



Published in final edited form as:

Int Symp Med Robot. 2023 April ; 2023: . doi:10.1109/ismr57123.2023.10130250.

Robotic Optical Coherence Tomography of Human Subjects with Posture-Invariant Head and Eye Alignment in Six Degrees of Freedom

Mark Draelos¹, Pablo Ortiz², Amit Narawane², Ryan P. McNabb³, Anthony N. Kuo^{3,2}, Joseph A. Izatt^{2,3}

¹Departments of Robotics and Ophthalmology, University of Michigan, 2505 Hayward St, Ann Arbor, MI USA.

²Department of Biomedical Engineering, Duke University, 101 Science Dr, Durham, NC USA.

³Department of Ophthalmology, Duke University Medical Center, 2351 Erwin Rd, Durham, NC USA.

Abstract

Ophthalmic optical coherence tomography (OCT) has achieved remarkable clinical success but remains sequestered in ophthalmology specialty offices. Recently introduced robotic OCT systems seek to expand patient access but fall short of their full potential due to significant imaging workspace and motion planning restrictions. Here, we present a next-generation robotic OCT system capable of imaging in any head orientation or posture that is mechanically reachable. This system overcomes prior restrictions by eliminating fixed-base tracking components, extending robot reach, and planning alignment in six degrees of freedom. With this robotic system, we show repeatable subject imaging independent of posture (standing, seated, reclined, and supine) under widely varying head orientations for multiple human subjects. For each subject, we obtained a consistent view of the retina, including the fovea, retinal vasculature, and edge of the optic nerve head. We believe this robotic approach can extend OCT as an eye disease screening, diagnosis, and monitoring tool to previously unreached patient populations.

Index Terms—

Medical robotics; optical coherence tomography; motion stabilization

I. INTRODUCTION

Optical coherence tomography (OCT) [1], a laser-based in-terferometric technique for volumetric imaging, has become an indispensable tool for ophthalmic diagnostics [2], [3]. Tabletop OCT systems for clinical use are now commonplace in ophthalmology offices, and practice patterns published by ophthalmic professional societies recommend OCT for diagnosis or management of many eye diseases, including diabetic retinopathy

[4], age-related macular degeneration [5], and glaucoma [6]. Despite two decades of OCT technology advancement [7] that have brought faster acquisitions, higher sensitivity, and new modalities (e.g., OCT angiography [8]), significant barriers to access and broad applicability remain, primarily due to the tabletop designs and operator presence requirements of commercially-available clinical systems. Consequently, OCT is largely unavailable for patients beyond ophthalmology specialty clinics, limiting its role in both primary care and hospital settings. This represents a missed opportunity for early detection of the above diseases that otherwise may progress to retinal blindness at a societal cost of billions of dollars [9], [10].

The use of mechanical head stabilization in tabletop designs underlies the first barrier to OCT beyond the specialist's office. As applied in ophthalmology, OCT is effectively *in vivo* volumetric microscopy. Indeed, widely available clinical systems often achieve axial and lateral resolutions of single and tens of micrometers in air, respectively, and seek to capture ocular microanatomy at similar scales. Combined with long acquisitions of high-resolution volumes for diagnostic purposes, this renders OCT highly susceptible to motion artifact. Whereas a microscopist could simply fix the sample to eliminate motion for a typical microscopy application, an ophthalmic photographer must contend with the physiologic eye motion and cooperation of an awake, breathing patient. In clinical practice, the most common mitigation strategy is bracing the patient's head against a tabletop OCT system using a chinrest and forehead strap, thereby mechanically attenuating motion. Patients who cannot use a chinrest, due to physical mobility limitations, altered states of consciousness, or difficulty following directions, are therefore ineligible for OCT imaging.

Complexity of operation produces the second barrier to broadly available OCT. To acquire high-quality images, careful optical alignment between the OCT instrument and the patient's eye is required. Although a handful of clinical OCT systems offer some form of automated alignment, usually with a motorized chinrest, this alignment is available over limited range and relies upon mechanical alignment guides. Consequently, ophthalmology offices provide dedicated imaging suites in which trained ophthalmic photographers perform diagnostic-quality OCT imaging. This involvement of additional space and personnel becomes prohibitive when the deployment of OCT to general medical settings, such as primary care offices, is considered.

We first introduced robotic OCT [11], [12], in which a robot-mounted scanner performs non-contact imaging, and demonstrated autonomous imaging of freestanding subjects to position OCT as more than an exclusive technique of the ophthalmic specialist. Our later work [13], [14] extended robotic OCT with gaze tracking for five degree-of-freedom (DoF) eye pose measurement and gaze optical aiming for elimination of fixation requirements. While suitable for imaging under controlled conditions, this system's dependence on fixed-base tracking cameras and need for forward-facing head orientations were incompatible with unstructured clinical environments where portability and flexibility are paramount. Outside of ophthalmic photography suites where a seated posture is mandatory and the head is mechanically stabilized, patients in the clinic, hospital ward, intensive care unit, and operating room are encountered in widely varying postures (e.g., seated vs. supine) with largely unrestricted head orientations.

We now present a next-generation robotic OCT system capable of imaging in any configuration that is mechanically reachable by our robot arm (Fig. 1a). By moving the face tracking cameras from a fixed-base to a scanner-integrated paradigm, we provide a tracking workspace that moves with the scanner and maintains a full frontal face view despite occlusion during alignment. By generalizing our robot motion planner to fuse face and eye tracking in six DoFs, we produce smooth scanner alignment trajectories that continuously respond to head and eye motion. With this next-generation robotic system, we show repeatable imaging of multiple human subjects under widely varying head orientations and postures, a key step towards general clinical utility that promises to enable OCT for large-scale screening and diagnosis of common retinal diseases. Our specific contributions in this paper are

1. a hardware-accelerated multiple-view eye-in-hand face tracking system,
2. a six DoF face alignment planner that maintains landmark visibility during plan execution,
3. an architecture for safe robotic eye imaging in postures that restrict mobility, and
4. a methodology for *in vivo* experiments that evaluate repeatable imaging under varying head orientations and postures.

