Techniques for robot-aided intraocular surgery using monocular vision

Sungwook Yang[1], **Joseph N Martel**[2], **Louis A Lobes Jr**[2], and **Cameron N Riviere**[3]

[1]Center for BioMicrosystems, Korea Institute of Science and Technology, Korea

[2]Department of Ophthalmology, University of Pittsburgh, Pittsburgh, USA

[3]Robotics Institute, Carnegie Mellon University, Pittsburgh, USA

Abstract

This paper presents techniques for robot-aided intraocular surgery using monocular vision in order to overcome erroneous stereo reconstruction in an intact eye. We propose a new retinal surface estimation method based on a structured-light approach. A handheld robot known as the Micron enables automatic scanning of a laser probe, creating projected beam patterns on the retinal surface. Geometric analysis of the patterns then allows planar reconstruction of the surface. To realize automated surgery in an intact eye, monocular hybrid visual servoing is accomplished through a scheme that incorporates surface reconstruction and partitioned visual servoing. We investigate the sensitivity of the estimation method according to relevant parameters and also evaluate its performance in both dry and wet conditions. The approach is validated through experiments for automated laser photocoagulation in a realistic eye phantom in vitro. Finally, we present the first demonstration of automated intraocular laser surgery in porcine eyes ex vivo.

Keywords

Medical robotics; surgery; reconstruction algorithms; visual servoing

1 Introduction

Robot-aided intraocular microsurgery is still challenging despite the large success of surgical robotics in a variety of medical applications (Fine et al., 2010). This is partly due to the small size of retinal structures. For instance, retinal vasculature is often less than 100 μm in diameter, and membranes such as the internal limiting membrane in the eye are around 5–10 μm thick (Brooks, 2000). In addition to such a high demand of precision, limited visualization and constrained access to the surgical lesion also hinder the application of conventional surgical robots in vitreoretinal surgery (Ida et al., 2012).

To overcome these limitations, new robotic platforms have been developed specifically for intraocular microsurgery. One of the first was the teleoperated robot-assisted microsurgery system of Das et al. (1999). Since then, a variety of types of teleoperated robots has been

Corresponding author: Cameron N Riviere, Robotics Institute, Carnegie Mellon University, Pittsburgh, USA., camr@ri.cmu.edu.

introduced for vitreoretinal surgery. Ueta et al., (2009) demonstrated intraocular surgery in an animal model using a newly developed teleoperated robot for retinal surgery. Later, this platform was improved to meet specific requirements for the degrees of freedom (DOF), accuracy, and workspace needed in vitreoretinal surgery (Ida et al., 2012). In addition, an interchangeable manipulator was also incorporated to facilitate use of various surgical tools. Rahimy et al., (2013) presented IRISS (intraocular robotic interventional surgical system), the first microsurgical platform capable of performing complete ophthalmic procedures, including both anterior and posterior segment surgery. More recently, a new IRISS robot system was developed by Wilson et al., (2017) in order to simultaneously manipulate multiple surgical instruments with a large range of motion. The systems could thus facilitate full procedures of surgical steps with simple tool changing. A compact type of teleoperated robots has also been developed (Nambi et al., 2016; Nasseri et al., 2013). These lightweight slave robots can be mounted on a patient's forehead to passively compensate for head movements. Nambi et al. (2016) developed a compact telemanipulated system for retinal surgery, incorporating a 6-DOF slave robot. The quick-change instrument adapter of the robot enables the system to use commercially available actuated instruments without modification during surgery. Meenink et al., (2012) also developed a compact and lightweight 6-DOF robot for robot-assisted vitreoretinal surgery.

Following these efforts to develop robotic platforms for vitreoretinal microsurgery, the world's first robot-assisted intraocular surgery was performed for membrane peeling using the Preceyes Surgical System in 2016 (MacLaren et al., 2017). The Steady Hand is a cooperative robot using shared control principles (Taylor et al., 1999; Üneri et al., 2010). The robot detects the operator's applied force or torque and selectively delivers fine motion to surgical tools while suppressing hand tremor (Üneri et al., 2010). A research group at the University of Leuven has presented robotic platforms that can be either telemanipulated or co-manipulated using the same robot manipulator; the robot can be controlled either remotely by a haptic joystick (Gijbels et al., 2014b) or directly by a surgeon's hand (Gijbels et al., 2014a; Willekens et al., 2017). The latter work showed the feasibility of robot-assisted retinal vein cannulation for retinal vein occlusion with cooperative manipulation of the robot; this has recently led to the first human trial to use a surgical robot for the treatment of retinal vein occlusion. As an alternative, to retain the surgeon's direct control and preserve the natural feel of manual operation, a handheld robot known as the Micron has been introduced (MacLachlan et al., 2012; Yang et al., 2015b). The Micron senses its own motion and filters erroneous motion, such as hand tremor, compensating unwanted motion using a micromanipulator built into its own tip.

For all these robotic platforms, image-guided intervention has increasingly been a focus of research, since it can improve safety, as well as increasing accuracy of operation during microsurgery (Buchs et al., 2013). For example, vision-based virtual fixtures can help to guide an end-effector to lesions to be treated, while preventing it from hitting normal and healthy tissue. Thus, collateral damage during vitreoretinal surgery can be minimized (Becker et al., 2013). A preliminary study of optical-coherence-tomography-guided intervention was also presented for robot-assisted ophthalmic surgery (Yu et al., 2013). Recently, Yu et al. (2016) presented a new virtual fixture scheme that integrates optical coherence tomography depth feedback with a vision-based virtual fixture for control. The

Micron has also demonstrated automated scanning with an intraocular optical coherence tomography probe scanning or a photocoagulation laser (Yang et al., 2012, 2014). The automated laser photocoagulation systems presented so far with the Micron necessitate registration of camera coordinates with global coordinates, followed by stereo reconstruction of the retinal surface, in order to utilize image-guidance in controlling the end-effector (Yang et al., 2014, 2015a, 2016). For the registration, we adopted a conventional camera calibration technique (direct linear transformation) and then reconstructed the retinal surface by triangulation of feature points on the surface using stereo cameras.

However, it is not feasible to apply the same techniques in an intact eyeball, including the cornea, the lens, and the vitreous (or the saline solution which replaces the vitreous during vitrectomy), since the camera calibration and resulting surface reconstruction are prone to failure due to considerable optical distortion and unreliable visual detection. Most calibration methods assume a classical perspective camera model in a single medium such as air, but this assumption does not hold in a complex eye entailing the refraction of the light (Bergeles et al., 2010). The optical path includes the cornea and lens; during surgery it also includes saline, with which the eye is filled after vitrectomy, and also a contact lens (or binocular indirect ophthalmomicroscope) to provide a wide-angle view during operation. Therefore, image-guided intraocular microsurgery that relies on conventional camera calibration and resulting surface reconstruction, may lead to failure of control in the realistic surgical environment; to date, only limited demonstration in dry eye phantoms has yet been shown.

Richa et al. (2012) proposed vision-based proximity detection of surgical tools in vitreoretinal surgery in order to prevent accidental collision between surgical tools and the retina. The proximity is detected by taking relative stereo disparity between surgical tools and the retinal surface, which is used to provide guidance to surgeons for safe operation. However, the minimal depth that can be resolved by this method was not discussed, although it could apply to a relatively large range of proximity change. Since the relationship between disparity and depth is approximated by 5.3 pixels/mm, fine depth manipulation for automated intraocular surgery would not be possible with the proximity detection. Accordingly, automated operation such as visual servoing of the surgical tools was not explored; the disparity was used only for offering proximity warning for motion below the 2-mm threshold (10 pixels).

This difficulty in intraocular surgery has led to the development of new 3D localization methods for controlling a microrobot inside the eye, taking the unique optical characteristics into account (Bergeles et al., 2010, 2012). Bergeles et al. introduced a focus-based method accounting for the optics of the human eye in the imaging and localization of the microrobot with a single stationary camera (Bergeles et al., 2010). They adopted an optical model called the Navarro schematic eye, based on biometric data. The study showed the feasibility of the technique on a variety of ophthalmic microscopes even with uncertainties of optical parameters used in modeling. However, the localization error is limited to a few hundred micrometers, and drift compensation and servoing of the microrobot were not demonstrated due to problems with real-time operation (Bergeles et al., 2012). Visual servoing of the microrobot was recently demonstrated in a phantom eye model, under a setup similar to

ophthalmic microscopy (Bergeles et al., 2012). A new intraocular projection model (*Raxel*-based projection model) was introduced to accurately localize the microrobot in real time. However, the servoing error remains at the level of a few hundred micrometers, which is still too large for microsurgery. In addition, it is complicated to model the eye with a large set of optical parameters, each of which brings its own uncertainty.

In contrast to the microrobot relying on vision feedback for localization, Micron always provides the tool tip pose with respect to the global coordinates using a custom-built optical tracking system ("Apparatus to Sense Accuracy of Position," or ASAP). Consequently, we do not require localizing the tool tip, but do still need to register the imaging system in the ASAP coordinates to utilize vision feedback for control. This unique feature of the Micron system allows us to avoid complex modeling of the imaging system in intraocular surgery, since we can partition degrees of freedom (DOF) in controlling the tool tip using techniques such as hybrid visual servoing (Yang et al., 2016); the 3-DOF tip control can be decomposed into 2-DOF visual servoing parallel to the retinal surface and 1-DOF control along the axis of the tool. If the retinal surface can be approximated in the ASAP coordinates, we can use a monocular camera for control, instead of using a stereo pair of cameras and considering the complex optical system to reconstruct the surface in 3D space.

This paper therefore proposes a new method to estimate the retinal surface with a monocular camera, utilizing the scanning capability of the 6-DOF Micron. We introduce a structured light approach for the estimation and then analyze the projective geometry of aiming beam trajectories on camera images. The new method is validated in a realistic phantom model, while addressing issues raised by erroneous stereo reconstruction in the complex eye model. Finally, using the monocular surface estimation method proposed, we provide the first demonstration of robot-aided intraocular laser surgery in an intact porcine eye *ex vivo*, which had not been feasible previously. A cursory description of the technique was reported previously in a preliminary report of work toward a virtual fixture for retinal vessel cannulation (Mukherjee et al., 2017); the present paper provides the first systematic description of the technique, including both dual- and single-cone methods.

2 Handheld robot for intraocular microsurgery

The handheld robot system for image-guided intraocular surgery primarily consists of the active handheld robot, Micron, the vision system, and a laser system (Figure 1). Micron incorporates a miniature micromanipulator, a custom-built optical tracking system ("Apparatus to Sense Accuracy of Position," or ASAP), and a real-time controller. The micromanipulator provides an end-effector with six degrees of freedom (6-DOF) in actuation and with a cylindrical workspace 4 mm in diameter and 4 mm high (Yang, MacLachlan, et al., 2015). Given the 6-DOF actuation, it can impose a virtual remote center of motion (RCM) at the point of entry through the sclera, while independently positioning the tool tip over the retinal surface. Accordingly, the 6-DOF pose of the tool tip, including the 3-DOF position and the 3-DOF orientation, can be set for control; the 3D positions of the tip and the RCM in Euclidean space are used as control inputs to our control system. For control of Micron, ASAP tracks the position and orientation of the end-effector and handle at a sampling rate of 1 kHz with less than 10 μm RMS noise (MacLachlan and Riviere,

2009). The laser probe attached to the micromanipulator is accurately controlled in the specified 3-D workspace, while regarding undesired handle motion as a disturbance in control.

The vision system comprises an operating microscope, two charge-coupled device (CCD) cameras (a single CCD is used herein for surface reconstruction and control), and a desktop PC for image processing. Given image streams to the PC, the position of the aiming beam emitted from the laser probe is found via an ellipse-detection algorithm in OpenCV (OpenCV Library, 2018). The position of the aiming beam is then used for monocular surface reconstruction and visual servoing of the laser tip. Furthermore, the vision system is capable of tracking the retinal surface in real time, using the "EyeSLA M" algorithm (Braun et al., 2018) to compensate for eye movement; the tracking capability is essential for automated operation since the eye is moved both intentionally (by the surgeon, to view different areas) and unintentionally (by the patient) during intraocular microsurgery (Wright et al., 2006).

An Iridex Iriderm Diolite 532 Laser is interfaced with Micron, with a 23-gauge Iridex Endoprobe adapted to fit within Micron. Hence, the aiming beam (red light) emitted from the laser probe is utilized first for the reconstruction of retinal surface and later for visual servoing of the laser tip. For demonstration of automated subtasks in robot-aided intraocular surgery, we adopt laser photocoagulation as an exemplary application because it is a common treatment of retinal disorders such as diabetic retinopathy (Diabetic Retinopathy Study Research Group, 1981). In an automated version of the operation, patterned targets are planned on a preoperative image. The laser probe is then deflected to correct error between the aiming beam and the given target, while maintaining a constant standoff distance of the laser probe from the retinal surface. Once the distance between the aiming beam and the target comes within a specified targeting threshold, the laser is triggered to create the target lesion. This procedure is repeated until completion of all targets. The operator is instructed to hold the instrument still over the retinal surface, while keeping the handle position within a certain range of the initial position; the initial position is set at the beginning of automated execution. The overall surgical procedures are depicted in Figure 2.

3 Retinal surface estimation using monocular vision

We propose a new method to estimate the retinal surface with a monocular camera, by introducing predefined beam patterns on the surface. To find the surface in the ASAP coordinate with a single camera, we use the projective geometry of a beam trajectory created by scanning a laser probe, because the aiming beam is highly detectable in the eye, regardless of illumination change. The retinal surface is assumed to be parallel or near parallel to the image plane. As depth of field in intraocular microscopic imaging is quite shallow, focused images can be obtained only when the surface is located within a very narrow region perpendicular to the optical axis of the microscope (Zhou and Nelson, 1999). Hence, the projective geometry of the aiming beam is allowed by assuming the retinal surface to be locally planar in the area of interest. For instance, the planar assumption would incur a range of $\pm$ 45-um depth error for the average human eye with a diameter of 25 mm, while estimating a field of view of 3 mm in diameter (typically shown through an operating

microscope). Accordingly, this would be a good-enough approximation for a significantly smaller region of interest as in vitreoretinal microsurgery.

Regardless of optical distortion, the ray of an aiming beam always intersects the surface. For example, the circular motion of the laser probe with a remote center of motion (RCM) results in a conic shape of the beam trajectory about the nominal axis of the tool, as depicted in Figure 3. The trajectory subject to a conic section cut by the retinal surface is shown as an ellipse in the image plane, which is assumed to be parallel to the retinal surface. It is then found that the shape of the ellipse, i.e., the aspect ratio, is related to the tilt angle of a plane that cuts the cone. Therefore, our goal is to find a plane regarded as the retinal surface, in terms of a plane normal and a point that lies in the plane, using the relationship of the projected ellipses.

