# Contactless optical coherence tomography of the eyes of freestanding individuals with a robotic scanner

**Mark Draelos**[1,✉], **Pablo Ortiz**[1], **Ruobing Qian**[1], **Christian Viehland**[1], **Ryan McNabb**[2], **Kris Hauser**[3], **Anthony N. Kuo**[1,2], **Joseph A. Izatt**[1,2]

[1]Department of Biomedical Engineering, Duke University, Durham, NC, USA.

[2]Department of Ophthalmology, Duke University Medical Center, Durham, NC, USA.

[3]Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA.

## Abstract

Clinical systems for optical coherence tomography (OCT) are used routinely to diagnose and monitor patients with a range of ocular diseases. They are large tabletop instruments operated by trained staff, and require mechanical stabilization of the head of the patient for positioning and motion reduction. Here we report the development and performance of a robot-mounted OCT scanner for the autonomous contactless imaging, at safe distances, of the eyes of freestanding individuals without the need for operator intervention or head stabilization. The scanner uses robotic positioning to align itself with the eye to be imaged, as well as optical active scanning to locate the pupil and to attenuate physiological eye motion. We show that the scanner enables the acquisition of OCT volumetric datasets, comparable in quality to those of clinical tabletop systems, that resolve key anatomic structures relevant for the management of common eye conditions. Robotic OCT scanners may enable the diagnosis and monitoring of patients with eye conditions in non-specialist clinics.

The use of optical coherence tomography (OCT)[1] has revolutionized diagnostic imaging[2–4] for medical[5,6] and surgical[7,8] care of the eye. Eye care providers now routinely use OCT to manage a variety of common ocular diseases, including age-related macular degeneration[9,10], diabetic retinopathy[11], glaucoma[12] and corneal dysfunction[13]. Indeed,

since its inception, OCT has had a prominent role in defining diagnostic criteria for and driving treatment decisions in many of these diseases. Unfortunately, clinical OCT systems designed for such purposes are commonly large tabletop instruments sequestered in dedicated imaging rooms of ophthalmology offices or large eye centres. Moreover, they require mechanical head stabilization (for example, a chinrest or forehead brace) for head positioning, eye alignment and motion suppression, as well as trained ophthalmic photographers for operation. OCT imaging is consequently restricted to non-urgent evaluations of cooperative patients in ophthalmology care settings due to the poor portability, stabilization needs and operator-skill requirements of modern systems. Thus, instead of being a ubiquitous tool available in both routine and emergent care environments, OCT is an exclusive imaging modality of the ocular specialist, restricted by imaging-workspace and operator barriers.

Efforts are underway to lower these barriers, albeit individually. From an operator-skill standpoint, manufacturers of many commercial OCT systems have incorporated limited self-alignment over a small working range into their tabletop scanners. While easier to operate, such systems still depend on chinrests for approximate alignment and motion stabilization. Furthermore, the optical and mechanical components necessary to achieve self-alignment add substantial bulk and weight to such a table-dependent scanner. From an imaging-workspace standpoint, handheld OCT[14–21] shrinks and lightens the scan head so that the operator can articulate it with ease. The OCT scan head is brought to the patient, as opposed to tabletop systems which do the reverse. Scans obtained in this manner become limited by manual alignment and motion artefacts[14,20–22], so that only highly skilled operators with steady hands reliably succeed. Image-registration techniques[23,24] effectively mitigate small motion artefacts in post-processing; however, these algorithms are fundamentally limited by the raw image quality. In those cases where the raw image is lost due to misalignment, no correction is possible. Neither self-aligning tabletop scanners nor handheld scanners overcome the workspace and operator-skill barriers that restrict routine OCT imaging to cooperative, ambulatory individuals. Self-aligning tabletop scanners still require mechanical head stabilization and handheld scanners still require trained operators. Moreover, these two approaches are incompatible: the additional bulk of automated alignment components renders handheld scanners even more unwieldy.

We foresee a greatly expanded role for OCT once imaging-workspace and operator barriers are eliminated. Rather than bringing the patient to the OCT scanner and bracing their head against it, we propose completely automated alignment that brings the scanner to the patient and actively tracks their movements for stabilization from a distance without any operator involvement. In primary care settings, we envision OCT-based eye exams during annual office visits that provide immediate results using machine learning algorithms for interpretation[25–27]. This is especially promising for chronic, progressive eye conditions such as age-related macular degeneration[28] and glaucoma[29], where the standard of care includes regular monitoring. Optometrists, who do not routinely use ophthalmic photographers, could even offer OCT eye exams for disease surveillance in individuals who they would otherwise see only for refractive error. In emergency departments, we envision turnkey OCT-based eye diagnostics to replace direct ophthalmoscopy, which has a low completion rate (14 %) in this setting even when medically indicated[30,31]. This would extend OCT imaging to

acutely ill patients who cannot physically sit upright due to injury, strict bedrest or loss of consciousness. Furthermore, we envision unattended OCT imaging stations in pharmacies or malls, which can complete an eye exam fully autonomously.

Transforming a tabletop OCT scanner to realize our vision eliminates its primary technique for motion suppression: mechanical head stabilization with chinrests and forehead braces. A scanner that is not designed to actively track the patient's eye and aim the scan in real time to compensate for movement is unlikely to yield acceptable images. This is especially the case when performing high-density volumetric acquisitions with point-scan swept-source OCT (SS-OCT), a next-generation OCT modality for both angiographic and densely sampled three-dimensional (3D) structural imaging[32]. For example, a 15 mm × 15 mm volume to scan the entire cornea at $20 \times 20$ μm lateral resolution (750 A-scans per B-scan, 750 B-scans per volume) requires more than 5 s to complete, with a 100 kHz swept-source laser, which affords ample time for patient movement, involuntary or otherwise. Indeed, our earlier work[33] demonstrated that automated alignment without scanner-based active tracking is insufficient to remove artefacts from physiologic motion. The new imaging scenarios we propose require scanner-integrated tracking that works in concert with the alignment mechanism. Although more technologically complex than a simple chinrest, this contactless approach offers immediate benefits, especially from the patient perspective. Avoiding chinrests and forehead braces not only improves patient ergonomics but also reduces the complexity of undergoing an OCT scan to that of posing for a photograph. Moreover, the absence of physical contact all but eliminates hygiene and infectious disease concerns associated with OCT systems as shared equipment. This is particularly relevant in recent months, where the proliferation of fomites has become a daily threat to patients and healthcare workers alike.

Here, we overcome these barriers by positioning a specially designed active-tracking scan head using a robot arm. With 3D cameras to find the patient in space and the robot arm to articulate the scan head, we grossly align the scan head to the patient's eye with a workspace comparable to the human arm, achieving the flexibility of handheld OCT without operator tremor. By using cameras integrated in the scan head to locate the patient's pupil and active-tracking components within the OCT system itself, we augment the OCT acquisition in real time to optically correct motion artefacts, obtaining tabletop-quality images without mechanical head stabilization. To illustrate this approach and extend our previous work[33], we demonstrate both operator-activated automatic alignment and operator-free autonomous imaging in freestanding individuals, a setting chosen to elicit worst-case performance due to its greater physiologic motion compared with seated or supine conditions. We show OCT volumes obtained in this manner that rival images from clinical OCT systems and then validate these volumes through expert grader analysis and extraction of anatomic measurements. This robot-articulated, active-tracking paradigm may advance OCT imaging towards primary care, emergency department and completely unattended settings, and may enable imaging of position-restricted individuals.

## Results

### Active-tracking scan head characterization in simulation and phantoms.

We designed separate anterior and retinal scan heads with active-tracking capabilities and a suitable working distance ( 8 cm) for robotic imaging of patients (Fig. 1a and Supplementary Video 1). In optical simulation, our anterior segment scan head (Fig. 1c) achieved a 28 mm × 28 mm maximum lateral field of view and theoretical lateral resolution of 43 μm (Fig. 1e). The objective lens supported a ±7.5 mm lateral scan, which was sufficient to cover the limbus[34] while affording 6.5 mm of lateral aiming range. In optical simulation, our retinal scan head (Fig. 1d) achieved a $16 \times 16°$ maximum field of view at the cornea and a theoretical lateral resolution of 8.5 μm (Fig. 1f). The fast steering mirror (FSM) for pupil pivot aiming provided up to ±5 mm lateral aiming range before vignetting at the objective lens. By filling the entire objective with the OCT scan, we obtained a retinal field of view sufficient to encompass the parafoveal region. The minimum pupil diameter of 3.2 mm (Fig. 1g) was appropriate for both scotopic and mesopic illumination conditions[35,36]. The anterior segment and retinal scan heads had design masses of 1.2 kg and 2.3 kg, respectively, within the 3 kg payload limit of our robot arm.

Next, we characterized the active-tracking capabilities of our anterior and retinal scan heads in phantoms to determine their performance limits (Table 1). Our phantom tracking tests (Fig. 2a) demonstrated lateral-tracking accuracy better than the anterior scan head's optical resolution (Fig. 2c) and much smaller than the retinal scan head's pupil beam diameter (Fig. 2f). The reduced lateral accuracy compared to our previous work[33] resulted from lens weight reduction and field of view enlargement that degraded camera resolution while still satisfying our aiming needs. Axial-tracking accuracy was within 6 pixels and 12 pixels along an A-scan for the anterior (Fig. 2i) and retinal (Fig. 2k) scan heads, respectively, given ideal resolution of the OCT engine. On the basis of latency and bandwidth estimates (Fig. 2e,h,m) from phantom step responses (Fig. 2b), we expected to attenuate eye motion at frequencies seen during robot pursuit (Supplementary Fig. 3), especially during anterior imaging. For our OCT imaging configuration, the maximum number of B-scans affected before the tracking correction initiated varied from 0.3 to 3.3 B-scans, based on the latency of the active-tracking elements.