II. RELATED WORK

Handheld ophthalmic OCT [15]–[18] has emerged as an alternative to tabletop OCT for clinical imaging in challenging environments, such as operating rooms [19] and intensive care units [20]. Under this arrangement, an expert operator positions the OCT probe by hand to image the patient's eye without need for a chinrest or patient cooperation. Although considerably more flexible than tabletop OCT, handheld OCT increases susceptibility to motion and requires a particularly skilled operator to align the probe. Moreover, the operator's physical stamina is a consideration and fatigue becomes a limiting factor for imaging, especially over prolonged periods [20]. Thus, moving the OCT probe from the tabletop to the hand strengthens barriers to accessible OCT.

An intermediate between tabletop and handheld OCT is microscope-integrated OCT [21]. Here, the OCT system is freed from the table and suspended over the patient through integration with a floor- or ceiling-mounted microscope. This scheme is subsequently found primarily in operating rooms dedicated to ophthalmic procedures and has recently become clinically available. Others have adopted this OCT positioning paradigm but have instead developed portable stands or booms for non-surgical applications [22]. Except for the latter, a suspended OCT system is even less accessible due to its likely location with an ophthalmic surgery center.

Maneuvering an OCT probe with a robot arm has recently gained traction in the research community. Sprenger et al. [23] and Marques et al. [24] used a robot arm to scan an OCT system's field of view over a larger surface. Huang et al. [25] extended this approach with feedback from the OCT system to improve data acquisition. Similarly, Jivraj et al. [26] combined an OCT system for axial ranging with a robotic laser ablation platform to

maintain a fixed distance from the tissue target. These systems either lack sensing necessary to track a target in six DoFs or react to target misalignment on a time-scale too slow for ophthalmic application to awake patients.

III. ROBOTIC OCT

A. Overview

Robotic OCT uses a multi-scale tracking and alignment strategy for obtaining OCT images without mechanical head stabilization. During coarse alignment, the robot arm brings the OCT scanner to the patient guided first by face tracking cameras and then by pupil tracking cameras, once they detect the eye. For fine alignment, pupil tracking cameras drive optical components within the OCT sample and reference arms to correct for lateral and axial error, respectively, and augment the scan waveforms to correct for gaze error. For this work, we performed hardware upgrades to the system previously reported in [13] to support scanner-integrated face tracking, detect contact near the scanner's objective lens, and increase reach with a UR5e robot arm (Universal Robots; Odense, Denmark) optionally mounted on a vertical translation stage (Ewellix; Utrecht, The Netherlands) (Fig. 1). Briefly, this system implements a galvanometer-scanned (Pangolin Laser Systems; Sanford, FL) 4F retinal telescope with 10 cm working distance, a fast-steering mirror (Optics In Motion; Long Beach, CA) in the retinal conjugate plane, and voice-coil motor (H2W Technologies; Santa Clara, CA) in the reference arm.

B. Multiple-View Face Tracking

Prior system generations relied upon fixed-base face cameras that could not provide the necessary tracking data for head orientations other than forward-facing or during alignment when the scanner occluded the face. This next-generation system instead integrated two RealSense D435 depth cameras (Intel; Santa Clara, CA) into the scanner itself (Fig. 1b–c), an arrangement which allows detection of faces pointed roughly within $\pm 45^\circ$ of the scanner. This setup allowed us to not only maintain face tracking throughout the imaging session but also dynamically reconfigure the face tracking workspace by reorienting the robot end-effector. At close range, however, such as when the scanner is aligned, neither camera obtained a full view of the face from their side-mounted vantage point. We therefore re-projected the depth images from both face cameras to the image plane of a synthetic camera with focal point at the robot's end-effector (Fig. 2).

Face detection and landmark labeling was subsequently performed on the GPU using dlib [27] and its bundled models (Fig. 2d). To assign unique identities to each detected face, we clustered face detections over space and time. Faces detected in a given frame needed to uniquely match the position of detections from a moving window of prior frames in order to maintain their identity. This allowed the system to operate the presence of multiple individuals and consistently align with the intended face. We fit a head pose for each face using the 3D positions of eyes and mouth as determined by mapping facial landmark regions to the point cloud. In our head coordinate system, the line between the eyes defined the horizontal axis whereas the orthogonal line through the mouth defined the vertical axis.

C. Face Alignment Planner

Equipped now with scanner-integrated face cameras, we extended our face alignment capabilities to six DoFs using a dedicated motion planner (Fig. 3). This planner generated a Cartesian motion plan to smoothly align the scanner with the pupil given the scanner pivot frame and a guide point along the scanner's optical axis. In these plans, both the scanner pivot and the guide point moved along the line segment in Cartesian space between their current and target positions at matching rates. The guide point, however, only defined the orientation of the scanner's optical axis (i.e., the vector from the guide point to the scanner pivot). In addition, the scanner rotated at a uniform rate around its optical axis to match the vertical axis of the head. This guaranteed that forward progress towards the eye would keep the face tracking cameras oriented towards the face for a wide variety of head poses (Fig. 4) and yielded smooth motions that would not startle or concern subjects. We generated time-optimal trajectories along these paths using TOPP-RA [28] that respected joint and Cartesian velocity limits. We checked trajectories for self-collision, but the structure of the imaging workspace avoided the need to check for environmental collisions.

Furthermore, path and trajectory generation updated for new face tracking results at 5–10 Hz, depending upon alignment distance-to-go and joint-space path complexity. When revising an existing trajectory, the planner chose a branch point along that trajectory 200 ms in the future as its initial conditions. After calculating the new path, we determined the interval of feasible initial path velocities using TOPPRA. The planner then sampled these initial velocities and generated corresponding trial trajectories using Ruckig [29] that matched the branch point initial conditions. The shortest duration trial trajectory was then used to blend in the new trajectory at the branch point.

D. Pupil Tracking and Alignment

During fine alignment based on five DoF pupil tracking [13], we generated on-the-fly trajectories using Ruckig [29] to first position the pupil pivot and then orient the scanner along the eye's optical axis, again respecting robot joint velocity limits (Fig. 5). Fine alignment trajectory generation updated at 125 Hz. Due to the eye's rotational symmetry, pupil tracking could not report eye roll (i.e., z -axis rotation in Fig. 5) so we substituted the head roll (i.e., z -axis rotation in Fig. 2d) as the missing DoF for the fine alignment orientation. When the pupil view exceeded a relative gaze angle of about 10° , as may happen for discordant head and eye orientations, we tracked only the pupil pivot and maintained the scanner orientation until gaze tracking resumed. With scanner-integrated face tracking cameras, we required full frontal face tracking during fine alignment for subject safety.

