

# Accurate real-time depth control for CP-SSOCT distal sensor based handheld microsurgery tools

Gyeong Woo Cheon,<sup>1,\*</sup> Yong Huang,<sup>1</sup> Jaepyeng Cha,<sup>1</sup> Peter L. Gehlbach,<sup>1,2</sup>  
and Jin U. Kang<sup>1</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD, 21218, USA

<sup>2</sup>Wilmer Eye Institute, Johns Hopkins School of Medicine, 600 N. Wolfe Street, Baltimore, MD, 21287, USA  
\*gwcheon@jhu.edu

**Abstract:** This paper presents a novel intuitive targeting and tracking scheme that utilizes a common-path swept source optical coherence tomography (CP-SSOCT) distal sensor integrated handheld microsurgical tool. To achieve micron-order precision control, a reliable and accurate OCT distal sensing method is required; simultaneously, a prediction algorithm is necessary to compensate for the system delay associated with the computational, mechanical and electronic latencies. Due to the multi-layered structure of retina, it is necessary to develop effective surface detection methods rather than simple peak detection. To achieve this, a shifted cross-correlation method is applied for surface detection in order to increase robustness and accuracy in distal sensing. A predictor based on Kalman filter was implemented for more precise motion compensation. The performance was first evaluated using an established dry phantom consisting of stacked cellophane tape. This was followed by evaluation in an *ex-vivo* bovine retina model to assess system accuracy and precision. The results demonstrate highly accurate depth targeting with less than 5  $\mu\text{m}$  RMSE depth locking.

©2015 Optical Society of America

**OCIS codes:** (060.2370) Fiber optics sensors; (170.4500) Optical coherence tomography; (150.5758) Robotic and machine control; (170.4460) Ophthalmic optics and devices.

## References and links

1. S. P. N. Singh and C. N. Riviere, "Physiological tremor amplitude during retinal microsurgery," in *Proceedings of the IEEE 28th Annual Northeast Bioengineering Conference* (2002), pp. 171–172.
2. M. Patkin, "Ergonomics applied to the practice of microsurgery," *Aust. N. Z. J. Surg.* **47**(3), 320–329 (1977).
3. R. H. Taylor, P. Jensen, L. Whitcomb, A. Barnes, R. Kumar, D. Stoianovici, P. Gupta, Z. Wang, E. DeJuan, and L. Kavoussi, "A Steady-Hand Robotic System for Microsurgical Augmentation," *Int. J. Robot. Res.* **18**(12), 1201–1210 (1999).
4. R. H. Taylor and D. Stoianovici, "Medical Robotics in Computer-Integrated Surgery," *IEEE Trans. Robot. Autom.* **19**(5), 765–781 (2003).
5. C. N. Riviere, J. Gangloff, and M. De Mathelin, "Robotic compensation of biological motion to enhance surgical accuracy," *Proc. IEEE* **94**(9), 1705–1716 (2006).
6. T. Ueta, Y. Yamaguchi, Y. Shirakawa, T. Nakano, R. Ideta, Y. Noda, A. Morita, R. Mochizuki, N. Sugita, M. Mitsuishi, and Y. Tamaki, "Robot-Assisted Vitreoretinal Surgery: Development of a Prototype and Feasibility Studies in an Animal Model," *Ophthalmology* **116**(8), 1543 (2009).
7. I. Kuru, B. Gonenc, M. Balicki, J. T. Handa, P. L. Gehlbach, R. H. Taylor, and I. I. Iordachita, "Force sensing micro-forceps for robot assisted retinal surgery," in *Proceedings of the 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2012), pp. 1401–1404.
8. B. Gonenc, J. T. Handa, P. L. Gehlbach, R. H. Taylor, and I. I. Iordachita, "Design of 3-DOF force sensing micro-forceps for robot assisted vitreoretinal surgery," in *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2013), pp. 5686–5689.
9. H. Yu, J.-H. Shen, R. J. Shah, N. Simaan, and K. M. Joos, "Evaluation of microsurgical tasks with OCT-guided and/or robot-assisted ophthalmic forceps," *Biomed. Opt. Express* **6**(2), 457–472 (2015).
10. C. N. Riviere and N. V. Thakor, "Modeling and canceling tremor in human-machine interfaces," *IEEE Eng. Med. Biol. Mag.* **15**(3), 29–36 (1996).
11. C. N. Riviere, W. T. Ang, and P. K. Khosla, "Toward Active Tremor Canceling in Handheld Microsurgical Instruments," *IEEE Trans. Robot. Autom.* **19**(5), 793–800 (2003).