### Automatic alignment for OCT imaging in freestanding individuals.

We then performed initial human testing of our robotic OCT system under Institutional Review Board approval by using automatic alignment mode, in which an operator activated the system with a foot pedal (Fig. 3). In all five participants, we obtained high- and intermediate-density OCT volumes over 8 s and 4 s, respectively, which revealed clinically relevant anatomy of the anterior segment (Fig. 3c,d). Residual motion artefacts were readily corrected in post-processing (Fig. 3a,b) despite the long acquisition intervals. To further explore the effect of active tracking during such acquisitions, we transiently suspended axial and lateral tracking during anterior imaging of one freestanding participant (Fig. 4a). Immediately once axial tracking was suspended ($t_1$ in volume 1), we observed large axial motion artefacts. The effect was especially pronounced in volume 2, where the axial motion arose from an interaction between the participant's sway towards and away from the

scanner while the robot alone attempted to compensate. When axial tracking was resumed ($t_2$ in volume 3), the motion artefact was immediately suppressed and a largely motion-free volume 4 was obtained. The same sequence of effects was observed when lateral tracking was paused ($t_3$ in volume 5). Volume 6, where no lateral tracking was present, exhibited artefactual distortion of the pupil's circular shape due to lateral motion, compared with fully tracked volume 4. Again, lateral motion was attenuated when tracking was resumed ($t_4$ in volume 7).

We investigated these qualitative observations in active-tracking telemetry using residual motion estimates obtained by registering the corresponding OCT data. The axial and lateral registration shifts (Fig. 4b,c) exhibited large signals primarily when active tracking was suspended (grey regions, approximately). Furthermore, when we aligned them with the pupil-tracking telemetry, we observed that the registration shifts recapitulated pupil motion waveforms when active tracking was suspended (Fig. 4d,e). Although non-identical due to blinks and lost A-scans during galvanometer flyback between volumes, the similarity of the registration shifts and tracking telemetry indicated that active tracking was effective in eliminating large motion artefacts.

### Fully autonomous OCT imaging in freestanding individuals.

To eliminate the operator entirely, we implemented a fully autonomous controller for our robotic OCT system that required no operator intervention during imaging. As with automatic imaging, the autonomous controller obtained anterior segment volumes in all five participants (Fig. 5a,c) with attenuated motion artefacts that revealed the same anatomic structures of interest as in Fig. 3d. The total interaction time (from participant entrance to exit) for all imaging sessions was 50 s or less, and the controller maintained imaging for 8–10 s per eye, recovering from initial tracking difficulties with participant 2 (Fig. 5b). Similarly, autonomous retinal imaging yielded motion-stabilized volumes of all five participants' foveae (Fig. 6a,c) with small residual motion. Although some vignetting and tilt was present, especially in participants 1 and 2, these volumes revealed the foveal pit, the surrounding parafoveal region and large retinal vessels. The striping artefact seen in most volumes was due to slight tilting of adjacent B-scans that translation-only registration did not correct. In B-scans through the macula (Fig. 6d), we identified major retinal tissue layers and the underlying choroid. As with anterior imaging, total interaction time was again under 50 s for all participants (Fig. 6b).

Finally, we conducted an analysis to assess the quality of these OCT volumes obtained autonomously by extracting anatomic measurements and using an expert reviewer for retinal B-scan evaluation. Both anterior segment and retinal measurements yielded results (Table 2) that were consistent with previous studies reporting central corneal thickness (CCT)[37–40], anterior chamber depth (ACD)[40–42], and parafoveal retinal thickness (PRT)[43–45] in healthy normal eyes, like those of our participants and the general population. Moreover, the expert reviewer rated all selected retinal B-scans as appropriate for formal grading, indicating that they met the quality standards necessary for disease evaluation. These results suggested that our autonomously acquired OCT data were suitable for clinical diagnostics.

## Discussion

As a structural eye imaging modality, OCT has greatly expanded diagnostic capabilities in ophthalmology. The focus on deploying clinical OCT systems as complex tabletop instruments, however, has hampered the use of OCT in other eye-examination settings, particularly in primary care clinics and emergency departments. The data presented here demonstrate that robot-mounted scan heads with active tracking can overcome the requirements of tabletop systems for trained operators and mechanical head stabilization. Thus, it is now conceivable to implement OCT in routine screening roles where neither a complete exam (for example, a dilated eye exam) nor an ophthalmic photographer's skill is needed. Moreover, the capability exists to bring the OCT scanner to the patient, rather than the patient to the scanner, without sacrificing motion stabilization as with handheld scanners. As long as the patient can be positioned within the robot's workspace, OCT imaging is possible, whether they are standing, seated or supine. Although we demonstrated motion-stabilized imaging of freestanding individuals to exaggerate motion artefacts, we anticipate applying robotic OCT to seated individuals when standing confers no mobility advantage, such as during outpatient office visits. Reduced eye movement under these circumstances will result in even less residual motion artefacts than seen in Figs. 5 and 6. Nevertheless, our results indicate suitable performance during freestanding imaging, as would be desirable for high-throughput screening and unattended imaging in public settings. In all cases, contactless operation is still preferable, even when patients are seated, due to improved imaging ergonomics and minimization of the decontamination requirements normally needed to prevent contact transfer of infectious agents.

A complete OCT-based eye screening system must both acquire images and interpret them. In this Article, we present a robotic solution that addresses the collection of OCT screening images. Recently developed machine learning techniques[25–29] for ophthalmic OCT analysis address the interpretation aspect. Pairing the two forms a fully autonomous point-of-care OCT screening tool capable of evaluating individuals for eye disease without an ophthalmic photographer or an ophthalmologist. We believe that such point-of-care systems will revolutionize eye care both inside and outside the traditional ophthalmology setting. In the traditional ophthalmology setting, point-of-care OCT could reduce staffing requirements and increase the number of patients seen by reducing visit times. Outside the traditional ophthalmology setting, primary care practitioners could incorporate a structural eye exam into annual physical examinations, especially with the developing role of OCT in systemic screening for diseases such as dementia[46,47], as well as monitor chronic ocular conditions. Patients with new or worsening disease would receive a referral with low false-positive rates and patients with stable disease could forward their latest images to their ophthalmologist for review. With such easy access, we expect point-of-care OCT diagnostics to improve follow-up and reduce unnecessary ophthalmologist visits for otherwise well-managed patients. Before automated analysis becomes clinically accepted, remote OCT reading centres can provide interpretation services for otherwise unattended OCT screening and acquisition. We find this arrangement beneficial for both patients and ophthalmologists, who receive more detailed care and expanded patient monitoring, respectively.

The biggest limitations of our system are its resolution and, for retinal imaging, its field of view. For point-scan OCT scanners such as these, the diffraction-limited resolution is determined by the diameter of the laser beam and numerical aperture of the final lens. Consequently, most OCT systems operate with large beams, close to the eye, and in dim conditions to promote pupil dilation. For our anterior scan head, we opted to increase the working distance at the expense of lateral resolution so that we could maintain a large separation for increased safety in this initial demonstration system. Similarly, for our retinal scan head, the large working distance limits the field of view and operation in lit environments limits the maximum beam diameter. Except for the retinal beam limit, these issues are addressable in redesign by reducing the working distance and/or increasing the scan head's objective diameter[48], at a financial and weight cost. Substantially reducing the working distance for both scan heads is readily achievable, although integration of safety controls (for example, scan head-integrated proximity or contact sensors) beyond what we describe here may become necessary. In addition, montaging multiple OCT scans from different angles can circumvent the limited retinal of view if longer imaging sessions are acceptable. We expect automation of the manual eye length and defocus adjustment required here for retinal imaging to be readily achieved with commercially available electrically tunable lenses and OCT image brightness metrics. Notably, autonomous imaging of paediatric, visually impaired or elderly patients or patients with movement disorders may prove more difficult than that of the healthy volunteers shown here.

Raw volumes (that is, unregistered) taken with active tracking enabled still exhibited small motion artefacts. Examples include volumes 4 and 8 in Fig. 3m and Supplementary Figs. 1 and 2. The two primary sources of this artefact are active-tracking limitations and rotational eye motion. The active-tracking system operates with a delay that can permit noticeable warping artefacts during rapid eye motions due to uncompensated relative motion between the eye and scanner. Furthermore, the active-tracking elements are in motion while the scan is ongoing, meaning that B-scans acquired during particularly large corrections may exhibit stretching that is not corrected with the rigid registration performed here. The tracking cameras also introduce noise into the tracking system (Fig. 2d,g,j,l), which manifests in the resultant volume (Supplementary Fig. 2a,c), although this effect has lower amplitude and does not exhibit any prominent frequency components (Supplementary Fig. 3). Finally, the eye exhibits rotational motions, such as micro-saccades, that our active-tracking system is not designed to correct. For both scan heads, correcting for a change in eye gaze direction would require the robot to reposition the scan head rather than aim the scan. Instead, the small gaze offsets of our participants manifested as tilt in our volumes, both for cornea and retina. Although active-tracking delay is potentially reducible with hardware upgrades, rotational correction will probably be difficult to eliminate with optical elements alone.

Point-of-care OCT is poised to radically change eye-examination practices in non-specialist settings, to the benefit of both patients and eye care providers. The system demonstrated here eliminates the imaging-workspace and operator-skill barriers that presently confine OCT to ophthalmology offices. Once the aforementioned refinements are addressed, spread to primary care offices and emergency departments is the logical next step. We envision a future where regular OCT-based eye exams improve care for all patients, for early detection in healthy eyes and for monitoring in diseased ones.

## Methods

### OCT data collection and processing.

We drove both tracking scan heads with a swept-source OCT engine (Fig. 1b) with ideal 5.4 μm axial resolution in air, 7.4 mm imaging depth, and a VCM (VCS05–060-LB-01-MCH, H2W Technologies) in the reference arm[49,50]. The VCM provided reference arm length adjustment of 6 mm with 1 μm accuracy. The engine used a 100 kHz swept-source laser centred at 1,050 nm with 100 nm spectral bandwidth (Axsun Technologies). OCT data was recorded from the balanced receiver using a 1.8 gigasamples s$^{-1}$ analogue–digital converter card (ATS9360, AlazarTech) and processed using custom C++ software with graphics processing unit acceleration. The raw data were processed according to standard frequency domain OCT techniques[51] (for example, DC subtraction, dispersion compensation and inverse discrete Fourier transform), and the resulting volumes were saved as log-scale measurements. We generated B-scans from the log-scale volumes by rescaling based on chosen black and white thresholds. We used identical thresholds for all anterior segment B-scans and for all retinal B-scans. In addition, retinal B-scans were cropped laterally to remove galvanometer flyback artefacts and axially to remove regions that did not contain the retina; the same cropping was used for all retinal B-scans.