E. System Integration

We implemented face tracking, motion planning, and robot control in Python on a Windows PC. The workload was distributed across multiple processes with message passing via Cyclone DDS to achieve true concurrency with Python. We implemented pupil tracking in C++ which communicated results to our control software via the local network. The robot control process streamed joint-space setpoints to the robot at 125 Hz with a lookahead of several setpoints to compensate for computational or network latency. It further monitored joint velocities and end-effector forces as part of our safety architecture (Section IV-D).

IV. METHODS & EXPERIMENTS

A. OCT System and Engine

We acquired retinal OCT data using the robotic OCT scanner described above and a custom swept-source OCT engine. This system featured a 100 kHz or 200 kHz swept-source (Axsun Technologies; Billerica, MA) at 1060 nm with 100 nm bandwidth and a balanced photoreceiver signal detection chain (Thorlabs; Newton, NJ) with a 1.8 GS s⁻¹ digitizer (AlazarTech; Quebec, Canada). For the 100 kHz source, we performed volumetric imaging at 500 vx × 500 vx × 2752 vx using 2 × spectral upsampling with ten interleaved aiming scans, yielding a volume interval of 4.5 s. For the 200 kHz source, we performed high-speed volumetric imaging at 200 vx × 500 vx × 1024 vx, yielding a volume interval of 0.87 s. Real-time OCT processing and display was accomplished in Python using the Vortex open-source OCT library [30]. We performed rigid registration of adjacent cross-sections axially and laterally using cross-correlation in post-processing.

B. Human Imaging Experiments

We performed imaging of human subjects to evaluate the flexibility and repeatability of our system under varying head orientations and postures (Fig. 4). We considered head orientation and postures separately because, prior to installation of the vertical translation stage partway through our study, the elevated base of our system had limited angular reachability for reclining or supine subjects.

For head orientation evaluation, we asked subjects to stand or sit on a stool while looking towards targets positioned around the room. These targets produced head positions that included neutral forward-facing (baseline), right and left yaw, right and left roll, and up and down tilt. For each configuration, the robot semi-automatically aligned coarsely and finely with the subject's chosen eye when triggered by the operator. Once fine alignment had been achieved, the operator acquired 4–6 OCT volumes with scanner active aiming enabled. A total of six identical configurations were used across all human subjects.

For posture evaluation, we asked subjects to stand, sit upright on a stool, and lean back or lay supine using a recliner. These seated or recliner configurations yielded head elevation angles of about 90°, 60°, and 30°, respectively. For each configuration, the robot automatically aligned coarsely and finely with the subject's left and right eyes when triggered by the operator. Once fine alignment had been achieved, the operator acquired 5–10 OCT volumes with scanner active aiming enabled. A total of four identical configurations were used across all human subjects with minor variations due to stature.

C. OCT Image Analysis

We evaluated the captured OCT volumes qualitatively and quantitatively. For qualitative evaluation, we visually inspected *en face* projections and foveal cross-sections to ensure consistent visibility of relevant structures, including retinal vessels and layers, respectively, across head orientations and postures. For quantitative evaluation, we computed the signal-to-noise ratio (SNR) of foveal cross-sections using the peak power of the retina and the average power of the transparent vitreous humor.

D. Safety Architecture

Experiments were performed under a Duke University Medical Center IRB-approved protocol. For optical safety under ANSI Z136.1, we limited OCT power to 1.59 mW at 1060 nm and pupil illumination power to 260 μ W at 850 nm. For robot safety, our control software generated trajectories that limited robot joint and Cartesian velocities to 45 $^{\circ}$ s $^{-1}$ and 100 mm s $^{-1}$, respectively. In addition, our control software monitored the force-torque sensor integrated into the robot's end-effector and halted motion if the non-gravitational force exceeded 20 N. We chose this force threshold due to its low likelihood of eliciting pain and very low risk of causing injury upon contact to the head, face, neck, or shoulders [31]. We further configured the robot's dedicated safety computer to enforce joint and Cartesian velocity limits of 60 $^{\circ}$ s $^{-1}$ and 250 mm s $^{-1}$, respectively, and an external force limit of 100 N.

For reclined or supine imaging, a pinning hazard existed since the subjects were less able to avoid potential robot collisions. We therefore integrated a mechanical contact sensor around the scanner's objective lens that triggered a system emergency stop via the robot's safety chain. We designed this sensor to require less than 20 N to activate. This provided an additional safeguard independent of our control software to disable the robot below the minimum force limit of 100 N supported by the safety computer. In addition, both the subject and operator had independent emergency stop stations accessible throughout the imaging sessions.

V. RESULTS & DISCUSSION

We evaluated our next-generation robotic OCT system through imaging of five human subjects with no known retinal pathology under varying head orientations and postures. The system successfully handled variation in head pose during coarse alignment and variation in eye gaze during fine alignment for all imaged subjects. Fig. 6 shows OCT summed voxel projections (SVPs), B-scans, and volumes obtained from two standing and seated subjects with varying head orientations. Fig. 7 shows OCT SVPs, B-scans, and volumes obtained from two standing, seated, reclined, and supine subjects in neutral head orientations. For each subject, we obtained a consistent view of the retina, including the fovea, retinal vasculature, and edge of the optic nerve head. The active tracking scanner obtained stabilized OCT volumes despite residual relative head and eye motion, as demonstrated by the *en face* projections. Furthermore, we obtained consistent retinal SNR across head orientations and postures except in isolated cases (Fig. 6b column 5 and Fig. 7a column 3) due to eye focus variations. The time required for alignment and imaging in each configuration was approximately one minute across all subjects. No tuning of face or pupil tracking parameters was required across these subjects.

Scanner-integrated face tracking and extended robot arm reach have enabled unconstrained and repeatable OCT imaging with our next-generation robotic OCT system. With fixed-base cameras eliminated, we imaged subjects in head orientations and postures that would have previously resulted in imaging failure due to loss of face tracking. With extended reach, we similarly imaged at large angular offsets and at more comfortable distances from the robot's base. We believe these upgrades are essential for imaging in unstructured clinical environments where the scanner is brought to the patient instead of the reverse. Based on the

results shown here, we can expect to image subjects ranging from outpatients seated in exam rooms to inpatients supine in hospital beds.