12. W. T. Ang, P. K. Pradeep, and C. N. Riviere, "Active tremor compensation in microsurgery," in *Proceedings of the 26th Annual International Conference of IEEE Engineering in Medicine and Biology Society, EMBS* (2004), 1, pp. 2738–2741.
13. R. A. MacLachlan, B. C. Becker, J. C. Tabarés, G. W. Podnar, L. A. Lobes, Jr., and C. N. Riviere, "Micron: An actively stabilized handheld tool for microsurgery," *IEEE Trans. Robot.* **28**(1), 195–212 (2012).
14. S. Yang, M. Balicki, R. a. MacLachlan, X. Liu, J. U. Kang, R. H. Taylor, and C. N. Riviere, "Optical coherence tomography scanning with a handheld vitreoretinal micromanipulator," in *Proceedings of the 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2012), pp. 948–951.
15. S. Yang, M. Balicki, T. S. Wells, R. A. MacLachlan, X. Liu, J. U. Kang, J. T. Handa, R. H. Taylor, and C. N. Riviere, "Improvement of Optical Coherence Tomography using Active Handheld Micromanipulator in Vitreoretinal Surgery," in *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2013), pp. 5674–5677.
16. C. J. Payne, H. J. Marcus, and G.-Z. Yang, "A smart haptic hand-held device for neurosurgical microdissection," *Ann. Biomed. Eng.* (2015).
17. J. P. Ehlers, Y. K. Tao, S. Farsiu, R. Maldonado, J. A. Izatt, and C. A. Toth, "Integration of a spectral domain optical coherence tomography system into a surgical microscope for intraoperative imaging," *Invest. Ophthalmol. Vis. Sci.* **52**(6), 3153–3159 (2011).
18. P. Hahn, J. Migacz, R. O'Connell, J. A. Izatt, and C. A. Toth, "Unprocessed real-time imaging of vitreoretinal surgical maneuvers using a microscope-integrated spectral-domain optical coherence tomography system," *Graefes Arch. Clin. Exp. Ophthalmol.* **251**(1), 213–220 (2013).
19. P. Hahn, J. Migacz, R. O'Donnell, S. Day, A. Lee, P. Lin, R. Vann, A. Kuo, S. Fekrat, P. Mruthyunjaya, E. A. Postel, J. A. Izatt, and C. A. Toth, "Preclinical evaluation and intraoperative human retinal imaging with a high-resolution microscope-integrated spectral domain optical coherence tomography device," *Retina* **33**(7), 1328–1337 (2013).
20. Y. K. Tao, S. K. Srivastava, and J. P. Ehlers, "Microscope-integrated intraoperative OCT with electrically tunable focus and heads-up display for imaging of ophthalmic surgical maneuvers," *Biomed. Opt. Express* **5**(6), 1877–1885 (2014).
21. J. P. Ehlers, S. K. Srivastava, D. Feiler, A. I. Noonan, A. M. Rollins, and Y. K. Tao, "Integrative advances for OCT-guided ophthalmic surgery and intraoperative OCT: microscope integration, surgical instrumentation, and heads-up display surgeon feedback," *PLoS ONE* **9**(8), e105224 (2014).
22. J. P. Ehlers, T. Tam, P. K. Kaiser, D. F. Martin, G. M. Smith, and S. K. Srivastava, "Utility of intraoperative optical coherence tomography during vitrectomy surgery for vitreomacular traction syndrome," *Retina* **34**(7), 1341–1346 (2014).
23. K. M. Joos and J.-H. Shen, "Miniature real-time intraoperative forward-imaging optical coherence tomography probe," *Biomed. Opt. Express* **4**(8), 1342–1350 (2013).
24. K. Zhang and J. U. Kang, "Graphics processing unit accelerated non-uniform fast Fourier transform for ultrahigh-speed, real-time Fourier-domain OCT," *Opt. Express* **18**(22), 23472–23487 (2010).
25. K. Zhang and J. U. Kang, "Real-time intraoperative 4D full-range FD-OCT based on the dual graphics processing units architecture for microsurgery guidance," *Biomed. Opt. Express* **2**(4), 764–770 (2011).
26. X. Liu, I. I. Iordachita, X. He, R. H. Taylor, and J. U. Kang, "Miniature fiber-optic force sensor based on low-coherence Fabry-Pérot interferometry for vitreoretinal microsurgery," *Biomed. Opt. Express* **3**(5), 1062–1076 (2012).
27. X. Liu, M. Balicki, R. H. Taylor, and J. U. Kang, "Towards automatic calibration of Fourier-Domain OCT for robot-assisted vitreoretinal surgery," *Opt. Express* **18**(23), 24331–24343 (2010).
28. J. U. Kang, Y. Huang, K. Zhang, Z. Ibrahim, J. Cha, W. P. Lee, G. Brandacher, and P. L. Gehlbach, "Real-time three-dimensional Fourier-domain optical coherence tomography video image guided microsurgies," *J. Biomed. Opt.* **17**(8), 081403 (2012).
29. K. Zhang, W. Wang, J. Han, and J. U. Kang, "A surface topology and motion compensation system for microsurgery guidance and intervention based on common-path optical coherence tomography," *IEEE Trans. Biomed. Eng.* **56**(9), 2318–2321 (2009).
30. J. U. Kang, J. H. Han, X. Liu, and K. Zhang, "Common-path optical coherence tomography for biomedical imaging and sensing," *J. Opt. Soc. Korea* **14**(1), 1–13 (2010).
31. K. Zhang and J. U. Kang, "Common-path low-coherence interferometry fiber-optic sensor guided microincision," *J. Biomed. Opt.* **16**(9), 095003 (2011).
32. Y. Huang, K. Zhang, C. Lin, and J. U. Kang, "Motion compensated fiber-optic confocal microscope based on a common-path optical coherence tomography distance sensor," *Opt. Eng.* **50**(8), 083201 (2011).
33. Y. Huang, X. Liu, C. Song, and J. U. Kang, "Motion-compensated hand-held common-path Fourier-domain optical coherence tomography probe for image-guided intervention," *Biomed. Opt. Express* **3**(12), 3105–3118 (2012).
34. C. Song, P. L. Gehlbach, and J. U. Kang, "Active tremor cancellation by a "smart" handheld vitreoretinal microsurgical tool using swept source optical coherence tomography," *Opt. Express* **20**(21), 23414–23421 (2012).
35. C. Song, D. Y. Park, P. L. Gehlbach, S. J. Park, and J. U. Kang, "Fiber-optic OCT sensor guided "SMART" micro-forceps for microsurgery," *Biomed. Opt. Express* **4**(7), 1045–1050 (2013).
36. K. Briechele and U. D. Hanebeck, "Template matching using fast normalized cross correlation," *Proc. SPIE* **4387**, 95–102 (2001).
37. R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME-J. Basic Eng.* **82**(1), 35–45 (1960).

38. S.-K. Weng, C.-M. Kuo, and S.-K. Tu, "Video object tracking using adaptive Kalman filter," *J. Vis. Commun. Image Represent.* **17**(6), 1190–1208 (2006).
  39. Y. Zheng, S. Chen, W. Tan, R. Kinnick, and J. F. Greenleaf, "Detection of tissue harmonic motion induced by ultrasonic radiation force using pulse-echo ultrasound and Kalman filter," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **54**(2), 290–300 (2007).
  40. S. H. P. Won, F. Golnaraghi, and W. W. Melek, "A fastening tool tracking system using an IMU and a position sensor with Kalman filters and a fuzzy expert system," *IEEE Trans. Ind. Electron.* **56**(5), 1782–1792 (2009).
  41. R. K. Mudi and N. R. Pal, "A robust self-tuning scheme for PI- and PD-type fuzzy controllers," *IEEE Trans. Fuzzy Syst.* **7**(1), 2–16 (1999).
  42. A. M. Bageci, M. Shahidi, R. Ansari, M. Blair, N. P. Blair, and R. Zelkha, "Thickness Profiles of Retinal Layers by Optical Coherence Tomography Image Segmentation," *Am. J. Ophthalmol.* **146**(5), 679–687 (2008).
  43. D. Cabrera Fernández, H. M. Salinas, and C. A. Puliafito, "Automated detection of retinal layer structures on optical coherence tomography images," *Opt. Express* **13**(25), 10200–10216 (2005).
- 

