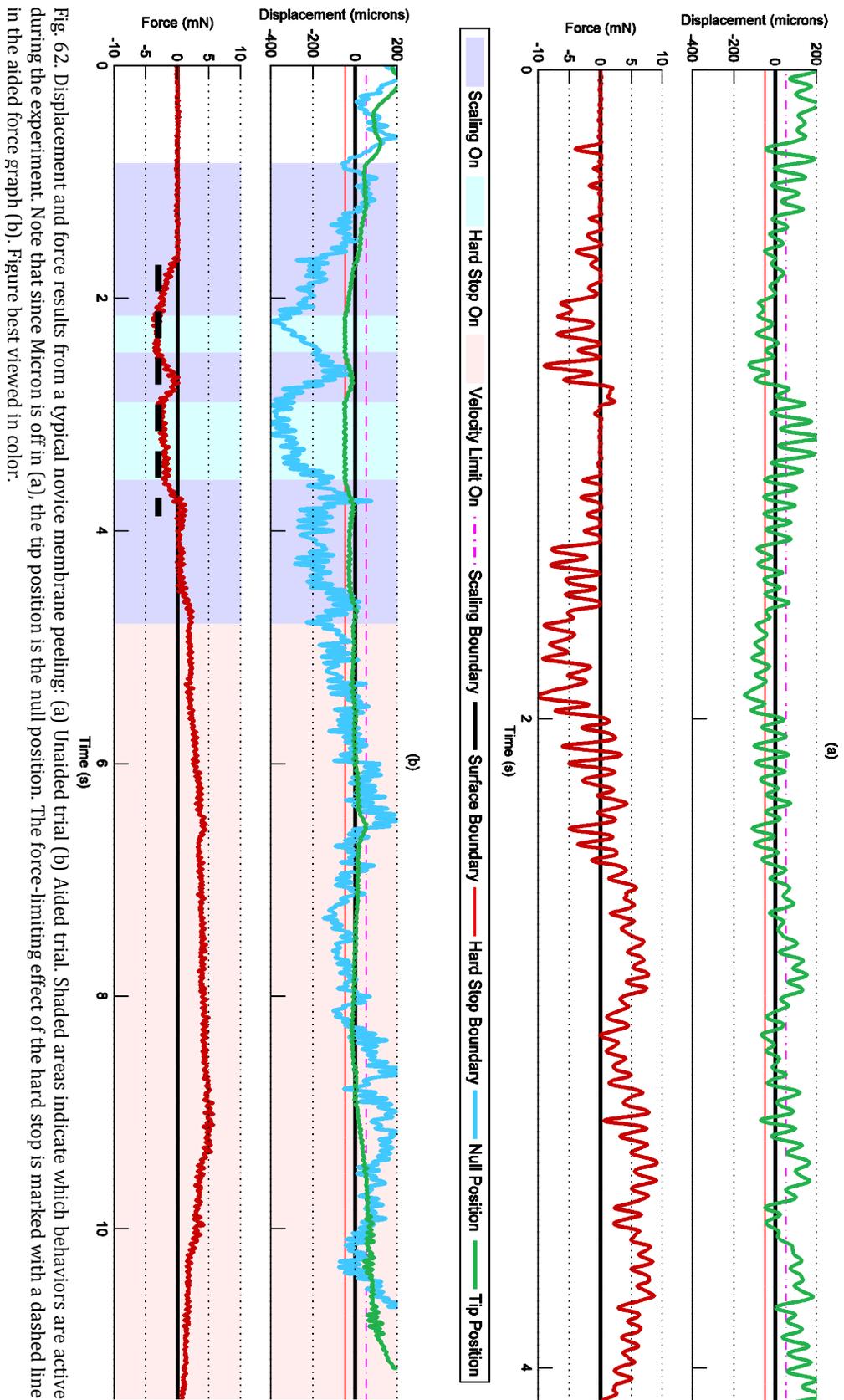


Fig 61. Surgeon results for membrane peeling with maximum upwards and downwards force per trial, averaged over 10 randomized unaided and aided runs. (a) Rubber slide (b) Sorbothane slide. *Significant results ($p < 0.05$).



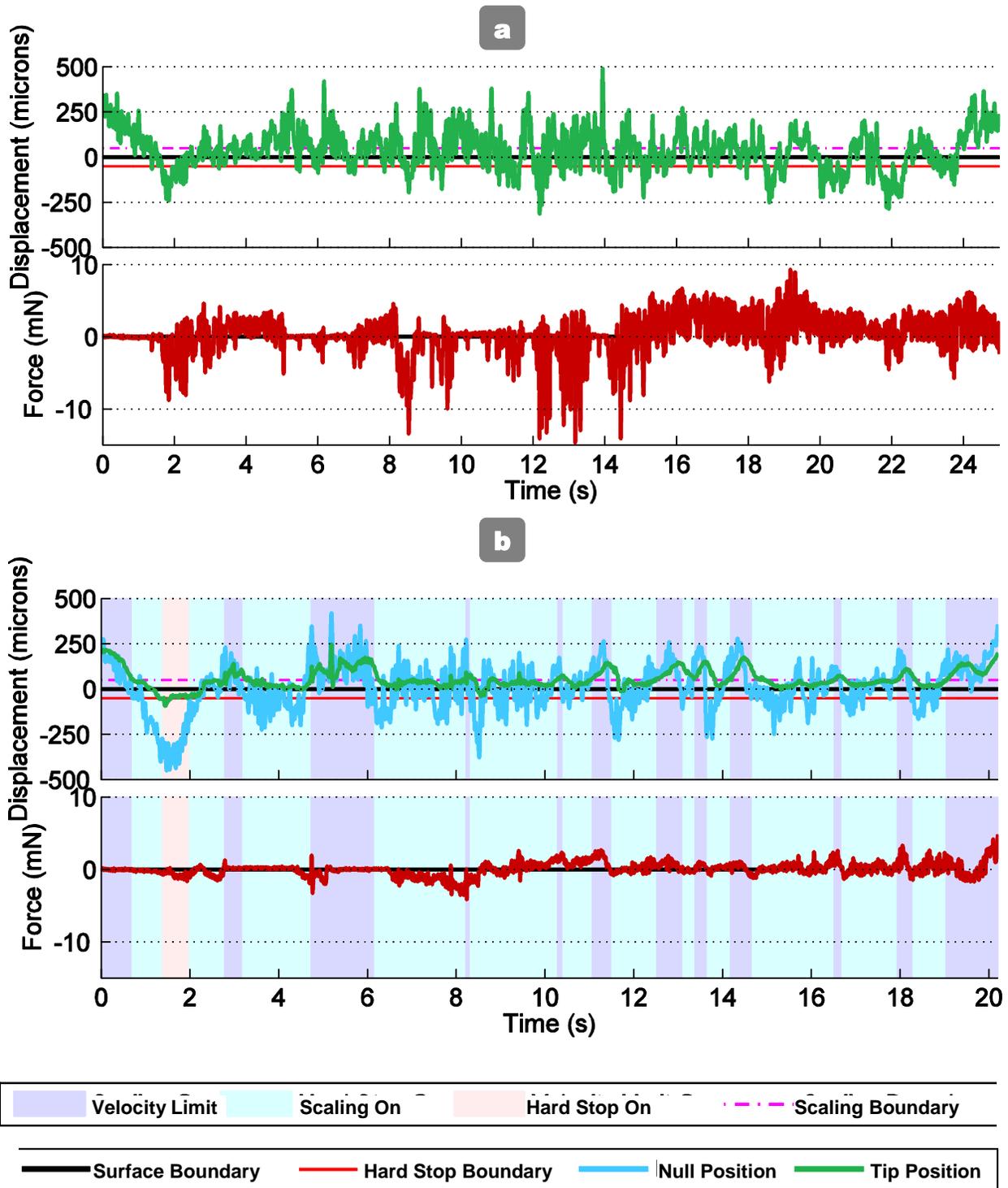


Fig 63. Displacement and force results from a typical aided membrane peel performed on sorbothane by a surgeon. (a) Unaided (b) Aided

6.8 DISCUSSION

We have described how Micron can be applied to retinal membrane peeling and demonstrated its effectiveness in regulating force in a realistic phantom by reducing forces by up to 80%. The novelty of our solution in aiding membrane peeling is three-fold. First, it utilizes a light-weight, fully handheld micromanipulator, avoiding the bulkiness of master/slave or cooperative robotic arms and capitalizing on surgeon's familiarity with handheld tools. Second, we encompasses the entire membrane peeling procedure (engage, lift, and delaminate) on a phantom with similar force characteristics to the retina under a surgical microscope. Third, our approach uses no direct force feedback. Tremor suppression, position-based virtual fixtures derived from visual information, and velocity limiting modes are shown to provide not only smooth micromanipulation, but also sufficient information to regulate forces on the membrane, even in the absence of direct force feedback in the control loop. This is particularly useful because no hardware modifications to existing tool tips or instrument attachments are required.

7 LASER PHOTOCOAGULATION

A recently introduced laser photocoagulation approach by Blumenkranz et al. utilizes a systems of mirrors mounted on a two axis galvanometric scanner attached to a slit lamp which deflects the laser beam in order to apply patterns of up to 50 pulses rapidly at a single command from the foot pedal [66]. Patterns demonstrated by this system include circular arcs, lines, and rectangular and circular grids. The semi-automatic application of laser spots has the potential for significant reduction of treatment time, although the system could not always avoid previous laser burns when applying a new pattern [64]. This approach has been commercialized as the Pascal Photocoagulator [64], [163]. However, it is designed for office use rather than for the operating room.

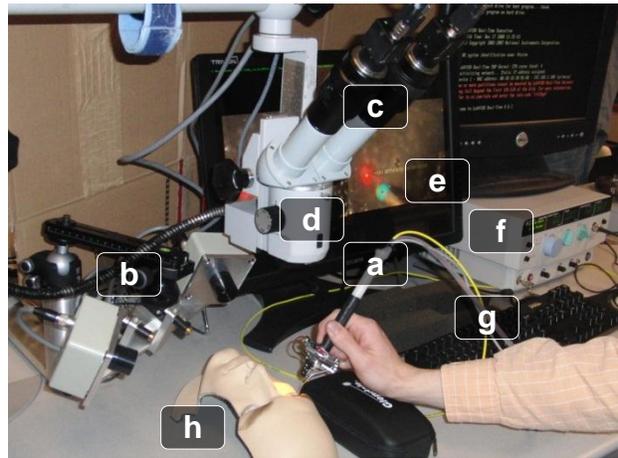


Fig. 64. The setup for laser photocoagulation with Micron, an active handheld micromanipulator. Micron (a) is shown held in the hand of the operator. The position of Micron is measured optically in 3D by the two ASAP PSDs (b) and stereoscopically in 2D via two cameras (c) mounted to the microscope (d). Real-time video with overlaid information is displayed on the 3D computer monitor (e). The laser (f) connects to Micron (a) via fiber optic (g). Targets are placed under the hollowed out eye of the face model (h).

With the goal to extend the benefits of semi-automated systems to the realm of intraocular surgical application, we present the initial phases in the development of an assisted intraocular laser system that will increase the speed and accuracy of the placement of laser burns. It avoids accidental burns in dangerous areas such as the macula and optic nerve and blocks firing of the laser when the distance between the target tissue and the laser is too small or too great. The robotic platform of this system for laser photocoagulation is Micron [27].

7.1 SYSTEM SETUP

An Iridex® Iriderm Diolite 532 Laser is attached to Micron, and the optical fiber from an Iridex® 20-gauge standard straight EndoProbe handpiece runs through the shaft of the instrument to the end-effector. The setup is shown in Fig. 64. When fired, the laser optic setup yields a laser spot size that is 200-400 μm in diameter. A retinal image captured through the microscope at any time before the start of photocoagulation is used as the background upon which the desired pattern can be planned. The operator selects predefined patterns such as rectangular grids, circles, arcs, and ovals to place on the retinal image. A complete set of burn locations can be specified by moving, scaling, and stretching these primitive shapes. The whole pattern of target burn locations consisting of individual targets can be loaded into the control software that runs during the operation. For the remainder of this paper, it is assumed that all 2D positions in the image are measured in units of pixels and are defined in homogenous coordinates.

7.2 TRACKING OF BURN LOCATIONS AND TIP POSITIONS

Registration of the preoperative image to the real-time video aligns the burn locations to the current frame, compensating for any movement of the eye. A number of possible algorithms can be

used [164–166], but a straightforward feature-based approach of aligning interest points between the current frame of the video and the preoperative image works well for feature-rich images. Interest points are detected using Speeded Up Robust Features (SURF) [159] and aligned with a planar homography that is estimated using the standard computer-vision RANSAC algorithm [130]. Thus, assuming an approximately planar scene, registration can be maintained throughout the procedure, keeping burn locations consistent with their initial placement on the preoperative image.

In addition to registering the target burn locations, the vision system also tracks the 2D position of the color-coded tip of Micron in real time in the left and right video views, respectively. A color-based algorithm [125] tracks the centroid of a blue patch painted on the end of the tip of Micron. A red finder beam generated by the laser while in treatment mode provides an aiming guide to the operator between firings of the laser. The aiming beam is important because it shows the location of the burn before the laser is activated, allowing for continuous targeting. The vision system also tracks the centroid of the aiming beam in both left and right views.

Finally, the vision system is responsible for removing noisy images. During the execution of the procedure, video frames captured while the laser is firing are automatically detected and removed to avoid low contrast images and blooming effects caused by the intensity of the laser. The popular Intel® OpenCV library is used to implement the computer vision techniques.

7.3 COORDINATE SYSTEM CALIBRATION

Only the vision system can sense the targets, the aiming beam, and the relative error between them. However, control occurs in the 3D space of ASAP by setting a 3D goal position and using a PID controller to reach that goal. This separation of sensing and control necessitates transformations between the image and the ASAP coordinate systems. This is possible by measuring the only variable that is observable in both coordinate systems and using it to calculate the transformations. The tip position of Micron is measured in three places: the left image p_L , the right image p_R , and in 3D space P . Measurements of p_L , and p_R are obtained from color trackers while P is measured by the PSD optical trackers in ASAP sensing the pulsed LEDs on the shaft of Micron.

The coordinate system transformation takes the form $p_c = M_c P \forall c \in (L, R)$, where P is the 3D homogenous coordinate of a point in space, p_c is the 2D homogenous coordinate of P imaged in camera $c \in (L, R)$, and M_c is a 3×4 camera projective matrix. A preoperative 30 s calibration is used to measure the 3D tip position P and the corresponding 2D tip positions p_L and p_R , which are used to estimate the transformations M_L and M_R using the Direct Linear Transform [129].

Because this estimation is performed with precision in the micrometer range, even very slight shifts or movements in the camera or PSD arrangement can introduce significant errors into the transformations, as can nonlinearities of the optical tracking sensors caused by axial rotations of the tool. To rectify these errors, Micron employs an online recursive least squares approach that minimizes the error of the transformation estimate for each frame during the entire operation, refining and adapting the coordinate system transformations over time for camera c :

$$M_c \leftarrow M_c + \eta(p_c - M_c P)P^T \quad (35)$$

where $(p_c - M_c P)P^T$ is the error of the transformation and η governs how closely the online calibration algorithm adapts the transformation to the measured data. This adaptive calibration is important for maintaining an accurate transformation between coordinate systems throughout the procedure.

Using these transformations, we define two mappings between the coordinate spaces: image projection $\phi(P, M_c) \rightarrow p_c$ and stereo triangulation $\Phi(p_L, p_R, M_L, M_R) \rightarrow P$. The first mapping ϕ is a projection that maps a point P in 3D space, defined by the coordinate system of ASAP, to a 2D point p_c on an image seen by camera c . The second mapping Φ involves solving for the intersection of two rays in space, and triangulates a pair of 2D points p_L and p_R , seen by the left and right cameras, to find the most likely 3D point P that corresponds to the observations $p_L = \phi(P, M_L)$ and $p_R = \phi(P, M_R)$.