Generalization to six DoF in our control software has further unlocked new tracking capabilities. Our system can now track any face that its detection models identify and align with any subject that it can reach without self-collision. To image a subject who is seated outside the scanner's initial field of view, the operator needs only to point the scanner towards that subject using the robot's hand-guiding mode. Provided the subject is within reach, the automated alignment system can continue from there unassisted. This recapitulates the familiar and effective scheme by which portable X-ray systems are wheeled into patient rooms and positioned via a boom to face the target anatomy before subsequent careful alignment.

A limitation of our tracking approach is the handling of discordant head orientations and gaze angles, such as maximal leftward or rightward gaze. At relative gaze angles greater than about 10° , the LED reflections required for gaze estimation may overlie the iris or sclera where they exhibit little contrast, leading to loss of gaze tracking. Since the face alignment planner initially orients the scanner to match the head pose, the pupil tracker may encounter the eye in such a configuration that fine angular alignment cannot proceed. Similarly, the face tracker models from [27] are trained for frontal face views. Under certain circumstances, the robot may follow the subject's extreme gaze to such an angle that face detection is lost. Thus, for these experiments shown here, we asked subjects to use a nominally neutral gaze. In practice, such a gaze requirement is not problematic for cooperative subjects, but investment in training side-view face models may be warranted.

As noted previously, we restricted head orientations to neutral during early reclined and supine imaging experiments due to robot reach limitations. This was because our system performed best when the robot mounting matched the elevation of the head. The large variation in head elevation between standing and supine subjects made it difficult to optimize for both, and for the purposes of this work, we targeted primarily subjects who were standing or sitting on a stool. In later experiments, this challenge was readily overcome by incorporating vertical translation stage into the system that we adjusted to an appropriate height. Since all tracking is now scanner-integrated, however, the control software need not know the robot base elevation for effective imaging as it requires no external references. Even manually-operated or counter-balanced stages without position feedback would consequently suffice.

VI. CONCLUSION

Robotic OCT with scanner-integrated face tracking and an extended reach robot arm provides the necessary hardware capability for imaging subjects in arbitrary postures and head orientations. Coupled with software designed to plan and execute alignment trajectories in six DoFs, these important hardware changes support future use in unstructured clinical environments. We believe the robotic approach demonstrated here can extend OCT as an eye disease screening, diagnosis, and monitoring tool to previously unreached patient populations.

Acknowledgments

This work was supported by grants K99-EY034200, U01-EY028079, and R01-EY029302 from the National Institutes of Health.

REFERENCES

- [1]. Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, Hee MR, Flotte T, Gregory K, Puliafito CA, and Fujimoto JG, "Optical coherence tomography," *Science*, vol. 254, no. 5035, pp. 1178–1181, 1991. [Online]. Available: <https://www.science.org/doi/abs/10.1126/science.1957169> [PubMed: 1957169]
- [2]. Schuman JS, Fujimoto JG, Duker JS, Ishikawa H, and Wollstein G, Eds., *Optical Coherence Tomography of Ocular Diseases*. New Jersey, USA: Slack, Inc., 2013.
- [3]. Schmidt-Erfurth U, Klimescha S, Waldstein SM, and Bogunovi H, "A view of the current and future role of optical coherence tomography in the management of age-related macular degeneration," *Eye*, vol. 31, no. 1, pp. 26–44, Jan 2017. [Online]. Available: 10.1038/eye.2016.227 [PubMed: 27886184]
- [4]. AAO PPP Retina/Vitreous Committee, "Diabetic retinopathy," in *American Academy of Ophthalmology Preferred Practice Pattern Collection*, 2019.
- [5]. —, "Age-related macular degeneration," in *American Academy of Ophthalmology Preferred Practice Pattern Collection*, 2019.
- [6]. AAO PPP Glaucoma Committee, "Primary open-angle glaucoma," in *American Academy of Ophthalmology Preferred Practice Pattern Collection*, 2020.
- [7]. Swanson EA and Fujimoto JG, "The ecosystem that powered the translation of OCT from fundamental research to clinical and commercial impact [Invited]," *Biomed. Opt. Express*, vol. 8, no. 3, pp. 1638–1664, Mar 2017. [Online]. Available: <https://opg.optica.org/boe/abstract.cfm?URI=boe-8-3-1638> [PubMed: 28663854]
- [8]. Spaide RF, Klancnik J, James M, and Cooney MJ, "Retinal vascular layers imaged by fluorescein angiography and optical coherence tomography angiography," *JAMA Ophthalmology*, vol. 133, no. 1, pp. 45–50, 01 2015. [Online]. Available: 10.1001/jamaophthalmol.2014.3616 [PubMed: 25317632]
- [9]. Rein DB, Zhang P, Wirth KE, Lee PP, Hoerger TJ, McCall N, Klein R, Tielsch JM, Vijan S, and Saaddine J, "The Economic Burden of Major Adult Visual Disorders in the United States," *Archives of Ophthalmology*, vol. 124, no. 12, pp. 1754–1760, 12 2006. [Online]. Available: 10.1001/archophth.124.12.1754 [PubMed: 17159036]
- [10]. Brown MM, Brown GC, Lieske HB, Tran I, Turpcu A, and Colman S, "Societal costs associated with neovascular age-related macular degeneration in the united states," *Retina*, vol. 36, no. 2, pp. 285–298, 2016. [PubMed: 26428606]
- [11]. Draelos M, Ortiz P, Qian R, Keller B, Hauser K, Kuo A, and Izatt J, "Automatic optical coherence tomography imaging of stationary and moving eyes with a robotically-aligned scanner," in *IEEE International Conference on Robotics and Automation (ICRA)*, May 2019, pp. 8897–8903.
- [12]. Draelos M, Ortiz P, Qian R, Viehland C, McNabb R, Hauser K, Kuo AN, and Izatt JA, "Contactless optical coherence tomography of the eyes of freestanding individuals with a robotic scanner," *Nature Biomedical Engineering*, vol. 5, no. 7, pp. 726–736, July 2021.
- [13]. Ortiz P, Draelos M, Viehland C, Qian R, McNabb RP, Kuo AN, and Izatt JA, "Robotically aligned optical coherence tomography with 5 degree of freedom eye tracking for subject motion and gaze compensation," *Biomedical Optics Express*, vol. 12, no. 12, pp. 7361–7376, Dec 2021. [PubMed: 35003839]
- [14]. Ortiz P, Draelos M, Narawane A, McNabb RP, Kuo AN, and Izatt JA, "Robotically-aligned optical coherence tomography with gaze tracking for image montaging of the retina," in *IEEE International Conference on Robotics and Automation (ICRA)*, May 2022, pp. 3783–3789.