## 1. Introduction

Accurate and precise tool tip manipulation is imperative when performing microsurgery. A well-known barrier to precision in microsurgery is, physiological hand tremor, occurring predominating in the 6-12 Hz frequency domain with several hundred-micron-order amplitude [1, 2]. Retinal microsurgery is performed by passing microsurgical tools through trocars that provide access through the eye wall (sclera). The trocar not only provides safe access into the eye but it also guides and stabilizes the surgical tools and acts as a remote center of motion (RCM). As a result, the lateral motion of the surgical tool is relatively well stabilized. In the case of axial motion, it is the surgeon that currently provides all of the guidance and control. The freehand tool movements are guided predominantly by visual information acquired from the surgeon's view in the operating microscope. Such tool tip visualization is limited in the axial direction by the ability to resolve small changes in position as well as by obstruction of the retinal surface by the tool shaft and tip. A variety of robot-assisted systems have been developed to actively overcome such a limitation and to achieve the surgical objectives while minimizing surgical risks [3–9]. Early in the evolution of this technology, it was the main point to develop a robot-assisted system with high accuracy and user-friendly operating manipulation [3–7]. Recently, diverse active tool-tips and sensors have been applied to the system to enhance its functionality [8,9].

In parallel, Riviere, *et al.* has been developing a handheld microsurgical system that has differentiated itself from traditional robotic arm based systems [10–15]. The handheld system operates independently without the assistance of external equipment such as a robotic arm by using its own embedded sensors and actuators. The advantages and disadvantages of the two assistance systems are numerous and are still being elucidated. However, advantages of the handheld systems include but are not limited to: lower cost, potential disposability, greater portability, ease of packaging, lower production costs, ease of use, adaptability, ease of customization and elimination of the large, costly cumbersome, robotic-assist systems that must now otherwise be clumsily implemented in the human clinical setting. A fundamental challenge of any handheld microsurgical tool is cancellation of hand tremor. A variety of motion sensors including accelerometer, magnetometer, position-sensitive detectors with infrared LED, etc. and small size of actuators such as piezo-electric motors and voice coils have been introduced. Moreover, a frequency analysis based tremor model has been previously built in accordance with the roughly sinusoidal nature of tremor [10].

Optical coherence tomography (OCT) has recently emerged as an effective imaging modality to support microsurgery due to its micron-order resolution and subsurface imaging capability [17–23]. Microscope mounted OCT systems now play a significant roles in intraoperative visualization for forward viewing during actual surgical procedures. Our research group has established a potential role for a real-time OCT intraoperative guidance system with fast computational speed using a graphic processing unit (GPU) that is forward viewing and incorporated into the actual handheld tools [24–35]. Unlike typical OCT imaging, this application uses OCT as a distal sensor and its potential effectiveness for ocular microsurgery has been demonstrated [29–35]. The motion compensation that is based on the feedback from the OCT distal sensor effectively suppresses undesired axial motion; as a

result, the frequency region related to hand tremor is significantly decreased [34]. This result was applied to microsurgical tool development and improved performance has been demonstrated [35]. We have strategically chosen to incorporate common-path OCT (CP-OCT) into handheld microsurgery tools that are comparatively simple in design, relatively affordable and have the potential to be used as disposable surgical products.

An important difference between intraoperative OCT imaging and OCT distal-sensing is that in application, OCT distal sensing information is used to measure the physical distance of the sensor from the retinal surface. Moreover, in order to utilize OCT as an effective distal sensor, additional signal processing is required to identify and track the surface of a complex target tissue in a complex environment. While surface detection can be simplified to simple peak detection for tissues having a single layer or a dominant reflective surface; the retina has a complex multi-layered anatomy with multiple optically reflecting interfaces that serve as its optical signature. As in a complex surgical environment there is no guarantee that the strongest reflected signal always comes from the retinal surface or surface pathology, the characteristic reflective signal of the retina serves as an advantage that is utilized in the present signal processing algorithms. If this were not anticipated in advance, erroneous reflective data could result in consequential false distance information. Furthermore, real-time operation increases the complexity of the analysis by introducing motion, reduced time for making decisions, and a reduced amount of information available to make real time decisions. Evaluation of multi-layered structures requires edge detection analysis rather than peak detection to correctly find the surface [29]. Edge detection further enhances the surface detection method as the A-scan signal from retina has multi-peak complexity. Therefore, we developed and incorporated a robust distance sensing method utilizing a shifted cross-correlation algorithm. It uses all of the signature information available in the A-scan to determine distance, rather than just the first peak; thus, the system produces robust distance data regardless of whether erroneous data is encountered.

In addition to achieve accurate and stable distal sensing, the handheld system must be designed to efficiently and effectively utilize the distance information to rapidly control the mechanical actuator for accurate and precise manipulation. It is important to appreciate that the time required for the OCT signal to be implemented into tool action requires heavy computational processing including averaging, Fourier transformation, surface detection, and that there is a communication delay that results in time gap between sampling and compensating motion. Therefore, we utilized a GPU to reduce the processing time in the design of this system. Additionally, a predictor was developed to achieve more effective compensation. The predictor based on Kalman filter is used for the system latency compensation. A performance evaluation of this system is presented using a dry phantom consisting of 20 layers of cellophane tape. This is followed by an evaluation in a biologically relevant model, the ex-vivo bovine retina.