We generated volumetric images using direct volume rendering[52] implemented in custom Python and C++ software with graphics processing unit acceleration. Volumes were thresholded, rescaled, $3 \times 3 \times 3$-median filtered, and Gaussian smoothed along the axial and fast scan dimensions before raycasting. All volumetric renders used the same pipeline and parameters, except for variation in the threshold, render angle and axial scale between anterior segment and retinal volumes. As with B-scans, an identical threshold was applied to all anterior segment volumes and applied to all retinal volumes. We performed translation-only registration of adjacent B-scans in post-processing according to their peak cross-correlation in the lateral and axial directions. We filtered the cross-correlation signal to reject large shifts (for example, blinks) and applied Gaussian smoothing to eliminate low-amplitude noise. Volumes were also cropped to remove galvanometer flyback and lens reflection artefacts, with the same copping applied to each eye segment.

### OCT scan head design and fabrication.

For the anterior segment, we designed a telecentric scanner (Fig. 1c) that performed lateral tracking by electrically offsetting the galvanometer angular scan waveform to yield a linear shift. For the retina, we designed a conventional 4*f* retinal telescope (Fig. 1d) with a FSM in the retinal conjugate plane[53]. Switching between anterior segment and retinal imaging consequently required changing out scan heads. We performed retinal lateral tracking by introducing scan tilt at the FSM in the telescope's Fourier plane, which produced a lateral shift of the pupil pivot. Both designs used 2 inch B-coated achromatic doublets (Thorlabs) with 200 mm focal lengths for the objective and relay lens pairs. Galvanometers (Saturn 1B, ScannerMax) were used to laterally scan the OCT A-scan position (that is, optical beam) across the sample (that is, the anterior eye or retina) to create B-scans and had small-angle step-response time of 90 μs. Short-pass dichroics inserted the OCT beam path over the inline pupil camera's field of view. The retinal scan head's $2 \times 3$ inch FSM (OIM202.3, Optics in

Motion) had ±3° angular range and had small-angle step-response time of under 7 ms. Scan head optical design was performed in OpticStudio (Zemax) using 1,000 nm, 1,050 nm and 1,100 nm wavelengths over the 3 × 3 matrix of configurations shown in Fig. 1e,f. We used the Navarro eye model[54] for retinal scan head optical optimization. Scan head mechanical design was performed in Inventor (Autodesk) according to the optical optimization results. We assembled the scan heads from 3D printed plastic parts derived from the mechanical models.

### Active tracking by linear triangulation.

Active tracking relied on three Blackfly S monochrome cameras (BFS-U3–042M, FLIR Systems) in each scan head's horizontal plane (Fig. 1c,d), one of which shared the OCT objective to obtain a lateral view coincident with the OCT scan. The left and right camera poses were calibrated in the inline camera's coordinate frame using chessboard calibration targets. Each camera was initialized using OpenCV stereo calibration[55], and then all camera parameters were refined with bundle adjustment. We detected the pupil in each camera view in parallel using custom C++ software that identified dark circles[33,53]. The pupil position in 3D space was estimated through linear triangulation at 350 Hz when at least two cameras reported pupil presence. To facilitate pupil detection, each scan head included an infrared illumination ring, using 850 nm light for the anterior scan head and 720 nm light for the retinal scan head. The cameras were fitted with filters to reject the OCT light. For lateral scan aiming, the galvanometer and FSM were calibrated to the inline pupil camera by sweeping the OCT A-scan across the sample plane at the working distance while the camera filters were temporarily removed. We detected the scan spot in custom Python software and performed linear regression to map galvanometer and FSM drive commands to physical position. For lateral scan aiming during anterior imaging, we adjusted the OCT scan waveform using a summing amplifier inserted before the input of each galvanometer axis to maintain lateral eye centration. For pupil pivot aiming during retinal imaging, we interfaced with the FSM directly to position the pupil pivot within the pupil laterally. For axial scan aiming (that is, reference arm adjustment) during both anterior segment and retinal imaging, we drove the VCM according to the pupil distance as determined from triangulation to dynamically compensate for the eye's axial displacement from the ideal imaging point. We applied a two-sample moving-average filter to reduce VCM jitter.

### Accuracy and precision characterization.

Quantitative tracking characterization in humans is difficult because another tracking modality is needed to serve as a gold standard. Instead, we opted for testing in phantoms to allow careful control of experimental conditions. To evaluate accuracy and precision, we thus recorded pupil-tracking telemetry at 350 Hz while we moved a paper eye phantom on a motorized stage (A-LSQ150D-E01, Zaber) with 4 μm repeatability and 100 μm accuracy (Fig. 2a). The paper eye phantom consisted of a black circle on a white background to satisfy the pupil-tracking algorithm. We did not record phantom displacements with OCT because the retinal scan head's converging, unfocused beams were not well-suited to measuring lateral shifts. Starting with the phantom grossly aligned with the scan head's optical axis at the working distance, we advanced the stage in 1 mm steps 5 times forward, 10 times in reverse, and 5 times forward. We therefore examined worst-case performance

for eye excursions of ±5 mm from the ideal imaging point at which the tracking camera calibration had been optimized. The experiment was performed axially and laterally, and yielded 21 total position measurements for each dimension. We considered the mean tracked position before the first step to be the zero position and then collected approximately 2 s of telemetry recordings after the stage had settled for approximately 1 s. To estimate accuracy (Fig. 2c,f,i,k), we subtracted the commanded stage position from the mean of each of these recordings and computed the root mean square value. To estimate precision (Fig. 2d,g,j,l), we subtracted the mean of each recording from itself and computed the root mean square value across all recordings together. Separate testing was done for the anterior and retinal scan heads because they required individual calibrations and therefore could exhibit different behaviour.

**Latency and bandwidth characterization.**

We used the OCT system to measure the active-tracking response for latency and bandwidth analysis because the goal of active tracking is to stabilize the OCT A-scan against motion. Moreover, this approach exploited the 10 μm temporal and 5.4 μm spatial resolution of our swept-source laser for measuring the response. We configured the OCT system to acquire successive A-scans without scanning the beam so that only active tracking affected the scan location and again used phantoms for careful control of experimental conditions (Fig. 2b). For axial (VCM) characterization, we mounted a retinal eye phantom (Rowe Eye, Rowe Technical Design) on a motorized stage (A-LSQ150D-E01, Zaber) and rapidly stepped the stage 1 mm axially. We read out axial-tracking error from the A-scan data by registering adjacent A-scans using cross-correlation (Fig. 2m). We applied the same procedure to determine anterior lateral (galvanometer) latency but tilted the paper eye phantom (from the accuracy and precision experiments) before rapidly stepping it 2 mm laterally instead. The paper eye phantom's tilt was such that lateral-tracking error manifested as axial displacement on OCT (Fig. 2e), allowing us to read out the tracking response as previously done. Extraction of axial-tracking error from OCT A-scans was performed in post-processing and was therefore not used for real-time tracking during imaging experiments. For retinal (FSM) characterization, we placed a spot occlusion on the retinal eye phantom, adjusted active tracking to position the scan on the occlusion, and rapidly stepped the phantom laterally 2 mm. When the OCT beam was blocked by the occlusion, the resulting A-scan had lower intensity and higher intensity otherwise. We were thus able to characterize active tracking by its ability to hold the beam on the occlusion, producing a lower intensity A-scan. We read out brightness using the maximum intensity from the A-scan data of an OCT acquisition during the lateral step (Fig. 2h). Use of the Rowe eye model was limited to these characterization experiments.

To estimate bandwidth of each active-tracking component (galvanometers, FSM and VCM), we measured their open-loop step responses $h(t)$ (Supplementary Fig. 4), their closed-loop transient responses $e(t)$ (Fig. 2e,h,m yellow), and the input step response of the motorized stage $x(t)$ (Fig. 2e,h,m blue). The closed-loop transient response was therefore written as

$$e(t) = x(t) - h(t) \times x(t) = [1 - h(t)] \times x(t) \tag{1}$$

because the active-tracking response $h(t) \times x(t)$ attempts to eliminate the step disturbance $x(t)$. We used each component's open-loop step response to fit transfer function models $\hat{h}(t)$ with MATLAB's System Identification Toolbox (MathWorks) using the lowest order models that provided a high-quality fit. For the galvanometers and FSM, first-order models provided an appropriate fit; for the VCM, a fourth-order model was needed. This provided a zero-latency model of each active-tracking component. We then determined the latency $\tau$ which best recapitulated the observed transient response $e(t)$ by minimizing the objective function

$$
\begin{aligned}
\tau &= \arg\min_{\hat{\tau}} \int_t \left| e(t) - \hat{e}(t) \right|^2 \, \mathrm{d}t \\
&= \arg\min_{\hat{\tau}} \int_t \left| e(t) - [1 - \hat{h}(t - \hat{\tau})] \times x(t) \right|^2 \, \mathrm{d}t,
\end{aligned}
\tag{2}
$$

which represents the sum of squared error between the observed $e(t)$ and predicted $\hat{e}(t)$ transient responses for a given time delay $\tau$. We determined the maximum motion attenuation frequency by analytically finding the lowest frequency at which the gain of the fit closed-loop transfer function exceeded unity (Supplementary Fig. 4). This was done by solving the closed-loop transfer function for the minimum $\omega$ such that

$$
\left| \frac{E(j\omega)}{X(j\omega)} \right| = \left| 1 - e^{-j\omega\tau} \widehat{H}(j\omega) \right| = 1,
\tag{3}
$$

where $E(j\omega)$, $X(j\omega)$ and $\hat{H}(j\omega)$ represent the continuous-time Fourier transforms of $e(t)$, $x(t)$ and $\hat{h}(t)$, respectively. The effect of control lag $\tau$ was determined by comparison to the solution for $\tau = 0$.