- [15]. Jung W, Kim J, Jeon M, Chaney EJ, Stewart CN, and Boppart SA, “Handheld optical coherence tomography scanner for primary care diagnostics,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 741–744, 2011. [PubMed: 21134801]
- [16]. LaRocca F, Nankivil D, Farsiu S, and Izatt JA, “Handheld simultaneous scanning laser ophthalmoscopy and optical coherence tomography system,” *Biomed. Opt. Express*, vol. 4, no. 11, pp. 2307–2321, Nov 2013. [Online]. Available: <https://opg.optica.org/boe/abstract.cfm?URI=boe-4-11-2307> [PubMed: 24298396]
- [17]. Nankivil D, Waterman G, LaRocca F, Keller B, Kuo AN, and Izatt JA, “Handheld, rapidly switchable, anterior/posterior segment swept source optical coherence tomography probe,” *Biomed. Opt. Express*, vol. 6, no. 11, pp. 4516–4528, Nov 2015. [Online]. Available: <https://opg.optica.org/boe/abstract.cfm?URI=boe-6-11-4516> [PubMed: 26601014]
- [18]. Yang J, Liu L, Campbell JP, Huang D, and Liu G, “Handheld optical coherence tomography angiography,” *Biomed. Opt. Express*, vol. 8, no. 4, pp. 2287–2300, Apr 2017. [Online]. Available: <https://opg.optica.org/boe/abstract.cfm?URI=boe-8-4-2287> [PubMed: 28736672]
- [19]. Scott AW, Farsiu S, Enyedi LB, Wallace DK, and Toth CA, “Imaging the infant retina with a hand-held spectral-domain optical coherence tomography device,” *American Journal of Ophthalmology*, vol. 147, no. 2, pp. 364–373.e2, 2009. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0002939408006533> [PubMed: 18848317]
- [20]. Imperio R, Tran-viet D, Mangalesh S, Viehland C, Izatt JA, Chen X, and Toth CA, “Enabling acquisition of optical coherence tomography angiography in cardiac surgery and intensive care,” *Investigative Ophthalmology & Visual Science*, vol. 63, no. 7, pp. 2962–F0116, 2022.
- [21]. Carrasco-Zevallos OM, Keller B, Viehland C, Shen L, Waterman G, Todorich B, Shieh C, Hahn P, Farsiu S, Kuo AN, Toth CA, and Izatt JA, “Live volumetric (4D) visualization and guidance of in vivo human ophthalmic surgery with intraoperative optical coherence tomography,” *Scientific Reports*, vol. 6, no. 1, p. 31689, Aug 2016. [Online]. Available: 10.1038/srep31689 [PubMed: 27538478]
- [22]. Duan Z, Huang K, Luo Z, Ma K, Wang G, Hu X, Zhang J, Luo X, Huang Y, Liu G, Ding X, Xiao P, and Yuan J, “Portable boom-type ultrahigh-resolution oct with an integrated imaging probe for supine position retinal imaging,” *Biomed. Opt. Express*, vol. 13, no. 6, pp. 3295–3310, Jun 2022. [Online]. Available: <https://opg.optica.org/boe/abstract.cfm?URI=boe-13-6-3295> [PubMed: 35781965]
- [23]. Sprenger J, Saathoff T, and Schlaefler A, “Automated robotic surface scanning with optical coherence tomography,” in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 2021, pp. 1137–1140.
- [24]. Marques MJ, Hughes MR, Vyas K, Thrapp A, Zhang H, Bradu A, Gelikonov G, Giataganas P, Payne CJ, Yang G-Z, and Podoleanu A, “En-face optical coherence tomography/fluorescence endomicroscopy for minimally invasive imaging using a robotic scanner,” *Journal of Biomedical Optics*, vol. 24, no. 6, p. 066006, 2019. [Online]. Available: 10.1117/1.JBO.24.6.066006 [PubMed: 31222989]
- [25]. Huang Y, Li X, Liu J, Qiao Z, Chen J, and Hao Q, “Robotic-arm-assisted flexible large field-of-view optical coherence tomography,” *Biomed. Opt. Express*, vol. 12, no. 7, pp. 4596–4609, Jul 2021. [Online]. Available: <https://opg.optica.org/boe/abstract.cfm?URI=boe-12-7-4596> [PubMed: 34457434]
- [26]. Jivraj J, Chen C, Barrows D, and Yang VXD, “Optical coherence tomography for dynamic axial correction of an optical end-effector for robot-guided surgical laser ablation,” *Optical Engineering*, vol. 58, no. 5, p. 054106, 2019. [Online]. Available: 10.1117/1.OE.58.5.054106
- [27]. King DE, “Dlib-ml: A machine learning toolkit,” *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.
- [28]. Pham H and Pham Q-C, “A new approach to time-optimal path parameterization based on reachability analysis,” *IEEE Transactions on Robotics*, vol. 34, no. 3, pp. 645–659, 2018.
- [29]. Berscheid L and Kröger T, “Jerk-limited real-time trajectory generation with arbitrary target states,” *Robotics: Science and Systems XVII*, 2021.
- [30]. Draelos M, “Vortex – An open-source library for building real-time OCT engines in C++ or Python,” 2022. [Online]. Available: <https://www.vortex-oct.dev/>

- [31]. Han D, Park M, Choi J, Shin H, Kim D, and Rhim S, "Assessment of pain onset and maximum bearable pain thresholds in physical contact situations," *Sensors*, vol. 22, no. 8, 2022. [Online]. Available: <https://www.mdpi.com/1424-8220/22/8/2996>