7.4 INITIAL CONTROL MECHANISM

One possible approach measures the error directly from the images as the relative distance between the finder beam (where the laser is currently aiming) and the closest aligned target i (where the laser should be aiming), transforming the error into the 3D ASAP coordinate system:

$$e = \Phi(l_L, l_R, M_L, M_R) - \Phi(a_L, a_R, M_L, M_R) \quad (36)$$

Where l_L, l_R and a_L, a_R are the burn target and laser aiming beam tracked locations in the left and right images, respectively. M_L, M_R are the camera calibrated project matrices that enable the mapping Φ to project 2D positions in the left and right images to 3D points in ASAP space. A control signal can be directly generated by using the error as an input to a PID controller. Since both the aiming beam and target are on a locally planar surface, this control signal drives the aiming beam to the target without affecting the distance of the instrument relative to the surface. Once the target location has been acquired by the aiming beam, the laser can be fired. While yielding very accurate results, control signals can only be updated at the camera capture rate of 30 Hz, resulting in slow convergence and laser burn rates less than 0.25 Hz.

Instead of waiting for a new frame from the camera to directly measure the error between the target and the aiming beam, an alternate method is to reconstruct the model of the targets in the 3D ASAP coordinate system and use a model of Micron to update control signals between camera frames. Since the 3D tip position and pose is measured by ASAP at 2 kHz, control signals to point the tip toward the target can be generated much faster, leading to quicker convergence rates and a higher overall burn rate. Since the tip of Micron generally moves more rapidly than the retinal surface, the 3D model of burn targets can be updated at 30 Hz and still remain largely valid in between camera frames. However, as this approach depends on good models of the targets and Micron, accuracy of burns is expected to be lower than using (36).

7.5 SURFACE RECONSTRUCTION

Developing a model of the targets can be accomplished by reconstructing the surface of the retina in 3D. Noting that that burn locations are applied on the surface of the retina, the 3D surface can be

reconstructed using dense stereo algorithms operating on the left and right images with each new set of frames [162]. However, assuming a non-deformable surface, a more relevant approach is to use structured light [167] since Micron is equipped with a red aiming beam laser. Because of the simplicity of the retinal surface being reconstructed (i.e., no discontinuities, no large spectral reflections, no opacity, no occlusions), complex coding of the structured light is unnecessary. By sweeping Micron back and forth above the surface and observing the intersection of the aiming beam with the surface in both images, the 3D structure of the surface S can be calculated by transforming each of the aiming beam points a_L and a_R to a 3D point P_s such that each point P_s belongs to surface S :

$$P_s = \Phi(a_L, a_R, M_L, M_R) \mid P_s \in S \quad (37)$$

While one could use any model for the surface S (plane, quadratic, splines, etc), our system uses a planar representation by fitting the observed 3D points $P_s \in S$ to a plane using a least squares algorithm. The planar assumption works for our experiments in vitro; however, a higher-order model such as a quadratic surface fit might be more suited for testing in vivo, in which the retina adheres to the curved shape of the eye. Only an initial 5-15 s calibration routine is necessary to collect enough correspondences from the structured light to reconstruct the planar surface. If desired, further refinement at time t of the surface S can be calculated iteratively during the procedure to yield a time-varying surface S^t . Once the surface has been reconstructed, the burn locations from the left and right images L_L and L_R are projected onto the reconstructed surface S to get the 3D burn locations:

$$l_{3D} = proj(\Phi(l_L, l_R, M_L, M_R), S) \in L_{3D} \quad \forall l_L \in L_L, l_R \in L_R \quad (38)$$

where, in the planar case, the projection maps the 3D burn location to the plane. The purpose of this projection is to reduce noise in 3D target calculations, since the Z-component of the reconstructed point is subject to the most uncertainty.

7.6 CONTROL SYSTEM

Mimicking the standard surgical procedure of burning target locations in sequence, the controller selects the nearest target within range, deflects the tip to aim at it, fires the laser while locking the aiming solution, and only moves on to the next nearest target location after the current burn is completed. If the next nearest target is not within range or all preoperatively specified burns have been applied, the Micron tip goes to its neutral position. As the operator executes a fly-over maneuver with the instrument, Micron can

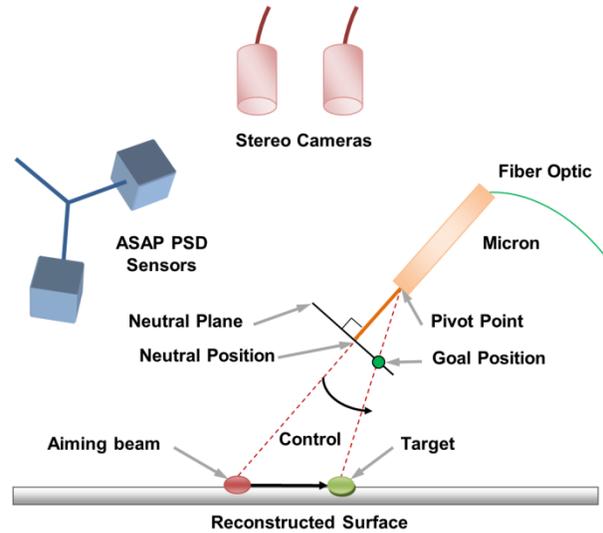


Fig. 65. Micron laser setup. The Micron handle attaches to the laser via a fiber optic cable. Cameras track the finder beam and target, while ASAP optically tracks the tip and pose of Micron. The goal location is determined by intersecting the ray connecting the pivot point and the target with the neutral plane. The piezoelectric actuators then deflect the tip from the neutral position to the goal position and fire the laser to burn the target.

quickly servo as it goes by to burn passing targets. With this semiautomated method, the operator performs the gross motion by sweeping Micron over clusters of targets and lets the control system handle the exact positioning of the tip and timing of the laser activation.

The controller uses the tip position P , instrument rotation R , and targets L_{3D} , all of which are known in the 3D ASAP space. The tip position and instrument rotation are sensed by ASAP at 2 kHz, while the 3D reconstructed targets are updated from the cameras at 30 Hz. This disparity in update rates is acceptable because the targets move slowly, if at all, during the procedure.

Selection of the current target is done by choosing the yet-untreated target that requires the least amount of movement from Micron's neutral position. The neutral position is defined by all actuators being at the zero position, thereby allowing for the greatest movement in any arbitrary direction; functionally, this state is equivalent to Micron in the off state. Since the arrangement of actuators defines a pivot point near the base of the handle, conceptually this optimization of least movement finds the target location that requires the minimum rotation of the tool about the pivot.

Micron uses a PID controller to reach specified 3D goal positions, so a goal position that causes the aiming beam to hit the target must be calculated. This calculation forms the core of the control system and is highly dependent on accurate estimated transformations and target models. While any goal position on the line connecting the pivot position with the target would result in the laser aiming at the target, but the best 3D goal position for Micron is the one that maximizes the available range of motion to account for tremor and the gross sweeping movements of the operator.

Since the workspace of Micron is shaped like a disc centered at the neutral position with its thickness tapering off towards the edges, the maximum transverse motion can be achieved on the

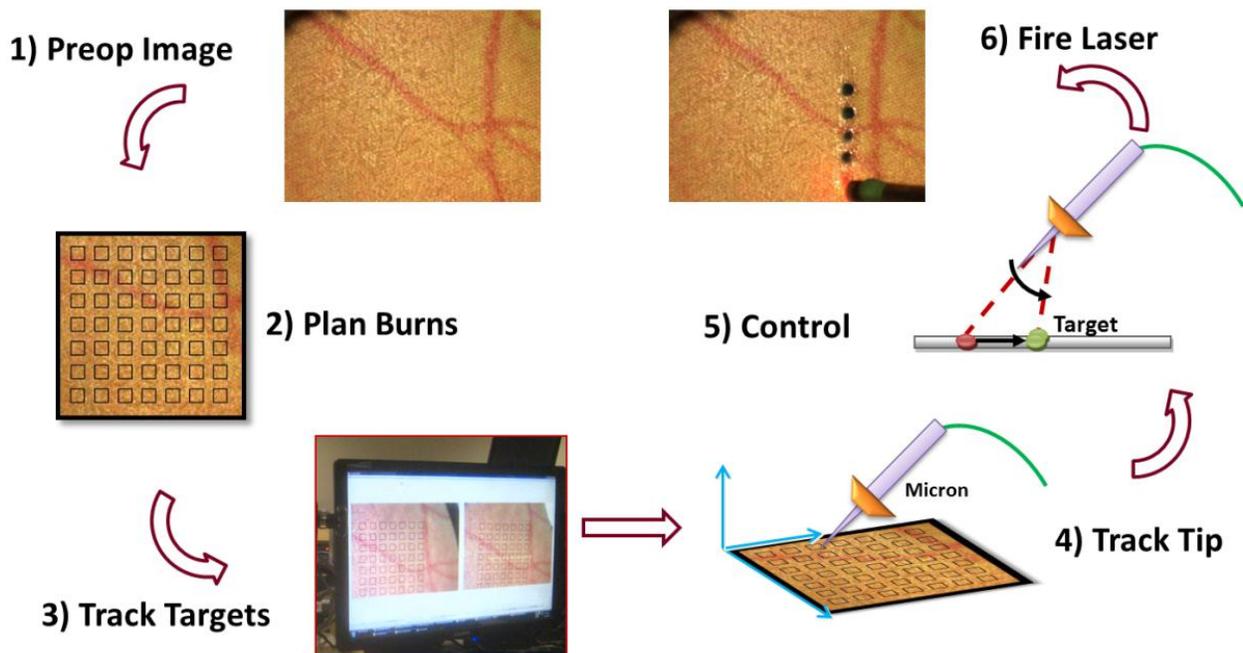


Fig. 66. Graphical depiction of the entire process of the proposed laser photocoagulation system.

plane normal to the shaft of the instrument and intersecting the neutral position. Because the axial range of motion is small and actuation in this direction greatly reduces available transverse movement, the controller leaves the job of depth stabilization to the operator. With this maximal maneuverability constraint, goal positions are restricted to this neutral plane N_p , yielding a single goal position P_G for any combination of tip position P , instrument rotation R , and target location l_{3D} . The 3D goal P_G is calculated at 1 kHz by intersecting the ray connecting the pivot point and the selected 3D target with the neutral plane:

$$P_G = \text{intersect}(N_p, l_{3D} - P_V) \quad (39)$$

The pivot point is defined by the tip position offset by the length of the shaft D and rotated by the pose R of the instrument in space: $P_V = P - R^{-1}D$. Using the PID controller to directly reach the goal point P_G , the system exhibits an underdamped behavior that has a tendency to overshoot the goal. Instead, a 30-ms minimum jerk trajectory is planned to the goal position, which results in much better tracking and significantly less overshoot. Once the goal position is reached, the laser is activated. The goal position is adjusted at 1 kHz to eliminate tremor or other motion until the laser has finished firing, at which point the tip returns to the neutral position and another untreated target is selected. The procedure terminates when all targets have been treated. See Fig. 66 for a graphical representation of the control system.

Small errors in the model of Micron or the calibration can cause large errors during the 3D reconstruction, so an additional 5-15 s calibration procedure is executed after the surface reconstruction. Errors in the calculated intersection of the aiming beam with

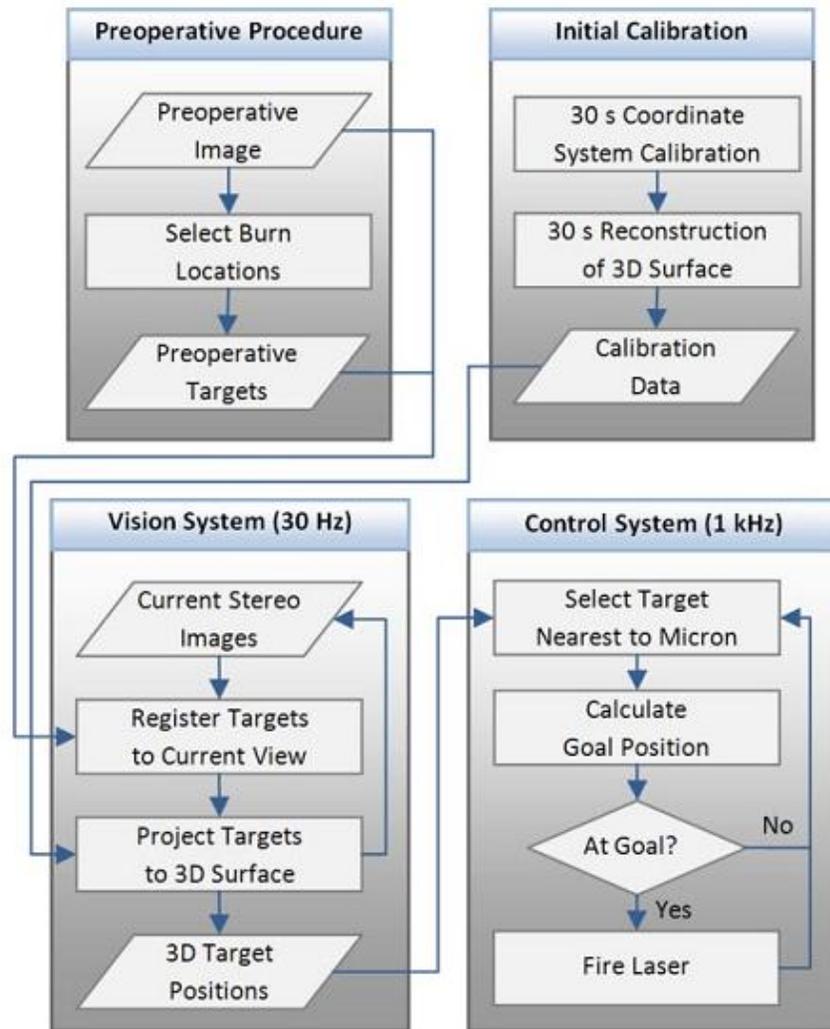


Fig. 67. Block diagram of system, showing data and execution flow. The preoperative procedure is executed first, followed by the initial calibration. Following these two steps, the vision and control systems are then executed and run simultaneously for the duration of the operation.

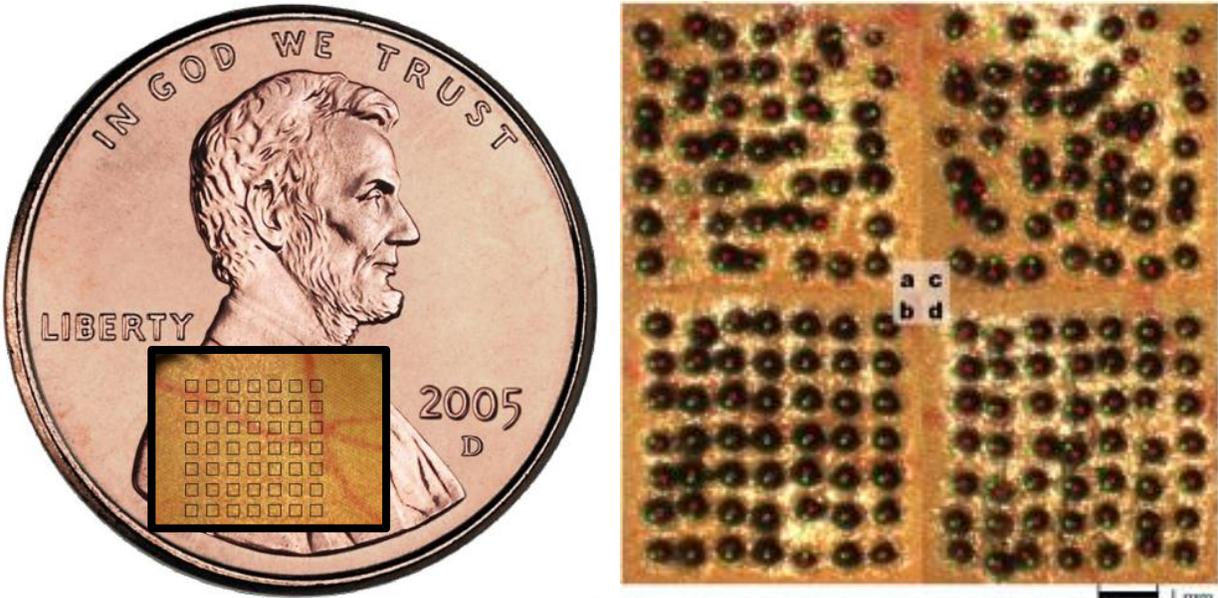


Fig. 68. (Left) Paper slide size in comparison to a US penny. (Right) Results from paper slide trials. Green solid dots indicated preoperatively specified burn locations, and red hollow circles indicate the centroid of the burn. (a) Top left: 0.5 Hz unaided. (b) Bottom left: corresponding aided trial from 0.5 Hz trial. (c) Top right: 1.0 Hz unaided. (d) Bottom right: corresponding aided from 1.0 Hz trial.

the reconstructed surface and the observed location of the aiming beam are recorded during this calibration. The control system then adjusts the surface position with the measured mean error to better align the reconstructed surface with the observations. This is usually then valid for the remainder of the procedure, provided the operator does not significantly rotate Micron.