## 2. Theory

### 2.1 Spatially shifted cross-correlation

As shown in Fig. 1, the series of A-scan data at T1, T2 and T3 show spatially shifted but mutually correlated waveforms. Because the A-scan data reflects the optical characteristics of the sample below the CP-OCT fiber sensor, it is natural that there is correlation between the spatially and temporally adjacent A-scan data. Thus, we can use this correlation to measure the distance variation. This method can be more effective than an edge detection algorithm to measure the distance variation when the A-scan data has a low signal-to-noise ratio and does not contain a dominant peak. Here is a brief overview of the spatially shifted cross-correlation scheme.

$$X_i * Y_{i+n} = \frac{\sum_i (X_i - m_X)(Y_{i+n} - m_Y)}{\sqrt{\sum_i (X_i - m_X)^2 \sum_i (Y_{i+n} - m_Y)^2}} = \rho_{-n} \quad (1)$$

where  $m_X$  and  $m_Y$  are the means of the sampled discrete A-scan data,  $X_i$  and  $Y_i$  and the subscript  $i$  and  $n$  are the index of A-scan data and shifted value respectively. The calculation of the cross-correlation between A-scan data generally requires a long computation time. Moreover, we need to compare the first A-scan data with every shifted version of the next A-scan data. This is a prohibitively time consuming operation and is not practical for a real-time system; thus, we need to calculate the equation more efficiently [36]. The numerator of the Eq. (1) can be modified as

$$\rho_n^{num} = \sum_i X_i^1 Y_{n-i}^2 \quad (2)$$

where  $X_i^1 = X_i - m_X$ ,  $Y_i^1 = Y_i - m_Y$ , and  $Y_i^2 = Y_{-i}^1$ . Considering the numerator is the convolution of the anterior A-scan data  $X^1$  with the reversed shifting posterior A-scan data  $Y^2$ , this equation can efficiently be calculated based on the discrete correlation theorem with the fast Fourier transform  $F$ :

$$\rho_n^{num} = \mathcal{F}^{-1}(\mathcal{F}(X^1) \cdot \mathcal{F}(Y^2)) = \mathcal{F}^{-1}(\mathcal{F}(X^1) \cdot \mathcal{F}^*(Y^1)) \quad (3)$$

where we used the convolution theorem: time reversal of the signal in time domain is equivalent to the complex conjugate in the Fourier transform. Additionally, the denominator of Eq. (1) is constant regardless of the shifting variable  $n$ . By searching  $n$  that achieves maximum  $\rho$ , we can rapidly find the maximum matching point and distance variation between the two sequential A-scan data with two FFT (forward and inverse) instead of N-times direct cross-correlation computations.

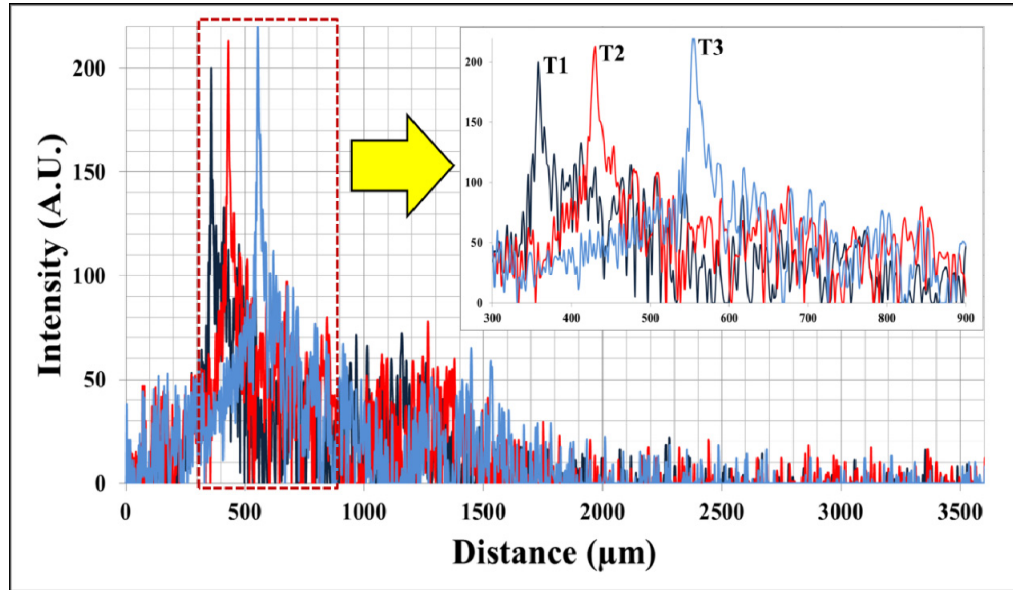


Fig. 1. Three consecutive A-scan data of 20-stacked layers of adherent cellophane tape near the DC line (multi-layered phantom sample); the upper right inserted graph is a partial magnified graph of area in dotted red square in lower entire graph.

## 2.2 Kalman filter

A faster system response is essential in decreasing involuntary hand tremor and to rapidly respond to target motion. The faster system response improves the accuracy of motion compensation over all frequency ranges [34]. However, there is an inevitable time delay between the distal measurement and subsequent compensation due to the processing time, and

also the communication delay between heterogeneous devices such as the workstation and the motor driver. Such a time delay can produce outdated and incorrect tool compensation signal especially in the case of a rapid target or hand movements. To avoid the potential risk of outdated compensation, a predictor is required. Here, we utilized a Kalman filter to adjust for the overall system latency. In this work, the Kalman filter was used as an algorithm that provides an efficient recursive solution to minimize the mean of the squared error [37]. Kalman filters have historically been extensively utilized in a broad area of tracking and assisted navigation systems [38–40]. The Kalman filter serves as an appropriate predictor in our case because it supports estimations of past, present, and future states even when the precise modeling of in our case tremor, is unknown. A portion of the unknown in this system e.g. results from potential encounters with a transiently obstructed imaging axis in the eye (such as blood); this would briefly present erroneous signal to the OCT distal sensor.

The Kalman filter consists of a set of mathematical equations. The equations can be divided into two main categories: 1) Prediction step and 2) Update step.