A similar interpretation utilizing such a circle-ellipse relationship has been investigated in computer vision and tomography (Chen et al., 2004; Noo et al., 2000). Chen *et al.* introduced a camera calibration method using two coplanar circles, analyzing the shape of ellipses shown on a camera (Chen et al., 2004). The relationship was also explored for the calibration of a cone-beam scanner used in both x-ray computed tomography and single-photon emission computed tomography (Noo et al., 2000). It also uses circular traces producing ellipses on the detector and the calibration geometry is determined analytically using the parametric description of these ellipses.

3.1 Projective geometry analysis

We use circular scanning of the laser probe to project an ellipse on the plane to be estimated. First, the laser probe is scanned to generate a circular pattern around a pivot point, which could be regarded as a RCM in vitreoretinal surgery. The ray from the laser probe results in a cone beam in 3D space as shown in Figure 3. Once the resulting trajectory is detected in a sequence of images, it is fitted as an ellipse. The fitted ellipse is then parameterized by the following descriptions:

c_e: The center of the ellipse;

m_a : The half length of the major axis;

m_b : The half length of the minor axis;

θ_e : The inclination angle from the *x*-axis of the image plane.

The ellipse is regarded as the conic section cut by the tilted plane to be estimated. Thus, the plane can be described by the rotation of the plane initially perpendicular to the axis of the cone. For instance, Figure 4 shows the tilted plane and corresponding ellipse on the conic section, given rotation about the *y*-axis of the cone represented by v_y. We then define the angle of rotation θ_{plane}, using the aspect ratio of the ellipse and the opening angle of the cone as in (1).

$$\theta_{plane} = \pm \operatorname{asin}\left(\sqrt{\frac{1 - r^2}{1 + \tan^2 \theta_{cone}}}\right), \quad (1)$$

where the aspect ratio γ and the opening angle θ_{cone} are defined in (2) and (3), respectively.

$$\gamma = \frac{m_b}{m_a} \quad (2)$$

$$\tan \theta_{cone} = \frac{r_{cone}}{h_{RCM}} \quad (3)$$

Given the tilt angle θ_{plane}, we estimate a point belonging to the plane as an offset d_{plane} from a vertex P_{RCM} along the axis of the cone. If the scale factor of the image s_{cam} is known, the point on the plane can be described as P_{plane} in the ASAP coordinates, using (4) and (5).

$$d_{plane} = \frac{\cos \theta_{plane}(1 - \tan^2 \theta_{plane} \tan^2 \theta_{cone})}{\tan \theta_{cone}} m_a s_{cam} \quad (4)$$

$$P_{plane} = P_{RCM} - d_{plane} v_{cone}, \quad (5)$$

where v_{cone} is a unit vector representing the axis of the cone. From the derivation, it is noted that the angle θ_{plane} related to the plane normal is scale-free to the images, whereas the point on the plane needs the image scale because of the value s_{cam}. Such an image scale can also be specified according to the zoom factor of an operating microscope.

However, the angle θ_{plane} itself does not uniquely describe the tilted plane in 3D, since the plane normal can be any vector, represents the tilting angle from the axis of the cone toward any direction; any arbitrary vector orthogonal to the axis of the cone can be taken as the axis of rotation, leading to an infinite set of normal vectors that yield the identical ellipse. This ambiguity is resolved as described below.

3.2 Retinal surface estimation

3.2.1. Surface normal estimation

3.2.1.1 Dual cone beam reconstruction: Since the surface normal would not uniquely be defined with a single ellipse itself, we propose dual cone beam reconstruction that utilizes two different ellipses resulting from two circle scans at the tip. Given two circular scans of

the tool tip, we obtain two projected ellipses on the image plane as shown in Figure 5. The resulting ellipses are then described by two sets of the angle and the point that describe the surface using (1) and (5).

$$\theta^i_{plane} \text{ and } P^i_{plane}, \text{ where } i \in \{1, 2\}. \quad (6)$$

Here, the absolute value of the angle is taken at this moment for further steps; the sign of the angle will be determined later. First, we set coordinate transformation R^{ASAP}_{Ci} for the ith cone that describes the coordinates of the cone with respect to the ASAP coordinates:

$$R^{ASAP}_{Ci} = \begin{bmatrix} v^i_x & v^i_y & v^i_z \end{bmatrix}, \quad (7)$$

where v^i_z is aligned with the axis of the ith cone, and is also identical to v_{cone} in (5). The tilted plane is then regarded as the rotation of the xy-plane defined in the coordinates of the ith cone, C_i. Hence, we initially define the normal vector of the plane as the rotation of the vector v^i_z about the y-axis v^i_y by the angle θ^i_{plane}:

$$n^i_{init} = R(\hat{v}^i_y, \theta^i_{plane})v^i_z. \quad (8)$$

As these normal vectors are not unique, the infinite sets of the normal vectors are interpreted as two 3D circles in a unit sphere as depicted in Figure 6. Therefore, the true normal vector must be located at the intersection of the two circles in the sphere. A degenerate case occurs if the axes of the two cones are parallel to each other; an infinite number of solutions would then exist for the normal.

To find the intersection of the two circles analytically, we introduce a supplementary plane that describes each circular trajectory in 3D. Each supplementary plane is then defined by a plane normal v^i_z and a point n^i_{init} in the unit sphere. Consequently, the intersection of the resulting planes is attained as a common line in 3D, which passes through the intersection of the two circles. The common line $l(s)$ is described by a unit vector u and a point p_0:

$$l(s) = p_0 + su, \quad (9)$$

where u is orthogonal to the two vectors, v^1_z and v^2_z, and p_0 is one of the points that lies along the line as in (10).

$$u = \frac{v_z^1 \times v_z^2}{\left\| v_z^1 \times v_z^2 \right\|} \text{ and } p_0 = \left[(n_{init}^1 \cdot v_z^1) v_z^2 - (n_{init}^2 \cdot v_z^2) v_z^1 \right] \times u \quad (10)$$

Since the common line also passes through the unit sphere, we calculate the intersection between the line and the unit sphere, instead of directly calculating the intersection of the two 3D circles:

$$\left\| p_0 + su \right\| = 1. \quad (11)$$

As a result, the solution subject to the surface normal is analytically derived in Cartesian coordinates, as in (12).

$$n_{plane} = p_0 + \left[-p_0{}^T u \pm \sqrt{\left(p_0{}^T u \right)^2 - \left(p_0{}^T p - 1 \right)} \right] u \quad (12)$$

Finally, we obtain one or two possible normal vectors n_{plane} for the estimated plane, depending on the number of intersections. If two normal vectors are found, one is identified as the true normal. The cross product of n_{plane} with a unit vector v is zero or nearly zero, where the vector v is defined by the two points on the plane as in (13).

$$n_{plane} \cdot v \approx 0, \text{ where } v = \frac{P_{plane}^1 - P_{plane}^2}{\left\| P_{plane}^1 - P_{plane}^2 \right\|}. \quad (13)$$

However, given real data, the method may fail in finding the intersection of two circles that represent sets of infinite normal vectors in 3D, as the estimated angle of the tilted plane in (1) could also be imprecise; such that imprecise estimation would yield incorrect circles in 3D. This could thus lead to either an incorrect surface normal or no solution on the unit sphere as depicted in Figure 7.

3.2.1.2 Single cone beam reconstruction: Therefore, we also propose an alternative method that incorporates point correspondences between 3D tip and 2D beam positions in a beam trajectory. For example, we assume that a data point is virtually projected on a plane from a ray resulting from the tip and RCM positions. We then utilize the fact that the phase angles of the data points with respect to the major axis of the virtual ellipse vary depending on which vector is chosen from the infinite set of normal vectors. Herein, the phase angle θ is an angle parameterizing a data point on the ellipse trajectory with the magnitudes (semi-major and semi minor axes) and the inclination angle of the ellipse from the x-axis of the image plane, as

$$[x, y] = [m_a \cos{(\theta + \theta_e)}, m_b \sin{(\theta + \theta_e)}].$$

Figure 8 visualizes each step of the surface estimation method (top) and resulting ellipse trajectory in the image plane (bottom). The first data point on each trajectory is marked as a circle. The first red line in Figure 8(a) shows the beam trajectory acquired from images. The green trajectory in Figure 8(b) is created by the initial estimation of the surface, assuming rotation of the xy-plane about the y-axis of the cone, where the green circle marker is located at the major axis of the ellipse in the image plane. Hence, it is noted that the phase angle of the green circle marker with respect to the major axis is different from the angle of the first data point on the red trajectory. If we could find a specific axis of rotation instead of the y-axis of the cone, the blue trajectory would be attained as in Figure 8(c), in which the phase angles of data points become identical to the actual angles from the red beam trajectory. Finally, the transverse coordinates of the plane are matched with the image coordinates as shown Figure 8(d). Therefore, our goal is to find the best normal vector that aligns the phase angles of projected data points with those angles of actual data points in the 2D beam trajectory acquired.

We use the projective geometry formulated in (1) and (5) for the initial surface estimation. Given the analysis of an ellipse trajectory in the image plane, an initial normal vector $\boldsymbol{n}_{init}$ of the surface is set, by rotating the axis of the cone $\boldsymbol{v}_z$ about the y-axis of the cone coordinates, $\boldsymbol{v}_y$, by the angle θ_{plane} as illustrated by the blue arrow in Figure 8(b):

$$\boldsymbol{n}_{init} = \boldsymbol{R}(\hat{\boldsymbol{v}}_y, \theta_{plane})\boldsymbol{v}_z, \quad (14)$$

where the coordinates of the cone beam are defined by coordinate transformation $\boldsymbol{R}_C^{ASAP} = [\boldsymbol{v}_x \ \boldsymbol{v}_y \ \boldsymbol{v}_z]$. We can then represent the infinite set of normal vectors resulting in the identical ellipse, as the rotation of the initial normal vector about the axis of the cone: a green line enclosing the axis of the cone in Figure 7(c). We thus regard the surface normal to be found $\boldsymbol{n}_{plane}$ as a rotation of the initial normal $\boldsymbol{n}_{init}$ about the axis of the cone as in (15),

$$\boldsymbol{n}_{plane} = \boldsymbol{R}(\hat{\boldsymbol{v}}_z, \theta_c)\boldsymbol{n}_{init}, \quad (15)$$

where the specific angle of rotation is denoted by θ_c.

Give the number k of data points on the trajectory, we define an objective function as in (16), in order to find the angle θ_c.

$$\underset{\theta_c \in [-\pi, +\pi]}{\arg\min} \sum_{k=1}^{n} \left(\theta_{beam}^k - \theta_{proj}^k\right)^2, \quad (16)$$

where the phase angles of the kth data point with respect to the major axis of each ellipse are denoted by θ^k_{beam} and θ^k_{proj} for the beam trajectory on the image and the estimated plane, respectively. First, we calculate the phase angle of each data point, θ^k_{beam}, using (17) and (18).

$$p^k_{beam} = p^k_{beam} - c_e \text{ and } \left[p^k_{bx} \; p^k_{by} \right]^T = R(-\theta_e)p^k_{beam}, \quad (17)$$

where p^k_{beam} is the kth data point of the beam trajectory in the image coordinates.

$$\theta^k_{beam} = \text{atan2}(p^k_{by}/m_b, p^k_{bx}/m_a) \quad (18)$$

In order to calculate the phase angle θ^k_{proj}, the projected point P^k_{proj} is found by intersection between the estimated plane and a ray created by the tip P^k_{tip} and RCM P_{RCM} positions.

$$P^k_{proj} = \Phi(P^k_{tip}, P_{RCM}, n_{plane}, P_{plane}) \quad (19)$$

Given a transformation, T^{plane}_{ASAP}, from the ASAP to the plane coordinates, the projected point belonging to the estimated plane is represented as a 2D point p^k_{proj}, by disregarding the z-component of the 3D vector p^k_{proj} on the plane:

$$p^k_{proj} \equiv \frac{1}{s_{cam}}T^{plane}_{ASAP}P^k_{proj}, \quad (20)$$

where the point is scaled down by the image scale s_{cam}. Once we are given a set of 2D points that forms a virtual ellipse on the estimated plane, the phase angle of the point, θ^k_{proj}, is calculated using the parameters of the projected ellipse, in the same manner as in (17) and (18).

As a result, we finally obtain the angle θ_c that minimizes the objective function in (16), using the Matlab™ function 'fminbnd.m,' which finds the minimum of a function within a fixed interval for a single variable. It should be noted that two possible solutions for the angle θ_c exist, due to the 'π' ambiguity of an ellipse; the inclination angle of the major axis is only defined within the range $[-\pi/2, \pi/2]$. In order to identify the true surface normal, we use the prior knowledge that the y-component of the surface normal should be negative, as the ASAP always sits head-down for receiving the LED light emitted from the Micron

handle; head-up orientation is impractical given that an operator would not hold the tool at an obtuse angle from the surface.

3.2.2. Coordinate mapping—The final step of the surface estimation is to find the principal vectors of the plane, u_{plane} and v_{plane}, which are aligned with the xy-image coordinates. Given the surface normal n_{plane} in (15), we define a 3D rotation R_v^n from the axis of the cone v_z to the normal vector n_{plane}. By applying the transformation R_v^n to the coordinate representation of the cone beam R_{C1}^{ASAP}, we attain a new coordinate representation $R_v^n R_{C1}^{ASAP}$; the resulting coordinates are depicted in the top of Fig. 4(c). Finally, the coordinate mapping is accomplished by applying rotation about the surface normal n_{plane} by the angle θ_{uv} as in (21),

$$R_{image}^{ASAP} = R(n_{plane}, \theta_{uv}) R_v^n R_{C1}^{ASAP}, \quad (21)$$

where $\theta_{uv} = \theta_{proj} - \theta_e$. The angle θ_{uv} is set by the difference between the inclination angles of the two ellipses, in order to align the major axes of those ellipses: θ_e from the actual beam trajectory, and θ_{proj} from projected ellipse. Finally, we obtain the retinal surface described with respect to the ASAP coordinates, in terms of the principal vectors of the plane, u_{plane} and v_{plane}, and the surface normal as (22).

$$R_{image}^{ASAP} = \begin{bmatrix} u_{plane} & v_{plane} & n_{plane} \end{bmatrix} \quad (22)$$

4 Evaluation of surface estimation

We first analyze the sensitivity of the surface reconstruction according to relevant parameters that may lead to error in estimation. The performance of the reconstruction is evaluated in the realistic eye models.

4.1 Characterization of surface estimation

We investigate how the tilt angle of the estimated plane varies upon the aspect ratio retrieved from the ellipse in the image plane, since the aspect ratio is a primary parameter that determines the angle of rotation for yielding the initial surface normal. In addition, the resulting angle does not change over the next procedures to find the final surface normal.