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

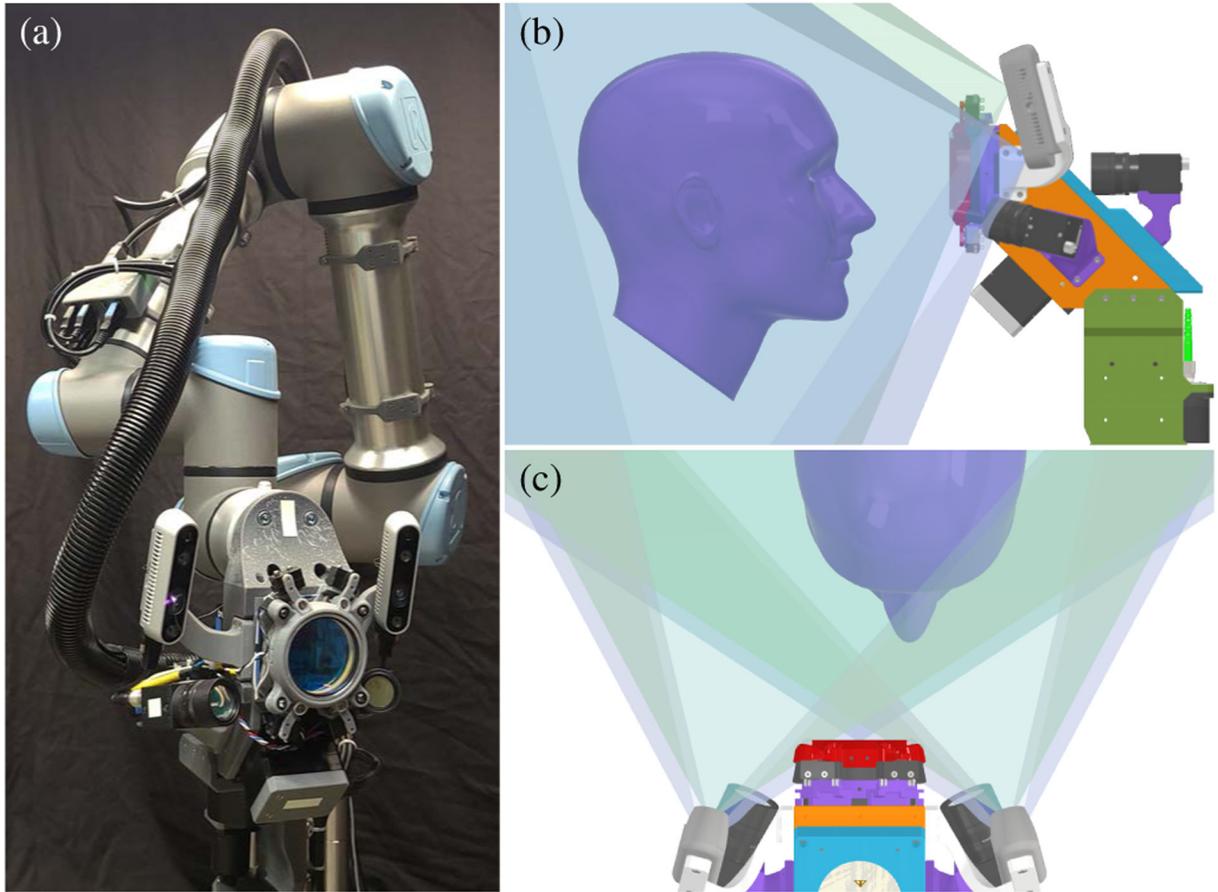


Fig. 1. Second-generation robotically-aligned OCT system (a) using a UR5e robot arm for 70% longer reach. Scanner-integrated face tracking cameras (b, c) and the tracking workspace (green/blue shaded) move with the robot to eliminate head orientation requirements.