### Robotic scan head positioning.

Robotic scan head positioning for automatic and autonomous imaging used a stepwise approach for obtaining eye images[33]. The controller used two 3D cameras (RealSense D415, Intel) positioned to view the imaging workspace from the left and from the right. We detected facial landmarks in the cameras' infrared images using OpenFace 2.0[56] and extracted the left and right eye positions in 3D space from the corresponding point clouds at approximately 30 Hz, subject to image processing times. During the initial 'position' stage, the controller held the scan head in a recovery position that prevented occlusion of the imaging workspace (Fig. 1a, frames 1 and 6) and waited to identify a face. During the 'align' stage, the controller grossly aligned the scan head with the target eye based on face tracking results (Fig. 1a, frames 2–5). This motion was performed open-loop without any scan head-based feedback. Eye positions were consequently smoothed with a 6-sample moving-average filter (approximately 200 ms) to reduce noise in the target position. In automatic imaging, the controller only entered the 'align' stage when the operator depressed the enabling foot pedal. In autonomous imaging, however, the controller entered the align stage once a face had been continuously detected within the imaging workspace for approximately 2 s.

Once the scan head detected the pupil, the controller entered the 'image' stage in which it finely aligned the scan head's optical axis with the eye using pupil-tracking results. We considered OCT images as valid during this stage. These fine alignment motions were performed closed-loop with a proportional control that eliminated the pupil-alignment error. Face tracking results continued to be used as a consistency check but not for servoing. In automatic imaging, the controller remained in the image stage until the operator released the foot pedal or the scan head reported a pupil-tracking loss. In autonomous imaging, the controller returned to the align stage to image the next eye or entered the 'done' stage after the preset imaging time had elapsed. The controller issued voice prompts during each stage to facilitate alignment and inform participants of its actions. We used the commercially available UR3 collaborative robot arm (Universal Robots) for manipulating the scan head with 100 μm repeatability. We chose the UR3 because it featured dedicated and redundant safety supervision computers, which we needed to protect our participants as described below. The controller downsampled the pupil-tracking results from 350 Hz to the UR3's command rate of 125 Hz by dropping outdated samples.

### Automatic and autonomous imaging schemes.

During automatic imaging, the operator minimally controlled the system through foot pedals that activated alignment and switched the target eye. They were also able to review the live imaging data as they were collected and conclude the session once satisfied with the resulting volumes. We imaged five freestanding individuals this way using the anterior scan head with high ($800 \times 800 \times 1,376$ voxel) and intermediate ($800 \times 400 \times 1,376$ voxel) density scans (Fig. 3a–l), requiring approximately 8 s and 4 s to acquire, respectively. For active-tracking characterization (Fig. 4b,c), we estimated residual motion as the lateral and axial shift required to achieve registration in post-processing because a motion-free volume should require little registration. Registration was performed using cross-correlation of adjacent B-scans as described above. To compensate for registration shift induced by the normal corneal shape, we took the mean per-B-scan shift across all volumes in the dimensions that were tracked as a baseline shift.

During autonomous imaging, the controller required no operator intervention. It detected the participant, issued audio prompts for workspace positioning, performed imaging of both eyes for a configurable duration, and dismissed the participant when done (Supplementary Video 1). With this controller, we autonomously imaged five freestanding individuals each using the anterior and retinal scan heads at 8–10 s per eye using a low density scan ($800 \times 200 \times 1,376$ voxel) requiring approximately 2 s to acquire (that is, 4–5 volumes per eye). For the anterior scan head, no per-participant configuration was performed; we initialized the system, asked the participant to approach, and did not intervene except to close the controller once it announced completion. For the retinal scan head, we manually adjusted the scan head to account for the defocus and length of each participant's eye before commencing autonomous operation. The retinal-imaging sessions were otherwise conducted identically to the anterior ones. Notably, we performed retinal imaging without pharmacologic pupil dilation and with ambient overhead lighting.

### Quality analysis of autonomously acquired OCT data.

We selected one autonomously acquired B-scan from each eye of each participant for each imaging mode (anterior segment or retinal) for quality analysis, yielding a total of ten anterior segment and retinal B-scans. We manually chose B-scans from volumes to avoid artefacts from tilt, which our system was not designed to correct, and blinking. For the anterior segment, we compared the CCT and ACD from our data with known values. When measuring CCT and ACD, we manually segmented the contact lens (if present), corneal epithelium, corneal endothelium and lens anterior surface. Before extracting measurements, we performed 2D index correction in the B-scan plane by using Snell's law, the segmentation boundaries, and the known indices of refraction for the anterior segment[57]. When B-scans included contact lenses, we used the known index of refraction of the appropriate lens material. For the retina, an unaffiliated, senior reviewer at the Duke Reading Center (a central OCT reading centre[58]) evaluated the gradability of B-scans through the fovea, which included inspection of retinal layers. This reviewer was a fellowship-trained retinal specialist who held relevant professional certifications for retinal OCT interpretation. Gradability evaluates whether or not a B-scan meets quality standards necessary for further interpretation and is commonly assessed before subjecting a B-scan to thorough analysis for signs of retinal disease. We also compared the mean PRT from these B-scans to published values. For measuring PRT, we manually segmented the internal limiting membrane (ILM) and retinal pigment epithelium (RPE) on the B-scan reviewed by the expert grader. To correct the tilt introduced by acquisition during reference arm adjustment, we rotated the segmented layers until horizontal before further analysis. We then performed 1D index correction along each A-scan and averaged the ILM-RPE distance within 0.5 mm laterally of the fovea. We performed these steps in accordance with the accepted convention for retinal thickness analysis. The ILM and RPE were visible in each segmented B-scan despite variations in imaging angle. We approximated the retinal index of refraction using that of the vitreous humour[57].

### Human participants and safety.

All experiments with human participants were performed under Duke University Medical Center Institutional Review Board approval and with informed consent of each participant, including consent to publish identifiable images. In particular, the authors affirm that the participant who is pictured in Supplementary Video 1 provided written informed consent for publication of their images. Optical power for OCT imaging and pupil illumination was set to comply with the ANSI Z136.1 standard, and compliance was verified for both light sources before imaging each participant. We mitigated the risk of unexpected contact of the scan head and/or robot arm with the participant by restricting the scan head's position in space, restricting the scan head's velocity through space, and providing the participant an easy and instinctive way to withdraw from the scan head. These three approaches were implemented in five distinct system levels. First, our software was written to never command the robot arm to move its end-effector outside a predefined rectangular prism in space (the allowed region) or faster than 20 cm s$^{-1}$. This confined the scan head to a small and easily managed region of space. Second, we configured the robot's safety computer to perform an emergency stop should the end-effector position violate the allowed region by more than 5 cm, should the end-effector velocity exceed 25 cm s$^{-1}$, or should the rotational velocity any

robot joint exceed $60° s^{-1}$. Third, we designed the scan heads with 8 cm working distance from the eye. At this distance, the worst-case stopping distance for the robot would not bring the scan head into contact with the participant should the safety computer detect a velocity limit violation. Fourth, we arranged the imaging workspace such that the participant could take one step in any direction away from the robot arm's base and move out of range. A simple and effective withdrawal manoeuvre was therefore always available to participants. Fifth, the participant and investigators had access to independent emergency stop stations during the entirety of each experiment.

Thus, we employed a combination of automatic, manual, and designed-in safety features to prevent unexpected contact of the scan head and/or the robot arm with the participant.

### Reporting Summary.

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The main data supporting the results in this work are available within the paper and its Supplementary Information. The raw data acquired during the study are available from the corresponding author on reasonable request, subject to approval from the Duke University Medical Center Institutional Review Board.

## Code availability

The custom software developed for this research is described in refs. [33,52]. This software is available from the authors upon reasonable request.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

## References

1. Huang D et al. Optical coherence tomography. Science 254, 1178–1181 (1991). [PubMed: 1957169]

2. Izatt JA et al. Micrometer-scale resolution imaging of the anterior eye in vivo with optical coherence tomography. JAMA Ophthalmol. 112, 1584–1589 (1994).

3. Hee MR et al. Optical coherence tomography of the human retina. JAMA Ophthalmol. 113, 325–332 (1995).

4. Puliafito CA et al. Imaging of macular diseases with optical coherence tomography. Ophthalmology 102, 217–229 (1995). [PubMed: 7862410]

5. Virgili G et al. Optical coherence tomography (OCT) for detection of macular oedema in patients with diabetic retinopathy. Cochrane Database Syst. Rev 1, CD008081 (2015). [PubMed: 25564068]

6. Miguel AIM, Silva AB & Azevedo LF Diagnostic performance of optical coherence tomography angiography in glaucoma: a systematic review and meta-analysis. Br. J. Ophthalmol. 103, 1677–1684 (2019). [PubMed: 30728123]

7. Carrasco-Zevallos OM et al. Review of intraoperative optical coherence tomography: technology and applications. Biomed. Opt. Express 8, 1607–1637 (2017). [PubMed: 28663853]

8. Ehlers JP et al. The DISCOVER study 3-year results: feasibility and usefulness of microscope-integrated intraoperative OCT during ophthalmic surgery. Ophthalmology 125, 1014–1027 (2018). [PubMed: 29409662]

9. American Academy of Ophthalmology-Preferred Practice Patterns Retina/Vitreous Panel, Hoskins Center for Quality Eye Care. Age-Related Macular Degeneration PPP (American Academy of Ophthalmology, 2015).

10. Schmidt-Erfurth U, Klimscha S, Waldstein S & Bogunovi H A view of the current and future role of optical coherence tomography in the management of age-related macular degeneration. Eye 31, 26–44 (2017). [PubMed: 27886184]

11. American Academy of Ophthalmology-Preferred Practice Patterns Retina/Vitreous Panel, Hoskins Center for Quality Eye Care. Diabetic Retinopathy PPP (American Academy of Ophthalmology, 2017).