As shown in Figure 9, the angle of the scanning tip with respect to the estimated plane rapidly increases, as the aspect ratio of the ellipse increases, getting closer to a circle. Hence, as the tip becomes perpendicular to the surface, estimation of the angle is less resolvable. While retrieving the tilt angle of the plane from (1), the equation itself is robust enough over such a range of operation and no numerical deficiency is involved in calculations except ellipse fitting. A degenerate case would occur only if the tool tip were parallel to the retinal

surface, which is impossible in operation. In practice, the angle at which an operator holds the Micron handle is typically in the range of 40–65°, which yields the aspect ratios that reside in a range of 0.643—0.866. For example, a natural angle at which the tool is held without the scleral constraint is about 45°. The angle becomes larger when holding the tool in vitreoretinal surgery due to the location of a trocar used; the typical angle in use of the eye phantom is about 60°. Perturbation of the aspect ratio by 0.001 leads to the change of the tilt angle by 0.081° when Micron is held at 45°. Given the same perturbation at 60°, a change of 0.12° in the tilt angle is observed. Therefore, we can conclude that the tilt angle of interest can be estimated without complication. Parameters used for the simulation are taken from a typical setup and summarized in Table 1.

We also evaluate how the resulting surface normal and depth are affected by uncertainty in measurement of the major and minor axes of an ellipse. For simulation, the major and minor axes vary by ±2 pixels from the initial values described in Table 1. To determine the amount of pixel variance, we first consider the positioning accuracy of the robot, which has been found to be 10—20 μm for automated circle-scanning (Yang, MacLachlan, et al., 2015). This amount of positioning error corresponds to about ±2 pixels under the 10× magnification of the operating microscope used; the 10 repeated experiments in the §4.2.1 exhibit variation from −0.81 to 0.97 pixel for the major axis and from −0.40 to 0.47 pixel for the minor axis. Figure 10 shows the resulting errors in estimation of the surface normal and depth, regarding variation in estimation of the ellipse. According to the simulation, the angle error is found to be between −3.63° and 4.15°, leading to depth error between −609 μm and 675 μm. From this analysis, we can determine an acceptable threshold for running the RANSAC algorithm (Fischler and Bolles, 1981) to eliminate outliers on a scanned ellipse trajectory; details will be discussed later in the test of real data. Since the depth error is relatively high, compared to the angle error, the depth error is further investigated below.

Although we assume the image plane is parallel to the retinal surface in a small area of interest, it is not necessarily so. In the configuration of a stereomicroscope, two CCD cameras have a certain distance offset with respect to each other, in order to offer disparity for stereo depth perception, which results in tilt of their optical axes. Accordingly, error in estimation of the angle is to be expected. Furthermore, the depth estimation error can also be exacerbated by error in the image scale used in (4). We thus set the acceptable range of the angle variation as ±5° and uncertainty on the image scale as a range of ±5% from the nominal value for simulation. Figure 11 presents the variation in depth error according to the angle error in estimation and uncertainty on the image scale. According to the simulation, the depth error is found to be between −2111 μm and 2402 μm overall. Without uncertainty in the image scale, the depth error in estimation is 275 μm per degree error. When the image scale deviates by 1% from the nominal value, the depth error is increased by 177 μm; the depth error in estimation is proportional to the amount of uncertainty on the image scale.

It is thus found that the depth estimation of the surface is relatively sensitive, and could yield a large error, because it is determined by the large offset value d_{plane} from the RCM in (4). For instance, the offset d_{plane} defining a plane should be greater than the distance h_{RCM} (e.g., 20 mm) between the RCM and the tool tip. Accordingly, any small error in the detection of

an ellipse, the alignment of the optics, and/or image scale can amplify error in depth estimation.

These error analyses for the relevant parameters indicate how parameters should be precisely managed in order to obtain a certain level of accuracy in the surface estimation. First, the angle estimation should be as accurate as possible to lower the overall depth error. The use of higher magnification may improve estimation in the lengths of the major and minor axes of the ellipse fitting, which could improve the accuracy of the angle estimation. In addition, precise calibration of the image scale to below 1% error is necessary, which can be achieved by considering both linear zoom-scale ratios and nonlinear correction factors.

4.2 Tests with real data

The retinal surface estimation was also evaluated on real data acquired by circular scanning of the laser probe. Tests were conducted under two test conditions: "open-sky" and eye phantom environments. First, we performed open-sky tasks, which were designed to primarily evaluate performance of the surface estimation itself, without involving optical distortion and any physical constraint on the RCM (herein regarded as a pivot point for scanning). The estimation was then evaluated in the eye phantom model with water inside and a contact lens on top. Hence, the final goal of these tests is to investigate the feasibility of the retinal surface estimation in a realistic environment.

4.2.1. "Open Sky" tests—First, we need to determine the most effective size of scan diameter at the tool tip. For instance, if a smaller size is used for the scan, the resulting ellipse in the image plane could also be small, and may form an inaccurate shape of the ellipse due to low signal-to-noise ratio. On the other hand, if too large a size is used, the resulting trajectory might be degraded as the manipulator reaches the edge of its workspace. Hence, it is important to determine an appropriate size of scan diameter, for attaining reliable results.

Scan diameters for these tests were set in a range of 300–1500 μm as presented in Figure 12. All tests were performed under 10× magnification, for which the image scale is 8.6 μm per pixel. To exclude other effects, Micron was firmly affixed to a solid base. For comparison, a reference plane was reconstructed using stereo-vision, prior to the tests. This reference plane was thus regarded as the ground truth for investigating the angle difference between two surface normal vectors: one from the reference and the other from the estimated plane. In addition, we also defined depth error of the estimated plane with respect to the reference plane.

As shown in Figure 13, the surface normal was reliably estimated at scanning diameters greater than 750 μm. In contrast, large errors were produced by small scanning diameters, such as 300 and 500 μm. The depth error also shows a similar trend as the scanning diameter increases. Small error was thus found with scanning diameters greater than 1000 μm. From these results, the effective size of scanning diameter was determined as 1000–1250 μm for further experiments.

We then evaluated how reliably the reconstruction method could estimate the surface for multiple measurements. These tests were repeated for a total of 10 trials, with the manipulator firmly fixed. In addition to the repeatability test, the results of the cone beam reconstruction were compared, given ellipse trajectories obtained in both the left and right image plane. We also applied the reconstruction method on the combined ellipse trajectory, by taking averages of the corresponding ellipse parameters, because a typical operating microscope incorporates stereo cameras for depth perception. The point on the plane was also calculated using the average angle.

The average angle errors over the total 10 trials were measured to be 3–4° for all cases: left, right, and combined data. The combined data show the average behavior of the results obtained in the left and right image plane, for both angle and depth errors, as presented in Figure 14. Figure 14 also shows that the angle error obtained in the left image plane is larger than the error from the right plane for all repeated measurements, whereas the depth error is lower in the right image plane. However, these trends did not hold in the other surface estimation. In addition, it is found experimentally that depth estimation from the combined data is slightly more robust than results from a single camera, yielding lower standard deviations.

The last test in the open-sky setting was to investigate the reconstruction accuracy under handheld conditions, whereas all former tests were conducted in firmly fixed conditions. Since data obtained by handheld scanning are likely to be noisy, the RANSAC algorithm is adopted to remove outliers from the data, and retrieve an actual elliptical trajectory as closely as possible. A threshold for determining outliers was set at 20 μm by taking into account the positioning accuracy of the micromanipulator (Yang, MacLachlan, et al., 2015).

In total, 10 trajectories were obtained for the analysis. After applying the RANSAC algorithm, about 60% of the data points were used for fitting ellipses. The levels of average errors in both surface normal and depth estimation are similar to the results from the clamped tests. However, a larger standard deviation is observed in handheld tasks, compared to the clamped tests, because various locations and orientations of the scanning probe were taken at the beginning of each scan. The average angle and depth error were measured as 4.05±1.83° and 38±464 μm for the total 10 trials, which are still acceptable in use, compared to the accuracy of stereo reconstruction.

4.2.2. Eye phantom tests—The performance of the new surface reconstruction was evaluated in the eye phantom model. Before testing the surface reconstruction, the surface of a paper slide sitting on the bottom of the eye phantom was reconstructed using a stereo camera, in order to offer the ground truth plane. First, we tested the reconstruction method in the dry phantom. In contrast to the experiments performed open-sky, the laser probe was inserted into the eye phantom through a trocar and supported at the point of entry. The final test was conducted in the water-filled eye phantom, including a contact lens on the top, as shown in Figure 15(a), in order to evaluate the re-construction accuracy under optically distorted imaging conditions, in particular for depth; the image scale under the wet setting was calibrated before experiments. For evaluation, we compared the estimated plane via the

structured-light approach with the reference plane given by stereo reconstruction, in terms of the angle and depth errors.

For the total 10 trials in the dry phantom, the average angle error was measured as 7.73±3.26°; the depth error was −487±371 μm. Both angle and depth errors were larger than those attained in the flat surface with the scleral constraint. Moreover, the standard deviations were also larger than the results from the flat surface.

In the wet phantom model, imaging quality was relatively poor, and also the depth of field became shallow due to the contact lens used, compared to imaging in air. As a result, the average angle error is slightly larger than in the dry phantom, whereas depth error is significantly reduced, as shown in Figure 16. It is noted that we disregarded certain erroneous reconstruction results in the analysis, in which the tip angle with the respect to the plane was greater than 70°; such a high incident angle is ergonomically unnatural in holding Micron. It occasionally occurred in eye phantom tasks, leading to erroneous surface normal estimation, while it was not observed in the open-sky tasks. It can be caused primarily by the horizontal alignment of the contact lens, which is primarily assumed to be parallel to the retinal surface and the image plane. However, the lens could get tilted when a large rotation is introduced to the eyeball, in order to aim off-center. Furthermore, it is found that the higher the estimated incident angle is, the larger the angle error is. This is also true in stereo reconstruction used for creating the reference plane, especially in the eye phantom, due to inaccuracy in detection of the tool tip for the camera calibration. Accordingly, it may be unjustifiable to conclude that the depth error is really smaller in the wet eye phantom, compared to the dry phantom. However, for both cases, the surface normal could be estimated to an acceptable level, yielding angle errors less than 10°; such an angular error would lead to a tilted surface in estimation, resulting in ±260 μm error in depth at the edge of the surface for the area of interest (3 mm in diameter).

5 Hybrid visual servoing using monocular vision

We have proposed a hybrid control scheme for robot-aided intraocular laser surgery (Yang et al., 2016), to address the issues raised by position-based visual servoing (Yang et al., 2014). For instance, inaccurate camera calibration results in erroneous stereo reconstruction of the retinal surface, aiming beam, and placed targets in 3D. Consequently, this may lead to failure of the visual servoing, which is a known weakness of such position-based visual servoing (Hutchinson et al., 1996). These issues can be addressed by the hybrid visual servoing, as a compromise between position-based and image-based visual servo control (Malis et al., 1999, Corke and Hutchinson, 2001). Specifically, Castaño and Hutchinson (1994) introduced a hybrid vision/position control structure called *visual compliance*. They proposed a servoing scheme to control the 2-DOF motion parallel to an image plane using visual feedback, and the remaining degree of freedom (perpendicular to the image plane) using position feedback psrovided by encoders. In our hybrid control, the 3-DOF motion of the tool tip is decoupled into the 2-DOF planar motion parallel to the retinal surface and the 1-DOF motion along the axis of the tool. The decoupled 2-DOF motion is then controlled via image-based visual servoing, to position the laser aiming beam at a target. The 1-DOF

axial motion is controlled to maintain a constant standoff distance from the estimated retinal surface using position feedback provided by ASAP.

5.1 Image Jacobian for visual servoing

The desired state vectors for image-based visual servoing are defined by the position of the tool tip as control inputs at the ASAP coordinate system. Since the 3D tip position is used as a control input to the Micron system, our goal for control is to obtain an instance of the goal tip position, subject to minimizing the beam error. To obtain the desired state vectors in 3D from a single image feature (herein, an aiming beam), we divide this procedure into two steps: from the single 2D image feature to the 3D planar motion of the aiming beam, and from the 3D planar motion to the corresponding 3D tool tip motion in the ASAP coordinates. First, the desired planar motion of the aiming beam is defined at the task plane, where the task plane is subject to the surface found by the estimation method proposed. Given the planar motion of the aiming beam, we finally obtain the 3D tip position resulting in the motion of the aiming beam. Herein, the 2-DOF planar motion of the laser tip is thus allowed above the task space, where the tip motion is assumed to be parallel to the task space within a small range of motion.

To compute the desired planar motion from the single image feature, we introduce an intermediate interaction matrix J_p that is full rank, which provides mapping between the image and the task spaces without a null space in control. Since we assume that the image plane is parallel to the retinal surface (regarded as a task plane), resulting in the interaction matrix $J_p \in \mathbb{R}^{2 \times 2}$ is defined by the relationship between two differential motions, as in (23).

$$\Delta x^{image} = J_p \Delta \Theta^{task}, \quad (23)$$

where x^{image} and Θ^{task} are differential motions in image and task planes, respectively. The interaction matrix J_p can then be analytically formulated as in (24), by taking differential motions, p_u and p_v, in the image plane, corresponding to motions along unit vectors, u_{plane} and v_{plane}, in the task plane.

$$\left[\Delta p_u \ \Delta p_v \right] = J_p I^{2 \times 2}, \quad (24)$$

where $I^{2 \times 2}$ is an identity matrix; the differential motions in the task space along, u_{plane} and v_{plane}, are subject to canonical bases with respect to the task plane coordinates. In the previous method using stereo reconstruction, the interaction matrix was derived by projecting points along the canonical bases of the task plane onto the image plane, given camera projection matrices. Instead, we decompose the inverse of the interaction matrix, J_p^{-1}, into the magnitude and direction of mapping between the image and task plane coordinates as in (25).

$$J_p^{-1} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix}^{-1} \begin{bmatrix} \dfrac{\Delta \boldsymbol{p}_u}{|\Delta \boldsymbol{p}_u|} & \dfrac{\Delta \boldsymbol{p}_v}{|\Delta \boldsymbol{p}_v|} \end{bmatrix}^{-1}, \quad (25)$$

where scale factors, s_x and s_y, are defined by $|\boldsymbol{p}_u|$ and $|\boldsymbol{p}_v|$, respectively. It is noted that the scale factors can be substituted by zoom factors for a corresponding magnification of the operating microscope, while preserving the direction of motion. This would be useful for accommodating zoom optics frequently used in intraocular surgery.

Given the new surface reconstruction method, the interaction matrix defined in (25) is then described by an image scale used, because the image coordinates are already matched with the principal vectors, $\boldsymbol{u}_{plane}$ and $\boldsymbol{v}_{plane}$ for the plane as in Section 3.2.2. Accordingly, we can simply formulate the inverse of the interaction matrix that describes a complete image Jacobian:

$$J_p^{-1} = \begin{bmatrix} s_{cam} & 0 \\ 0 & s_{cam} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad (26)$$

where the image scale s_{cam} is determined by the zoom factor of the operating microscope. It should be noted that the negative value is set for the second column vector of the right-side matrix in (26). This is because ASAP takes a right-handed coordinate system, whereas the camera image is described by a left-handed coordinate system.