Several safety measures are designed to prevent firing errors and misplaced burns. The first safety check limits the maximum distance the Micron tip can move when deflecting to aim at a target. A conservative threshold is used to allow enough range of motion to accommodate the user's tremor or gross hand movements during targeting and burning. The second safety check allows activation of the laser only after targeting is complete and if enough reserve manipulator range is available. In the case where Micron is close to the limits of its movement, it aborts and returns to the neutral position to avoid being unable to continue targeting the burn location while the laser is activated. The third check ensures proper Z-distance between the tip and target before firing the laser, because if the laser is too close or too far away, the tissue will burn unevenly or not at all. If at any point Micron decides the target is unreachable, the tip is returned to the neutral position and Micron begins the target selection process again. A block diagram of the entire system can be seen in Fig. 67.

7.7 EXPERIMENTAL PROCEDURE

For ease of initial development and testing, experiments were performed on pictures of a retina printed on a color, high-resolution printer. The yellow/orange tones of the ink on the paper served as a good absorptive material for the laser, yielding distinct, solid black burns. The paper slide was placed under a formed face mask with the eyes hollowed out to give a more realistic operating

environment. The Micron shaft is then inserted through the eye slot during the procedure to place laser burns on the paper. The surgeon's hand rested on a block next to the face mask; Micron was not braced against anything, thus the procedure was performed open-sky.

A 7x7 grid with the preoperative dot locations placed approximately 650 μm apart was selected as a good test pattern. Two different cases were tested during the experiment: unaided and aided. During the unaided case, the laser fires at a fixed frequency when the pedal is depressed. In this scenario, Micron is switched off, and it is the operator's responsibility to move the tip of the instrument to accurately place the burns in time with the laser repeat rate.

During the aided case, Micron is turned on and actively helps the user. The operator is responsible only for the gross movement, while Micron handles the precise targeting and firing of the laser. As with the unaided case, for safety reasons, the laser fires only when the pedal is depressed; however, the laser firing mechanism is under software control instead of being pulsed at a fixed frequency. This allows Micron to operate as fast or as slow as the operator feels comfortable moving the instrument.

During the paper trials, the individual laser pulses were set for a duration of 45 ms and a power of 3.0 W. The laser repeat rate for the unaided case was set to 0.5 Hz in the first set and 1.0 Hz in the second set of experiments. During the 0.5 Hz set, the surgeon was instructed to go slowly and steadily during the aided case. In the 1.0 Hz set, the surgeon was asked to proceed as fast as he felt comfortable during the aided task.

Tests with porcine retina in vitro were performed by the surgeon in a similar fashion to the paper slide procedure. Excised pig eyes were refrigerated until used and dissected immediately before the experiment. The retina was lifted out, placed on black felt, and smoothed out to form an even thickness. The prepared retina was mounted under the microscope (without the face model for simplicity). Preparation and mounting were done immediately before experimentation, to avoid drying of the tissue. The same 7x7 grid with 650 μm spacing was used with both unaided and aided scenarios. Laser pulses were 2.2 W for 27 ms, causing 200-300 μm diameter burns that appear on the retina as milky white spots. Because the surgeon was more familiar with the procedure by this time and the aided cases were averaging near 2.0 Hz for the paper slide trials, we fixed the laser repeat rate at 2.0 Hz for a more equitable mean error comparison during the porcine experiments.

7.8 EFFECTS OF VISUALIZATION AND TOOL ERGONOMICS

Two potentially significant differences exist between the setup used to perform the paper slide and porcine tissue experiments and the setup a surgeon typically uses: the visualization of the workspace and the ergonomics of the tool. We devised two experiments to test each of these factors to determine the potential impact on accuracy of burn placement.

While the computer monitor can display informative overlays, such as the targets in the video stream, the quality is not as high as viewing the workspace directly with the microscope. Specifically, the cameras capture images with a lower dynamic range and at lower resolution than the human eye. Additionally, a 60-ms lag is introduced in order to capture, process, and display the real-time video. Of particular interest is how these factors affect performance. To investigate this

question, we executed an additional test similar to those mentioned earlier with a 7x7 grid, except that the grid targets were printed directly onto the paper slides and the surgeon used the normal microscope eyepieces instead of the cameras and video. Only the unaided trials with an inert Micron could be run under such conditions. The surgeon executed one trial at the 0.5 Hz repeat rate, and one at the 1.0 Hz repeat rate.

The other potential factor we investigated was the ergonomics of the tool. The IRIDEX EndoProbe standard straight 20-gauge handpiece typically used in practice is significantly lighter and thinner than Micron. To determine if the ergonomics of Micron were significantly impacting performance, we ran the same two experiments again at the 0.5 and 1.0 Hz repeat rates with the IRIDEX EndoProbe instead of an inert Micron.

7.9 EXPERIMENTAL PROTOCOL

Experiments were performed by a vitreoretinal surgeon with more than 20 years of experience. For each set of experiments, eight trials were performed sequentially by the surgeon during a one-hour period, alternating between unaided and aided. Roughly one minute of rest was provided between the interleaved unaided and aided trials. Video of each trial was recorded, along with the target locations on the preoperative image. From this information, the mean error and duration for each trial were measured. Error was measured by taking the final frame of the video sequence and hand-marking the centroid of each burn. Individual error was calculated as the distance between the centroid of the burn and the target. Nearest-neighbor matching was used to match burns to targets. To avoid spurious matches, errors greater than the spacing between the target burn locations were discarded and noted as missed targets. Thus mean error was calculated only on the subset of targets where the burn location was in the neighborhood of the target. Statistical significance was determined by calculating p-values assuming a two-tailed test.

7.10 RESULTS

Fig. 68 and Fig. 69 present sample trials of the aided and unaided cases for both paper slide and porcine retina experiments. Mean error and mean duration for all the sets of trials are shown in Fig. 71 and Fig. 72. P-values for the mean error of all trials were less than 0.005 for all three sets of experiments. P-values for the trial durations were less than 0.005 only for the 0.5 and 1.0 Hz sets of experiments. Table 7 lists the mean errors for the three different sets of experiments, along with the overall reduction in mean error between the unaided and aided cases. To investigate the trade-off between speed and error, Fig. 70 presents a scatter plot of mean error vs. duration for each

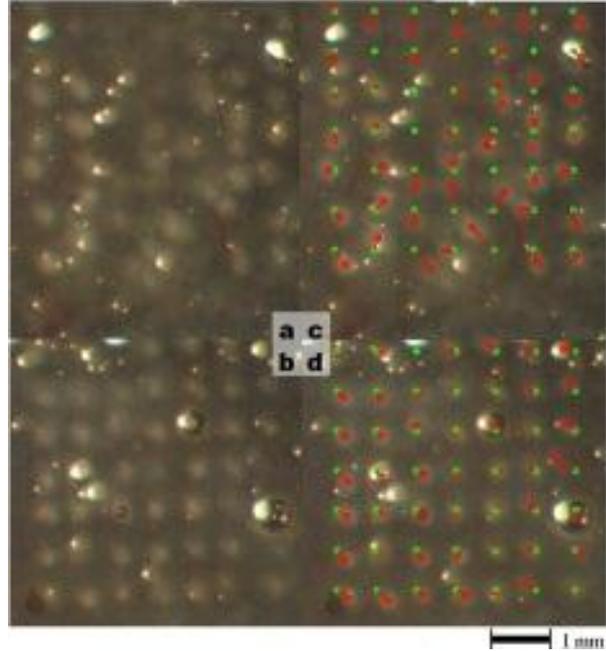


Fig. 69. Results from porcine retina trials. Green solid dots indicated preoperatively specified burn locations, and red hollow circles indicate the centroid of the burn. (a) Top left: 2.0 Hz unaided. (b) Bottom left: corresponding aided trial from 2.0 Hz trial. (c) Top right: 2.0 Hz unaided with errors depicted. (d) Bottom right: corresponding aided trial with errors depicted.

operation. While the surgeon can only perform accurately or speedily in a mutually exclusive way, the aided case with Micron is both simultaneously fast and accurate.

Trial/error (Hz)	Unaided (μm)	Aided (μm)	Reduction (%)
0.5	149	125	16.1
1.0	166	129	22.3
2.0	203	123	39.4

Table 7: Reduction in overall mean error of laser placements between unaided and aided trials.

In the investigation of the effects of different visualization systems, the mean error with the microscope view, when compared with the cameras and computer display view, was 58% lower at the 0.5 Hz repeat rate and 60% lower at the 1.0 Hz repeat rate. When running the same experiments with the normal probe instead of the inert Micron, the surgeon achieved error reductions of 72% and 68% in the 0.5 and 1.0 Hz cases, respectively.

7.11 DISCUSSION

The results presented demonstrate the general feasibility of an active handheld surgical device to decrease spot placement error and procedure duration in intraocular laser retinal photocoagulation. In both the paper slide and porcine retina trials, Micron significantly reduced the positioning error. The mean error for the unaided cases shows an interesting, but expected trend. At lower repeat rates, the surgeon has enough time to move and accurately place laser burns. At higher laser repeat rates, the mean error increases as the surgeon's ability to precisely control the placement of burns is exceeded. The position error for the aided cases is generally consistent across experiments, indicating that Micron can handle precision targeting even with fast gross movements of the instrument. Mean duration for the procedure was consistently shorter for the aided cases. Excepting the 0.5 Hz case, in which the surgeon was instructed to execute the procedure slowly and methodically, mean duration was similar for all aided cases. Although the present demonstration is limited to simplified conditions in vitro, the experimental results show that a surgeon using the current prototype of Micron can perform the photocoagulation procedure with an accuracy improvement of 15% or more, compared to the control case, and up to four times faster, depending on the laser repeat rate selected.

Furthermore, this approach has three attractive side effects. First, because targets are specified a priori and Micron incorporates safety features to prevent firing at non-target locations, accidental extraneous movements do not risk laser burns being applied inadvertently to dangerous areas, which is risk present if the laser is in repeat mode and the surgeon experiences unintentional movement.

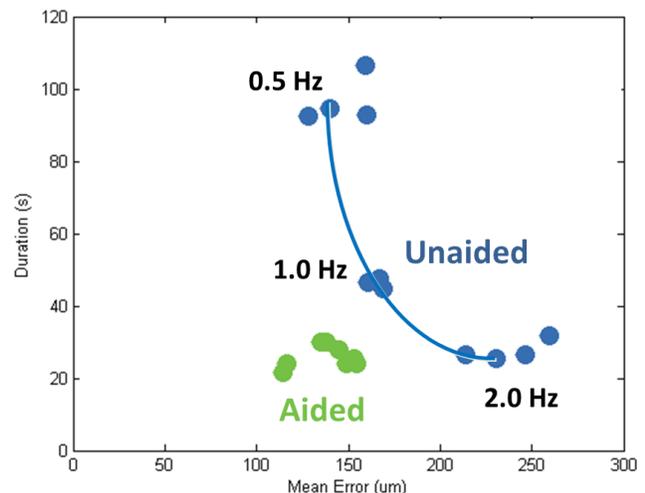


Fig. 70. Mean error vs. duration for all experiments (except the Aided case where the surgeon was asked to perform the operation slowly and deliberately) shows that the surgeon can be either slow and accurate or fast but with lots of error. When the procedure is performed Aided with Micron, the results are both fast and accurate.

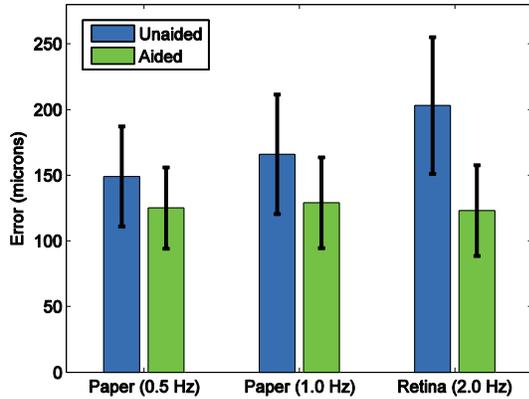


Fig. 71. Mean error bar chart with error bars representing the standard deviation for all three trials with laser repeat set to 0.5, 1.0, 2.0 Hz from left to right. Each scenario consisted of four unaided trials and four aided trials interleaved.

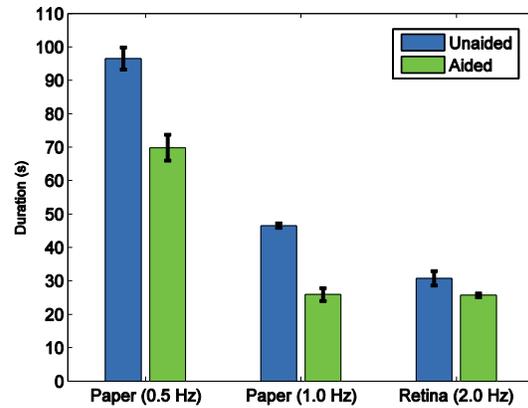


Fig. 72. Mean duration bar chart with error bars representing the standard deviation for all three trials with laser repeat set to 0.5, 1.0, 2.0 Hz from left to right. Each set consisted of four unaided trials and four aided trials interleaved. The apparent discrepancy in the 0.5 Hz case is because the surgeon was instructed to take his time during the aided trials.

Second, the surgeon is relieved of explicitly targeting burn locations and avoiding critical

areas such as the fovea or vasculature. Using our system, the surgeon is responsible for only the gross movements of the instrument while the robotic assistance performs the precise targeting and firing behaviors. This level of automation reduces the surgeon’s cognitive workload, which is generally thought of as a worthwhile goal [168]. Third, Micron’s pose information combined with the 3D reconstruction of the surface enables a safety mechanism that only fires the laser when the distance from tip to retina is within a pre-specified range. This prevents ineffectual burns when the tip is too far away from the surface and “too hot” laser burns when the probe moves too close to the retina for the specified duration and power of the laser. Avoiding burns that are “too hot” has the desirable effect of lowering the risk of unintended choroidal and retinal neovascularization as a late complication of the laser surgery [169].

The present study is a proof of concept; substantial further work is necessary in order to bring the technique to clinical feasibility. The comparisons dealing with visualization and instrument ergonomics show that both have an effect on performance. Viewing the workspace via cameras and computer monitor led to degradation in performance compared to viewing directly through the stereo operating microscope. How much of this degradation could be eliminated by training is not known. In any case, a system that injects graphics directly into the optics of the microscope would likely alleviate the problem. As for ergonomic factors, the 11% decrease in unaided accuracy with Micron when compared with the IRIDEX EndoProbe confirms that the size and weight of the existing prototype, although small enough to be usable, nonetheless measurably degrade performance. To minimize this effect, a smaller and lighter instrument is presently under development. Future research must also include techniques for proper control in the presence of the fulcrum imposed by the entry point at the sclerotomy, and suitable calibration for accurate tracking and control of Micron when viewed through the cornea and lens of the eye. Because pre-operatively selecting burn locations can be a tedious procedure, a semi-automatic method that can suggest burn locations from learned examples of previous surgeons’ patterns on the retina could

further expedite the operation. For extended procedures, the capability to re-image and re-register views of the eye to plan or re-plan new burn patterns intraoperatively also could be added and may be beneficial. Future evaluation in vivo in animal models will undoubtedly clarify the need for additional refinements.