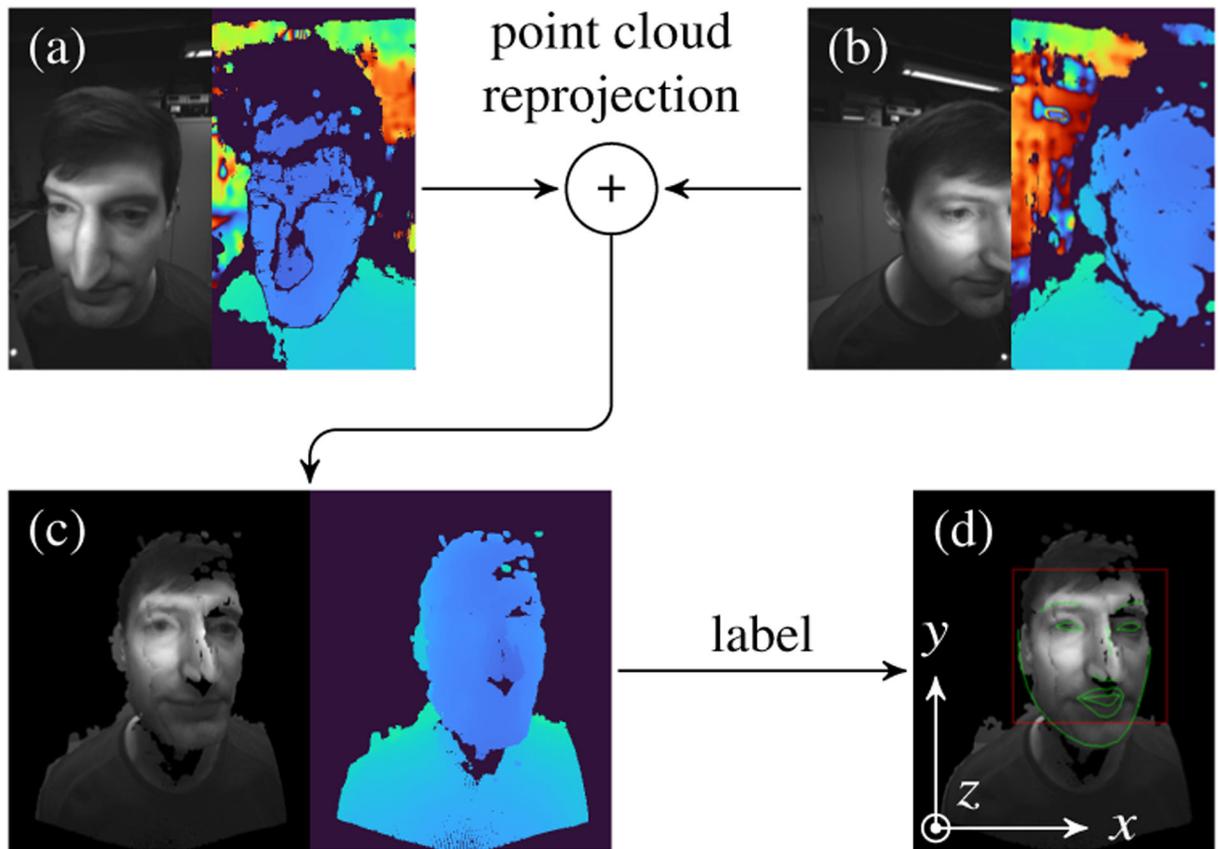


Fig. 2. Face tracking using synthetic frontal face view. Point clouds derived from left (a) and right (b) grayscale and depth images are projected onto the synthetic camera image plane to form composite grayscale and depth images (c). Face detection and landmarking (d) is then performed on the composite grayscale image.

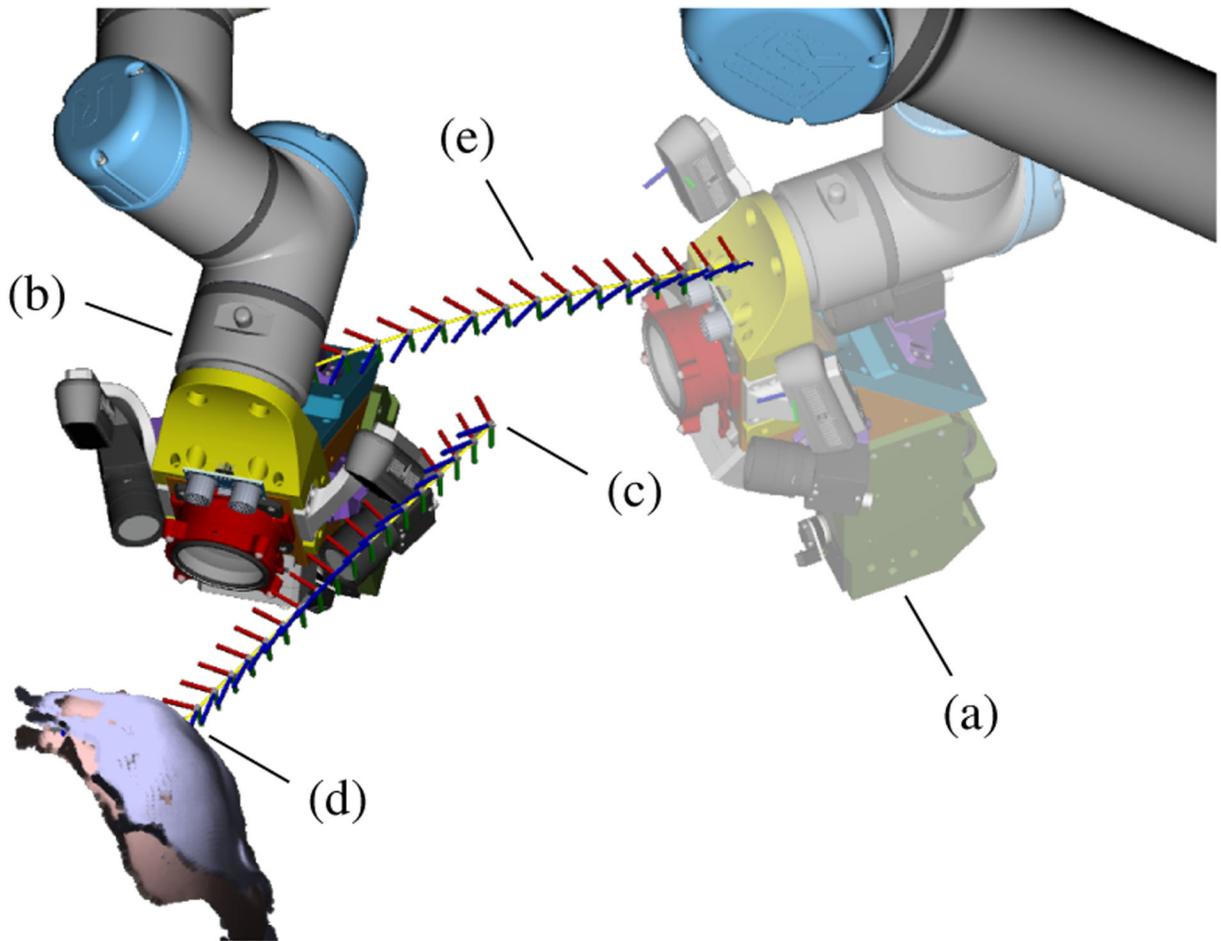


Fig. 3. Illustration of face alignment planning from the current (a) to the target (b) scanner position. The path moves the retinal scanner's pupil pivot in a straight line (c) towards the estimated pupil position (d) while the robot end-effector smoothly re-orientates (e) towards the front of the face.



Fig. 4. Robotic alignment of OCT scanner with diverse postures (standing, seated, supine) and head orientations (roll, pitch, yaw). OCT imaging is possible for any reachable and collision-free alignment pose.

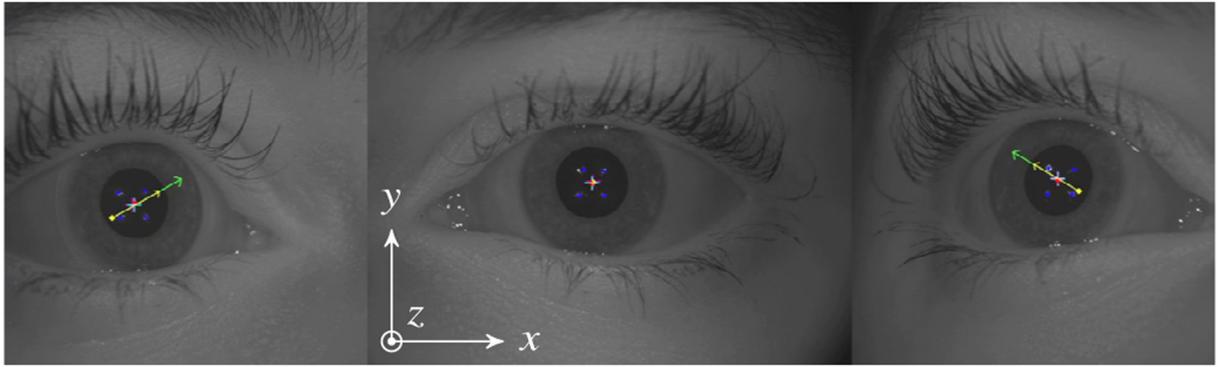


Fig. 5.