5.2 Formulation of hybrid visual servoing

To control the tool tip in the 3D space (ASAP coordinates), we extend the 2D vector Θ^{task} to the 3D vector X^{plane} using the orthonormal bases of the plane described in the ASAP coordinates as in (27), where $\boldsymbol{u}_{plane}$ and $\boldsymbol{v}_{plane} \in \mathbb{R}^{3 \times 1}$

$$\Delta X^{plane} = \begin{bmatrix} \boldsymbol{u}_{plane} & \boldsymbol{v}_{plane} \end{bmatrix} \Delta \Theta^{task} \quad (27)$$

Using (15), we can then define the planar motion of the 3D vector X^{plane} corresponding to differential motion in the image space

$$\Delta X^{plane} = \begin{bmatrix} \boldsymbol{u}_{plane} & \boldsymbol{v}_{plane} \end{bmatrix} J_p^{-1} \Delta x^{image} \quad (28)$$

Since the tool tip is located above the plane by d_{surf}, the actual displacement of the tool tip ΔX_{tip}^{actual}, corresponding to the planar motion of the aiming beam on the plane ΔX_{beam}^{plane}, is scaled down by the ratio of the lever arms,

$$\Delta X_{tip}^{actual} = r_{lever} \Delta X_{beam}^{plane}, \quad (29)$$

where $r_{lever} = d_{RCM}/(d_{RCM} + d_{surf})$ and d_{RCM} is the distance of the tool tip from the RCM. Herein, small angular motion pivoting around an RCM is assumed, as the displacement of the aiming beam is much smaller than the distance of the tool tip from the RCM.

Finally, the inverse of an image Jacobian $J^{-1} \in \mathbb{R}^{3 \times 2}$ is derived for visual servoing by substituting (28) into (29), given a 2D error Δe_{beam}^{image} between target and current beam positions on the image plane and the corresponding 3D displacement of the laser probe to correct the error in the ASAP coordinates:

$$\Delta E_{tip}^{ASAP} = J^{-1} \Delta e_{beam}^{image}; \quad (30)$$

$$J^{-1} = r_{lever} \left[u_{plane} \ v_{plane} \right] J_p^{-1}. \quad (31)$$

By substituting (26) into (31), the inverse of the image Jacobian for control is completely defined in (32).

$$J^{-1} = r_{lever} s_{cam} \left[u_{plane} \ -v_{plane} \right] \quad (32)$$

The position of the tool tip is regulated by a PD controller as in (33), to minimize the error Δe_{beam}^{image} between the current aiming beam and the target positions.

$$P_{tip} \equiv P_{tip} + p_k \Delta E_{tip} + p_d \Delta \dot{E}_{tip} \quad (33)$$

In addition, a remaining degree of freedom along the axis of the tool is regulated to maintain a specific distance d_{lim} between the tool tip and the retinal surface. Consequently, we incorporate a depth-limiting feature with image-based visual servoing in control to fully define the 3-DOF motion of the tool tip, as in (34).

$$P_{goal} \equiv P_{tip} + (d_{lim} - d_{surf}) \cdot n_{tool}, \quad (34)$$

where $\boldsymbol{n}_{tool}$ is a unit vector describing the axis of the tool. As a result, the 2D error is minimized via the visual servoing loop, while the distance of the tool tip from the retinal surface is regulated by the position control loop.

A switching control alternates between open-loop and closed-loop controllers in order to speed up the automated operation. The goal position for the tip is thus set initially by the image Jacobian (30), which acts as an open-loop controller. The feedback controller (33) then corrects the remaining error between the target and current beam positions. This control scheme is also beneficial in addressing issues raised by unreliable beam detection, especially due to saturation in images at the instant of laser firing. The control scheme is completed by incorporating an image Jacobian update as introduced in Yang et al. (2016) to compensate any error involved in such an analytical Jacobian, primarily due to the assumption that the image plane is nearly parallel to the retinal surface.

6 Experiments

We performed robot-aided intraocular laser surgery in a wet eye phantom and in porcine eyes *ex vivo*, which had not been feasible previously due to erroneous stereo reconstruction of the retinal surface. The new surface reconstruction method was thus applied in place of stereo reconstruction for demonstration of hybrid visual servoing in realistic eye models. The experimental settings and results for robot-aided intraocular laser photocoagulation are summarized in Table 2.

6.1 Eye phantom

We first tested the hybrid visual servoing in the wet eye phantom, while incorporating the newly introduced surface reconstruction. The eye phantom was filled with water in place of the vitreous humor, as is done in vitrectomy in the human eye, and was covered with a contact lens on the portion representing the cornea as shown in Figure 15(a). A slip of paper was attached to the inner surface of the eye phantom, as a target surface printed with artificial vasculature. Multiple ring patterns were introduced for targets around the circumference of rings of diameter 1, 2, and 3 mm with a spacing of 600 μm, as a typical arrangement for the treatment of proliferative retinopathy (Bandello et al., 1999). Similar patterns are also used intraoperatively to seal retinal breaks. The arrangement thus provided a total of 32 targets. In order to compensate eye movement during operation, the eyeSLAM algorithm was also used to track artificial blood vessels, as shown in Figure 17 (indicated with green dots). For demonstration of laser photocoagulation, the power of the 532nm laser was set at 3.0W with a duration of 20 ms. We also set a targeting threshold of 100 μm for firing the laser, as this was used as a typical setting for automated operation; the resulting speed and accuracy for this threshold value were found to be 1.22 target/s and 48 μm, respectively (Yang et al., 2016).

However, it was not possible to make distinct black burns on the paper slide with the laser, since the paper was wet from the water (black burns were made clearly in previous experiments with dry phantoms (Yang et al., 2016; Yang, Lobes, et al., 2015)). We thus analyzed error between targets and aiming beam positions just after the instant of laser firing in each image sequence of the resulting video (figure 18). The average error was found to be

49±31 μm for the entire pattern, which is similar to the error obtained by hybrid visual servoing with stereo reconstruction in a dry phantom (Yang et al., 2016). In addition, we investigated execution time as a measure of the control performance. For example, erroneous surface estimation and/or incorrect formulation of the image Jacobian may lead to longer execution time or failure of servoing. The execution time was measured to be 21.6 s (1.48 target/s), which is comparable to previous results (Yang et al., 2016); the execution time was 26.2 s for the same 100-μm threshold in the dry eye phantom. The average steady-state error of the tool tip position in 3D was measured to be 26±17 μm for the 32 targets, while the depth of the tool varies by ±21 μm over the entire operation.

6.2 Porcine eye

We also performed robot-aided intraocular laser surgery in porcine eyes *ex vivo* as shown in Figure 15(b–c). First, automated laser photocoagulation was evaluated on fixed targets, in order to primarily validate the new surface reconstruction and hybrid visual servoing in the real tissue, while disabling the eyeSLAM tracking. Although the surface estimation result could not be evaluated quantitatively with any ground truth, it did provide a reasonable value of the surface normal for a typical ASAP sensor orientation. For demonstration of laser photocoagulation in the real tissue, the laser power was set at 1.0W with a duration of 50 ms. The total execution time was 36.0 s for 32 targets with the 100-μm targeting threshold, while 14.2 of 36.0 s were spent on two specific targets (2 of 32). Without these outliers, the execution time would have been 21.8 s for the other 30 targets (1.37 target/s). This was due to unreliable detection of the aiming beam, caused by bubbles on the contact lens. Hence, speed of operation is still comparable to previous experiments in the eye phantom (1.22 target/s) (Yang et al., 2016).

However, it should be noted that the cornea of porcine cadaver eyes gets cloudy due to apoptosis in the corneal epithelium, which hindered clear imaging of the posterior segment of the eyes during the tests. The tests *ex vivo* thus led to a primary difficulty in simultaneously detecting both the aiming beam and the posterior segment of the eye. For instance, in order to reliably detect the aiming beam, less illumination is required, whereas the brighter light is preferable for clearly imaging the fundus and tracking blood vessels. Hence, it was challenging to find a single setting to reliably detect both the aiming beam for visual servoing and blood vessels for tracking. Detection of blood vessels could also be interfered with by the intense light of the aiming beam; the aiming beam diffuses rather widely on the translucent retina surface, instead of creating a distinct red spot. Furthermore, we could observe only a few blood vessels (one or two branches) in the porcine eyes; this sometimes resulted in jitter or drift in tracking.

We therefore modified experimental settings for further evaluation, in order to address the issues in imaging *ex vivo*. The eyes were dissected to remove the cloudy cornea, and the eyes cut in half were placed inside the eye phantom as shown in Figure 15(b). The phantom was then filled with water and also covered with a contact lens on top, as in surgery. We also utilized both the left and right images: one for detecting the aiming beam and the other for blood vessels. Thus, different camera settings were introduced for the two cameras. The lower gain was set for the left camera to reliably detect the aiming beam, whereas the higher

gain was used for clearly imaging the posterior segments in the right camera, as shown in Figure 19. The tracking result from the right images was then transformed to the left images via a planar homography constructed during the surface reconstruction; Figure 19 shows the transformed blood vessels in the large main images.

Figure 19 demonstrates automated laser photocoagulation in the porcine eye over the time of operation. The total execution time to burn 32 targets was 60 s. This is considerably longer than the time of automated operation without burning targets on the same eye: 25.2 s, which was performed for verification of visual servoing prior to firing the laser, using the 100-μm targeting threshold. During the experiment with laser firing, unreliable detection of the aiming beam was rather frequently observed, as the retina became detached and floated in the dissected eye; 17.7 of 60 s were spent on specific targets (2 of 32) due to the problem. This thus led to a lower speed of operation, 0.71 target/s, even for the other 30 targets. The accuracy of operation was also investigated by measuring error between target and aiming beam positions just after the instant of laser firing, because the contrast of the resulting burns in the images was low. The average error increased to 62 ± 36 μm, larger than that obtained in the wet eye phantom. This was because unreliable detection of the aiming beam not only delayed time of operation but also degraded control performance. Nevertheless, the error by automated operation is still lower than errors at all frequencies (burns/s) under manual operation, all of which produced average error greater than 100 μm (Yang, Lobes, et al., 2015).

Finally, we tested robot-aided intraocular laser surgery in an intact porcine eye *ex vivo* as shown in Figure 15(c). Here, the structure of the eyeball was preserved intact, including the cornea, and the lens. Saline solution replaced the vitreous, as is standard in vitrectomy surgery. The number of targets was adjusted from 32 to 16 and placed around the circumference of rings of diameter 1 and 2 mm. This was because the eyeSLAM tracking was not robust enough yet in intact cadaver eyes for a long duration of the laser treatment. For the reduced number of targets, automated laser photocoagulation was successfully demonstrated in the intact porcine eye as shown in Figure 20, while addressing a variety of challenges encountered in the tests *ex vivo.* The average error for the 16 targets was 69 ± 36 μm. Although the error slightly increased, compared to errors from automated operation in the other eye models, it was still well below the error produced by manual operation; the minimum average error by manual operation was 102 μm at an operating frequency of 1.0 burns/s. The total execution time was 15.8 s for the 16 targets (1.01 target/s), which is comparable to the other experiments.

7 Discussion

We demonstrated robot-aided intraocular laser surgery using monocular vision in realistic eye models and porcine eyes *ex vivo*. To address erroneous stereo reconstruction in the complex eye model, a novel monocular surface estimation method was proposed, utilizing the automatic scanning of Micron and the geometric analysis of projected beam patterns. The retinal surface estimation was tested in various conditions, in order to evaluate its performance and the feasibility in the realistic eye models involving different media and optical distortion by lenses. We also investigated the accuracy of the reconstruction method

in the presence of uncertainties in estimation. Robot-aided intra-ocular laser surgery was accomplished by combination of the new monocular surface estimation with the hybrid visual servoing scheme, while accounting for practical issues in automation of microsurgery, such as camera calibration, 3D reconstruction, and zoom optics.

A complete demonstration in an intact eye has not been shown yet due to instability in tracking the eye. However, it is found that the new surface estimation works for the intact eye, regardless of its optical distortion, and can be used for automation of robot-aided intraocular surgery. We would be able to perform a complete demonstration in intact eyes, by improving imaging quality for the reliable detection of both aiming beam and blood vessels and by refining a relevant control scheme to tackle erroneous beam detection.

We proposed structured-light reconstruction to estimate the retinal surface and demonstrated the feasibility of the concept in an optically distorted imaging system. In the future, several refinements should be made for better outcomes, in terms of accuracy and robustness. First, it is found that depth estimation is relatively sensitive to uncertainties in the corresponding parameters. It could possibly be improved by a more sophisticated way of using both the left and right images from the stereo cameras, instead of just taking the average of the ellipse parameters, since even the average method yielded lower deviation in depth estimation than a single camera provided. For example, disparity of resulting ellipses can be utilized to refine the depth estimation (Richa et al., 2012), while considering it as a confidence parameter in fusing multiple measurements into a single outcome. In addition, the zoom factor of the microscope used for depth estimation and visual servoing may be estimated intraoperatively, in order to be used under unknown optics, rather than using a pre-calibrated scale factor of the optics. The known size of the tool tip could be a cue to find such a scale factor. Furthermore, it would be preferable for microsurgery if arbitrary beam patterns created by manual scanning could also allow estimation of the surface, utilizing projective geometry similar to that used for an elliptical trajectory.

In this work, the shape of the retina is assumed not to change due to the applied laser burns. This assumption generally holds due to the unusual nature of the retina; the energy is absorbed by the retinal pigment epithelial and choroidal layers, which have unusually high blood flow, and little heat reaches the retina, which has been found to maintain its contour under laser photocoagulation.

In this paper, stereo reconstruction is used as ground truth for comparison of the cone beam reconstruction; this adds some noise to the process. One alternative to obtaining a plane described in 3D in the ASAP coordinate system would be running a plane fitting with several points of contact made by the Micron tool tip. The important factor in this method is to determine when the tip touches the surface, which can be accomplished either automatically or manually. For the automatic method, the tip should incorporate a force sensor that senses the time of contact. For the manual operation, the tool tip should be manipulated by means of manual stages to gently touch the surface while visually monitoring the tool tip position.

For application to other domains of robot-assisted surgery, the cone beam generated by automated scanning of the handheld micromanipulator could be substituted with a diverging beam source to create a similar cone beam in space, in place of a highly collimated laser. Thus, this approach may be used to locally estimate orientation of a target lesion and proximity of an operating tool in a small area of interest, such as in endoscopic surgery. In order to cover a larger area with automated operation, higher-order modeling of the target surface may be considered for such a large area. For example, we could build a spherical surface by mosaicking small planar patches obtained by the reconstruction.

Acknowledgments

References

Bandello F, Lanzetta P, Menchini U. 1999; When and how to do a grid laser for diabetic macular edema. Documenta Ophthalmologica. 97(3–4):415–419. [PubMed: 10896358]

Becker BC, MacLachlan RA, Lobes LA, CN, et al. 2013; Vision-based control of a handheld surgical micromanipulator with virtual fixtures. IEEE Transactions on Robotics. 29(3):674–683. [PubMed: 24639624]

Bergeles C, Kratochvil BE, Nelson BJ. 2012; Visually servoing magnetic intraocular microdevices. IEEE Transactions on Robotics. 28(4):798–809.