8 VESSEL CANNULATION

Although successful cannulations in a variety of animal and human models have been reported, there are difficulties involved in manipulating such tiny vessels. These reports are not surprising considering that physiological tremor of vitreoretinal surgeons has been measured at over 100 μm in amplitude [11], which is greater than the typical 40-120 μm diameter of retinal vessels. Recognizing the need for reduced tremor and increased precision, [36] describes cooperative control between a surgeon and a Steady-Hand robot for *in vivo* cannulation of 80 μm vessels in the chorioallantoic membrane of chicken embryos [92]. Though the approach did not improve the cannulation success rate, the experiment did show improved ability to maintain the cannula in the vessel. We propose a vessel cannulation system based on a handheld robot, Micron [27], using vision-based control. Preliminary results *ex vivo* and *in vivo*, performed by an experienced retinal surgeon, validate the system.

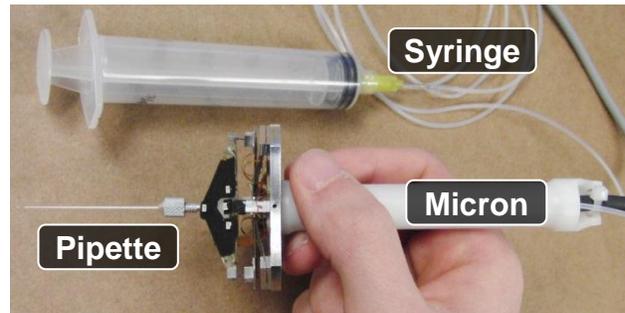


Fig. 73. Micron handheld micromanipulator with glass pipette attached with tubing to a 20 mL syringe.

8.1 SYSTEM SETUP

As shown in Fig. 73 and Fig. 74, the surgeon uses Micron in a 2x3 mm workspace under a Zeiss OPMI® 1 microscope at 25X magnification. An optical bridge splits the microscope view between the microscope eyepieces and two Flea®2 cameras (Point Grey Research, Richmond, BC, Canada). Mounted as a stereo pair, the cameras capture 800x600 resolution video at 30 Hz. For cannulation procedures, TIP1TW1 micropipettes (World Precision Instruments, Saratoga, FL, US) with an inner diameter (ID) of 1 μm are secured to the output plate of Micron. Tubing threads through the hollow center of Micron's shaft and connects to a 20 mL syringe, which can be depressed to force gas or fluid through the micropipette and into the vessel.

8.2 VISION SYSTEM

The stereo cameras are responsible for tracking both the tip of the micropipette and the vessels in the image to ascertain the proximity of the micropipette to the vessel as depicted in Fig. 75. The relative distance is used by the control system, which adjusts the scale factor that determines how much hand motion is transferred from the handle to the tip of the instrument.

Since the micropipettes are formed from clear glass and the tip tapers off to a nearly invisible point, directly tracking the tip with

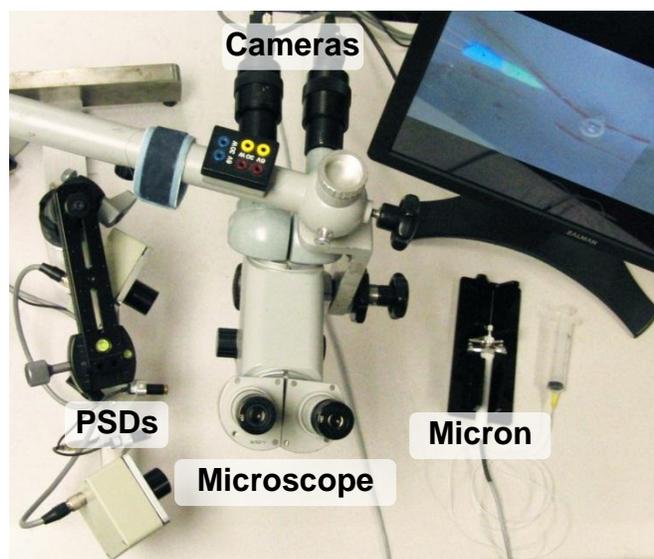


Fig. 74. Cannulation setup with Micron micromanipulator, microscope, stereo cameras, and PSD optical sensors.

vision poses a significant challenge. As an effective alternative, two unique colored fiducials are painted in stripes around the micropipette near the tip. Fast color segmentation [126] tracks the centroids of the fiducials, and the true micropipette tip position is estimated with interpolation. A quick 30 s calibration period registers the camera views with the world coordinate system as defined by the PSDs measuring the 6-DOF pose of the instrument; registration is maintained by an online adaptive least squares calibration procedure that adjusts the camera coordinate system to the world frame with each new camera image.

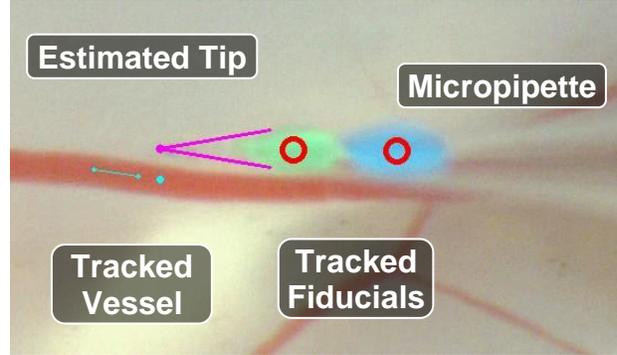


Fig. 75. Stereo cameras track 3D positions of green and blue fiducials on the micropipette to estimate the tip position. The vision system also detects vessels in the image to calculate the center of the vessel nearest the micropipette tip.

Vessels are similarly tracked with a Gaussian Mixture Model of learned vessel colors [125]. By searching locally around the estimated position of the micropipette tip, the XY center of the nearest vessel is located. Because vessels appear as locally invariant structures, i.e., one section of the vessel does not look significantly different from another section, correspondences for 3D reconstruction of the vessel are difficult to find. Instead, a global estimate of the mean depth Z of all blood vessels in view is calculated by running a simplified Lucas-Kanade tracker [170], which estimates the translation between the left and right camera views of the blood vessels. The 3D distance between the micropipette tip and the center of the nearest vessel as calculated by the stereo cameras is then given to the control system.

8.3 TREMOR COMPENSATION AND MOTION SCALING

Two key components form basis of the control system: tremor compensation and motion scaling. Hand tremor is compensated by commanding the micromanipulator tip $P_T \in \mathbb{R}^3$ to the the null position $P_N \in \mathbb{R}^3$ that has first been filtered by a lowpass filter F_L . Since the null position is mechanically tied to the handle, the micropipette tip output is the smoothed input of the operator's hand movements:

$$P_T = F_L(P_N) \quad (40)$$

Since physiological hand tremor generally lies in a frequency band of 8-12 Hz, a first-order 1.5 Hz lowpass Butterworth filter was selected for F_L , which greatly attenuates tremor while only introducing a minor lag and settling time into the eye-hand coordination feedback loop.

The second key component used by the control system is motion scaling, which gives the surgeon increased precision. Motion scaling takes the handle motion as calculated by the change in null position ΔP_N and transfers only a fraction $1/s \in \mathbb{R}$ of the movement to the micromanipulator tip ΔP_T :

$$\Delta P_T = \frac{1}{s} \Delta P_N \quad (41)$$

For example, if the scale factor is 2 and the surgeon moves the handle 100 μm , the tip only moves 50 μm . Any uncompensated tremor, drift, or deliberate movement is scaled down, granting the surgeon the additional comfort and safety of using larger movements to manipulate tiny anatomy. A scale factor of unity is equivalent to running no motion scaling (i.e., Micron turned off) while a scale factor of ∞ corresponds to fixing the tip in space (i.e., disregarding all operator's movement).

One important limitation of motion scaling is the decreased range of motion. During tremor compensation, the lowpass filter gain rolls off to unity at 0 Hz so the manipulator range of motion is only used temporarily to smooth out sudden jerks and/or tremor. However, since motion scaling is applied all the way down to 0 Hz, the effective range of motion of the device is reduced by a factor s . For this reason, tremor compensation can be performed everywhere, whereas motion scaling can only be applied in a limited volume, generally near the vessel where it is most effective in aiding the cannulation procedure.

Additionally, because the 0.5 mm axial range of Micron is significantly smaller than the 1.0 mm transverse range, motion scaling is applied anisotropically. While tremor compensation is applied in all directions at all times, motion scaling is only activated as discussed earlier in the vicinity of the vessel and only in the transverse directions of the tool. In addition to preserving manipulator range along the axis of the tool where the range of motion is most limited, anisotropic scaling has an additional benefit. Since the surgeon approaches the vessel along a roughly parallel trajectory, unity motion scaling axially allows for quick thrusts into the vessel while motion scaling in the transverse direction increases the ability of the surgeon to keep the micropipette centered on the vessel.

8.4 CONTROL SYSTEM

The control system combines tremor compensation with anisotropic motion scaling to aid the surgeon in cannulation procedures. Tremor compensation is applied at all times for smooth movements. Motion scaling is activated only when the micropipette is closer than a certain threshold to the vessel, as measured by the vision system. A threshold value of 500 μm was selected empirically. Although motion scaling could be implemented with a velocity controller, estimation of the velocities from measured positions is noisy. A preferable approach in our system is to select some reference point $P_R \in \mathbb{R}^3$ for motion scaling, measure the offset between the tremor-compensated null position $\mathbf{F}_L(P_N)$ and the reference point P_R , and use a position controller to drive the micromanipulator tip P_T to the measured offset position, scaled by $1/s$. Anisotropic motion scaling is achieved by taking into account the rotation $R \in \mathbb{R}^{3 \times 3}$ of the tool, and introducing individual scale factors $s_X, s_Y, s_Z \in [1, \infty]$:

$$P_T = S(R\mathbf{F}_L(P_N) - RP_R) + RP_R \quad \text{where } S = \begin{bmatrix} 1/s_X & 0 & 0 \\ 0 & 1/s_Y & 0 \\ 0 & 0 & 1/s_Z \end{bmatrix} \quad (42)$$

The reference point P_R is selected as the tremor compensated null position at the time of activating motion scaling $P_R = \mathbf{F}_L(P_N)$ and held constant until scaling is deactivated.

As the operator moves Micron during the procedure, the motion scaling displaces the tip away from the null position, reducing the range of motion. At some point, the micromanipulator tip must return to the null position, otherwise the actuators will saturate and Micron will be unable to

provide any assistance. Ideally, we would prefer sufficient range of motion to complete the cannulation and withdraw from the vessel before returning the tip to the null position, thus maximizing the range for other operations.

Turning off motion scaling is done when the vision system senses the tip has left the vicinity of the vessel (i.e., the distance between them exceeds a certain threshold). However, directly commanding the micromanipulator tip to the null position causes a rapid and unpleasant twitch in the tool tip. Furthermore, transient tracking errors may falsely trigger turning off the motion scaling. Since tracking is not entirely reliable, the problem of determining when and how to return to pure tremor compensation is exacerbated.

The proposed solution involves a graceful transition from motion scaling to tremor compensation while at the same time gradually re-centering the manipulator at the null position. To account for noisy tracking, the scale is decreased exponentially over time, providing robustness to intermittently noisy or incorrect distance measurements. Denoting t as the number of successive time steps the vision system has detected the tip outside the vicinity of the vessel, the scale factor for each direction is reduced at each time step:

$$s_D \leftarrow s_D - \frac{\alpha}{(1 + \gamma)^t} \quad \forall D = (X, Y, Z) \quad (43)$$

where α specifies how much to reduce the scale at each time-step. An exponential discount factor γ encodes a sense of how confidence over time increases the convergence rate. Thus, several intermittent noisy measurements will only decrease motion scaling slightly. In contrast, a sequence of measurements indicating that motion scaling is no longer needed signifies a much higher confidence and effects a higher convergence rate back to tremor compensation. Equation (43) terminates for each direction when the scale factor is unity, at which point the manipulator tip has returned to the null position. Alternatively, (43) terminates if the tip is re-detected within the vicinity of the vessel, in which case the reference point P_R is recalculated as:

$$P_R = (S - I)^{-1}(SRP_T - RF_L(P_N)) \quad (44)$$

and the scale factor is reset to avoid discontinuous jumps in tip position due to changing scale factors. Thus, the control system accomplishes a gentle transition from motion scaling to tremor compensation in a way that robustly deals with noise. Empirical testing indicates values of $\alpha = 3 * 10^{-5}$ and $\gamma = 1 * 10^{-4}$ are effective for a smooth 4–8 s transition.

8.5 EXPERIMENTS

Cannulation experiments were performed *ex vivo* by an experienced retinal surgeon under two scenarios: Aided (with the assistance of Micron) and Unaided (freehand, with Micron powered off). Porcine eyes were used as the animal model for all experiments. Sections of the back eye around the optic nerve were removed with the retina and vessels still attached. Vessel diameters were measured before each cannulation attempt, and total cannulation duration was recorded. Total duration excludes time spent clearing any bubbles in the vitreous caused by failed injections. Injections were performed by depressing a 20 mL syringe of air connected to the micropipette. The surgeon performed the insertion of the cannula into the vessel and orally indicated when to depress

the syringe. Multiple injection attempts were allowed per trial. Successful cannulation was defined as air entering the vessel and displacing the blood; a cannulation was recorded as a failure when the vessel became too damaged to continue. Trials alternated between several Unaided attempts and several Aided attempts to limit ordering effects.

8.6 RESULTS

As seen in Table 8, for vessels with diameters between 40 and 60 μm , the cannulation success rate was higher with similar durations in successful trials. Vessels larger than 60 μm exhibit more equal success rates. Visually, Fig. 76 presents the paths inscribed by the micropipette tip during cannulation as traces on the image for both the Aided and Unaided cases.

From these images, it can be seen that tremor is lessened and the surgeon can better track the vessel in the Aided case, thus reducing trauma to the vessel and surrounding tissue.

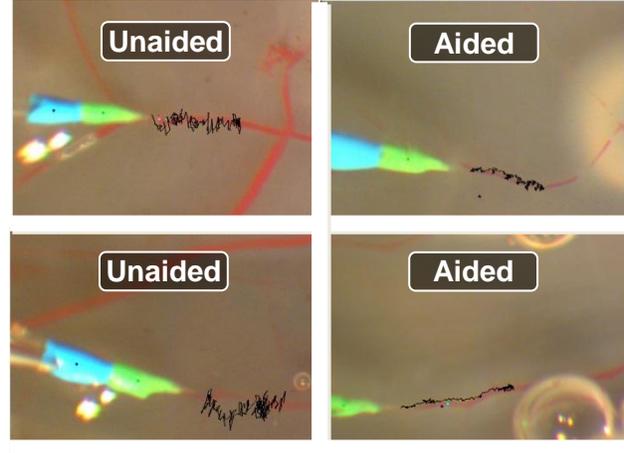


Fig. 76. Example traces of the micropipette tip during cannulation on porcine retina overlaid onto microscope images, showing tremor during the procedure without Micron (Unaided) and with Micron (Aided).

Trials	Total Experiments	Successful Experiments	Successful Durations (s)
Unaided	7	2 (29%)	76±31
Aided	8	5 (63%)	72±19

Table 8: Cannulation results for 40-60 μm diameter vessels

8.7 CHICKEN EGG MEMBRANE

In addition to our experiments with *ex vivo* porcine eyes, we performed further testing *in vivo* in chick embryo chorioallantoic membrane [92]. The chicken eggs were incubated to allow the vessel to grow, and several open-sky cannulation procedures were performed by an experienced surgeon. We found after one initial session, the surgeon was able to very consistently cannulate the targeted vessel in both Aided and Unaided cases, although the Aided cases exhibited less tremor. Fig. 77 shows typical unaided and aided traces during these chicken egg procedures.



Fig. 77. Example traces of the micropipette tip during cannulation on chicken eggs overlaid onto microscope images, showing tremor during the procedure without Micron (Unaided) and with Micron (Aided).

8.8 LEARNING CURVES

To examine the effect of task learning curves as the surgeon becomes more proficient, we analyzed success rate for each session. As shown by the Aided and Unaided test scenario success rates plotted over time in Fig. 78, we discovered the surgeon eventually became equally proficient under either scenario. It is clear, however, that the Aided case is useful to ease the learning curve. Further testing is necessary to examine how the learning curve generalizes to more realistic scenarios and other surgeons.

