Light-evoked deformations in rod photoreceptors, pigment epithelium and subretinal space revealed by prolonged and multilayered optoretinography

Bingyao Tan[a,b,1], **Huakun Li**[c,1], **Yueming Zhuo**[d,e], **Le Han**[a,b], **Rajeshkumar Mupparapu**[a,b], **Davide Nanni**[c], **Veluchamy Amutha Barathi**[a,f,g], **Daniel Palanker**[d,h,2], **Leopold Schmetterer**[a,b,c,f,i,j,k,l,2], **and Tong Ling**[a,b,c,m,2]

[a]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore; [b]SERI-NTU Advanced Ocular Engineering (STANCE) Program, Singapore; [c]School of Chemistry, Chemical Engineering and Biotechnology, Nanyang Technological University, Singapore; [d]Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA 94305, USA; [e]Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA; [f]Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore; [g]Academic Clinical Program in Ophthalmology, Duke-NUS Graduate Medical School, Singapore; [h]Department of Ophthalmology, Stanford University, Stanford, CA 94305, USA; [i]Department of Ophthalmology, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore; [j]Department of Clinical Pharmacology, Medical University of Vienna, Austria; [k]Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, Austria; [l]Institute of Molecular and Clinical Ophthalmology, Basel, Switzerland; [m]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

Phototransduction involves changes in concentration of ions and other solutes within photoreceptors and in subretinal space, which affect osmotic pressure and the associated water flow. Corresponding expansion and contraction of cellular layers can be imaged using optoretinography (ORG), based on phase-resolved optical coherence tomography (OCT). Until now, ORG could reliably detect only photoisomerization and phototransduction in photoreceptors, primarily in cones under very bright stimuli. By employing a novel subpixel bulk motion correction algorithm, which enabled imaging of the nanometer-scale tissue dynamics during minute-long recordings, and unsupervised learning of spatiotemporal patterns, we discovered optical signatures of the other retinal structures' response to visual stimuli. These include inner and outer segments of rod photoreceptors, retinal pigment epithelium, and subretinal space in general. High sensitivity of our technique enabled detection of the retinal responses to very dim stimuli: down to 0.01% bleach level, corresponding to natural levels of scotopic illumination. We also demonstrated that with a single flash, the optoretinogram can map retinal responses across a 12° field of view, potentially replacing multifocal electroretinography, with its long acquisition time and low spatial resolution. This new technique expands the diagnostic capabilities and practical applicability of optoretinography, providing a more complete replacement of electroretinography, while combining structural and functional retinal imaging in the same OCT machine.

Introduction

Phototransduction involves changes in concentration of ions and other solutes within photoreceptors and in subretinal space (SRS), which affect osmotic pressure and the associated water flow. Ion homeostasis of SRS is regulated by retinal pigment epithelium (RPE), a monolayer of cells between the photoreceptor outer segment and the choroid, and a variety of ocular diseases, such as age-related macular degeneration, retinitis pigmentosa and Best's vitelliform macular dystrophy, are associated with dysfunction of epithelial transport of RPE cells (1–3). *In-vivo* assessment of RPE ion transport relies mostly on c-wave in electroretinography (ERG) (4) or electrooculography (5), which require contact electrodes and provide rather low spatial resolution.

Optoretinography (ORG) is an emerging imaging technology for non-invasive optical probing of retinal physiology *in vivo*. It usually utilizes phase-resolved optical coherence tomography (OCT) to detect mechanical deformations of retinal cells in response to visual stimuli, for example, the deformation of photoreceptors' outer segments (OS) (6–11), which is associated with photoisomerization and phototransduction. Unlike conventional OCT, where its axial motion sensitivity is constrained by the bandwidth-limited axial resolution (typically $> 1\,\mu$m), axial sensitivity of the phase-resolved OCT is defined by the signal-to-noise ratio, potentially enabling the detection of nanometer-scale tissue dynamics (12). ORG was recently employed to classify chromatic cone subtypes in human subjects based on their responses to flashes of various colors (13), and its clinical potential was demonstrated in assessing the progression of retinitis pigmentosa (14). Furthermore, considering the similarity in photoreceptors' structure and function across mammals, small animals such as rodents, serve as convenient and cost-effective models for studying the mechanisms of photoreceptor degeneration using ORG (6, 15). In genetically modified animals, for example *Gnat1* and *Gnat2* mutants and knockouts, ORG can provide new insights into the various stages of the phototransduction cascade (16, 17). In addition, unlike human patients, animals are usually free of other ocular pathologies, and their imaging under anesthesia experiences fewer motion artifacts.