Bergeles C, Shamaei K, Abbott JJ, et al. 2010; Single-camera focus-based localization of intraocular devices. IEEE Transactions on Biomedical Engineering. 57(8):2064–2074. [PubMed: 20442042]

Braun D, Yang S, Martel JN, Riviere CN, Becker BC. 2017; EyeSLAM: real-time localization and mapping of retinal vessels during intraocular microsurgery. International Journal of Medical Robotics and Computer Assisted Surgery. 14(1):e1848.

Brooks HL. 2000; Macular hole surgery with and without internal limiting membrane peeling. Ophthalmology. 107(10):1939–1948. [PubMed: 11013203]

Buchs NC, Volonte F, Pugin F, et al. 2013; Augmented environments for the targeting of hepatic lesions during image-guided robotic liver surgery. Journal of Surgical Research. 184(2):825–831. [PubMed: 23684617]

Castaño A, Hutchinson S. 1994; Visual compliance: Task-directed visual servo control. IEEE Transactions on Robotics and Automation. 10(3):334–342.

Chen, Q, Wu, H, Wada, T. Camera calibration with two arbitrary coplanar circles. In: Pajdla, T, Matas, J, editorsComputer Vision—ECCV 2004. (Lecture Notes in Computer Science). Vol. 3023. Berlin: Springer; 2004. 521–532.

Corke PI, Hutchinson SA. 2001; A new partitioned approach to image-based visual servo control. IEEE Transactions on Robotics and Automation. 17(4):507–515.

Das H, Zak H, Johnson J, et al. 1999; Evaluation of a telerobotic system to assist surgeons in microsurgery. Computer Aided Surgery. 4(1):15–25. [PubMed: 10417827]

Diabetic Retinopathy Study Research Group. 1981; Photocoagulation treatment of proliferative diabetic retinopathy. Clinical application of Diabetic Retinopathy Study (DRS) findings. DRS report number 8. Ophthalmology. 88:583–600. [PubMed: 7196564]

Fine HF, Wei W, Goldman R, et al. 2010; Robot-assisted ophthalmic surgery. Canadian Journal of Ophthalmology. 45(6):581–584. [PubMed: 21135893]

Fischler MA, Bolles RC. 1981; Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM. 24(6):381–395.

Gijbels, A; Vander Poorten, EB; Gorissen, B; , et al. Experimental validation of a robotic comanipulation and tele-manipulation system for retinal surgery. 5th IEEE RAS & EMBS international conference on biomedical robotics and biomechatronics; Sao Paulo, Brazil. 12–15 August 2014; Piscataway, NJ: IEEE; 2014a. 144–150.

Gijbels, A; Vander Poorten, EB; Stalmans, P; , et al. Design of a teleoperated robotic system for retinal surgery. IEEE international conference on robotics and automation; Hong Kong, China. 31 May–7 June 2014; Piscataway, NJ: IEEE; 2014b. 2357–2363.

Hutchinson S, Hager GD, Corke PI. 1996; A tutorial on visual servo control. IEEE Transactions on Robotics and Automation. 12(5):651–670.

Ida Y, Sugita N, Ueta T, et al. 2012; Microsurgical robotic system for vitreoretinal surgery. International Journal of Computer Assisted Radiology and Surgery. 7(1):27–34. [PubMed: 21573828]

MacLachlan RA, Riviere CN. 2009; High-speed microscale optical tracking using digital frequency-domain multiplexing. IEEE Transactions on Instrumentation and Measurement. 58(6):1991–2001. [PubMed: 20428484]

Maclachlan RA, Becker BC, Tabarés JC, et al. 2012; Micron : an actively stabilized handheld tool for microsurgery. IEEE Transactions on Robotics. 28(1):195–212. [PubMed: 23028266]

MacLaren RE, Edwards T, Xue K, et al. 2017; Results from the first use of a robot to operate inside the human eye. Investigative Ophthalmology & Visual Science. 58(8):1185.

Malis E, Chaumette F, Boudet S. 1999; 2 1/2D visual servoing. IEEE Transactions on Robotics and Automation. 15(2):238–250.

Meenink, HCM, Hendrix, R, Naus, GJL. , et al. Robot assisted vitreoretinal surgery. In: Gomes, P, editorMedical Robotics: Minimally Invasive Surgery. Oxford: Woodhead Publishing; 2012. 185–209.

Mukherjee, S; Yang, S; MacLachlan, RA; , et al. Toward monocular camera-guided retinal vein cannulation with an actively stabilized handheld robot. IEEE international conference on robotics and automation; Singapore. 29 May–3 June 2017; Piscataway, NJ: IEEE; 2017. 2951–2956.

Nambi M, Bernstein PS, Abbott JJ. 2016; A compact tele-manipulated retinal-surgery system that uses commercially available instruments with a quick-change adapter. Journal of Medical Robotics Research. 1(2):1630001.

Nasseri, MA; Eder, M; Nair, S; , et al. The introduction of a new robot for assistance in ophthalmic surgery. 35th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Osaka, Japan. 3–7 July 2013; Piscataway, NJ: IEEE; 2013. 5682–5685.

Noo F, Clackdoyle R, Mennessier C, et al. 2000; Analytic method based on identification of ellipse parameters for scanner calibration in cone-beam tomography. Physics in Medicine and Biology. 45(11):3489–3508. [PubMed: 11098919]

OpenCV Library. OpenCV; 2018. Available at: https://opencv.org/

Rahimy E, Wilson J, Tsao T-C, et al. 2013; Robot-assisted intraocular surgery: development of the IRISS and feasibility studies in an animal model. Eye (London, England). 27:972–978.

Richa R, Balicki M, Sznitman R, Meisner E, Taylor R, Hager G. 2012; Vision-based proximity detection in retinal surgery. IEEE Transactions on Biomedical Engineering. 59(8):2291–301. [PubMed: 22692869]

Taylor R, Jensen P, Whitcomb L, Barnes A, Kumar R, Stoianovici, Gupta D, Wang, De Juan E, Kavoussi L. 1999; A steady-hand robotic system for microsurgical augmentation. The International Journal of Robotics Research. 18(12):1201–1210.

Ueta T, Yamaguchi Y, Shirakawa Y, et al. 2009; Robot-assisted vitreoretinal surgery: Development of a prototype and feasibility studies in an animal model. Ophthalmology. 116(8):1538–1543. [PubMed: 19545902]

Üneri, A; Balicki, MA; Handa, J; , et al. New steady-hand eye robot with micro-force sensing for vitreoretinal surgery. 3rd IEEE RAS and EMBS international conference on biomedical robotics and biomechatronics (BioRob); Tokyo, Japan. 26–29 September 2010; Piscataway, NJ: IEEE; 2010. 814–819.

Willekens K, Gijbels A, Schoevaerdts L, et al. 2017; Robotassisted retinal vein cannulation in an in vivo porcine retinal vein occlusion model. Acta Ophthalmologica. 95(3):270–275. [PubMed: 28084059]

Wilson JT, Gerber MJ, Prince SW, et al. 2018; Intraocular robotic interventional surgical system (IRISS): Mechanical design, evaluation, and master-slave manipulation. International Journal of Medical Robotics and Computer Assisted Surgery. 14(1):e1842.

Wright CHG, Barrett SF, Welch AJ. 2006; Design and development of a computer-assisted retinal laser surgery system. Journal of Biomedical Optics. 11(4):41127.

Yang, S; Balicki, M; MacLachlan, RA; , et al. Optical coherence tomography scanning with a handheld vitreoretinal micromanipulator. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); San Diego, CA, USA. 28 August–1 September 2012; Piscataway, NJ: IEEE; 2012. 948–951.

Yang S, Lobes LA, Martel JN, et al. 2015a; Handheld-automated microsurgical instrumentation for intraocular laser surgery. Lasers in Surgery and Medicine. 47(8):658–668. [PubMed: 26287813]

Yang S, Maclachlan RA, Martel JN, Lobes LA, Riviere CN, Member S. 2016; Comparative Evaluation of Handheld Robot-Aided Intraocular Laser Surgery. IEEE Transactions on Robotics. 32(1):246–251. [PubMed: 27019653]

Yang, S; MacLachlan, RA; Riviere, CN. Toward automated intraocular laser surgery using a handheld micromanipulator. IEEE/RSJ international conference on intelligent robots and systems; Chicago, IL, USA. 14–18 September 2014; Piscataway, NJ: IEEE; 2014. 1302–1307.

Yang S, MacLachlan RA, Riviere CN. 2015b; Manipulator design and operation of a six-degree-of-freedom handheld tremor-canceling microsurgical instrument. IEEE/ASME Transactions on Mechatronics. 20(2):761–772. [PubMed: 25419103]

Yu, H; Shen, JH; Joos, KM; , et al. Design, calibration and preliminary testing of a robotic telemanipulator for OCT guided retinal surgery. IEEE international conference on robotics and automation (ICRA); Karlsruhe, Germany. 6–10 May 2013; Piscataway, NJ: IEEE; 2013. 225–231.

Yu H, Shen JH, Joos KM, et al. 2016; Calibration and integration of B-Mode optical coherence tomography for assistive control in robotic microsurgery. IEEE/ASME Transactions on Mechatronics. 21(6):2613–2623.

Zhou Y, Nelson BJ. 1999; Calibration of a parametric model of an optical microscope. Optical Engineering. 38(12):1989–1995.

Appendix A Derivation of cone beam reconstruction

A.1 Derivation of tilt angle

The cross-section of the cone beam in Figure 4 is depicted in Figure A1 (in the xz-plane). The lengths of the semimajor and major axes are defined along the line A_0A_3 as in (A1).

$$A_0A_1 = m_a \text{ and } A_0A_3 = 2m_a \quad \text{(A1)}$$

From (A1) and the tilt angle of the plane θ_p, the length A_0C_1 is derived:

$$A_0C_1 = A_0A_3 \cos \theta_p = 2m_a \cos \theta_p. \quad \text{(A2)}$$

The length C_1C_2 is defined by the length A_3C_1 and the opening angle of the cone beam θ_c as in (A3).

$$C_1 C_2 = A_3 C_1 \tan \theta_c = 2m_a \sin \theta_p \tan \theta_c, \quad (A3)$$

where $A_3 C_1 = 2m_a \sin \theta_p$.

Since the length $A_0 C_0$ is the half of the length $A_0 C_2$, the length $A_0 C_0$ is parameterized by m_a, θ_p, and θ_c.

$$
\begin{aligned}
A_0 C_0 &= \frac{1}{2}(A_0 C_1 + C_1 C_2) \qquad\qquad (A4) \\
&= \frac{1}{2}(2m_a \cos \theta_p + 2m_a \sin \theta_p \tan \theta_c) \\
&= m_a \cos \theta_p + m_a \sin \theta_p \tan \theta_c
\end{aligned}
$$

The length $A_0 C_0$ can also be defined along the tilted plane as in (A5). Herein, the length $A_1 A_2$ is denoted by a.

$$A_0 C_0 = (A_0 A_1 + A_1 A_2) \cos \theta_p = (m_a + a) \cos \theta_p \quad (A5)$$
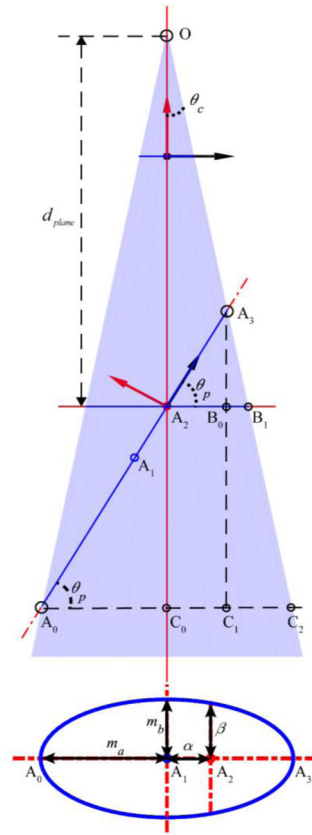
Fig. A1.
Parameters for derivation of tilt angle and depth.

Using (A4) and (A5), The length α is thus described as in (A6).

$$\alpha = m_a \tan \theta_p \tan \theta_c. \quad (A6)$$

The conic section resulted by the tilted plane is subject to having the interaction between the resulting ellipse and the circular trajectory with the radius A_2B_1 in 3D. The radius is denoted by β. We then define A_2B_1 with the sum of A_2B_0 and B_0B_1:

$$A_2B_1 = \beta = A_2B_0 + B_0B_1. \quad (A7)$$

Given the length α in (A6), the lengths, A_2B_0 and B_0B_1 are defined as in (A8).

$$A_2B_0 = (m_a - \alpha) \cos \theta_p \quad (A8)$$
$$B_0B_1 = (m_a - \alpha) \sin \theta_p \tan \theta_c$$

By substituting (A8) into (A7), we describe β in terms of m_a, θ_p, and θ_c:

$$\beta = m_a \cos \theta_p (1 - \tan^2 \theta_p \tan^2 \theta_c). \quad (A9)$$

A point described by α and β lies on the trajectory of the ellipse (A10).

$$\frac{\alpha^2}{m_a^2} + \frac{\beta^2}{m_b^2} = 1 \quad (A10)$$

By substituting (A6) and (A9) into (A10), the aspect of the ellipse $r = \frac{m_b}{m_a}$ is then described by the opening angle and the tilt angle of the plane in (A10).

$$\left(\frac{m_b}{m_a}\right)^2 = \cos^2 \theta_p (1 - \tan^2 \theta_p \tan^2 \theta_c) \quad (A11)$$

From (A10), we finally obtain the tilt angle θ_p:

$$\theta_p = \pm \sin^{-1} \left(\sqrt{\frac{1 - r^2}{1 + \tan^2 \theta_c}} \right), \quad (A12)$$

where $r = \frac{m_b}{m_a}$ and $\tan \theta_c = \frac{r_{cone}}{h_{RCM}}$.