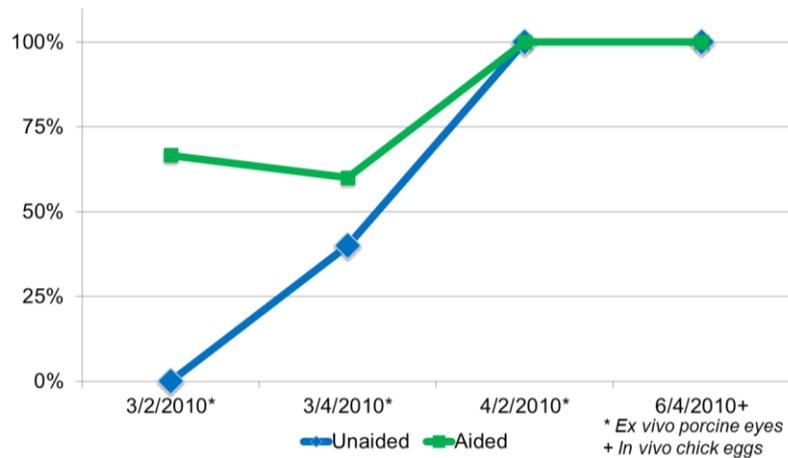
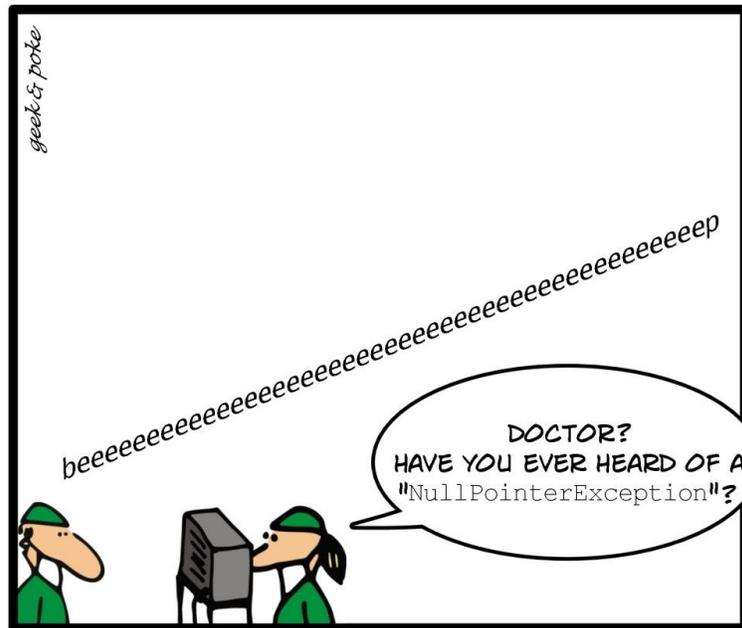


Fig. 78. Learning curve of cannulation success for 40-60 μm vessels over four sessions spaced over a 3 month period.

8.9 DISCUSSION

In these preliminary surgeon experiments, the robot aid of tremor compensation with motion scaling applied by the active handheld micromanipulator increased the success rate of small vessels while maintaining similar procedure duration. After several sessions, the surgeon learned to cannulate without the aid of Micron, demonstrating that additional testing is necessary to determine how Micron can aid in more realistic scenarios. For instance, a 6-DOF prototype of Micron is being built that can enforce a remote center of motion; this will allow testing of Micron to move from open-sky to the modified Johns Hopkins eye phantom described in Section 2.6.6. Future work should include evaluations of cannulations with the new 6-DOF prototype of Micron in more realistic scenarios.

DISCUSSION



RECENTLY IN THE OPERATING ROOM

Photo credit: <http://geekandpoke.typepad.com/>. Reprinted under the Creative Commons license.

In this thesis, we have addressed challenges in designing micromanipulator behavioral aids for Micron, a fully handheld micromanipulator. Motivated by the difficulties in performing retinal microsurgery, we have developed a general and extensible a framework that incorporates tremor suppression, motion-scaling, and virtual fixtures. Work in improving control system rejection of tremor and real-time visual analysis of anatomy, including a new localization and mapping algorithm for high-magnification microscope views of retinal operations, has been presented. Task specific behaviors have been implemented using the developed framework and evaluations have been performed on *ex vivo* and *in vivo* tissue, demonstrating the applicability of the research to microsurgical operations.

In this thesis, we have worked to validate the central hypothesis that the development of more effective behaviors can be achieved with the fusion of vision and control to provide real-time, contextual aids to the surgeon. A conclusion summarizing what has been accomplished in this thesis and highlighting the core contributions of our research are presented, followed by some discussion regarding future directions and how the capabilities of handheld micromanipulators such of Micron can be expanded and improved in further research.

9 CONCLUSIONS

With microsurgery techniques opening up new avenues of treatment for previously untreatable conditions, it is no wonder that surgeons are finding themselves needing increasingly precise control at extremely fine scales. Coupled with the success of commercial robots like Intuitive Surgical's da Vinci platform for macroscopic surgeries, it is unsurprising that a number of micromanipulator robotic solutions have been proposed [18], [19], [21–23], [25], [26], [31], [32]. With the goal of aiding and supporting surgeons during microsurgical procedures, such robots have been held back by confounding factors such as bulkiness, price, difficult control mechanisms, difficulties in achieving micrometer-precise movements, and familiarity. Our handheld robot, Micron [27], is a small, lightweight robot that addresses a number of these issues while providing a tremor-free platform from which more intelligent behaviors can be implemented.

9.1 SUMMARY

In this thesis, we advance the capabilities of Micron from generally cancelling tremor with frequency-based methods described in [27] to providing context-specific, intelligent behaviors based on real-time visual analysis of the anatomy and the surgeon's actions during the operation. We choose to focus on procedures in the demanding field of vitreoretinal microsurgery (Chapter 2). With the types of behaviors needed for these procedures in mind, we developed a virtual fixture framework for handheld micromanipulation (Chapter 3) that is enhanced with a feedforward control system for improved tremor rejection and constraining the instrument tip (Chapter 4). We couple the control system with real-time vision algorithms that analyze the anatomy and surgeon's movements (Chapter 5). Finally, as a systems thesis, a significant amount of time is spent on designing and testing context-specific behaviors for each procedure to show the practical application of the theoretical work presented. Micron has been shown to aid three different retinal procedures under several different metrics: membrane peeling because of the incredible delicateness required, prevalence of operation, and higher chance for complications (Chapter 6); laser photocoagulation because of its prevalence, tedious nature, and propensity for automation (Chapter 7); and vessel cannulation because it is a promising new treatment for a currently untreatable condition that is too difficult to perform by the majority of vitreoretinal surgeons (Chapter 8).

9.2 CONTRIBUTIONS

During the course of researching how vision-based control can be combined with Micron to help surgeons during retinal operations, we have made several key contributions to the field: a virtual fixtures framework based on position rather than force input, a feed-forward control system to cancel tremor more effectively, a fast EyeSLAM algorithm that localizes and maps the retina, and a full vision-based control system for Micron that runs in real-time. Finally, to validate our contributions, we have developed easy-to-use, yet realistic eye models and performed extensive evaluations of Micron.

9.2.1 VIRTUAL FIXTURE FRAMEWORK FOR HANDHELD MICROMANIPULATION

Existing virtual fixture formulations are designed for master/slave configurations [56], [95], [96], [100], [171] or use applied forces to guide the tip [39], [41], [94], [97], [172]; such approaches are

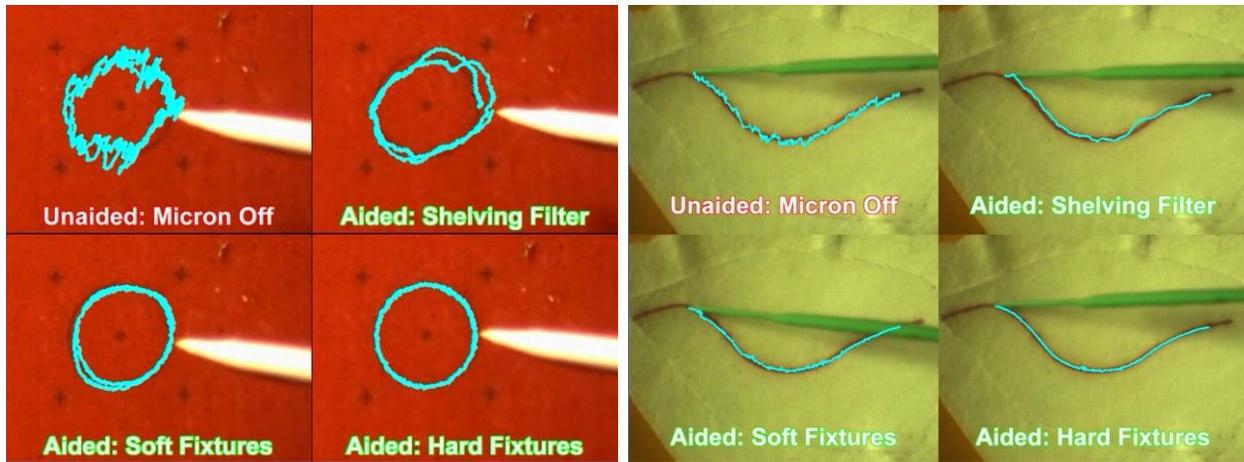


Fig. 79. Surgeon's pointing accuracy increases with virtual fixtures in circle and vein tracing tasks.

not well suited to fully handheld platforms such as Micron. In Chapter 3, we present the novel theoretical contribution of a virtual fixture framework for fully handheld robots that (a) uses the position of the robot's null position to constrain the tip and (b) very naturally integrates tremor suppression, motion scaling, and virtual fixture compositing. Similar to [39], [54], virtual fixtures are created from visual analysis of anatomy. As seen in Fig. 79, the virtual fixtures framework shows excellent performance in synthetic circle and vein tracing tasks and outperforms both unaided and state-of-the-art shelving tremor compensation techniques.

9.2.2 MODEL-PREDICTIVE, FEED-FORWARD CONTROL FOR TREMOR REJECTION

Analysis and modeling tremor is a well-researched topic [29], [32], [103], [112–114], [117], [123], [173], [174], [175–177]. Although some research has integrated tremor models into real-time cancellation systems [32], [103], [175–177], all of these approaches use tremor models to estimate tremor and isolate voluntary motion, which is then used to drive the micromanipulator. However, none use the tremor models to account for latency in actuation for

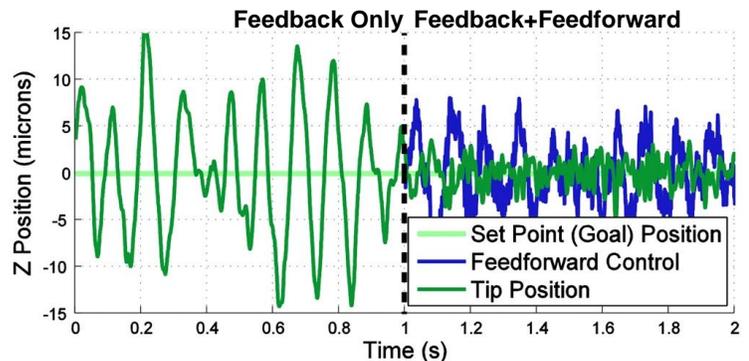


Fig. 80. Feedforward control uses tremor models to anticipate, predict, and correct error caused by latencies in actuation for improved tremor rejection.

improved performance. While operating Micron, several milliseconds of actuation delay is enough for tremor to be evident in the output. In Chapter 4, one of our contributions is to develop a feedforward controller that predicts tremor into the future and adjusts the goal position with a feedforward controller. When combined with the existing PID feedback controller, our model-predictive, feedforward control approach allows an additional 50% rejection of tremor (Fig. 80).

9.2.3 LOCALIZATION AND MAPPING OF THE RETINA (EYESLAM)

To construct virtual fixtures around anatomy on the eye, an accurate map of the vasculature and a model of the eyeball motion is required. While a wide variety of research on retinal mosaicking and registration exist [146–154], only several approaches can run real-time [149], [150], [152]. Unfortunately, none are well suited to high-magnification, low-texture, dynamic-illumination conditions experienced during vitreoretinal surgery. In Section 5.5, we propose a robust simultaneous localization and mapping of the retinal vessels on the retina. It uses the noisy vessel detection of Can et al. [131] for feature detection, occupancy grids for mapping, and iterative closest point (ICP) for localization. EyeSLAM

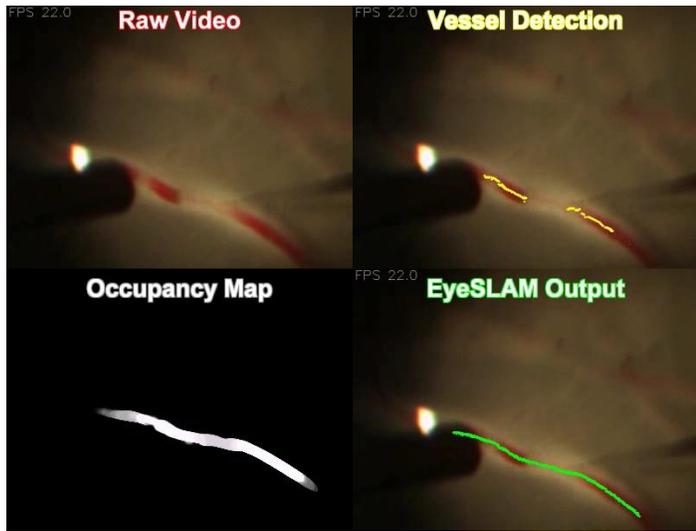


Fig. 81. EyeSLAM performing real-time localization and mapping of the retinal vasculature in a difficult video. Notice the high magnification, very sparse vasculature tree, noisy features, and occlusions present.

operates in real-time at 30-40 Hz on half-size images, converges rapidly within seconds, is robust to occlusions by the instrument shaft, and handles varying illumination caused by shifts in the handheld light-pipe. Even in areas of sparse vasculature with nearly straight vessels, one of the worst-case scenarios for ICP, EyeSLAM performs well (Fig. 81).

9.2.4 RETINAL EYE PHANTOM MODEL WITH PORCINE RETINA

A major problem in evaluating retinal procedures is finding a good model that is easy-to-use but realistic. For preliminary work, obtaining live human or animal retina is a challenge, so alternative models are often used. Chicken egg membranes have been used [23], [92], but laser photocoagulation is difficult in this model and vessel cannulation is not very realistic due to the exaggerated elasticity compared to the retina. Ueta et al. [19] perform a vitrectomy on ex vivo pig eyes, but this is challenging without expensive vitrectomy equipment; furthermore, because the lens of the eye clouds shortly after death, finding fresh enough pig eyes is problematic. One of the contributions of our work is the development of a fairly realistic, yet relatively simple model for the

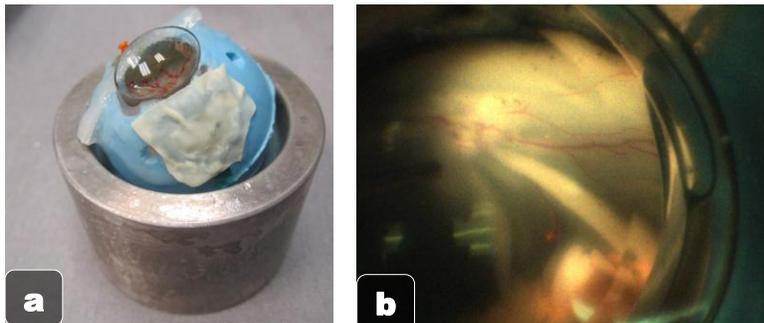


Fig. 82. Retinal eye model with porcine retina in a modified Johns Hopkins eye phantom.

retina that includes the sclerotomy and illumination constraints of real retinal surgery. By excising porcine retina and placing it in the bottom of a modified version of the Johns Hopkins phantom filled with silicone oil, we have developed a fairly realistic model of vitreoretinal surgery that is cheap and easy to create (Fig. 82).