- Prediction step is

$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \hat{\mathbf{x}}_{t-1|t-1} + \mathbf{B}_t \mathbf{u}_t$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^T + \mathbf{Q}_t$$

- Update step is

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t (\mathbf{y}_t - \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1})$$

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^T + \mathbf{R}_t)^{-1}$$

$$\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1}$$

where  $\hat{\mathbf{x}}$  is an estimated state,  $y$  is a measurement variable, and  $u$  is an control variable. The matrices  $\mathbf{F}$ ,  $\mathbf{B}$ ,  $\mathbf{P}$ ,  $\mathbf{Q}$ ,  $\mathbf{H}$ ,  $\mathbf{R}$ , and  $\mathbf{K}$  describe process equation, control variable mapping, error covariance, process noise covariance, measurement variable mapping, measurement noise covariance, and Kalman gain, respectively. If the tool tip is assumed to move along a straight line in axial direction, 1D motion equation can be applied for the modelling of Kalman filter. Additionally, we assume that the acceleration is a random variable having Gaussian distribution with zero mean and variance  $\sigma^2$  when the tool tip gets to the target position. Then, we can derive the matrices.

$$\mathbf{x}_t = \begin{bmatrix} x_t \\ v_t \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ v_{t-1} \end{bmatrix} + a_t \begin{bmatrix} \Delta t^2 / 2 \\ \Delta t \end{bmatrix} = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{w}_t$$

$$\mathbf{y}_t = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \mathbf{x}_t + \begin{bmatrix} r_t \\ 0 \end{bmatrix} = \mathbf{H}_t \mathbf{x}_t + \mathbf{b}_t$$

where  $x$ ,  $v$ ,  $a$ ,  $b$  are the position, velocity, acceleration of the tool tip, and observation noise. The process and measurement noise covariance are defined respectively as

$$\mathbf{Q}_t = \text{var}(\mathbf{w}_t) = E[\mathbf{w}_t \mathbf{w}_t^T] = \sigma^2 \begin{bmatrix} \Delta t^4 / 4 & \Delta t^3 / 2 \\ \Delta t^3 / 2 & \Delta t^2 \end{bmatrix}$$

$$\mathbf{R}_t = \text{var}(\mathbf{b}_t) = E[\mathbf{b}_t \mathbf{b}_t^T] = \begin{bmatrix} r_t^2 & 0 \\ 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} r_t^2 & 0 \\ 0 & c_1 \end{bmatrix}$$

where  $c_1$  is small constant and the reason to change the form of  $R_t$  is to prevent zero value determinant during Kalman gain calculation. When it comes to error covariance, we set

$$P_0 = \begin{bmatrix} c_2 & 0 \\ 0 & c_2 \end{bmatrix}$$

where  $c_2$  is suitably large constant because we do not know the starting position. Because our system actively compensates the deviation distance from the target position, we can set the control variable and its mapping matrix as

$$B_t u_t = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ y_{tgt} \end{bmatrix}$$

where  $y$  is the measured position and  $y_{tgt}$  is the target position.

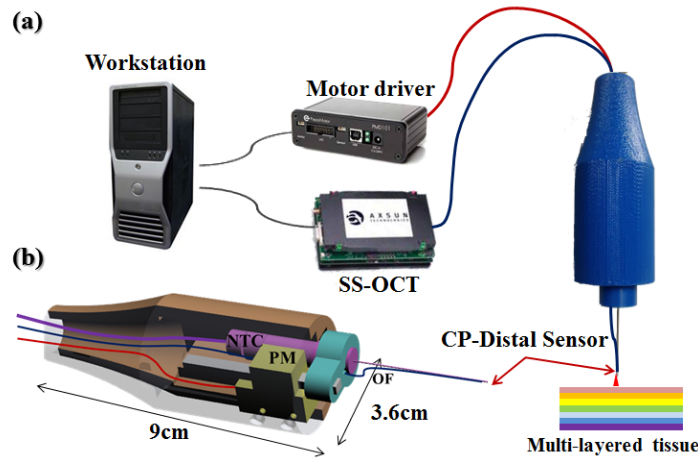


Fig. 2. a) Active depth targeting and locking microsurgery system, b) CAD cross-sectional image of handheld tool (NTC: needle-to-tube connector, PM: piezo-motor, OF: optical fiber).

### 3. Materials and methods

#### 3.1 System

The handheld CP-OCT distal sensor system consists of three main modules: SS-OCT, a workstation, and a piezoelectric linear motor driven handheld tool as shown in Fig. 2(a). Figure 2(b) shows the CAD design of the handheld tool. The OCT system consists of a swept source OEM engine (AXSUN, central wavelength  $\lambda_0$ : 1060nm, sweeping rate: 100kHz, scan range: 3.7mm in air), a photo-detector and a digitizer with a sampling rate of up to 500MSPS with 12-bit resolution, a Camera Link DAQ Board, and a Camera Link frame grabber (PCIe-1433, National Instruments). The workstation (Precision T7500, Dell) with general-purpose computing on graphics processing units (GPGPU, GeForce GTX590, Nvidia) processes the sampled spectral data and control the linear motor. The parallel processing (CUDA, Nvidia) of the GPGPU can reduce the processing time considerably in FFT, background noise subtraction and averaging, which is mainly responsible for the processing delay [24, 25]. The handheld tool uses a piezoelectric linear motor (LEGS-LL1011A, PiezoMotor) which has a maximum speed of 15 mm/s, maximum stroke of 80 mm, stall force of 6.5 N, and a resolution of less than 1 nm. The communication protocol between the motor driver and the workstation is RS232 and the protocol is implemented virtually on top of USB protocol. For this reason, it is important to configure the USB settings suitable for the short and frequent message transmission to minimize the communication delay, which is another factor of the system

delay. Specifically, buffer size and timeout for transmission are the parameters that need to be adjusted. The buffer size was fitted for the most frequent message size and the timeout was set at the minimum value. The housing of the handheld tool as shown in Fig. 2(b) was made by a 3D printing machine. The dimensions of the handheld tool are 3.6 cm diameter and 9 cm length. The optical fiber for distal sensing is attached to the end of the needle and the needle-to-tube connector (RN Coupler-p/n, Hamilton) is mounted into the handheld tool to give a greater load to the linear motor similar to realistic application.