12. Schuman JS, Puliafito CA, Fujimoto JG & Duker JS Optical Coherence Tomography of Ocular Diseases 3rd edn (Slack, 2012).

13. American Academy of Ophthalmology Cornea/External Disease Preferred Practice Patterns Panel, Hoskins Center for Quality Eye Care. Corneal Edema and Opacification PPP (American Academy of Ophthalmology, 2013).

14. Radhakrishnan S et al. Real-time optical coherence tomography of the anterior segment at 1310 nm. Arch. Ophthalmol. 119, 1179–1185 (2001). [PubMed: 11483086]

15. Jung W et al. Handheld optical coherence tomography scanner for primary care diagnostics. IEEE Trans. Biomed. Eng. 58, 741–744 (2011). [PubMed: 21134801]

16. Lu CD et al. Handheld ultrahigh speed swept source optical coherence tomography instrument using a MEMS scanning mirror. Biomed. Opt. Express 5, 293–311 (2013). [PubMed: 24466495]

17. Nankivil D et al. Handheld, rapidly switchable, anterior/posterior segment swept source optical coherence tomography probe. Biomed. Opt. Express 6, 4516–4528 (2015). [PubMed: 26601014]

18. LaRocca F et al. In vivo cellular-resolution retinal imaging in infants and children using an ultracompact handheld probe. Nat. Photon. 10, 580–584 (2016).

19. Yang J, Liu L, Campbell JP, Huang D & Liu G Handheld optical coherence tomography angiography. Biomed. Opt. Express 8, 2287–2300 (2017). [PubMed: 28736672]

20. Viehland C et al. Ergonomic handheld OCT angiography probe optimized for pediatric and supine imaging. Biomed. Opt. Express 10, 2623–2638 (2019). [PubMed: 31143506]

21. Song S et al. Development of a clinical prototype of a miniature hand-held optical coherence tomography probe for prematurity and pediatric ophthalmic imaging. Biomed. Opt. Express 10, 2383–2398 (2019). [PubMed: 31143494]

22. Spaide RF, Fujimoto JG & Waheed NK Image artifacts in optical coherence angiography. Retina 35, 2163–2180 (2015). [PubMed: 26428607]

23. Ricco S, Chen M, Ishikawa H, Wollstein G & Schuman J Correcting motion artifacts in retinal spectral domain optical coherence tomography via image registration. In International Conference on Medical Image Computing and Computer-Assisted Intervention (eds Yang GZ et al.) 100–107 (Springer, 2009).

24. Kraus MF et al. Motion correction in optical coherence tomography volumes on a per A-scan basis using orthogonal scan patterns. Biomed. Opt. Express 3, 1182–1199 (2012). [PubMed: 22741067]

25. Lu W et al. Deep learning-based automated classification of multi-categorical abnormalities from optical coherence tomography images. Transl. Vis. Sci. Technol. 7, 41 (2018). [PubMed: 30619661]

26. Loo J, Fang L, Cunefare D, Jaffe GJ & Farsiu S Deep longitudinal transfer learning-based automatic segmentation of photoreceptor ellipsoid zone defects on optical coherence tomography images of macular telangiectasia type 2. Biomed. Opt. Express 9, 2681–2698 (2018). [PubMed: 30258683]

27. dos Santos VA et al. CorneaNet: Fast segmentation of cornea OCT scans of healthy and keratoconic eyes using deep learning. Biomed. Opt. Express 10, 622–641 (2019). [PubMed: 30800504]

28. De Fauw J et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat. Med. 24, 1342–1350 (2018). [PubMed: 30104768]

29. Asaoka R et al. Using deep learning and transfer learning to accurately diagnose early-onset glaucoma from macular optical coherence tomography images. Am. J. Ophthalmol. 198, 136–145 (2019). [PubMed: 30316669]

30. Bruce BB et al. Nonmydriatic ocular fundus photography in the emergency department. N. Engl. J. Med. 364, 387–389 (2011). [PubMed: 21268749]

31. Bruce BB et al. Feasibility of nonmydriatic ocular fundus photography in the emergency department: phase I of the FOTO-ED study. Acad. Emerg. Med. 18, 928–933 (2011). [PubMed: 21906202]

32. Drexler W et al. Optical coherence tomography today: speed, contrast, and multimodality. J. Biomed. Opt. 19, 071412 (2014). [PubMed: 25079820]

33. Draelos M et al. Automatic optical coherence tomography imaging of stationary and moving eyes with a robotically-aligned scanner. In 2019 International Conference on Robotics and Automation 8897–8903 (IEEE, 2019).

34. Van Buskirk EM The anatomy of the limbus. Eye 3, 101–108 (1989). [PubMed: 2695343]

35. MacLachlan C & Howland HC Normal values and standard deviations for pupil diameter and interpupillary distance in subjects aged 1 month to 19 years. Ophthalmic Physiol. Opt. 22, 175–182 (2002). [PubMed: 12090630]

36. Schröder S, Chashchina E, Janunts E, Cayless A & Langenbucher A Reproducibility and normal values of static pupil diameters. Eur. J. Ophthalmol. 28, 150–156 (2018). [PubMed: 28885673]

37. Herndon LW et al. Central corneal thickness in normal, glaucomatous, and ocular hypertensive eyes. Arch. Ophthalmol. 115, 1137–1141 (1997). [PubMed: 9298054]

38. La Rosa FA, Gross RL & Orengo-Nania S Central corneal thickness of Caucasians and African Americans in glaucomatous and nonglaucomatous populations. Arch. Ophthalmol. 119, 23–27 (2001). [PubMed: 11146722]

39. Yap TE, Archer TJ, Gobbe M & Reinstein DZ Comparison of central corneal thickness between fourier-domain OCT, very high-frequency digital ultrasound, and Scheimpflug imaging systems. J. Refract. Surg. 32, 110–116 (2016). [PubMed: 26856428]

40. Buehl W, Stojanac D, Sacu S, Drexler W & Findl O Comparison of three methods of measuring corneal thickness and anterior chamber depth. Am. J. Ophthalmol. 141, 7–12 (2006). [PubMed: 16386970]

41. Lam AKC, Chan R & Pang PCK The repeatability and accuracy of axial length and anterior chamber depth measurements from the IOLMaster. Ophthalmic Physiol. Opt. 21, 477–483 (2001). [PubMed: 11727876]

42. Fotedar R et al. Distribution of axial length and ocular biometry measured using partial coherence laser interferometry (IOL Master) in an older white population. Ophthalmology 117, 417–423 (2010). [PubMed: 20031227]

43. Sánchez-Tocino H, Alvarez-Vidal A, Maldonado MJ, Moreno-Montañés J & García-Layana A Retinal thickness study with optical coherence tomography in patients with diabetes. Invest. Ophthalmol. Vis. Sci. 43, 1588–1594 (2002). [PubMed: 11980878]

44. Goebel W & Kretzchmar-Gross T Retinal thickness in diabetic retinopathy: a study using optical coherence tomography (OCT). Retina 22, 759–767 (2002). [PubMed: 12476103]

45. Alamouti B & Funk J Retinal thickness decreases with age: an OCT study. Br. J. Ophthalmol. 87, 899–901 (2003). [PubMed: 12812895]

46. Ko F et al. Association of retinal nerve fiber layer thinning with current and future cognitive decline: a study using optical coherence tomography. JAMA Neurol. 75, 1198–1205 (2018). [PubMed: 29946685]

47. Mutlu U et al. Association of retinal neurodegeneration on optical coherence tomography with dementia: a population-based study. JAMA Neurol. 75, 1256–1263 (2018). [PubMed: 29946702]

48. Qian R et al. Characterization of long working distance optical coherence tomography for imaging of pediatric retinal pathology. Transl. Vis. Sci. Technol. 6, 12 (2017).

49. Pircher M, Baumann B, Götzinger E, Sattmann H & Hitzenberger CK Simultaneous SLO/OCT imaging of the human retina with axial eye motion correction. Opt. Express 15, 16922–16932 (2007). [PubMed: 19550983]

50. Mecê P, Scholler J, Groux K & Boccara C High-resolution in-vivo human retinal imaging using full-field OCT with optical stabilization of axial motion. Biomed. Opt. Express 11, 492–504 (2020). [PubMed: 32010530]

51. Wojtkowski M, Leitgeb R, Kowalczyk A, Bajraszewski T & Fercher AF In vivo human retinal imaging by Fourier domain optical coherence tomography. J. Biomed. Opt. 7, 457–463 (2002). [PubMed: 12175297]

52. Viehland C et al. Enhanced volumetric visualization for real time 4D intraoperative ophthalmic swept-source OCT. Biomed. Opt. Express 7, 1815–1829 (2016). [PubMed: 27231623]

53. Carrasco-Zevallos O et al. Pupil tracking optical coherence tomography for precise control of pupil entry position. Biomed. Opt. Express 6, 3405–3419 (2015). [PubMed: 26417510]

54. Escudero-Sanz I & Navarro R Off-axis aberrations of a wide-angle schematic eye model. J. Opt. Soc. Am. A 16, 1881–1891 (1999).

55. Bradski G The OpenCV Library. Dr Dobb's J. Softw. Tools 25, 120–125 (2000).

56. Zadeh A, Lim YC, Baltrušaitis T & Morency L Convolutional experts constrained local model for 3D facial landmark detection. In 2017 IEEE International Conference on Computer Vision Workshops 2519–2528 (IEEE, 2017).