Illustration of left, include, and right pupil camera view with tracking results overlaid. The pupil position is triangulated whereas eye gaze is derived from Purkinje reflections of the eye illumination LEDs.

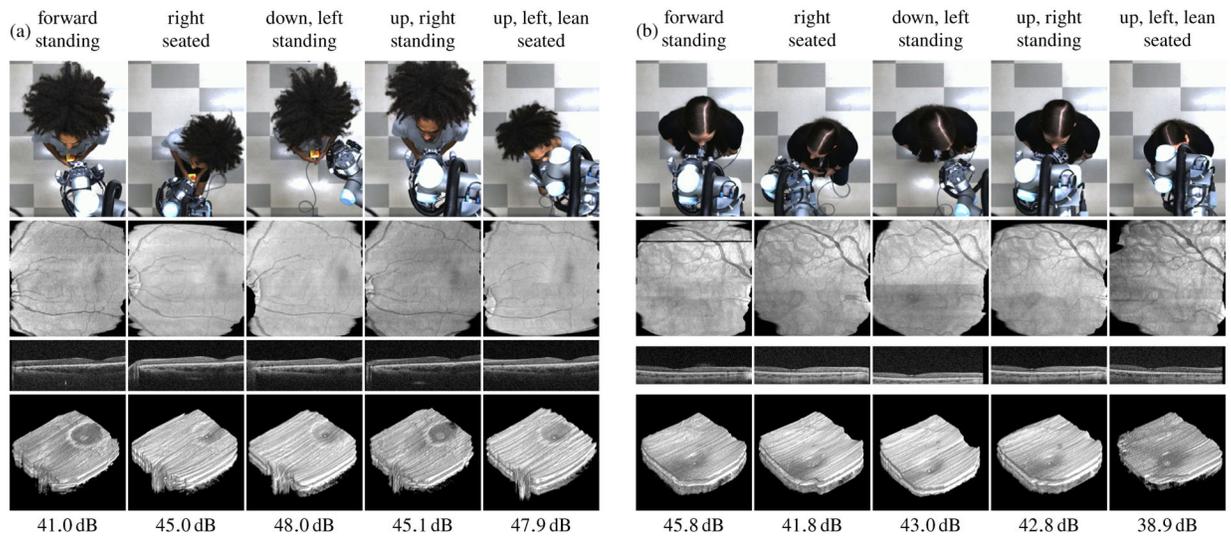


Fig. 6. Registered summed voxel projections, non-averaged B-scans through the fovea, registered volumes, and retinal SNR of the eyes of two subjects (a and b) with five varied postures and head poses. From left to right, imaging configurations progressively add head yaw, pitch, and roll. Retinal anatomy is imaged consistently despite large variations in head pose and the presence of significant subject motion.

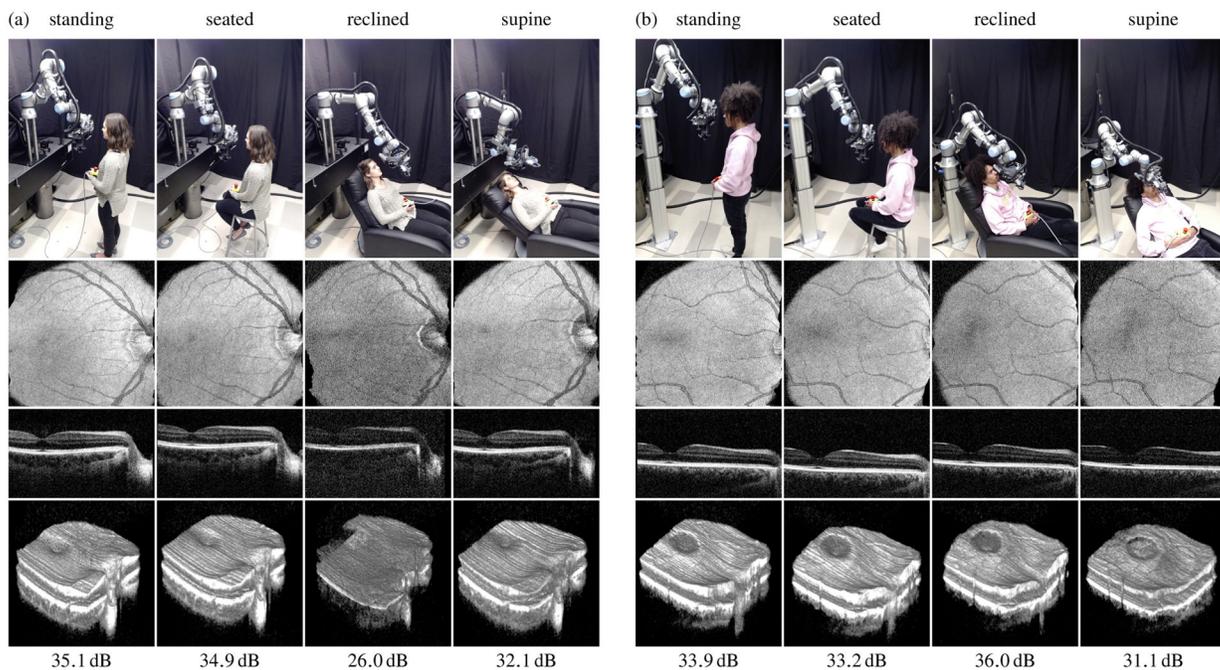


Fig. 7. Registered summed voxel projections, averaged B-scans through the fovea, registered volumes, and retinal SNR of the eyes of two subjects (a and b) with four varied postures. Retinal anatomy is imaged consistently despite large variations in posture and, for the standing case, the presence of significant subject motion. Note that these imaging sessions occur before (a) and after (b) installation of the vertical translation stage for the robot.