A.2 Derivation of plane depth

We define an intersection A_2 between the resulting ellipse and the axis of the cone in 3D. Herein, the distance OA_2 is denoted by d_{plane}. Since the real size of the ellipse needs to be used, the camera scale s_{cam} is multiplied on the length β obtained in the image coordinates as in (A13).

$$OA_2 = d_{plane} = \frac{A_2 B_1}{\tan \theta_c} = \frac{\beta s_{cam}}{\tan \theta_c} \quad (A15)$$

Using (A9) and (A15), the d_{plane} is finally defined:

$$d_{plane} = \frac{\cos \theta_p (1 - \tan^2 \theta_p \tan^2 \theta_c)}{\tan \theta_c} m_a s_{cam}. \quad \text{(A16)}$$

The derivation in (A16) can be further explored by using the trigonometric relationships (A17) from (A12).

$$\tan^2 \theta_p = \frac{1 - r^2}{r^2 + \tan^2 \theta_c} \quad \text{(A17)}$$

$$\cos^2 \theta_p = \frac{r^2 + \tan^2 \theta_c}{1 + \tan^2 \theta_c}$$

By substituting (A17) into (A16), the distance d_{plane} is rewritten as in (A18).

$$d_{plane} = \frac{m_b^2 s_{cam}}{\sin \theta_c \sqrt{m_b^2 + m_a^2 \tan^2 \theta_c}} \quad \text{(A18)}$$
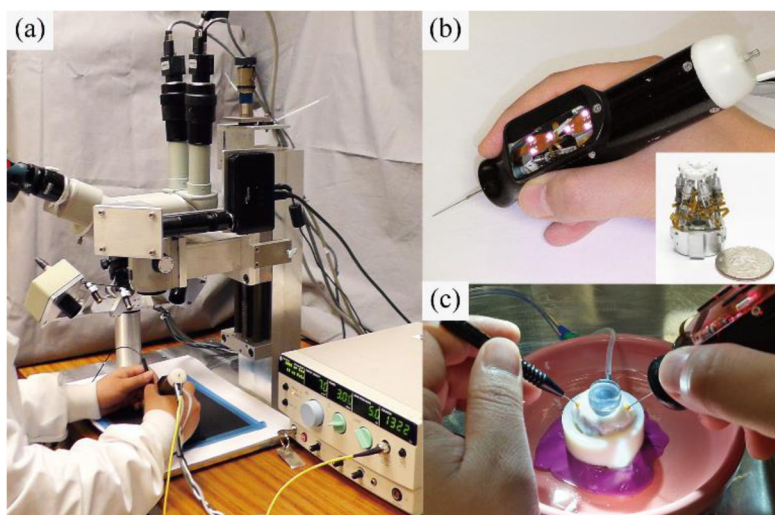
Fig. 1.
System setup for robot-aided intraocular laser surgery: (a) overall system setup. (b) Micron with the miniature micromanipulator; (c) Porcine eye and intraocular surgical setup for the test ex vivo test.
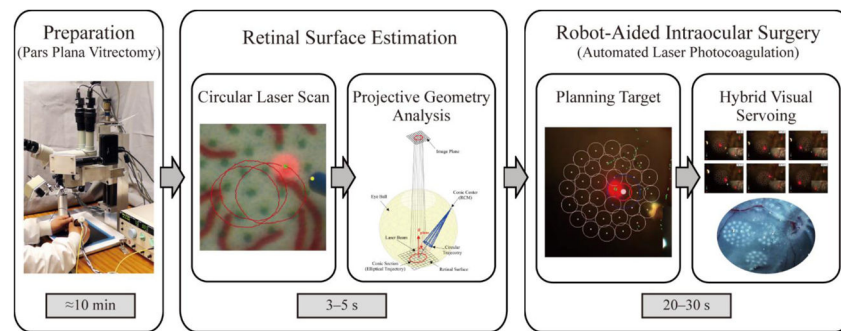
Fig. 2.
The first step to the robot-aided intraocular surgery is to make three small incisions (≈23-gauge) in the sclera of the eye: one for an illumination light pipe, one for a surgical tool, and the remaining port to maintain intraocular pressure during surgery. This procedure may involve the removal of the vitreous humor from the eye for clear visualization (replacing it with saline), which is typically completed within 10 min. Once the ASAP sensor head is properly set, the surface reconstruction is performed. For the surface reconstruction, we make a circular scan of the tool tip incorporating the laser aiming beam over the retinal surface with a typical speed of 1 mm/s, which results in an elliptical trace showing the images acquired. We analyze the projective geometry with the corresponding ellipse in order to reconstruct the retinal surface with respect to the control coordinates of the robot. These steps for the retinal surface estimation take less than 5 s, including ellipse fitting and optimization. Then a laser spot pattern is designed in terms of shape (e.g., circular or grid), spacing, and number of spots to be treated, taking into account the location and size of the lesion. The planned target is then placed and aligned with the lesion. Finally, automated laser surgery is performed. During operation, the robot system maintains a constant standoff distance from the retinal surface using hybrid visual servoing of the laser probe. The procedure would take 20–30 s for a pattern of 32 targets.
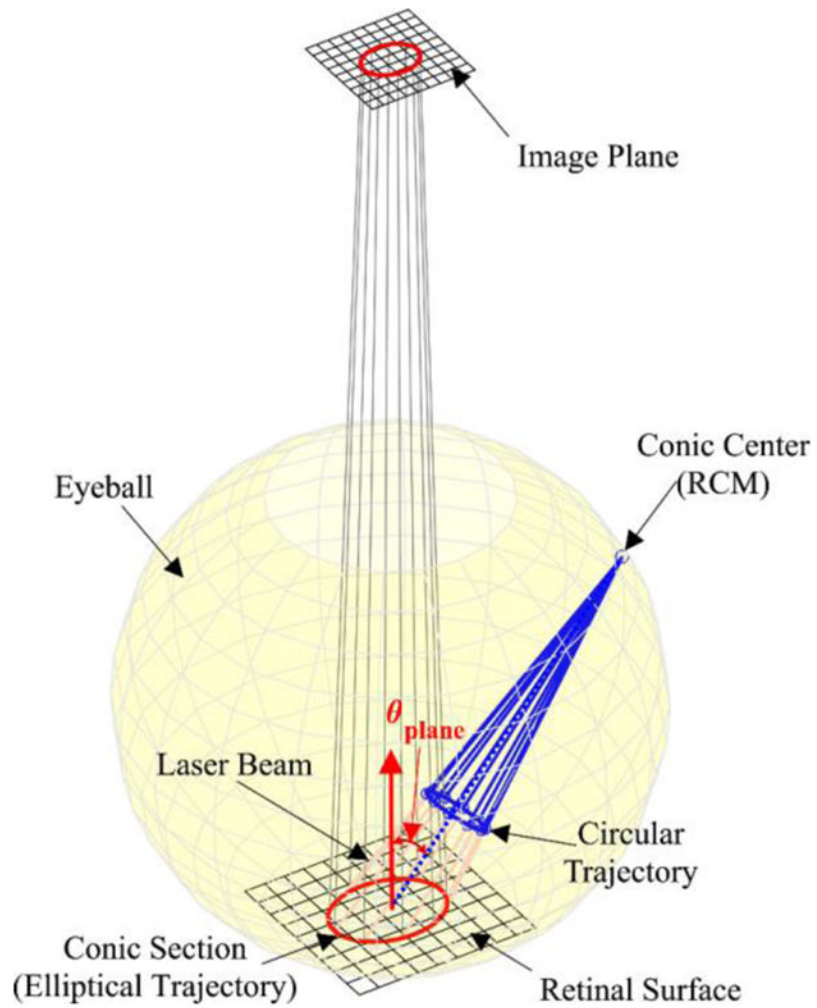
Fig. 3.
A conic section shown as an ellipse on the retinal surface. RCM, remote center of motion.
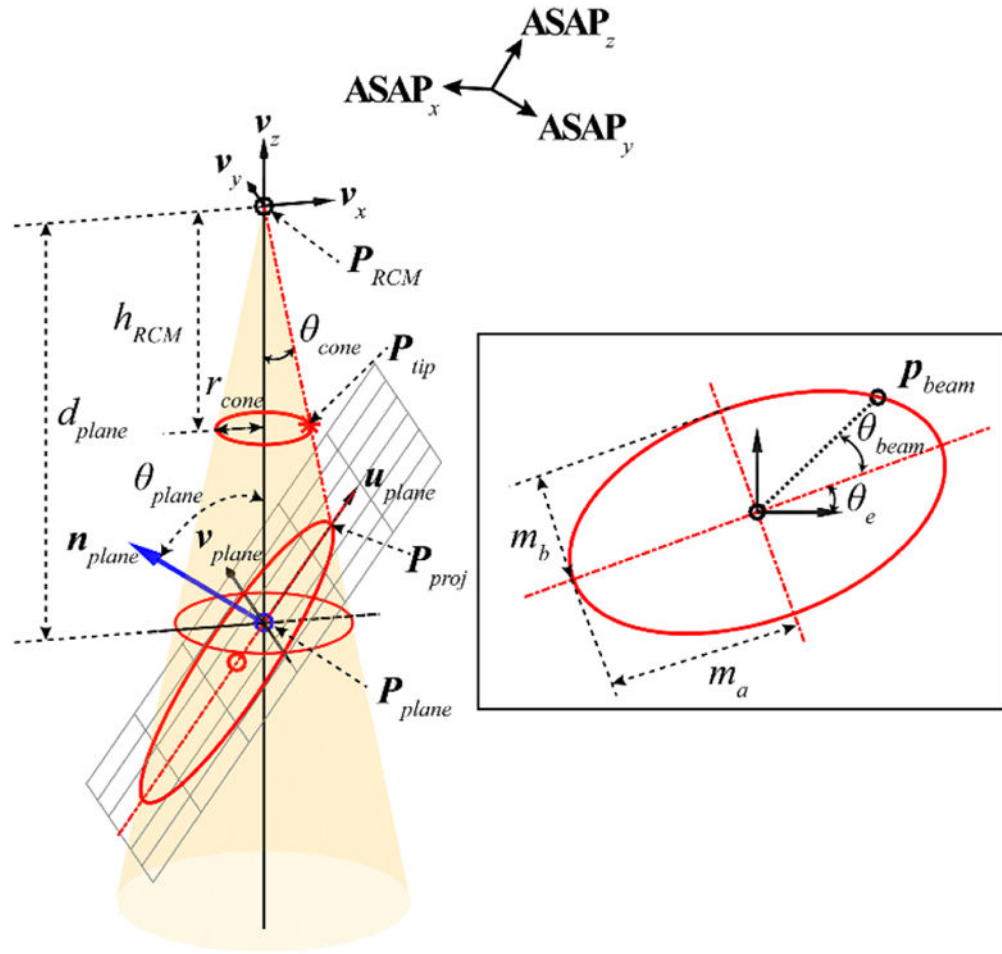
Fig. 4.