9.2.5 EXTENSIVE EVALUATION OF MICRON WITH RETINAL PROCEDURES

In order to evaluate the theoretical contributions of this thesis, significant effort was placed on systems design to extensively evaluate the proposed methods with retinal procedures. Heavy use of software engineering with multi-threaded, highly synchronized C++ modules made real-time visual analysis and control of Micron possible. A new network interface was introduced to enable the high-level computer vision and virtual fixture framework systems to run on a consumer-grade quad-core desktop. A full system for vision-based control of Micron was developed and tested with multiple surgeons in many different environments.

Throughout the practical evaluation section, we contribute extensive evaluation of three retinal procedures that have never been reported with Micron before this thesis. Retinal membrane peeling was simulated with peeling plastic wrap off of rubber and sorbothane slides. Retinal laser photocoagulation was simulated on paper slides and ex vivo open-sky porcine retina. Retinal vessel cannulation was performed on ex vivo open-sky porcine retina. All three experimental procedures were run with at least one experienced vitreoretinal surgeon to ensure the results are medically relevant. In membrane peeling and laser photocoagulation experiments, sufficient results were performed to show the effectiveness of vision-based control with Micron with statistical significance ($p < 0.05$).

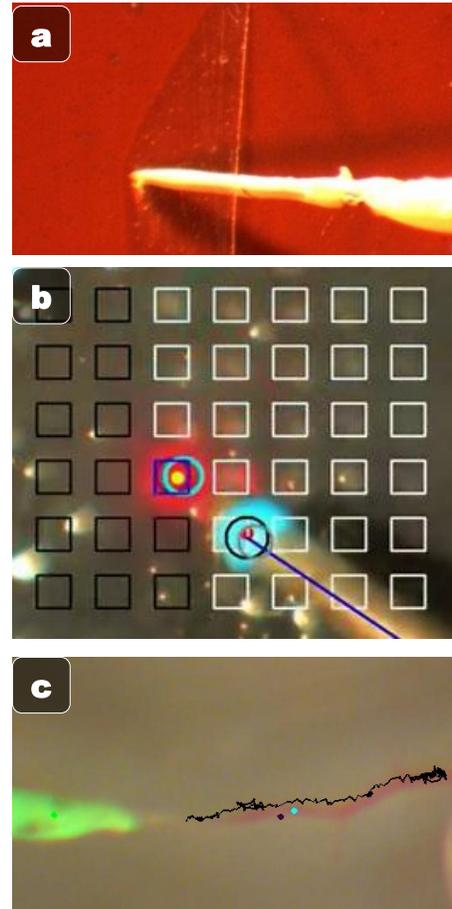


Fig. 83. Three retinal procedures reported: (a) membrane peeling (b) laser photocoagulation (c) vessel cannulation.

9.3 VALIDATION OF THESIS STATEMENT

In Section 1.5, we proposed three hypotheses. We now revisit those hypotheses and judge them based on the results obtained in this thesis.

1. *Micromanipulation tasks can be decomposed into steps, during which applying tip position constraints (virtual fixtures) with a handheld tool increase operator precision and safety.*

In Chapter 3, we propose the virtual fixture framework, which constrains the tip using visual information. Evaluation in Section 3.3.3 with synthetic tasks that mimic medically relevant steps in a procedure such as point, move/hold, and various tracing tasks show increased operator precision. In a multi-surgeon series of experiments performed in an open-sky environment, 3D RMS error is significantly reduced by 93% for Hold Still, 94% for Circle Tracing, 84% for Move/Hold, and 77% for Vein Tracing compared to both unaided and state-of-the-art tremor cancellation algorithms that do not depend on vision [27] ($p < 0.05$). Likewise, maximum 3D error during these trials is reduced by 45-85%.

2. *Real-time visual analysis of anatomy and the surgeon's movements provides the context that allows the controller to select and customize the virtual fixtures.*

In Chapter 5, we present a comprehensive vision system for tip tracking, multi-view geometric reconstruction, and mapping and locating retinal anatomy. All algorithms run in real-time for fast, low-latency control of Micron. Adaptive camera calibration (Section 5.2) maintains a consistently valid registration between visually detected anatomy and the ASAP control coordinate frame for high accuracy. In Section 3.2.3 and 3.3.2, we show that generalizable splines can model anatomy and be used to generate 3D spline virtual fixtures to significantly decrease 3D RMS tracing error by 77% ($p < 0.05$). Dense stereo vision described in Section 5.3 has been used to reconstructs the surface of the eye phantoms in real-time. It has been successfully employed to regulate depth of Micron's tip in vessel tracing experiments inside a curved eye phantom (Sections 2.6.2, 3.3.2, and 3.3.3) and on membrane peeling experiments on sorbothane slide (Sections 2.6.1, 6.4, and 6.7). Finally Section 5.5 describes how a novel algorithm for mapping and localizing retinal vessels in the eye can be used for laser photocoagulation and vessel cannulation procedures.

3. *Vision-based virtual fixtures aid surgeons during microsurgical operations.*

As the culmination of all contributions, the essential question this thesis tries to answer is: does vision-based control of Micron actually work? In the latter half of this thesis, we attempt to answer this question by combining all the theoretical work into practical implementations and evaluate performance on several retinal procedures. In membrane peeling experiments (Chapter 6), we show that using nothing but vision-based virtual fixtures (i.e. no force control), surgeons using Micron can peel plastic wrap from rubber and sorbothane with a 77-83% reduction in applied downward force ($p < 0.05$). In laser photocoagulation of paper slides and porcine retina (Chapter 7), visual control of the laser with Micron's aid increases the surgeon's speed by 1-4X while simultaneously reducing error 16-39% ($p < 0.05$). In vessel cannulation of open-sky porcine retina (Chapter 8), vision-based control with Micron aids a surgeon's initial success rates, lessens the learning curve, reduces tremor, and feels more ergonomic. Both quantitatively and qualitatively over a number of metrics, vision-based control of Micron has shown improvements in simulated retinal surgeries.

These results validate our three core hypotheses and lend evidence to our stated thesis statement:

Fusing real-time visual information into the control system of a fully handheld robot enables a rich set of task-specific behaviors known as virtual fixtures that can be applied to increase surgeon precision and performance during micromanipulation, such as vitreoretinal procedures.

10 FUTURE DIRECTIONS

The proposed work makes a progressive stride towards the goal of synergistic behavior aids that analyze, understand, and augment a surgeon's operation, but this thesis only represents a small step in the direction of using robots as intelligent aids for surgeons. In this chapter, we outline several next steps that could advance Micron's capabilities further.

10.1 NEW 6-DOF MICRON MICROMANIPULATOR

The greatest limiting factors in the current 3-DOF Micron manipulator are the small range of motion and lack of a remote center of motion. While we show significant performance improvements with Micron, we have had to make a number of design decisions to account for the small range of motion, such as anisotropic motion scaling, lower-than-desired scale factors, and performing procedures in open-sky environments rather than realistic eye phantoms. Furthermore, the majority of error in many tasks occurs while the device is in saturation, or when the range of motion has been exceeded. Moving from open-sky procedures to more realistic retinal environments with sclerotomy constraints has been especially problematic as the pivot point of the current Micron manipulator is near the grip rather than where the shaft passes through the sclera of the eye. A remote center of motion moves this pivot point arbitrarily down the shaft, enabling tremor cancellation and other behaviors at the tip without any motion where the instrument enters the eye. Another artifact of moving from open-sky to more realistic eye models is more motion is required to maneuver the eyeball so the light-pipe and microscope view are positioned properly

A new 6-DOF freedom prototype of Micron has been designed and is currently under construction and optimization for handheld use [178]. The new version, depicted in Fig. 84, is built upon a Stewart platform that expands the range of motion from $\sim 0.6 \times 1.5 \times 1.5$ mm to $\sim 4.0 \times 4.0 \times 4.0$ mm, a substantial increase that should make a significant difference in the types of behaviors that Micron can perform and the scale factors it can achieve. Furthermore, with the 6-DOF prototype, the manipulator can enact a remote center of motion, which should greatly aid procedures performed in an eye phantom or real eyeball.

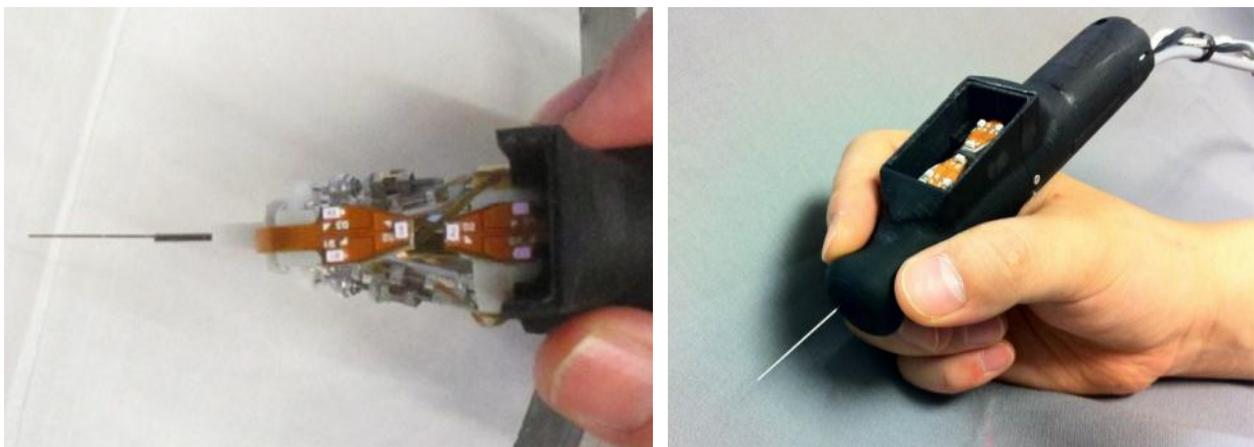


Fig. 84. New 6-DOF Micron manipulator prototype using a Stewart-platform to achieve a larger range of motion with a remote center of motion.

10.2 MORE REALISTIC EVALUATION

Evaluation in this thesis has been primarily limited to synthetically-created phantoms and *ex vivo* porcine retina in an open-sky environment. While the experiments presented here highlight the possible advantages of using handheld micromanipulation devices in surgery, Micron is quite far from clinical trials or widespread use in an operating room. One contribution of our research is the development of a more realistic eye model with the *ex vivo* porcine retina inside an eye phantom filled with saline or silicone oil and capped with a vitrectomy lens. Further work on adapting the new 6-DOF version of Micron and more advanced computer vision algorithms will allow more realistic evaluation of Micron in the near future. In the farther future, *in vivo* evaluation in chicken eggs or rabbit eyes is necessary to more fully test how Micron works in increasingly realistic surgical scenarios.

10.3 NEW AREAS OF SURGICAL APPLICABILITY

The three retinal surgeries this this addresses are only a very small subset of the possible procedures Micron could be applied towards. In the retina, other procedures such as arteriovenous sheathotomy operations could be investigated. Also, use of imaging modalities such as Optical Coherence Tomography (OCT) to allow sub-micron accuracy in 3-D tissue reconstruction and to see below the surface of the retina is gaining interest [179]. Micron could be used to visually target an area and perform a spiral scan around the target with the OCT probe. In other fields such as Ear/Nose/Throat (ENT) which require fine, demanding micromanipulation, Micron could be a promising tool to remove tremor and define sensitive forbidden areas for increased safety. With some hardware adaptation to the shaft and tip, Micron has shown encouraging results in preliminary experiments (see Fig. 85) with simulated stapes microsurgery [180].

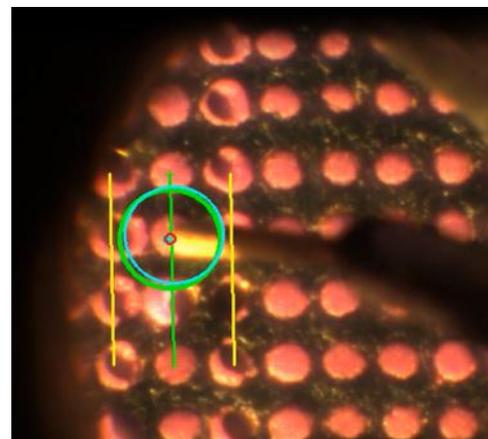


Fig. 85. Example of using the virtual fixture framework in simulated stapes footplate surgery.

10.4 EXPANDED VIRTUAL FIXTURES FRAMEWORK

While the virtual fixture framework in its current formulation has proven itself adequate for the tasks described in this thesis, there are some expansions that could improve the algorithm. One current omission of the framework is the lack of an orientation fixture, caused in part by the fact that the 3-DOF version of Micron cannot control orientation independently of tip position. However, as work with Micron transitions to the new 6-DOF Micron, orientation fixtures will become more important. This is especially true for laser photocoagulation procedures where the important constraint is not on the tip position but where the shaft (and thus the laser) is pointing (see Fig. 86). Other improvements to the virtual fixture framework could include a more advanced compositing mechanism to enforce multiple fixtures, either simultaneously or over a temporal sequence.

10.5 MORE ROBUST VISION SYSTEMS

Currently, there is a delicate balance in computer vision when choosing an algorithm: often-times there is a trade-off between real-time performance and the sophistication/accuracy. Because of the high-speed nature of retinal surgery and the handheld nature of Micron, very fast responses are needed, which necessitates choosing algorithms that are fast, but perhaps not as accurate or brittle. The current system tries to be as robust as possible, but many procedures require tedious tuning of illumination, careful placement of fiducials, or on-the-fly optimization of algorithm parameters. This level of brittleness is not always practical in an operating room. New, effective, and robust computer algorithms developed specifically for the types of problems encountered in microsurgery are needed. Furthermore, as CPU abilities rise and GPU-acceleration becomes more powerful and widespread, more sophisticated algorithms should help increase the robustness of the Micron system. Such improvements in tracking, dense depth reconstruction, calibration, and anatomy detection will aid safety, reduce the need for tuning, and better performance.

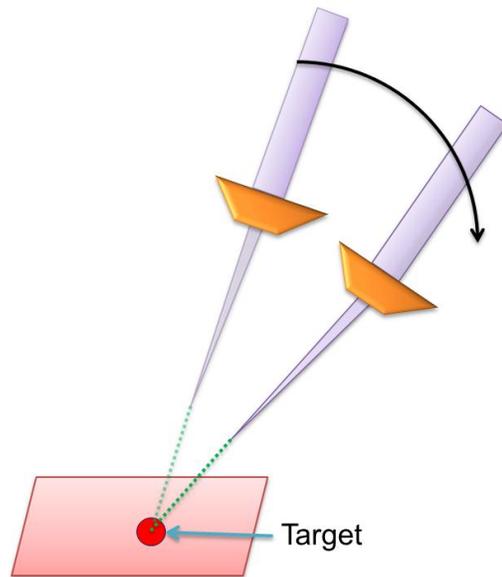


Fig. 86. Orientation virtual fixtures constrain not the tip position, but where the tip is pointing. An example of this is laser targeting.

10.6 MORE SOPHISTICATED HARDWARE

Although ASAP provides very accurate and high-bandwidth position measurements, the line of sight restriction is particularly frustrating in a surgical environment where the surgeon often changes the position and orientation of the instrument. Frequent adjustments to the ASAP head are necessary to keep the handheld Micron device within the workspace. There are also concerns of space in the operating room, occlusions of the line of sight, and possible contaminants falling on the sensor or LEDs. A magnetic solution that provides high-resolution, high-bandwidth position information would be significantly more convenient to use and increase safety.

Likewise, much of the medical hardware used in this thesis is outdated, simply because it was donated or could be obtained cheaply. As a consequence, some of the more realistic vessel cannulation experiments are limited not only by system performance, but poor equipment and procedures. Higher illumination levels, better microscope equipment, improved visibility, and a vitrectomy machine to correctly process porcine eyes could make an impactful difference in performance during evaluation. Using actual equipment in a hospital operating room merges nicely with the future work of more realistic evaluation and should go a long way towards advancing Micron's capabilities.

10.7 INTEGRATING FORCE SENSING

For vessel cannulation and especially retinal membrane peeling, regulating force and preventing excessive forces to the retina is critical to the outcome of the procedure. Our virtual fixture framework combined with stereo vision has proved beneficial for enforcing approximate constraints on forces using nothing more than computer vision. However, a more practical and effective solution could be to employ a sensor that directly measures the force at the end-effector similar to [38]. In this case, full control methodologies can be developed that close the loop directly around applied forces rather than approximately guessing the forces via visual observation. This could allow much more sophisticated behavior that not only reacts to visual analysis of the anatomy and surgeon's actions but also the forces being applied at the tip of the instrument.