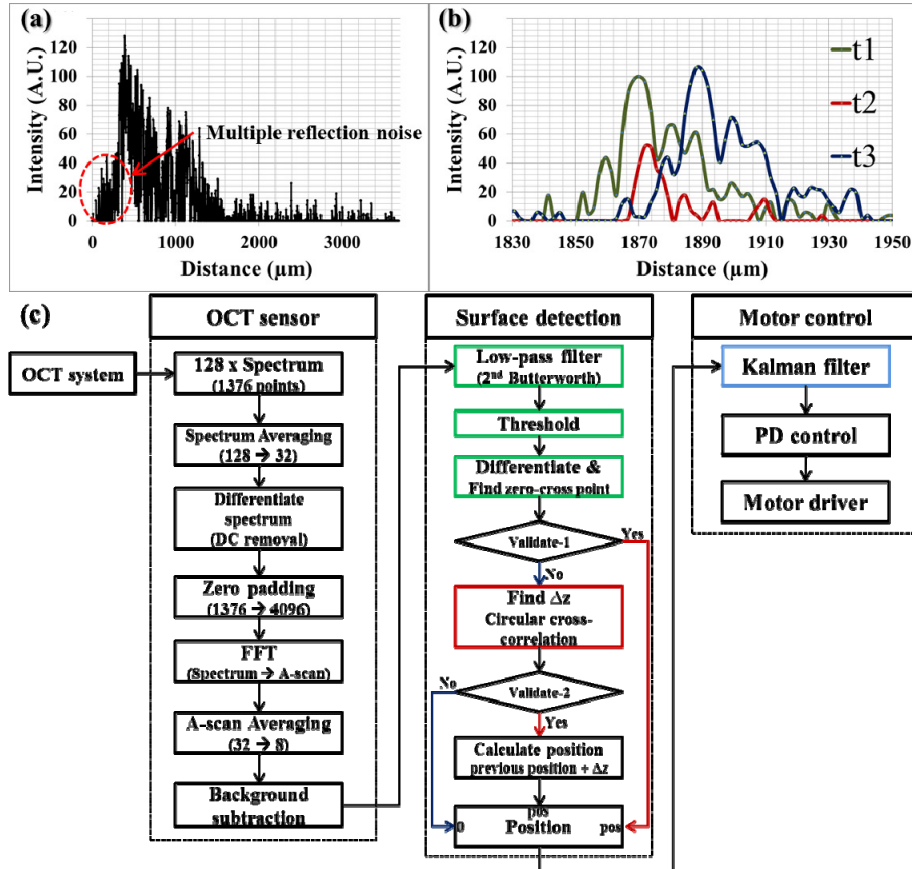


Fig. 3. a) A-scan data of *ex-vivo* open bovine retina, b) three consecutive A-scan data of stacked layers of adherent cellophane tape, c) data processing flowchart consisting of 1) OCT data process, 2) Surface detection, and 3) Motor control.

### 3.2 The control algorithm

The sampled spectral data is sequentially processed in the following three steps as shown in Fig. 3(c). The first step is to transform the spectral data to A-scan data similar to that in typical OCT imaging. Specifically, 128 spectra each consisting of 1376 data points are transmitted from the frame grabber in every 1.28 ms. The system response or sampling rate should be determined based on the computational performance. Comparing to the previous work in our group [34], the system response can be improved from 500 Hz to 780 Hz with the help of parallel processing with GPU. The total computational and communication delay is 650  $\mu$ s on average and we reserved the other half of time (= 630  $\mu$ s) for variable system performance. Four sequential spectra are averaged to increase the signal-to-noise ratio (SNR) followed by a discrete differentiation to remove the DC component in frequency domain. Then, zero padding is applied to increase the data points of the averaged spectra from 1376 to



4096 before fast Fourier transform (FFT). After the transformation, additional averaging and background subtraction are conducted on the A-scan data. As a result, we have 8 A-scan data and each A-scan data has 2048 data points that covers a 3686.4  $\mu\text{m}$  axial distance in the air (1.8  $\mu\text{m}$  between each data point). Note that it is important to remove the background noise in A-scan data for the subsequent A-scan matching process using shifted cross-correlation in order to find the distance variations since this process is susceptible to the stationary background noise. The final processed A-scan data are used for the next step, the surface detection.

As shown in Fig. 3(a), peak detection is not suitable for the surface detection in a sample with a complex multi-layered structure. Alternatively, an edge detection scheme can be used. For the edge detection, first, 2nd order Butterworth low-pass filtering is applied to extract the envelope shape of the A-scan data because the A-scan data contains multiple local maxima and minima that hinder the edge detection. Then, a thresholding is applied to discriminate the signal from the noise. The noise level without sample (= background noise) determines the minimum thresholding value. Whereas, the maximum value is set to be adaptive to the maximum intensity of the current signal because the noise level, as indicated by the area with red dotted circle in Fig. 3(a) between the sample and the zero delay line, increases as the fiber sensor tip approaches the sample. This is because multiple reflections in close proximity induce stronger interference; as a result, the noise level also increases. Next, the first zero-cross point is identified to find the surface after a discrete differentiation.

Two processes validate the calculated position: checking the value itself (if the position is zero, the detection is failed) and comparing it with the previous position. According to the failure of position detection, there are two possible cases. If the intensity of A-scan data is smaller than the threshold value, the surface detection is considered to be failed even though the A-scan data still has sufficient features to find the surface as shown in Fig. 3(b). In another case, we see that the distance variation between the two consecutive positions has an abnormally large value as compared to the presumed maximum value. This is mainly because the predominant reflective layer within a sample makes the adaptive thresholding too high so that it erases the surface peak as well as the background noise. To make up for the limitations of simple edge detection, we used a spatially shifted cross-correlation as described in the previous section. As shown in Fig. 3(b), the A-scan data at  $t_2$  has similar features as the A-scan data at  $t_1$  and  $t_3$  in spite of its low intensity. In this case, spatially shifted cross-correlation can find the matching point and distance variation as a result. The result is validated based on the normalized cross-correlation value, which reflects how the two A-scan data look similar. In the real implementation, we set 60% as a threshold.

Finally, the calculated position data is transmitted to the third step, motor control module. Instead of using the position data directly to control the motor, a Kalman filter is applied as a predictor for more accurate motion compensation. The predictor is needed to make up for the time delay between the distal measurement and the actuator movement due to the processing and communication delay, which is 650  $\mu\text{s}$  on average in the implemented system. The previous prediction value is alternatively used as an input data when position detection is failed in the previous step as long as the number of consecutive failure is smaller than the predetermined value. This predicted position is then used to calculate the actual compensating length through modified proportional-derivative (PD) control. Typical proportional-integral-derivative (PID) control is not appropriate because the distance between sample surface and tool tip is continuously varying. At the same time, the compensating length should be calculated in accordance with the separate length between target and tool tip and current speed of the tool. The parameters of PD control are modified adaptively based on the current trend of the process [41]. Finally, the compensating value is transmitted to the motor driver.