57. Atchison DA Optical models for human myopic eyes. Vis. Res. 46, 2236–2250 (2006). [PubMed: 16494919]

58. Tan CS & Sadda SR The role of central reading centers–current practices and future directions. Indian J. Ophthalmol. 63, 404–405 (2015). [PubMed: 26139801]
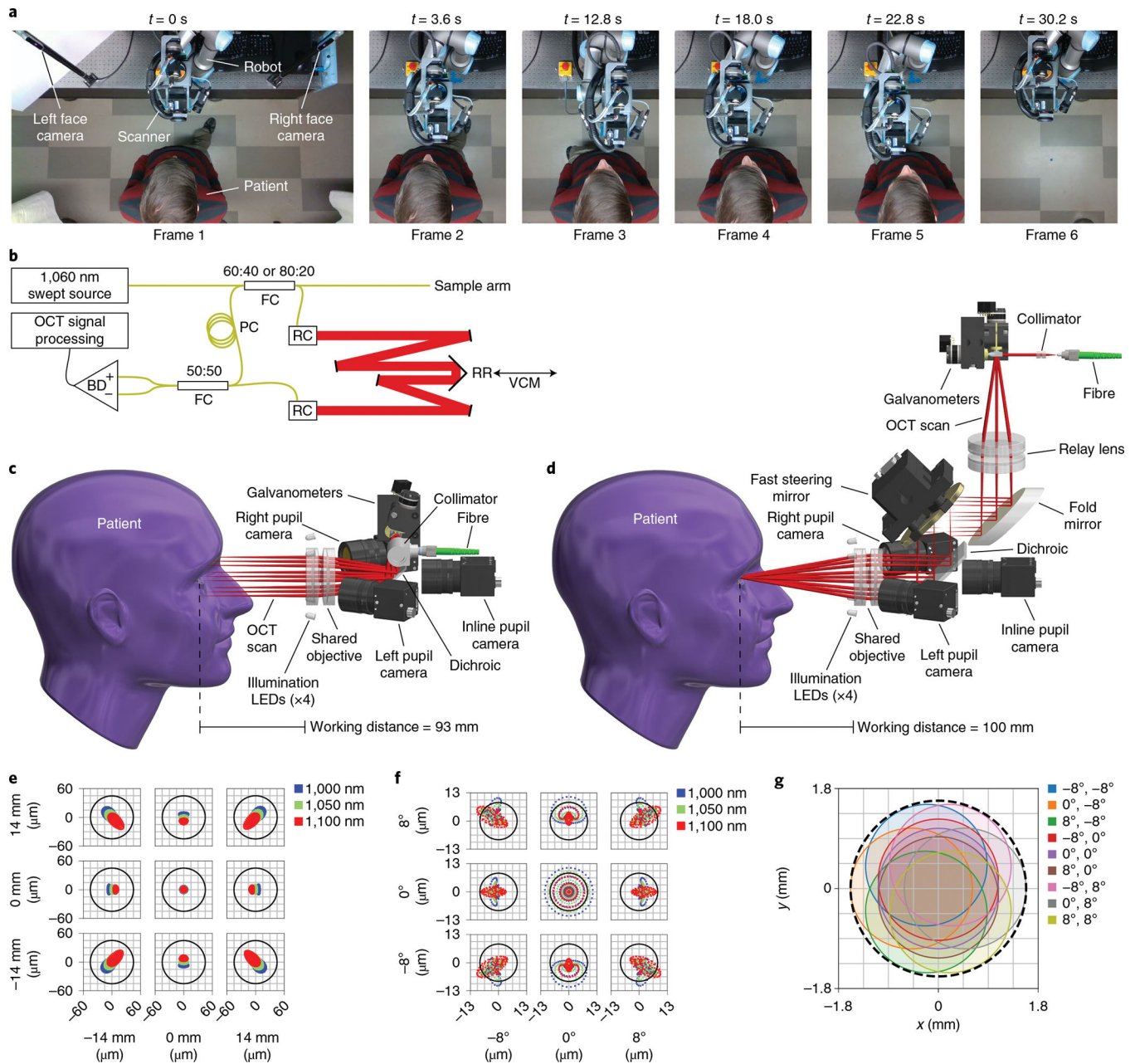
**Fig. 1 |. OCt system and scan heads for active tracking.**
**a**, Robotic scanner positioning to approach the patient, align with the right (frames 2 and 3) and left (frames 4 and 5) eyes, and retract once done (frame 6). Images show a top-down view of the robot's workspace during the imaging session in sequence from left to right. The robot arm follows the patient to keep the eye centred within the working range for optical active tracking (Supplementary Video 1). **b**, OCT engine with Mach–Zehnder topology, transmissive reference arm with adjustable length, and balanced detection. The retroreflector is mounted on the voice coil motor (VCM) to provide high-frequency compensation in path length for axial active tracking. BD, balanced detector; FC, fibre coupler; PC, polarization controller; RC, reflective collimator; RR, retroreflector; black lines, electrical wires; yellow

lines, optical fibre; red lines, free-space optical path of reference arm. **c**, Anterior scan head model with optical ray trace, showing a telecentric scan with a 93 mm working distance. Lateral active tracking is achieved by altering the galvanometer scan angles, which shifts the scan laterally. LEDs, light-emitting diodes. **d**, Retinal scan head model with optical ray trace, showing a 4*f* retinal telescope with a 100 mm working distance. Lateral active tracking is achieved by tilting the FSM located in the telescope's Fourier plane, such that tilt manifests as lateral translation at the pupil pivot. The scan pattern completely fills the objective so galvanometer scan angles remain fixed despite tracking. **e**, Anterior scan head optical spot diagram for the three design wavelengths over a ±14 mm scan, indicating diffraction-limited performance. The Airy radius is 43 μm. **f**, Retinal scan head optical spot diagram for the three design wavelengths over a ±8° scan, indicating nearly diffraction-limited performance due to residual spherical aberration. The Airy radius is 8.5 μm for the entrance beam diameter of 2.5 mm. **g**, Retinal scan head pupil wobble diagram indicating minimum acceptable diameter of 3.2 mm (dashed black line).
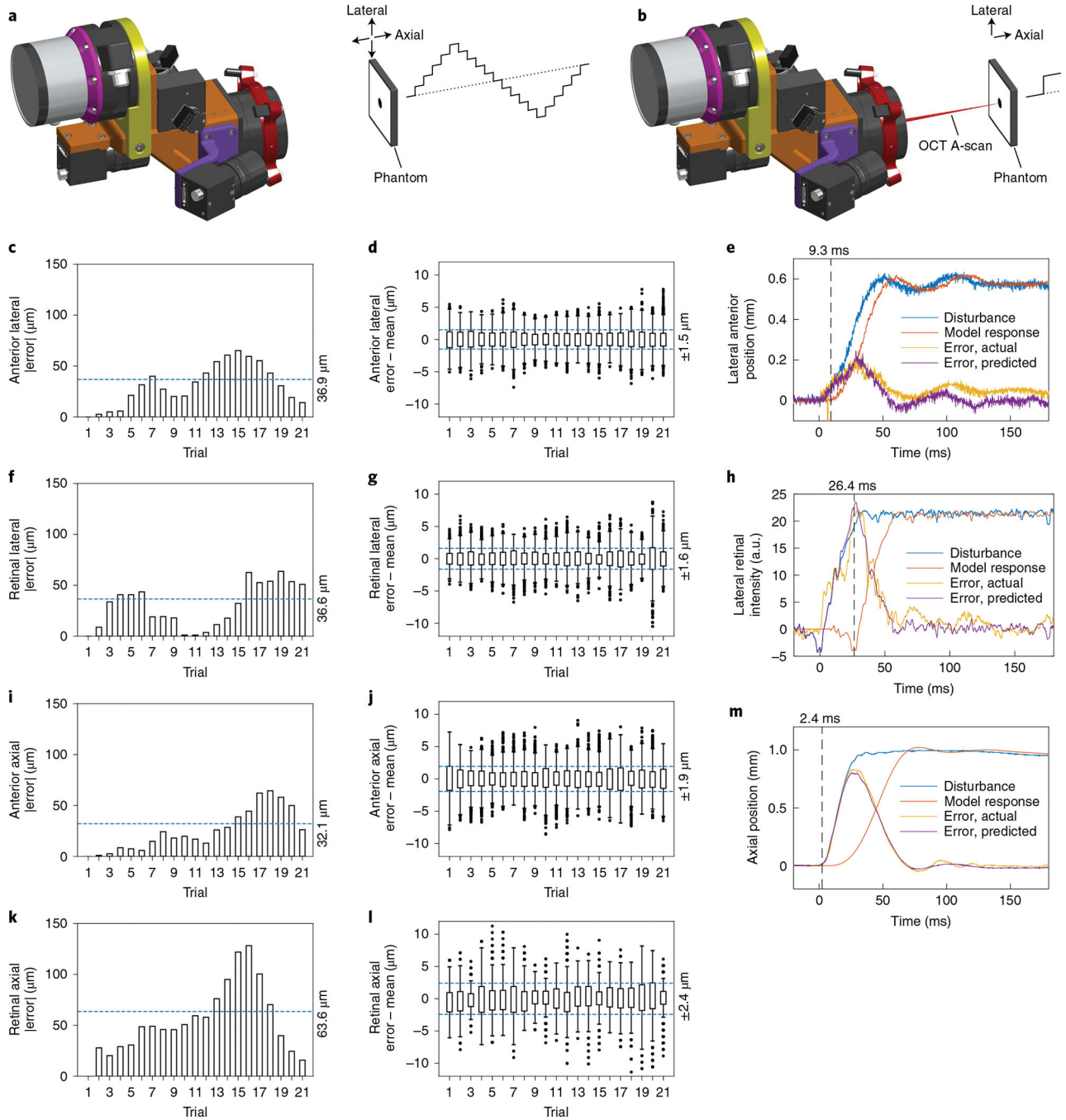
**Fig. 2 |. Accuracy, precision and transient response for anterior and retinal active-tracking scan heads.**
**a**, Setup for accuracy and precision measurements, in which the stationary tracking cameras (left) measure phantom position through a sequence of lateral and axial 1 mm steps as illustrated (right). **b**, Setup for latency measurements, in which the stationary scan head actively tracked the phantom to maintain the OCT A-scan position during single lateral 2 mm and axial 1 mm steps as shown. **c**, Absolute mean error for anterior lateral steps, indicating an accuracy of 36.9 μm (*n* 1,369 samples per trial). **d**, Position distribution