10.8 CONCLUSION

We have briefly mentioned several avenues of future work that would improve the research presented in this thesis. Ideally, research in these and related directions will yield more robust robotic platforms that couple surgeons' expert knowledge with the positioning accuracy of micromanipulators to make the operating room a less scary place.

11 WORKS CITED

- [1] M. Patkin, "Ergonomics applied to the practice of microsurgery," *Australian and New Zealand Journal of Surgery*, vol. 47, no. 3, pp. 320–329, 1977.
- [2] H. L. Brooks, "Macular hole surgery with and without internal limiting membrane peeling," *Ophthalmology*, vol. 107, no. 10, pp. 1939–1948, 2000.
- [3] H. Ahmadiéh, H. Sajjadi, M. Azarmina, M. Soheilian, and N. Baharivand, "Surgical management of intraretinal foreign bodies," *Retina*, vol. 14, no. 5, pp. 397–403, 1994.
- [4] F. Bandello, P. Lanzetta, and U. Menchini, "When and how to do a grid laser for diabetic macular edema," *Doc. Ophthalmologica*, vol. 97, pp. 415–419, 1999.
- [5] 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.
- [6] B. Gupta, M. Elagouz, D. McHugh, V. Chong, and S. Sivaprasad, "Micropulse diode laser photocoagulation for central serous chorio-retinopathy," *Clinical & Experimental Ophthalmology*, vol. 37, no. 8, pp. 801–805, 2009.
- [7] M. B. Parodi and F. Bandello, "Branch retinal vein occlusion: Classification and treatment," *Ophthalmologica*, vol. 223, no. 5, pp. 298–305, 2009.
- [8] I. U. Scott, M. S. Ip, P. C. VanVeldhuisen, N. L. Oden, B. A. Blodi, M. Fisher, C. K. Chan, V. H. Gonzalez, L. J. Singerman, and M. Tolentino, "A randomized trial comparing the efficacy and safety of intravitreal triamcinolone with observation to treat vision loss associated with macular edema secondary to central retinal vein occlusion: the Standard Care vs Corticosteroid for Retinal Vein Occlus," *Arch. Ophthalmol.*, vol. 127, no. 9, pp. 1115–1128, 2009.
- [9] C. P. Shah, "A randomized trial comparing intravitreal triamcinolone acetate and focal/grid photocoagulation for diabetic macular edema," *Evidence-Based Ophthalmology*, vol. 10, no. 1, p. 29, 2009.
- [10] J. N. Weiss, "Retinal surgery for treatment of central retinal vein occlusion," *Ophthalmic Surgery and Lasers*, vol. 31, no. 2, pp. 162–165, 2000.
- [11] S. P. N. Singh and C. N. Riviere, "Physiological tremor amplitude during retinal microsurgery," in *Proc. IEEE Northeast Bioeng. Conf.*, 2002, pp. 171–2.
- [12] G. H. Ballantyne, "Robotic surgery, telerobotic surgery, telepresence, and telerobotics. Review of early clinical results," *Surgical endoscopy*, vol. 16, no. 10, pp. 1389–402, Oct. 2002.
- [13] S. Horgan and D. Vanuno, "Robots in laparoscopic surgery," *Journal of Laparoendoscopic & Advanced Surgical Techniques*, vol. 11, no. 6, pp. 415–419, 2001.
- [14] R. H. Taylor and D. Stoianovici, "Medical robotics in computer-integrated surgery," *IEEE transactions on robotics and automation*, vol. 19, no. 5, pp. 765–781, 2003.
- [15] B. Davies, "A review of robotics in surgery," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 214, no. 1, pp. 129–140, 2000.
- [16] A. R. Lanfranco, A. E. Castellanos, J. P. Desai, and W. C. Meyers, "Robotic surgery: a current perspective," *Annals of Surgery*, vol. 239, no. 1, p. 14, 2004.
- [17] K. Ikuta, T. Kato, and S. Nagata, "Micro active forceps with optical fiber scope for intra-ocular microsurgery," in *IEEE Micro. Electro. Mech. Syst.*, 1996, pp. 456–461.
- [18] H. Das, H. Zak, J. Johnson, J. Crouch, and D. Frambach, "Evaluation of a telerobotic system to assist surgeons in microsurgery," *Computer Aided Surgery*, vol. 4, no. 1, pp. 15–25, 1999.
- [19] 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.
- [20] 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.
- [21] 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.

- [22] 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. Health Technol. Inform.*, vol. 142, no. 1, pp. 215–7, 2009.
- [23] 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.
- [24] M. Balicki, A. Uneri, I. Iordachita, J. Handa, P. Gehlbach, and R. Taylor, "Micro-force sensing in robot assisted membrane peeling for vitreoretinal surgery.," in *Proc. Medical Image Computing and Computer-Assisted Intervention*, 2010, vol. 13, no. Pt 3, pp. 303–310.
- [25] 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.
- [26] 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.
- [27] R. A. MacLachlan, B. C. Becker, J. Cuevas Tabares, G. W. Podnar, L. A. Lobes Jr., and C. N. Riviere, "Micron: An actively stabilized handheld tool for microsurgery," *IEEE Trans. Robot.*, vol. 28, no. 1, pp. 195–212, 2012.
- [28] K. Frischmann, "Micromanipulator," .
- [29] R. C. Harwell and R. L. Ferguson, "Physiologic tremor and microsurgery," *Microsurgery*, vol. 4, no. 3, pp. 187–192, 1983.
- [30] J. Binder and W. Kramer, "Robotically assisted laparoscopic radical prostatectomy," *BJU International*, vol. 87, no. 4, pp. 408–410, 2001.
- [31] B. Mitchell, J. Koo, M. Iordachita, P. Kazanzides, A. Kapoor, J. Handa, G. Hager, and R. Taylor, "Development and application of a new steady-hand manipulator for retinal surgery," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2007, pp. 623–629.
- [32] C. N. Riviere, W. T. Ang, and P. K. Khosla, "Toward active tremor canceling in handheld microsurgical instruments," *IEEE Trans. Robot. Autom.*, vol. 19, no. 5, pp. 793–800, 2003.
- [33] I. W. Hunter, T. D. Doukoglou, S. R. Lafontaine, P. G. Charette, L. A. Jones, M. A. Sagar, G. D. Mallinson, and P. J. Hunter, "A teleoperated microsurgical robot and associated virtual environment for eye surgery," *Presence*, vol. 2, no. 4, pp. 265–280, 1993.
- [34] P. S. Jensen, K. W. Grace, R. Attariwala, J. E. Colgate, and M. R. Glucksberg, "Toward robot-assisted vascular microsurgery in the retina," *Graefes Arch. Clin. Exp. Ophthalmol.*, vol. 235, no. 11, pp. 696–701, 1997.
- [35] S. E. Salcudean and J. Yan, "Towards a force-reflecting motion-scale system for microsurgery," in *IEEE International Conference on Robotics and Automation*, pp. 2296–2301.
- [36] I. Fleming, M. Balicki, J. Koo, I. Iordachita, B. Mitchell, J. Handa, G. Hager, and R. Taylor, "Cooperative robot assistant for retinal microsurgery," *Proc. Med. Image Comput. Comput. Assist. Interv.*, vol. 5242, pp. 543–50, 2008.
- [37] G. Stetten, B. Wu, R. Klatzky, J. Galeotti, M. Siegel, R. Lee, F. Mah, A. Eller, J. Schuman, and R. Hollis, "Hand-held force magnifier for surgical instruments," in *Information Processing in Computer-Assisted Interventions*, vol. 6689, R. Taylor and G.-Z. Yang, Eds. Springer Berlin / Heidelberg, 2011, pp. 90–100.
- [38] I. Iordachita, Z. Sun, M. Balicki, J. Kang, S. Phee, J. Handa, P. Gehlbach, and R. Taylor, "A sub-millimetric, 0.25 mN resolution fully integrated fiber-optic force-sensing tool for retinal microsurgery," *International Journal of Computer Assisted Radiology and Surgery*, vol. 4, no. 4, pp. 383–390, 2009.
- [39] 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.
- [40] J. J. Abbott, G. D. Hager, and A. M. Okamura, "Steady-hand teleoperation with virtual fixtures," in *Proc. IEEE International Workshop on Robot and Human Interactive Communication*, 2003, pp. 145–151.

- [41] A. Kapoor, M. Li, and R. H. Taylor, "Constrained control for surgical assistant robots," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2006, pp. 231–236.
- [42] 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.
- [43] C. N. Riviere, R. S. Rader, and N. V. Thakor, "Adaptive cancelling of physiological tremor for improved precision in microsurgery," *IEEE Trans. on Biomedical Engineering*, vol. 45, no. 7, pp. 839–846, 1998.
- [44] C. N. Riviere and N. V. Thakor, "Modeling and canceling tremor in human-machine interfaces," *IEEE Eng. Med. Biol. Mag.*, vol. 15, no. 3, pp. 29–36, 1996.
- [45] B. C. Becker, R. A. MacLachlan, L. A. Lobes Jr, and C. N. Riviere, "Semiautomated intraocular laser surgery using handheld instruments," *Lasers Sur. Med.*, vol. 42, no. 3, pp. 264–273, 2010.
- [46] 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.
- [47] W. T. Latt, U. X. Tan, C. Y. Shee, C. N. Riviere, and W. T. Ang, "Compact sensing design of a handheld active tremor compensation instrument," *IEEE Sensors Journal*, vol. 9, no. 12, pp. 1864–1871, 2009.
- [48] R. A. MacLachlan and C. N. Riviere, "High-speed microscale optical tracking using digital frequency-domain multiplexing," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 6, pp. 1991–2001, 2009.
- [49] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *J. Opt. Soc. of Am. A*, vol. 4, no. 4, pp. 629–642, 1987.
- [50] C. Shi, B. Becker, and C. Riviere, "Inexpensive monocular pico-projector-based augmented reality display for surgical microscope," in *Intl. Symp. Comp. Based Med. Sys.*, 2012.
- [51] Z. Pezzementi, S. Voros, and G. D. Hager, "Articulated object tracking by rendering consistent appearance parts," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2009, pp. 3940–3947.
- [52] N. Ahmidi, G. Hager, L. Ishii, G. Fichtinger, G. Gallia, and M. Ishii, "Surgical task and skill classification from eye tracking and tool motion in minimally invasive surgery," *Proc. Med. Image Comput. Comput. Assist. Interv.*, pp. 295–302, 2010.
- [53] C. Reiley, H. Lin, D. Yuh, and G. Hager, "Review of methods for objective surgical skill evaluation," *Surgical Endoscopy*, vol. 25, no. 2, pp. 1–11, 2011.
- [54] M. Li, M. Ishii, and R. H. Taylor, "Spatial motion constraints using virtual fixtures generated by anatomy," *IEEE Trans. Robot.*, vol. 23, no. 1, pp. 4–19, 2007.
- [55] M. Li, A. Kapoor, and R. Taylor, "Telerobotic control by virtual fixtures for surgical applications," *Advances in Telerobotics*, vol. 31, pp. 381–401, 2007.
- [56] S. Park, R. Howe, and D. Torchiana, "Virtual fixtures for robotic cardiac surgery," in *Proc. Med. Image Comput. Comput. Assist. Interv.*, 2001, pp. 1419–1420.
- [57] S. Rizzo, F. Patelli, and D. R. Chow, *Vitreo-Retinal Surgery: Progress III*. Springer, 2008.
- [58] F. Holtz and R. F. Spaide, *Medical Retina*. Berlin: Springer, 2007.
- [59] D. Kasner, G. R. Miller, W. H. Taylor, R. J. Sever, and W. Norton, "Surgical treatment of amyloidosis of the vitreous," *Trans Am Acad Ophthal Otolaryngol*, vol. 72, no. 3, p. 410, 1968.
- [60] R. Macherner, "The development of pars plana vitrectomy: A personal account," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 233, no. 8, pp. 453–468, 1995.
- [61] D. A. Infeld and J. G. O'Shea, "Diabetic retinopathy," *Postgrad. Med. J.*, vol. 74, no. 869, pp. 129–133, 1998.
- [62] R. N. Frank, "Retinal laser photocoagulation: Benefits and risks," *Vision Research*, vol. 20, no. 12, pp. 1073–1081, 1980.
- [63] P. Leaver and C. Williams, "Argon laser photocoagulation in the treatment of central serous retinopathy," *The British Journal of Ophthalmology*, vol. 63, no. 10, p. 674, 1979.
- [64] C. Sanghvi, R. McLauchlan, C. Delgado, L. Young, S. J. Charles, G. Marcellino, and P. E. Stanga, "Initial experience with the Pascal photocoagulator: A pilot study of 75 procedures," *British Journal of Ophthalmology*, vol. 92, no. 8, p. 1061, 2008.

- [65] M. S. Markow, Y. Yang, A. J. Welch, H. G. Rylander III, and W. S. Weinberg, "An automated laser system for eye surgery," *IEEE Eng Med Biol Mag*, vol. 8, pp. 24–29, 1989.
- [66] M. S. Blumenkranz, D. Yellachich, D. E. Andersen, M. W. Wiltberger, D. Mordaunt, G. R. Marcellino, and D. Palanker, "Semiautomated pattern scanning laser for retinal photocoagulation," *Retina*, vol. 26, no. 3, pp. 370–376, 2006.
- [67] A. J. Packer, *Manual of Retinal Surgery*. Butterworth-Heinemann Medical, 2001.
- [68] S. Rogers, R. L. McIntosh, N. Cheung, L. Lim, J. J. Wang, P. Mitchell, J. W. Kowalski, H. Nguyen, and T. Y. Wong, "The prevalence of retinal vein occlusion: Pooled data from population studies from the United States, Europe, Asia, and Australia," *Ophthalmology*, vol. 117, no. 2, pp. 313–19.e1, 2010.
- [69] B. Nilufer and B. Cosar, "Surgical treatment of central retinal vein occlusion," *Acta Ophthalmologica*, vol. 86, no. 3, pp. 245–52, 2008.
- [70] 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.
- [71] N. Feltgen, B. Junker, H. Agostini, and L. L. Hansen, "Retinal endovascular lysis in ischemic central retinal vein occlusion:: one-year results of a pilot study," *Ophthalmology*, vol. 114, no. 4, pp. 716–723. e1, 2007.
- [72] B. Alif and E. de Juan Jr., "In vivo cannulation of retinal vessels," *Graefes Archive for Clinical and Experimental Ophthalmology*, vol. 225, no. 3, pp. 221–225, 1987.
- [73] M. K. Tsilimbaris, E. S. Lit, and D. J. D'Amico, "Retinal microvascular surgery: A feasibility study," *Invest. Ophthalmol. Vis. Sci.*, vol. 45, no. 6, pp. 1963–1968, 2004.
- [74] W. M. Tang and D. P. Han, "A study of surgical approaches to retinal vascular occlusions," *Arch. Ophthalmol.*, vol. 118, no. 1, pp. 138–43, 2000.
- [75] J. N. Weiss, "Apparatus and method for cannulating retinal blood vessels," .
- [76] A. M. Jousseaume, T. W. Gardner, and B. Kirchhof, *Retinal Vascular Disease*. Berlin, Heidelberg: Springer Verlag, 2007.
- [77] M. R. Glucksberg, R. Dunn, and C. P. Giebs, "In vivo micropuncture of retinal vessels," *Graefes Archive for Clinical and Experimental Ophthalmology*, vol. 231, no. 7, pp. 405–407, 1993.
- [78] N. E. Kelly and R. T. Wendel, "Vitreous surgery for idiopathic macular holes: Results of a pilot study," *Arch. Ophthalmol.*, vol. 109, no. 5, p. 654, 1991.
- [79] 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.
- [80] W. E. Smiddy, W. Feuer, and G. Cordahi, "Internal limiting membrane peeling in macular hole surgery," *Ophthalmology*, vol. 108, no. 8, pp. 1471–1476, 2001.
- [81] C. Haritoglou, C. A. Gass, M. Schaumberger, A. Gandorfer, M. W. Ulbig, and A. Kampik, "Long-term follow-up after macular hole surgery with internal limiting membrane peeling," *Am. J. Ophthalmol.*, vol. 134, no. 5, pp. 661–666, 2002.
- [82] D. Weinberg, D. G. Dodwell, and S. A. Fern, "Anatomy of arteriovenous crossings in branch retinal vein occlusion," *Am. J. Ophthalmol.*, vol. 109, no. 3, pp. 298–302, 1990.
- [83] J. Zhao, S. M. Sastry, R. D. Sperduto, E. Y. Chew, and N. A. Remaley, "Arteriovenous crossing patterns in branch retinal vein occlusion," *Ophthalmology*, vol. 100, no. 3, p. 423, 1993.
- [84] M. D. Osterloh and S. Charles, "Surgical decompression of branch retinal vein occlusions," *Arch. Ophthalmol.*, vol. 106, no. 10, pp. 1469–1471, 1988.
- [85] R. R. Lakhanpal, M. Javaheri, H. Ruiz-Garcia, and J. DE, "Transvitreal limited arteriovenous-crossing manipulation without vitrectomy for complicated branch retinal vein occlusion using 25-gauge instrumentation," *Retina*, vol. 25, no. 3, p. 272, 2005.
- [86] M. E. Opremcak and R. A. Bruce, "Surgical decompression of branch retinal vein occlusion via arteriovenous crossing sheathotomy: A prospective review of 15 cases," *Retina*, vol. 19, no. 1, p. 1, 1999.
- [87] F. Becquet, J. F. Le Rouic, X. Zanlonghi, P. Peronnet, E. Hermouet-Leclair, C. Pousset-Decre, and D. Ducournau, "Efficiency of surgical treatment for chronic macular edema due to branch retinal vein occlusion," *Journal Français d'Ophthalmologie*, vol. 26, no. 6, p. 570, 2003.