Cone beam analysis and corresponding parameters to describe a target plane.
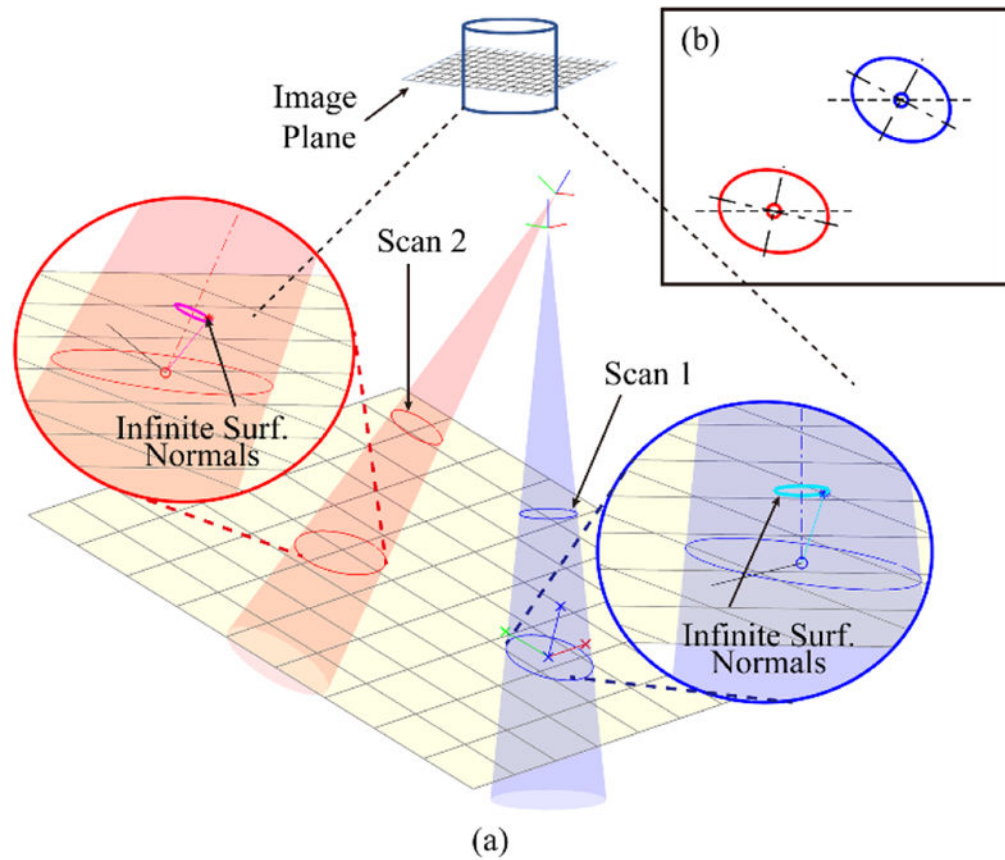
Fig. 5.
(a) Conic sections forming two ellipses on the retinal surface by circular scanning of the laser probe about the axis of the tool. (b) Resulting ellipses shown in the image plane.
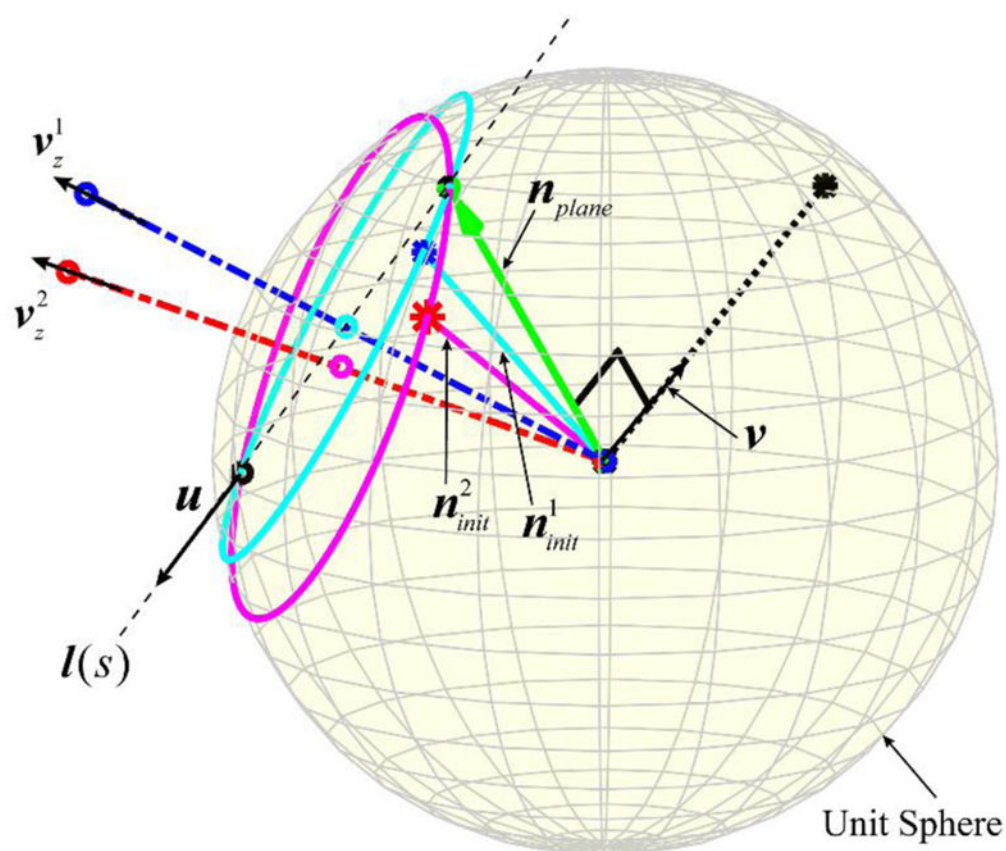
Fig. 6.
Infinite sets of two initial normal vectors in the unit sphere. The true surface normal is indicated as the green arrow.
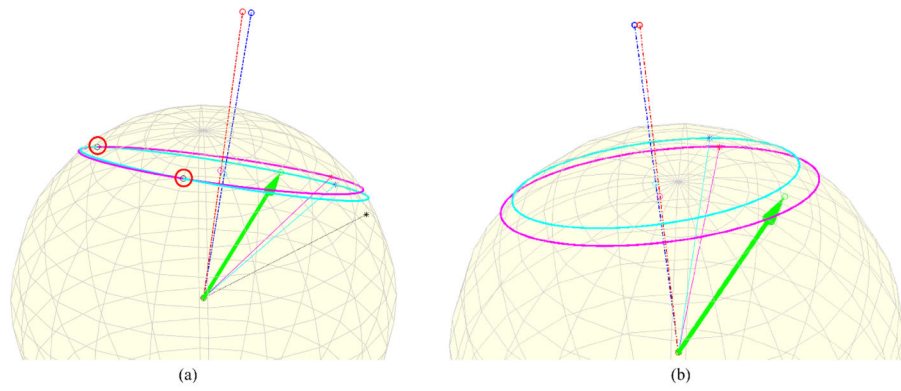
Fig. 7.
Failure cases of dual cone beam reconstruction with real data. The green arrow indicates the true normal vector that should be estimated. The red circles indicate the estimated surface normal vectors. (a) Incorrect intersections are found. (b) No solution exists.
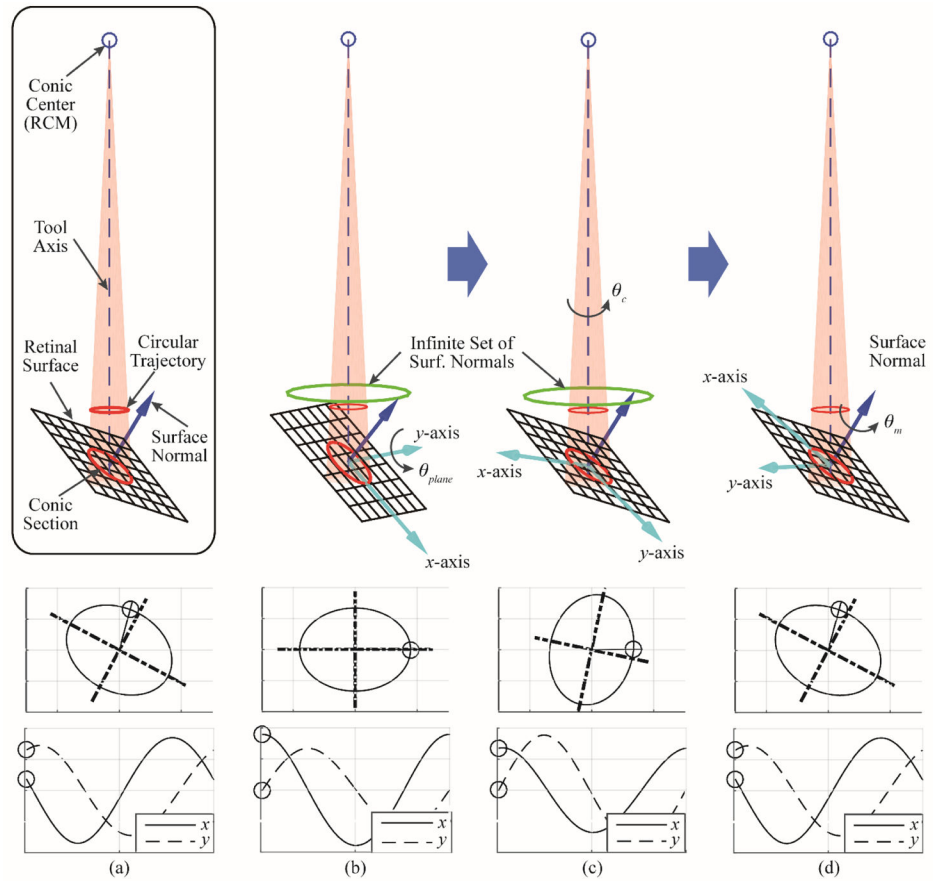
Fig. 8.
Each step of the retinal surface estimation with ellipse trajectories and corresponding data points in the image plane. The first data point on each trajectory is marked by a circle. (a) Circular scanning of the laser probe over the retinal surface (top) and the resulting beam trajectory acquired in the image plane (bottom). (b) Initial estimation of the surface normal and the resulting plane by rotation of the x–y plane of the cone about its y-axis (top). The projected ellipse from the initial estimation, shown in the image plane (bottom). (c) Rotation of the initial plane about the axis of the cone (top). The trajectory undergoes optimization for matching the phase angles of data points (bottom). (d) Rotation of the estimated planes about the surface normal vector (top). The transverse vectors of the estimated plane are aligned with the image coordinates (bottom). RCM, remote center of motion.
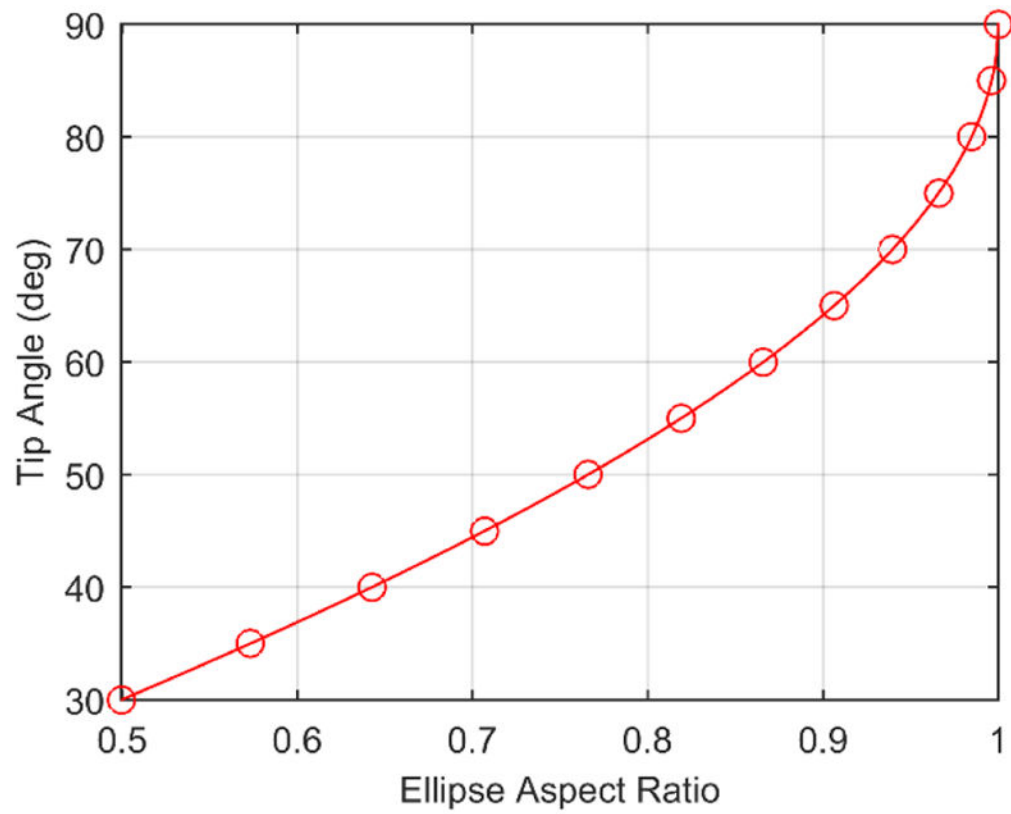
Fig. 9.
The tip angle with respect to estimated plane, according to the aspect ratio of an ellipse.
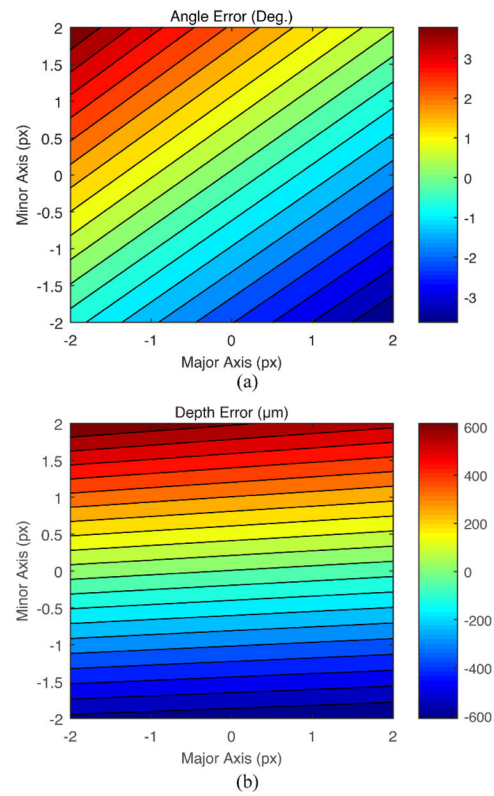
Fig. 10.
(a) Angle error in surface normal estimation, led by variation in lengths of the major and minor axes in pixels. (b) Surface depth error by variation in lengths of the major and minor axes.
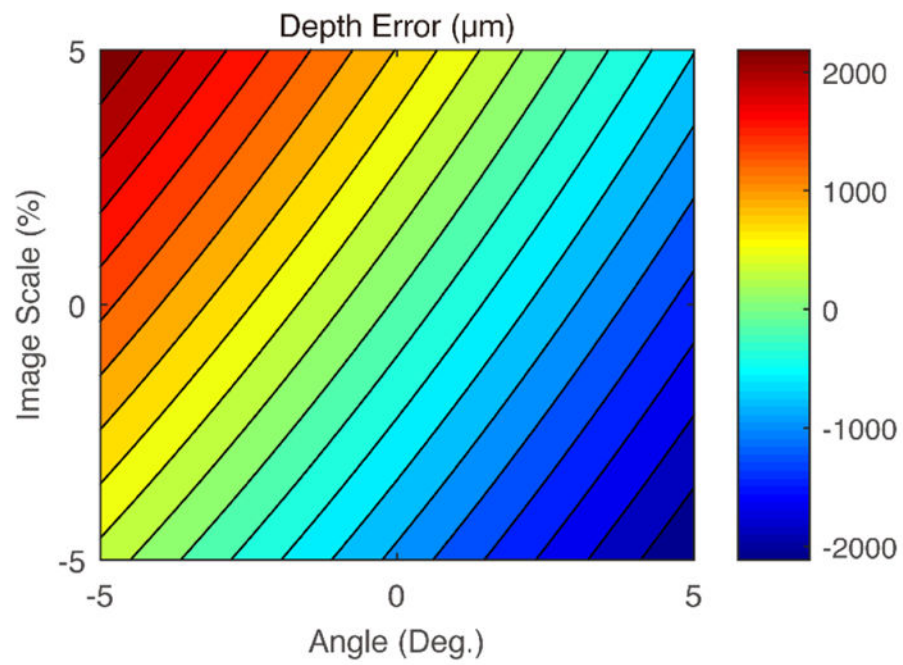
Fig. 11.
Depth error in surface estimation due to angle error in estimation and uncertainty on the image scale.
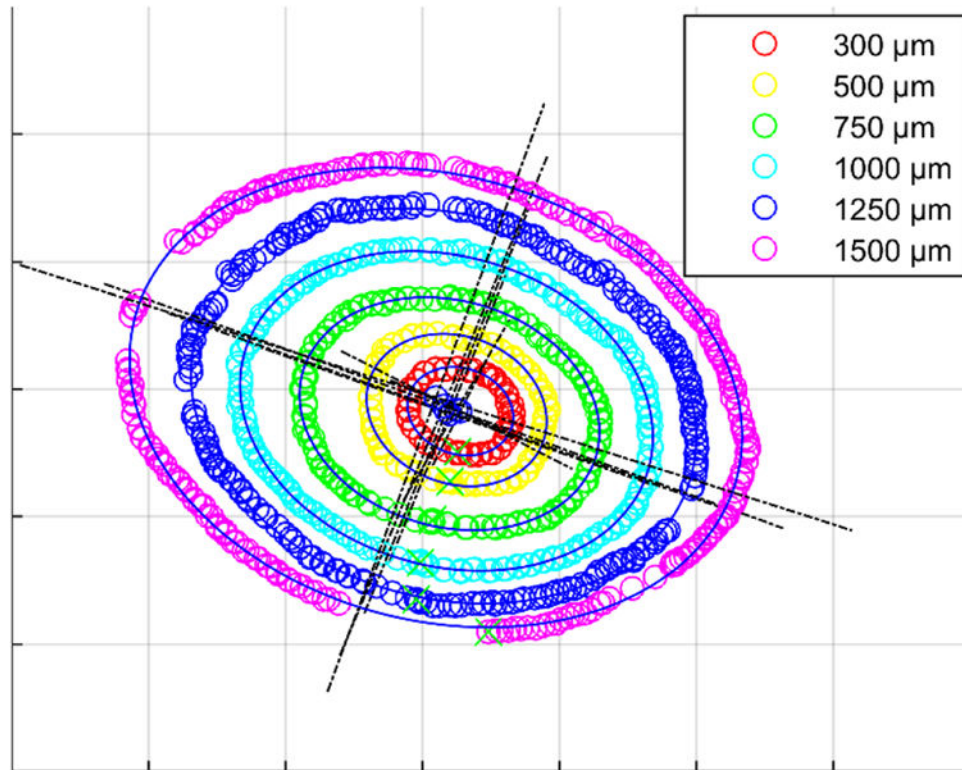
Fig. 12.
Ellipse trajectories with respect to the size of scanning diameter. Valid data points are only presented after removal of outliners using the RANSAC algorithm.
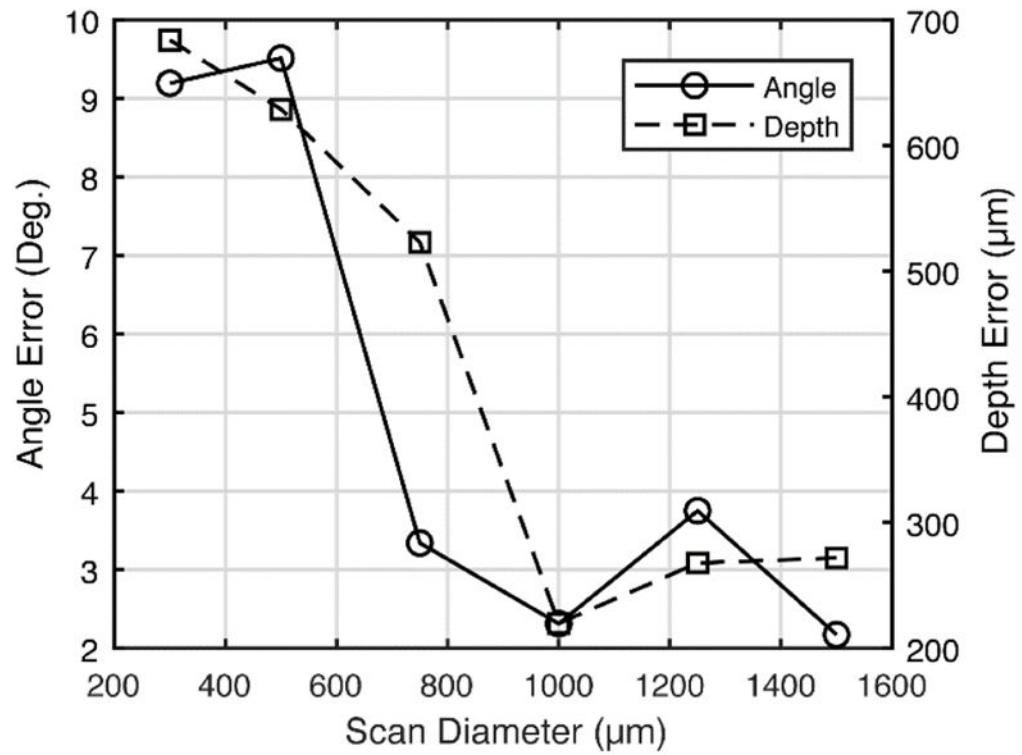
Fig. 13.
Angle and depth errors in surface estimation with respect to scan diameter: angle error marked as circle and depth error marked as square.
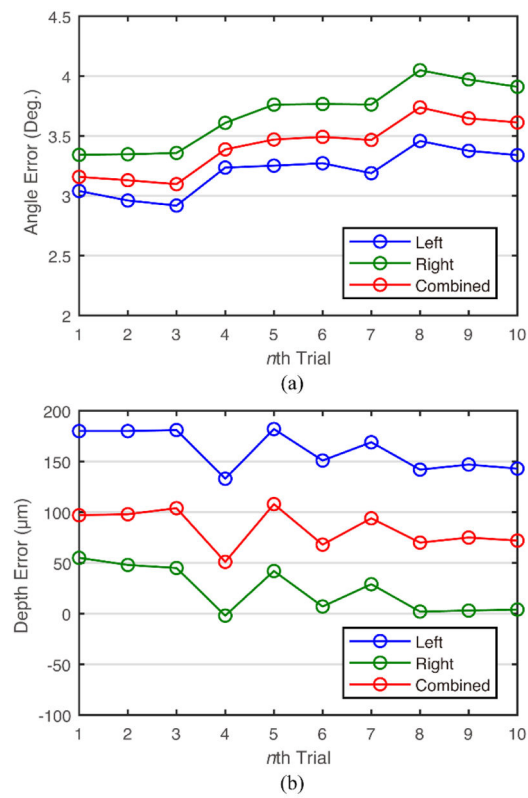
Fig. 14.
(a) Repeated measurements of angle errors for 10 trials. (b) Depth errors for 10 trials.
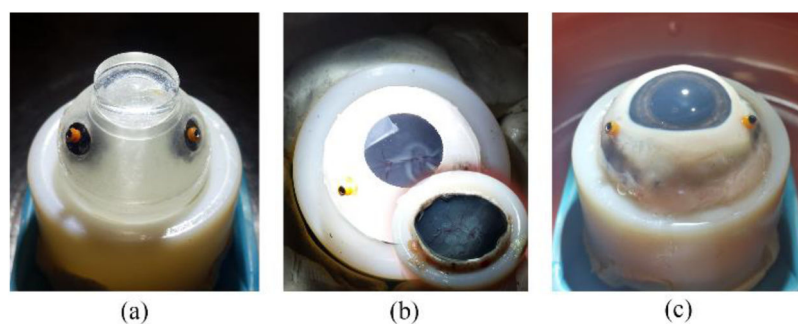
Fig. 15.
(a) Eye phantom filled with water and covered with a contact lens. (b) Dissected porcine eye placed inside the eye phantom; the contact lens is uncovered for visualization. (c) Intact porcine eye with the cornea for ex vivo test.
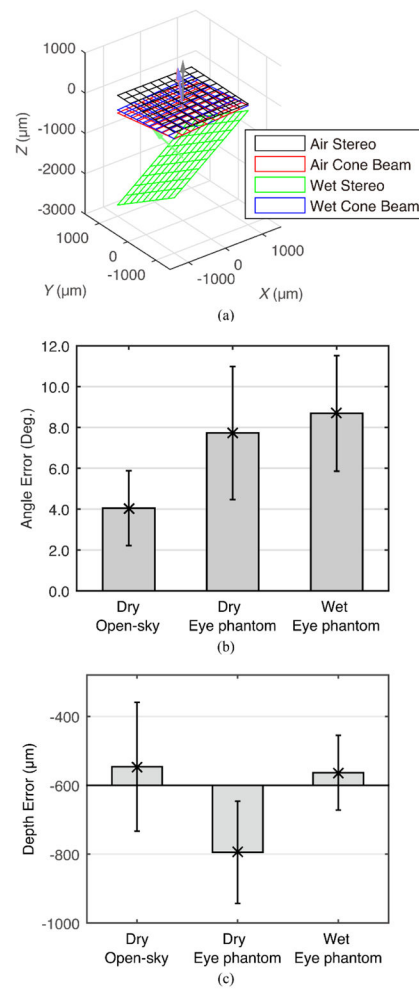
Fig. 16.