## 4. Experimental results and discussion

### 4.1 Phantom experiment

Two different experiments were conducted to evaluate the performance of the proposed distal sensing algorithm with the dry phantom in each case consisting of 20-stacked layers of adhesive cellophane tape. The phantom was set on the table at a 10-degree tilt angle to better represent the non-perpendicular orientation of instruments in application. First, we tested the effectiveness of the edge detection with the shifted cross-correlation method by comparing it with the edge detection only. The experiment was conducted with a random delay for a blind test and the shifted cross-correlation method was repeatedly turned on and off in every 1.28s. The offset value ( $= 921.6\mu\text{m}$ ) was intentionally added to the data from the test using the edge detection with shifted cross-correlation (yellow region) to easily distinguish these two different cases as shown in Fig. 4(a). The dark red vertical lines in Fig. 4(a) indicates detection failure and the number of vertical lines or the failure rate in the yellow region is much lower than in the other region (white region). In fact, only a single case of failure at around 8.1 second was observed over the 20-second observation when the shifted cross-correlation method was used. The failure rate without the shifted cross-correlation method was 28 incidents/s and 0.3 incidents/s with the shifted cross-correlation method. These results clearly indicate that the addition of the shifted cross-correlation methodology can greatly improve the stability of the OCT distal sensor.

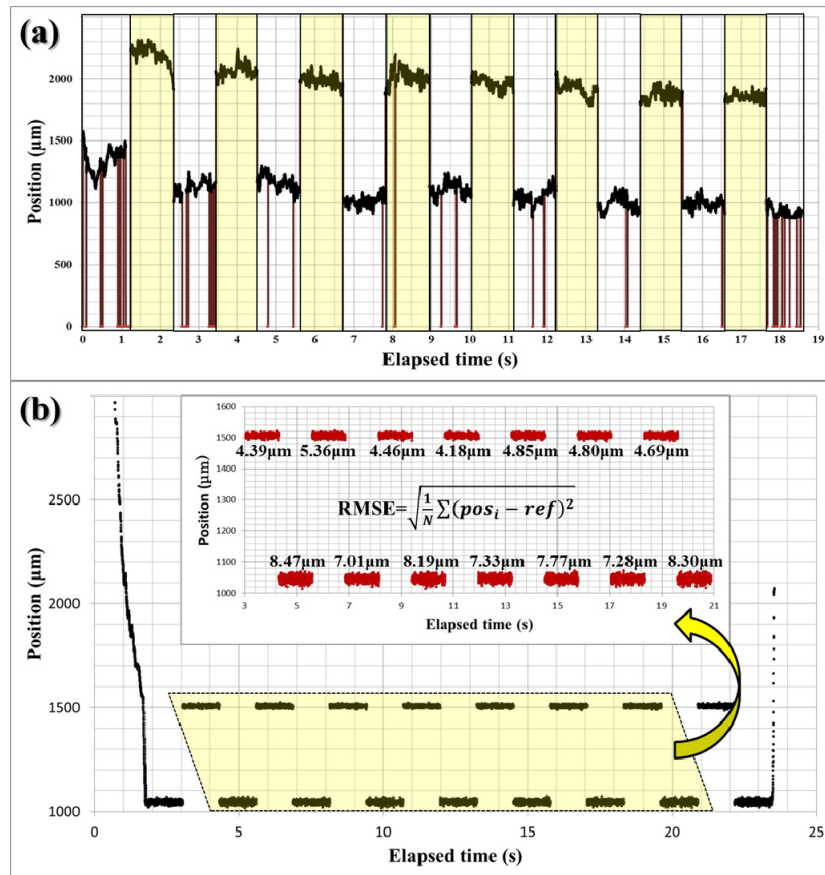


Fig. 4. Dry phantom experiment: a) evaluation test for the surface detection algorithm with the auxiliary shifted cross-correlation in non-depth-locking freehand, b) evaluation test for the depth-locking with the predictor based on Kalman filter: the upper inserted graph is a magnified graph of area in dotted yellow trapezoid in the entire graph.

In the second experiment, we evaluate the effectiveness of the predictor based on a Kalman filter. During the experiment, the predictor method was also repeatedly turned on and off in each 1.28s increment. The offset value ( $= 460.8\mu\text{m}$ ) was added to the data with a predictor to easily differentiate the result from the data without the predictor. In fourteen consecutive trials were conducted in total to maintain a fixed position and the results are shown in Fig. 4(b). The upper graph in Fig. 4(b) is a magnified view of the area in the dotted yellow trapezoid. The magnified view includes the RMSE for each trial. The average RMSE without the predictor is  $7.76\mu\text{m}$  and the average RMSE with the predictor is  $4.68\mu\text{m}$ . This result shows that the predictor operated effectively to make up for the processing time delay by deducing a more probable position when the linear motor actually operated.

#### 4.2 Bovine retina experiment

To evaluate the performance of the overall compensation system, we used an *ex-vivo* bovine retina model. Figure 5(a) shows the *ex-vivo* bovine retina after removing cornea, lens, and vitreous humor. The experiment was conducted on the same day as death thereby preserving the multi-layered structure of bovine retina, which gradually degrades beginning after the first few hours of the post mortem. We designed the experiment to evaluate the performance of the system in terms of dynamic depth targeting and precise depth locking by repeatedly holding and changing the position in a prescribed manner. Specifically, the user carried the tool tip to the surface of retina as shown in Fig. 5(b) and the experiment starts when the tool tip reached the predetermined offset distance. The offset distances were predetermined to be six different jumping distances from  $10\mu\text{m}$  to  $60\mu\text{m}$  in order to maintain the same reference line. Moreover, the system was programed to hold its current position for 1.28s ( $= 1000$  data points) to evaluate its depth-locking performance.