for anterior lateral steps with per-step mean removed, indicating a precision of 1.5 μm ($n$ 1,369 samples per trial). **e**, Anterior lateral step disturbance (blue) from laterally stepping a tilted phantom with resultant galvanometer tracking error (yellow). The time-delayed model response (orange) optimally recapitulates the tracking error (purple) at 9.3 ms latency. **f**, Absolute mean error for retinal lateral steps, indicating an accuracy of 36.6 μm ($n$ 637 samples per trial). **g**, Position distribution for retinal lateral steps with per-step mean removed, indicating a precision of 1.6 μm ($n$ 637 samples per trial). **h**, Retinal lateral step disturbance (blue) from laterally stepping an occluded phantom with resultant FSM tracking error (yellow). The time-delayed model response (orange) optimally recapitulates the tracking error (purple) at 26.4 ms latency. a.u., arbitrary units. **i**, Absolute mean error for anterior axial steps, indicating an accuracy of 32.1 μm ($n$ 1,295 samples per trial). **j**, Position distribution for anterior axial steps with per-step mean removed, indicating a precision of 1.9 μm ($n$ 1,295 samples per trial). **k**, Absolute mean error for retinal axial steps, indicating an accuracy of 63.6 μm ($n$ 700 samples per trial). **l**, Position distribution for retinal axial steps with per-step mean removed, indicating a precision of 2.4 μm ($n$ 700 samples per trial). **m**, Anterior and retinal axial step disturbance (blue) from axially stepping a phantom with resultant VCM tracking error (yellow). The time-delayed model response (orange) optimally recapitulates the tracking error (purple) at 2.4 ms latency. For all box plots, centre lines indicate the median, boxes extend between the lower and upper quartiles, whiskers are 1.5× of interquartile range and points indicate outliers.
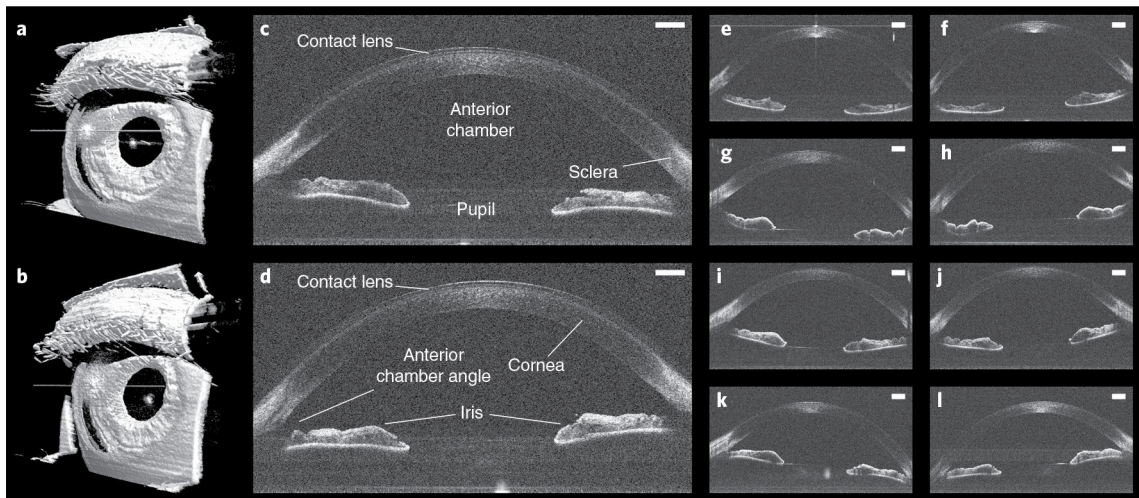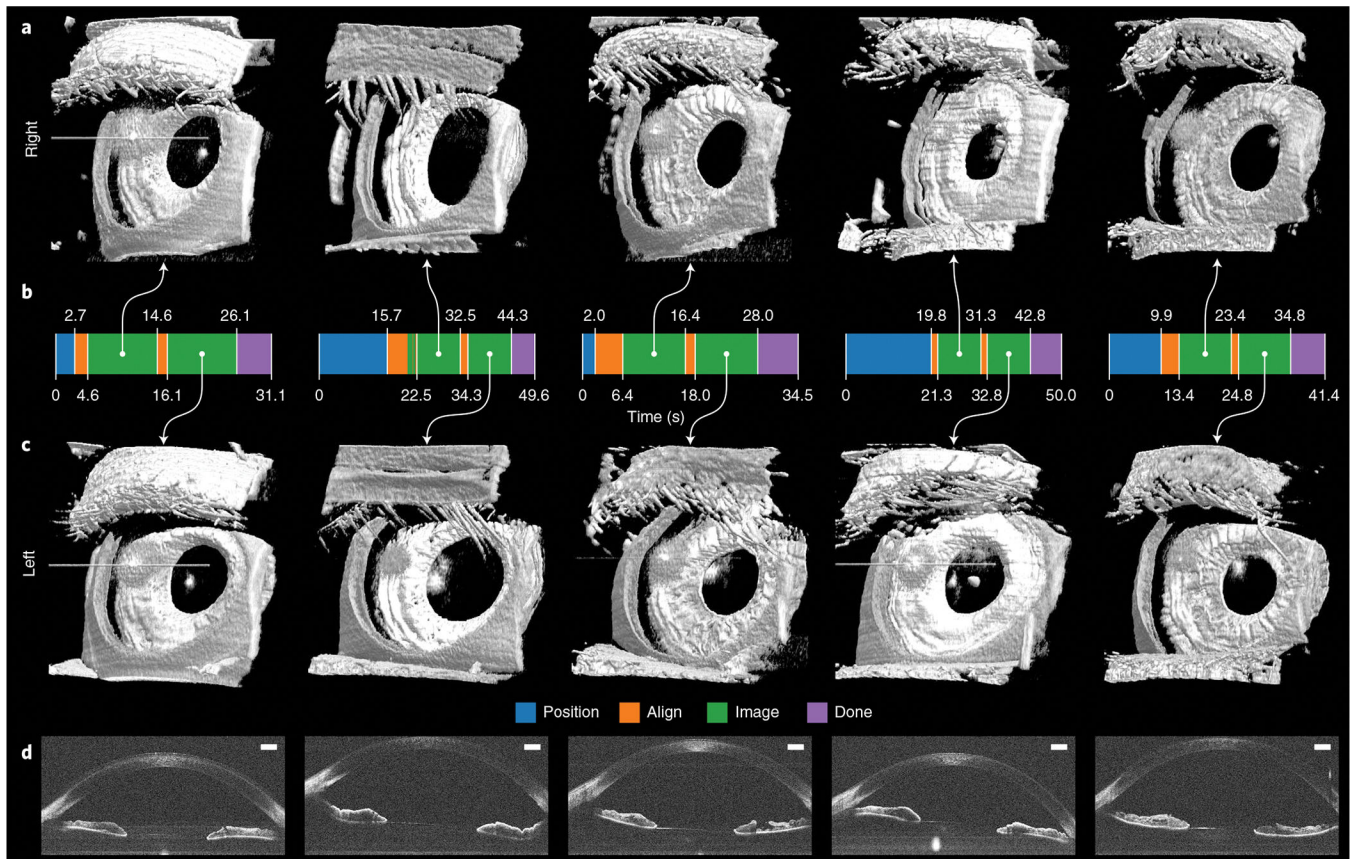
**Fig. 3 |. Anterior imaging results in freestanding individuals.**
**a,b,** Left (**a**) and right (**b**) 800 × 800 × 1,376 voxel anterior segment volumes obtained in a freestanding participant with automatic OCT imaging and registration in post-processing. **c,d,** Corresponding left (**c**) and right (**d**) anterior segment un-averaged B-scans from the same participant, revealing contact lens and relevant ocular anatomy. Scale bars, 1 mm. **e–l,** Left (**e,g,i,k**) and right (**f,h,j,l**) anterior segment un-averaged B-scans obtained with automatic OCT imaging from four additional participants, demonstrating reproducibility across participants. Scale bars, 1 mm.
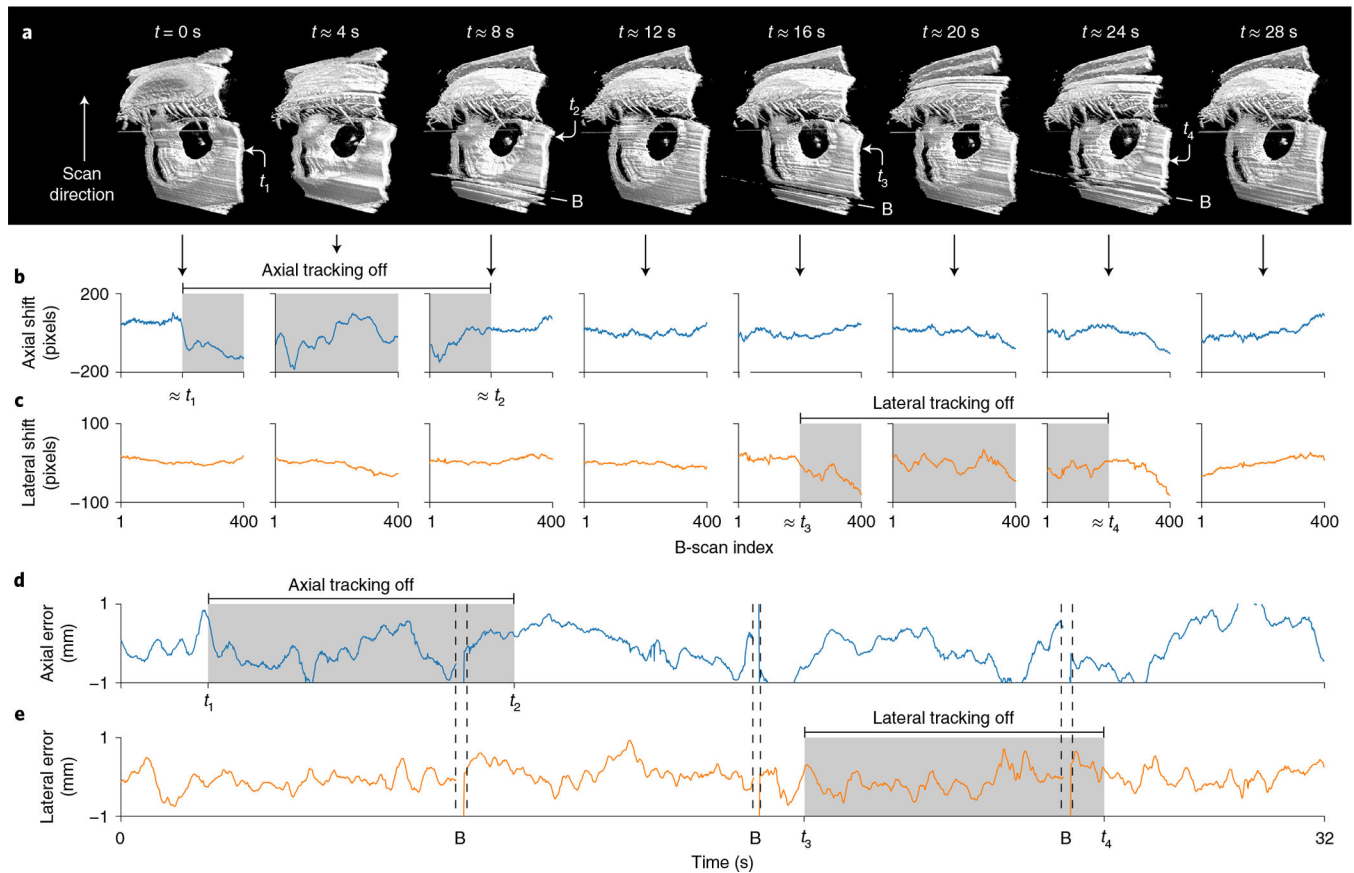
**Fig. 4 |. Volumetric imaging with and without active tracking in a freestanding individual.**
**a**, Sequential anterior segment $800 \times 400 \times 1{,}376$ voxel volumes 1–8 (from left to right)
without registration during transient suspension of axial tracking for $t_1 < t < t_2$ (grey) and
lateral tracking for $t_3 < t < t_4$ (grey). Loss of either active-tracking dimension produced
large disturbances in the raw data, compared with the fully tracked volumes 4 and 8. B,
artefacts from blink. **b**, Axial shift required to register each volume from **a**, demonstrating
increased axial motion artefact only for $t_1 < t < t_2$ (grey) when axial tracking was suspended.
Mean shift from volumes 4–8 taken as background. **c**, Lateral shift required to register
each volume from **a**, demonstrating increased lateral motion artefact only for $t_3 < t < t_4$
(grey) when lateral tracking was suspended. Mean shift from volumes 1–4 and 8 taken as
background. **d**,**e**, Axial (**d**) and lateral (**e**) alignment error from pupil tracking that aligns
approximately with **b**,**c**. The physical eye motion correlates with the registration shift during
the two tracking suspension periods, revealing the motion-stabilization effect of active
tracking. B, tracking loss during blink.