(a) Demonstration of the surface reconstruction results using the structured-light reconstruction and stereo reconstruction under dry and wet conditions. Angle and depth errors in surface estimation by handheld scanning of the laser probe: (b) angle errors and (c) depth errors.
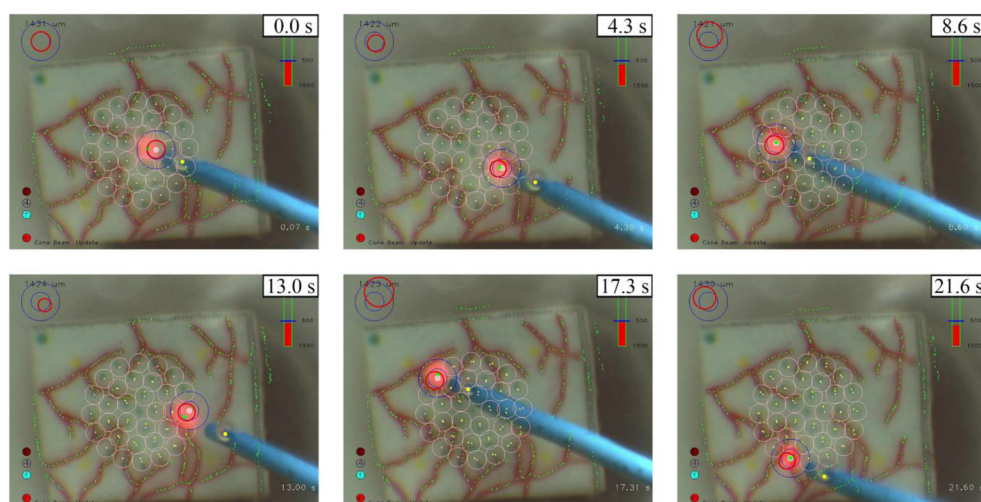
Fig. 17.
Demonstration of hybrid visual servoing based on structured-light reconstruction in the wet eye phantom. Pink circles indicate preplanned targets lying on the inner surface of the eye phantom. Green dots represent artificial blood vessels, as detected by the EyeSLAM algorithm, which tracks the vessels throughout the operation. The red solid circle on a current target represents a visual cue to maintain hand–eye coordination.
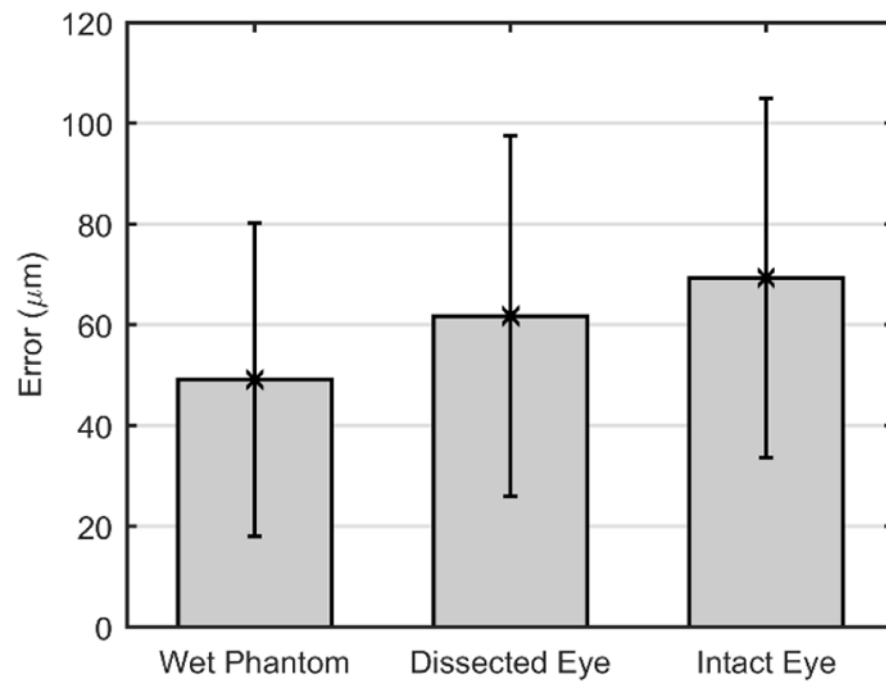
Fig. 18.
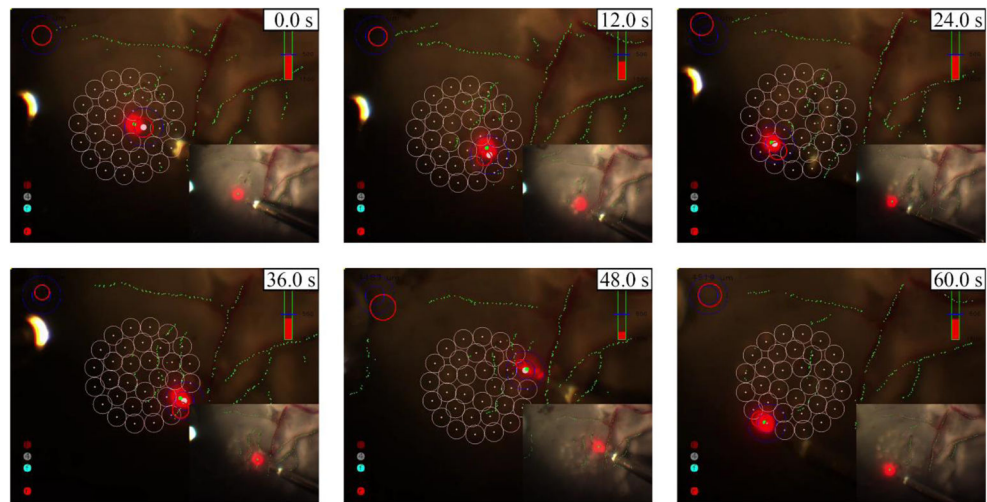Average errors of automated laser photocoagulation for different eye models, with error bars.

Fig. 19.
Demonstration of automated intraocular laser surgery in dissected porcine eye. The dissected eye was placed inside the eye phantom, using water and a lens on top for clear imaging. The main images show the left camera image with aiming beam detection and moving targets. The brighter inset images show the right camera image used to track blood vessels.
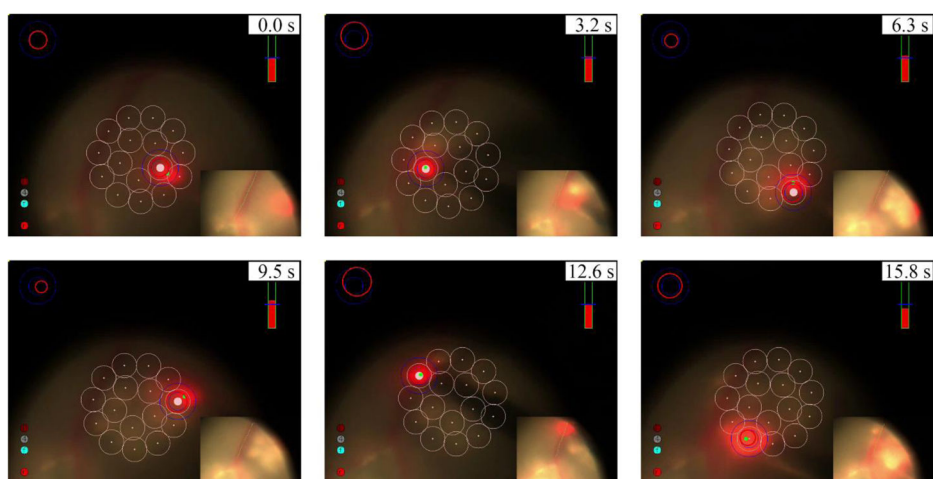
Fig. 20.
Demonstration of automated intraocular laser surgery in the intact porcine eye *ex vivo*.

Table 1

Parameters for evaluation of surface estimation.

Description	Values
Major axis	82.49 pixels
Minor axis	63.65 pixels
Aspect ratio	1.296
Scan diameter	1.0 mm
Remote center of motion distance from the tool tip	21.0 mm
Image scale	8.59 μm/pixel
Tip angle from the retinal surface	50.52°
Tip distance from retinal surface	1.509 mm

Table 2

Settings and results for robot-aided intraocular laser photocoagulation.

Experiment	Laser setting (duration, power)	Medium	Cornea	Contact Lens	EyeSLAM (tracking)	Burn Error (µm)	Speed (target/s)
Wet phantom	20 ms, 3.0 W	Water	X	O	O	49±31	1.48
Dissected porcine eye	50 ms, 1.0 W	Water	X	O	O	62±36	0.71 *
Intact porcine eye	50 ms, 1.0 W	Vitreous humor	O	O	X	NA	1.37 *
Intact porcine eye	50 ms, 1.0 W	Vitreous humor	O	O	O	69±36	1.01

*
without two outliers of the total 32 targets