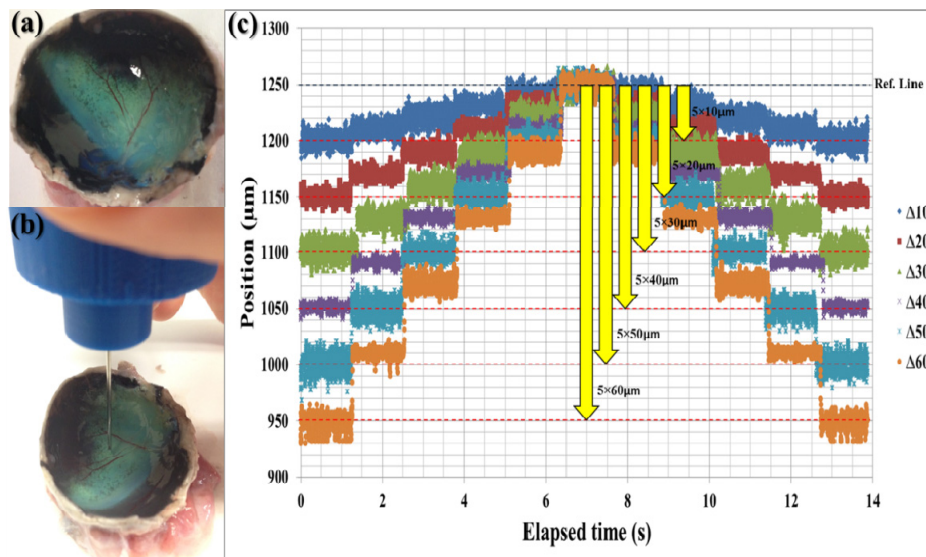


Fig. 5. Bovine retina experiment: a) bovine retina after removing cornea, lens, and vitreous humor, b) snapshot of the experiment, c) depth targeting and locking experiment result with six different jumping distance from  $10\mu\text{m}$  to  $60\mu\text{m}$ .

As shown in Fig. 5(c), the system consistently adjusted the needle position in accordance with the predetermined jumping distance. In total, there were 60 depth-targeting and 66 depth-locking trials. Table 1 shows the specific RMSE values of each depth-locking trial in Fig. 5(c). Figure 6 shows the average RMSE and standard deviation error bar with different jumping distance. The total average, maximum, and minimum of RMSE were  $4.996\mu\text{m}$ ,  $8.129\mu\text{m}$  and  $2.104\mu\text{m}$  respectively. The average depth locking RMSE,  $4.996\mu\text{m}$ , is smaller than  $1/5$  thickness of inner nuclear layer (INL),  $27 \pm 8\mu\text{m}$  (Mean  $\pm$  Standard deviation),

which is thinnest layer in human retina [42, 43]. Even worst RMSE, 8.129  $\mu\text{m}$ , is about 1/3 of INL thickness. These figures demonstrate the potential for this microsurgery system to be used for targeting and delivering a therapeutic agent into specific layer of the human retina.

**Table 1. Optical Constants of Thin Films of Materials [ $\mu\text{m}$ ]**

|           | $\Delta 10$ | $\Delta 20$ | $\Delta 30$ | $\Delta 40$ | $\Delta 50$ | $\Delta 60$ |
|-----------|-------------|-------------|-------------|-------------|-------------|-------------|
| <b>1</b>  | 5.264       | 4.562       | 6.750       | 2.944       | 8.129       | 5.939       |
| <b>2</b>  | 5.080       | 3.774       | 6.222       | 3.108       | 7.404       | 3.950       |
| <b>3</b>  | 6.473       | 4.297       | 7.161       | 3.038       | 5.447       | 5.488       |
| <b>4</b>  | 4.683       | 3.416       | 5.494       | 3.376       | 5.630       | 3.141       |
| <b>5</b>  | 5.233       | 5.017       | 6.114       | 3.891       | 5.061       | 4.823       |
| <b>6</b>  | 6.286       | 4.007       | 5.838       | 3.170       | 5.761       | 4.981       |
| <b>7</b>  | 6.367       | 4.078       | 6.427       | 2.600       | 5.513       | 4.344       |
| <b>8</b>  | 6.830       | 3.938       | 6.377       | 2.887       | 5.642       | 3.937       |
| <b>9</b>  | 5.148       | 4.187       | 6.685       | 3.181       | 5.275       | 5.118       |
| <b>10</b> | 5.259       | 3.200       | 7.342       | 2.104       | 6.826       | 2.941       |
| <b>11</b> | 6.063       | 4.255       | 6.603       | 2.982       | 7.277       | 5.388       |

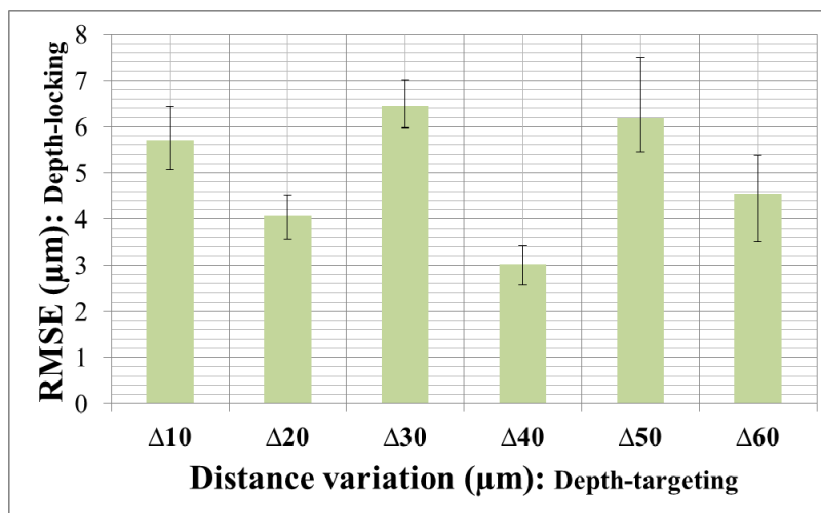


Fig. 6. Averages of RMSE with standard deviation error bar of the bovine retina experiment. The x-axis labels mean jumping distance and the y-axis indicates the means and standard deviations of the measurement specified in Table 1.

## 5. Conclusion

In conclusion, we have demonstrated a novel motion compensation control methods that allows accurate depth targeting and precise depth locking with micron-order RMSE. We have evaluated the technology in a dry phantom and in biological specimens possessing a multi-layered tissue structure. Future work will focus on implementing the potential for precise tool tip actions that this technology enables. The retina is an ideal model in which to assess the ability to e.g. use a derivative technology such as a high precision micro-injector system to accurately place potentially therapeutic agents in previously inaccessible retinal layers and micro-forceps system to precisely peel the epithelial layer of retina.

## Acknowledgments

This research was supported by the U.S. National Institutes of Health and the National Eye Institute (NIH/NEI) Grant R01EY021540-01. Research to Prevent Blindness; The J. Willard and Alice S. Marriott Foundation. The Gale Trust and a generous gift by Mr. Herb Ehlers and Mr. Bill Wilbur. Jaepyeong Cha is a Howard Hughes Medical Institute International Student Research Fellow.