**Fig. 5 |. Autonomous anterior segment results in freestanding individuals.**

**a,c,** Paired right (**a**) and left (**c**) $800 \times 200 \times 1{,}376$ voxel anterior segment volumes obtained with fully autonomous OCT imaging and registration in post-processing. Each right–left pair is from the same individual and imaging session. These scans captured the full cornea and iris surfaces, except those portions that the superior eyelid covered. Eyelashes are seen facing towards the eye in some volumes due to their large axial extent that produced wraparound complex-conjugate OCT artefacts. Raw volumes are shown in Supplementary Fig. 1. **b,** Autonomous system mode from initial to last participant detection. For all participants, the system reliably performed 10 s of imaging per eye after aligning with the right and then the left eye for a total session time of under 50 s. Moreover, the system recovered from eye-tracking loss, as seen with the second participant, and subsequently imaged for an uninterrupted 10 s. **d,** Un-averaged B-scans from volumes in **c**, revealing the cornea, anterior chamber, iris, angle of cornea and pupil as illustrated in Fig. 3. Scale bars, 1 mm.
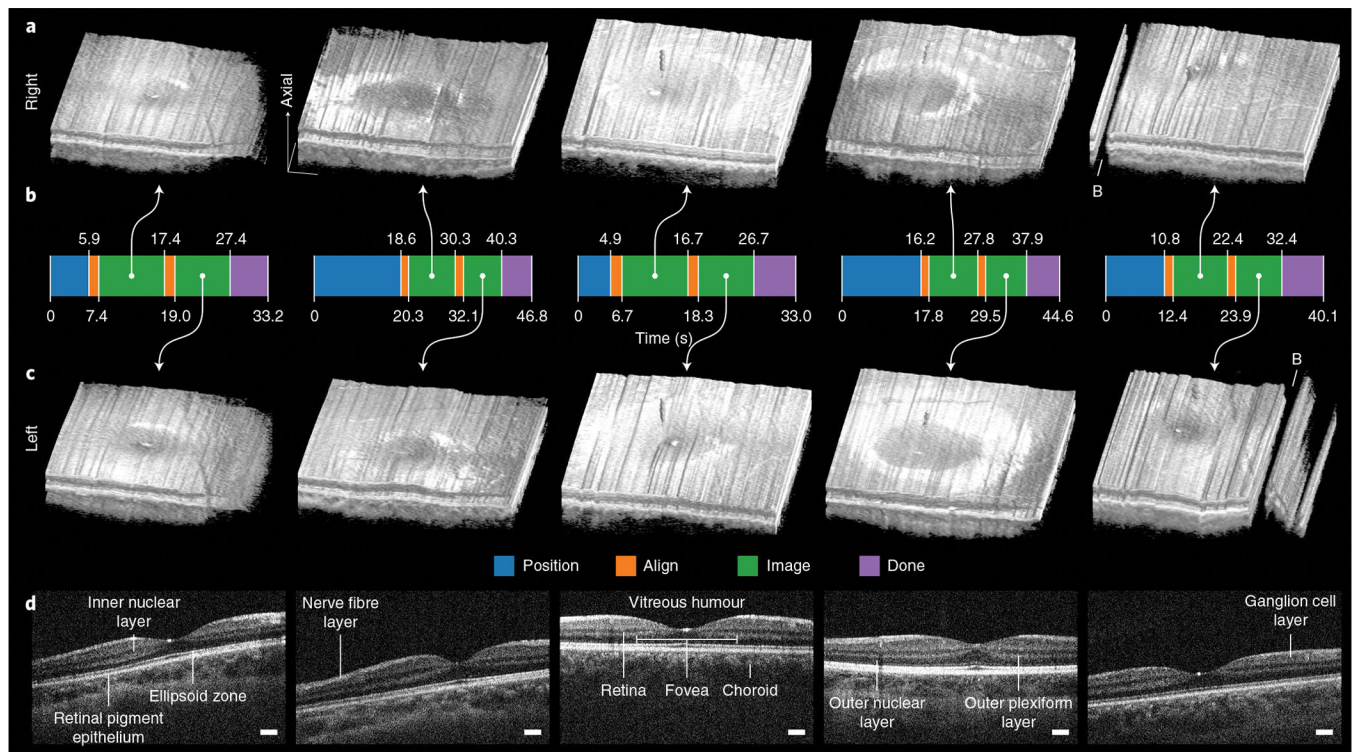
**Fig. 6 |. Autonomous retinal-imaging results in freestanding individuals with undilated eyes.**
**a,c,** Right and left $800 \times 200 \times 1{,}376$ voxel retinal volumes obtained with fully autonomous OCT imaging and registration in post-processing. These scans penetrate all retinal layers into the choroid and reveal the foveal pit within the surrounding parafoveal region. Volumes are axially stretched by a factor of two to reveal structure. Raw volumes are shown in Supplementary Fig. 2. B, artefacts from blink. **b,** Autonomous system mode from initial to last participant detection. For all participants, the system reliably performed 8–10 s of imaging per eye after aligning with the right and then the left eye with a total session time of under 50 s. The time required in advance to manually adjust the scan head for each participant's eye defocus and length is not included. **d,** Un-averaged B-scans through the fovea for each of five participants (from left to right), revealing the fovea centralis, constituent retinal layers and underlying choroid. B-scans are axially stretched by a factor of three to reveal structure. Scale bars, 250 μm (laterally).

**Table 1 |**

Active-tracking performance results

| Dimension | Scan head | Accuracy (μm) | Precision (μm) | latency (ms) | latency (B-scans)[a] | Bandwidth (Hz) |
|---|---|---|---|---|---|---|
| Lateral | Anterior | 36.9 | ±1.5 | 9.3 | 1.2 | 17.7 |
| | Retinal | 36.6 | ±1.6 | 26.4 | 3.3 | 6.2 |
| Axial | Anterior | 32.1 | ±1.9 | 2.4 | 0.3 | 6.1 |
| | Retinal | 63.6 | ±2.4 | 2.4 | 0.3 | 6.1 |

[a]For 800 A-scans per B-scan at 100 kHz A-scan rate, assuming instantaneous galvanometer flyback between B-scans.

**Table 2 |**

Overview of studies reporting CCT, ACD and PRT

| Study | N | CCt (mm) | ACD (mm) | PRt (μm) |
|---|---|---|---|---|
| Present study (mean ± s.d.) | 10 | 0.539 ± 0.026 | 3.78 ± 0.40 | 250 ± 34 |
| Ref. [37] | 59 | 0.561 | – | – |
| Ref. [38] | 77 | 0.548 | – | – |
| Ref. [39] | 81 | 0.508 | – | – |
| Ref. [41] | 26 | – | 3.60 | – |
| Ref. [40] | 88 | 0.528 | 3.32 | – |
| Ref. [42] | 1,321 | – | 3.24 | – |
| Ref. [43] | 44 | – | – | 258 |
| Ref. [44] | 30 | – | – | 259 |
| Ref. [45] | 100 | – | – | 249 |