- [88] N. Feltgen, C. Auw-Haedrich, R. Buchen, and L. L. Hansen, "Arteriovenous dissection in a living human eye: Clinicopathologic correlation," *Arch. Ophthalmol.*, vol. 123, no. 4, p. 571, 2005.
- [89] N. Feltgen, J. Herrmann, H. Agostini, A. Sammain, and L. L. Hansen, "Arterio-venous dissection after isovolaemic haemodilution in branch retinal vein occlusion: a non-randomised prospective study," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 244, no. 7, pp. 829–835, 2006.
- [90] J. Cinats, D. C. REID, and J. B. HADDOW, "A biomechanical evaluation of sorbothane," *Clinical Orthopaedics and Related Research*, vol. 222, p. 281, 1987.
- [91] Y. R. Bulsara and I. R. Matthew, "Forces transmitted through a laminated mouthguard material with a Sorbothane insert," *Dental Traumatology*, vol. 14, no. 1, pp. 45–47, 1998.
- [92] 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*, vol. 24, no. 3, pp. 427–434, 2004.
- [93] Y. Ida, N. Sugita, T. Ueta, Y. Tamaki, K. Tanimoto, and M. Mitsuishi, "Microsurgical robotic system for vitreoretinal surgery," *Intl. J. Comp. Asst. Rad. Surg.*, vol. 7, no. 1, pp. 27–34, 2011.
- [94] S. E. Everett and R. V. Dubey, "Vision-based end-effector alignment assistance for teleoperation," in *Proc. IEEE Int. Conf. Robot. Autom.*, 1999, pp. 543–549.
- [95] L. B. Rosenberg, "Virtual fixtures: Perceptual tools for telerobotic manipulation," in *IEEE Virt. Reality Ann. Int. Symp.*, 1993, pp. 76–82.
- [96] J. Funda, R. H. Taylor, B. Eldridge, S. Gomory, and K. G. Gruben, "Constrained cartesian motion control for teleoperated surgical robots," *IEEE Trans. Robot. Autom.*, vol. 12, no. 3, pp. 453–465, 1996.
- [97] B. L. Davies, S. J. Harris, W. J. Lin, R. D. Hibberd, R. Middleton, and J. C. Cobb, "Active compliance in robotic surgery—the use of force control as a dynamic constraint," *Proc. Inst. Mech. Eng., Part H: J. Eng. Med.*, vol. 211, no. 4, pp. 285–292, 1997.
- [98] R. Kumar, G. D. Hager, A. Barnes, P. Jensen, and R. H. Taylor, "An augmentation system for fine manipulation," in *Conference on Medical Image Computing and Computer-Assisted Intervention*, 2000, pp. 956–965.
- [99] M. Dewan, P. Marayong, A. M. Okamura, and G. D. Hager, "Vision-based assistance for ophthalmic micro-surgery," in *Proc. Med. Image Comput. Comput. Assist. Interv.*, 2004, pp. 49–57.
- [100] J. Ren, R. V. Patel, K. A. McIsaac, G. Guiraudon, and T. M. Peters, "Dynamic 3-D virtual fixtures for minimally invasive beating heart procedures," *IEEE Trans. Med. Imag.*, vol. 27, no. 8, pp. 1061–1070, 2008.
- [101] 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.
- [102] B. C. Becker, R. A. MacLachlan, and C. N. Riviere, "State estimation and feedforward tremor suppression for a handheld micromanipulator with a Kalman filter," in *Proc. IEEE Intl. Conf. Intell. Robot. Syst.*, 2011, pp. 5160–5165.
- [103] K. C. Veluvolu, W. T. Latt, and W. T. Ang, "Double adaptive bandlimited multiple Fourier linear combiner for real-time estimation/filtering of physiological tremor," *Biomed. Sig. Proces. Con.*, vol. 5, no. 1, pp. 37–44, 2010.
- [104] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 651–670, 1996.
- [105] Y. L. Ma and W. T. Hewitt, "Point inversion and projection for nurbs curve and surface: Control polygon approach," *Computer Aided Geometric Design*, vol. 20, no. 2, pp. 79–99, 2003.
- [106] X. D. Chen, J. H. Yong, G. Wang, J. C. Paul, and G. Xu, "Computing the minimum distance between a point and a nurbs curve," *Computer-Aided Design*, vol. 40, no. 10–11, pp. 1051–1054, 2008.
- [107] R. Richa, P. Poinet, and Chao Liu, "Three-dimensional motion tracking for beating heart surgery using a thin-plate spline deformable model," *Intl. J. Robot. Res.*, vol. 29, no. 2–3, pp. 218–230, 2010.
- [108] "OpenCV Library." [Online]. Available: <http://opencv.willowgarage.com/>.
- [109] H. Hirschmuller, "Stereo processing by semiglobal matching and mutual information," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 328–341, 2008.

- [110] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, 1992.
- [111] "GNU Scientific Library." [Online]. Available: <http://www.gnu.org/s/gsl/>.
- [112] J. Timmer, "Modeling noisy time series: Physiological tremor," *Intl. J. Bifurcation and Chaos*, vol. 8, no. 7, pp. 1505–1516, 1998.
- [113] J. Zhang and F. Chu, "Real-time modeling and prediction of physiological hand tremor," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2005, vol. 5.
- [114] J. M. Spyers-Ashby, P. G. Bain, and S. J. Roberts, "A comparison of fast fourier transform (FFT) and autoregressive (AR) spectral estimation techniques for the analysis of tremor data," *Journal of Neuroscience Methods*, vol. 83, no. 1, pp. 35–43, Aug. 1998.
- [115] A. Neumaier and T. Schneider, "Estimation of parameters and eigenmodes of multivariate autoregressive models," *ACM Transactions on Mathematical Software*, vol. 27, no. 1, pp. 27–57, Mar. 2001.
- [116] J. Makhoul, "Linear prediction: A tutorial review," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 561–580, 1975.
- [117] C. Gantert, J. Honerkamp, and J. Timmer, "Analyzing the dynamics of hand tremor time series," *Biological Cybernetics*, vol. 66, no. 6, pp. 479–484, Apr. 1992.
- [118] G. G. Walter and Y. Miao, "Wavelet neural networks for function learning," *IEEE Transactions on Signal Processing*, vol. 43, no. 6, pp. 1485–1497, Jun. 1995.
- [119] D. Simon, *Optimal State Estimation: Kalman, H [infinity] and Nonlinear Approaches*. John Wiley and Sons, 2006.
- [120] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. of Bas. Engr.*, vol. 82, no. 1, pp. 35–45, 1960.
- [121] G. Welch and G. Bishop, "An introduction to the Kalman filter," University of North Carolina, Chapel Hill, NC, 1995.
- [122] K. Dorfmüller-Ulhaas, "Robust optical user motion tracking using a Kalman filter," in *Proc. ACM Symp. on Virt. Real. Soft. Tech.*, 2003.
- [123] A. P. L. Bo, P. Poignet, F. Widjaja, and W. T. Ang, "Online pathological tremor characterization using extended Kalman filtering," in *Proc. Conf. IEEE Eng. Med. Biol. Soc.*, 2008, pp. 1753–1756.
- [124] R. Richa, M. Balicki, E. Meisner, R. Sznitman, R. Taylor, and G. Hager, "Visual tracking of surgical tools for proximity detection in retinal surgery," in *Proc. Info. Process. Comp. Assist. Interv.*, vol. 6689, Springer Berlin / Heidelberg, 2011, pp. 55–66.
- [125] M.-H. Yang and N. Ahuja, "Gaussian mixture model for human skin color and its applications in image and video databases," *Proc. SPIE:Storage and Retrieval for Image and Video Databases*, vol. 3656, pp. 458–466, 1998.
- [126] J. Bruce, T. Balch, and M. Veloso, "Fast and inexpensive color image segmentation for interactive robots," in *Proc. IEEE Intl. Conf. Intell. Robot. Syst.*, 2000, pp. 2061–2066.
- [127] M. K. Hu, "Visual pattern recognition by moment invariants," *Information Theory, IRE Transactions on*, vol. 8, no. 2, pp. 179–187, 1962.
- [128] F. Chaumette and S. Hutchinson, "Visual servo control, part I: Basic approaches," *IEEE Rob. Autom. Mag.*, vol. 13, no. 4, pp. 82–90, 2006.
- [129] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, UK: Cambridge University Press, 2003.
- [130] M. A. Fischler and R. C. Bolles, "Random sample consensus: Paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [131] 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.
- [132] 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.

- [133] 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.
- [134] 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.
- [135] 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.
- [136] C. Alonso-Montes, D. L. Vilariño, P. Dudek, and M. G. Penedo, "Fast retinal vessel tree extraction: A pixel parallel approach," *International Journal of Circuit Theory and Applications*, vol. 36, no. 5–6, pp. 641–651, Jul. 2008.
- [137] 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.
- [138] 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.
- [139] 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.
- [140] 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.
- [141] 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.
- [142] R. Perfetti, E. Ricci, D. Casali, and G. Costantini, "Cellular neural networks with virtual template expansion for retinal vessel segmentation," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 54, no. 2, pp. 141–145, Feb. 2007.
- [143] 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.
- [144] 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.
- [145] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [146] 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.
- [147] Y. Wang, J. Shen, W. Liao, and L. Zhou, "Automatic fundus images mosaic based on SIFT feature," *Journal of Image and Graphics*, vol. 6, no. 4, pp. 533–537, 2011.
- [148] 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.
- [149] 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.
- [150] 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 Transactions on Biomedical Engineering*, vol. 45, no. 1, pp. 105–118, 1998.
- [151] 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.

- [152] 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 Transactions on Biomedical Engineering*, vol. 58, no. 10, pp. 2816–24, Oct. 2011.
- [153] 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.
- [154] 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.
- [155] 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 Transactions on Biomedical Engineering*, vol. 57, no. 7, pp. 1707–18, Jul. 2010.
- [156] 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 transactions on bio-medical engineering*, vol. 58, no. 10, pp. 2816–24, Oct. 2011.
- [157] 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.
- [158] 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.
- [159] 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.
- [160] P. Gupta, P. Jensen, and E. de Juan, "Surgical forces and tactile perception during retinal microsurgery," in *Proc. Med. Image Comput. Comput. Assist. Interv.*, 1999, pp. 1218–1225.
- [161] N. Huynh, M. Akbari, and J. I. Loewenstein, "Tactile feedback in cataract and retinal surgery: A survey-based study," *J Acad Ophthalmol*, vol. 1, no. 2, pp. 79–85, 2008.
- [162] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Intl. J Comp. Vision*, vol. 47, no. 1, pp. 7–42, 2002.
- [163] D. Modi, P. Chirand, and L. Akduman, "Efficacy of patterned scan laser in treatment of macular edema and retinal neovascularization," *Clin Ophthalmol.*, vol. 3, pp. 465–470, 2009.
- [164] B. Zitova and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.
- [165] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.
- [166] J. V. Hajnal, D. J. Hawkes, and D. L. G. Hill, *Medical Image Registration*. CRC, 2001.
- [167] J. Batlle, E. Mouaddib, and J. Salvi, "Recent progress in coded structured light as a technique to solve the correspondence problem: a survey," *Pattern Recognition*, vol. 31, no. 7, p. 977, 1998.
- [168] C. Carswell, D. Clarke, and W. B. Seales, "Assessing mental workload during laparoscopic surgery," *Surgical Innovation*, vol. 12, no. 1, pp. 80–90, 2005.
- [169] R. Sandra, W. Joan, and O. Semin, "Review of the ocular angiogenesis animal models," *Semin Ophthalmol*, vol. 24, no. 2, pp. 52–61, 2009.
- [170] S. Baker and I. Matthews, "Lucas-kanade 20 years on: A unifying framework," *Intl J of Computer Vision*, vol. 56, no. 3, pp. 221–255, 2004.
- [171] L. B. Rosenberg, "The use of virtual fixtures as perceptual overlays to enhance operator performance in remote environments," USAF Armstrong Laboratory, 1993.
- [172] C. A. Moore Jr., M. A. Peshkin, and J. E. Colgate, "Cobot implementation of virtual paths and 3D virtual surfaces," *IEEE Trans Robot. Autom.*, vol. 19, no. 2, pp. 347–351, 2003.
- [173] A. P. L. Bo, P. Poignet, and C. Geny, "Filtering voluntary motion for pathological tremor compensation," in *Proc. IEEE Intl. Conf. Intell. Robot. Syst.*, 2009, pp. 55–60.
- [174] K. C. Veluvolu, U. X. Tan, W. T. Latt, C. Y. Shee, and W. T. Ang, "Bandlimited multiple Fourier linear combiner for real-time tremor compensation.," in *Proc. Conf. IEEE Eng. Med. Biol. Soc.*, 2007, pp. 2847–50.

- [175] A. Ghorbanian, M. Zareinejad, S. M. Rezaei, H. Sheikhzadeh, and K. Baghestan, "A novel control architecture for physiological tremor compensation in teleoperated systems," *International Journal of Medical Robotics and Computer Assisted Surgery*.
- [176] S. Tatinati, K. C. Veluvolu, and W. T. Ang, "Autoregressive model with Kalman filter for estimation of physiological tremor in surgical robotic applications," in *International Conference on Control, Automation and Systems*, 2011, pp. 454–459.
- [177] F. Zhe, W. Yubo, and K. C. Veluvolu, "Adaptive estimation of bandlimited physiological signals in real-time," in *International Conference on Control, Automation and Systems*, 2011, pp. 167–171.
- [178] Y. Sungwook, R. A. MacLachlan, and C. N. Riviere, "Design and analysis of 6 DOF handheld micromanipulator," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2012, pp. 1946–1951.
- [179] Y. Sungwook, M. Balicki, R. A. MacLachlan, X. Liu, J. U. Kang, T. R. H, and C. N. Riviere, "Optical coherence tomography scanning with a handheld vitreoretinal micromanipulator," in *Proc. Conf. IEEE Eng. Med. Biol. Soc.*, 2012, p. in press.
- [180] G. M. Grande, A. J. Knisely, B. C. Becker, S. Yang, B. E. Hirsch, and C. N. Riviere, "Handheld micromanipulator for robot-assisted stapes footplate surgery," in *Proc. Conf. IEEE Eng. Med. Biol. Soc.*, 2012, p. in press.
- [181] S. J. Garg and J. Vander, "Surgery for Retinal Venous Occlusions," *Review of Ophthalmology*, 2003.