Despite these potential benefits of conducting ORG studies in rodents, the axial resolution of near-infrared OCT (NIR-OCT) systems, typically limited to several micrometers in tissue, hinders the clear differentiation of rodent photoreceptor OS terminals and RPE, particularly given the interdigitation of microvilli with the OS. In a typical structural image captured by NIR-OCT, the OS and RPE are compounded in a speckled layer (Fig. 1A). Moreover, since the backscattered light is coherently convolved with the point spread function of the optical system, a single pixel in the OCT image can contain a mixture of phase signals from multiple cells. As a result, separating the ORG signals from the OS and RPE presents a significant challenge.

B.T., H.L., D.P., L.S. and T.L. designed the study. B.T. and L.S. built the NIR-OCT setup. R.M. and B.T. built the VIS-OCT setup. B.T., H.L., D.N. and L.H. conducted the experiments. V.A.B. supported the animal preparation. B.T., H.L., L.H. and T.L. analyzed the data. B.T., H.L., Y.Z., D.P., L.S. and T.L. wrote the draft. All work was supervised by T.L., D.P. and L.S. T.L. and L.S. obtained funding. All authors contributed to the final manuscript.

All authors declare no competing financial interests.

[1]B.T. and H.L. contributed equally to this work.

[2]To whom correspondence should be addressed.
E-mail: palanker@stanford.edu, leopold.schmetterer@seri.com.sg, tong.ling@ntu.edu.sg

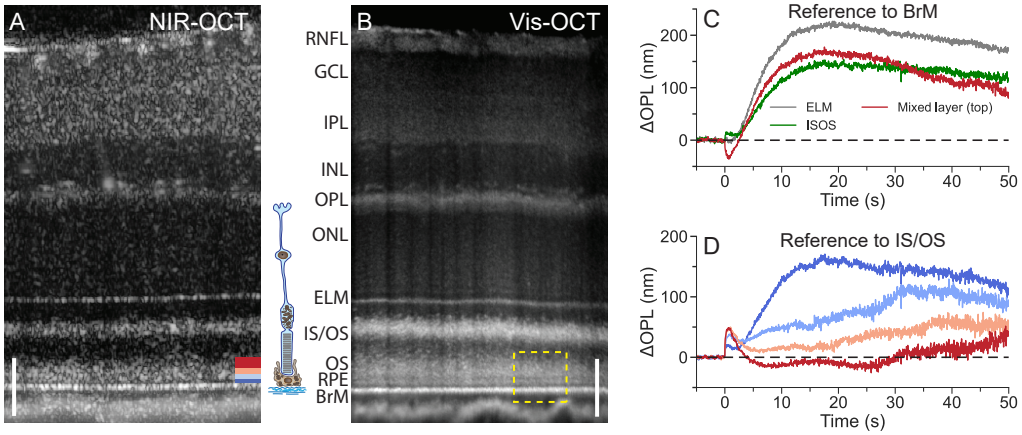

Fig. 1. Retinal layers and their dynamics in response to visual stimuli. (A) Averaged retinal B-scan from NIR-OCT. (B) Ultrahigh-resolution vis-OCT image helps to distinguish photoreceptors' OS, RPE, and BrM layers. An enlarged view of the dashed box is shown in Fig. 3B. Scale bar: 50 µm. (C) NIR-OCT was used for ORG imaging experiments. ORG signals obtained from various hyperreflective bands in the outer retina by taking the BrM as the reference. (D) ORG signals obtained at various depths in the mixed layer relative to IS/OS, with colors matching the corresponding bars at the bottom right of Fig. 1A. RNFL: retinal nerve fiber layer; GCL: ganglion cell layer; IPL: inner plexiform layer; INL: inner nuclear layer; OPL: outer plexiform layer; ONL: outer nuclear layer; ELM: external limiting membrane; IS/OS: inner segment/outer segment junction; OS: outer segment; RPE: retinal pigment epithelium; BrM: Bruch's membrane.

In this paper, we present a robust, unsupervised learning approach to finding the hidden spatiotemporal patterns in the phase signals measured from speckles, which reveals novel signatures of the SRS, photoreceptor inner and outer segments, and RPE responses to light. Notably, image registration using our phase-restoring subpixel motion correction algorithm enables ORG recordings down to nanometer-scale for tens of seconds, as opposed to the typical few seconds in previous studies, allowing for detection of much slower and more subtle phenomena in retinal responses to visual stimuli. We demonstrate these light-evoked responses under various scotopic and photopic conditions and map the OS and SRS dynamics across a wide field using a single flash.

Results

Outer retina dynamics in response to visual stimuli. All ORG imaging experiments were performed *in-vivo* using a custom-built NIR-OCT with an axial resolution of 2.0 µm in tissue. Cross-sectional (Fig. 1A) and volumetric scans were acquired in time sequences from a 12° field of view (FOV) on wild-type rat retinas. A custom-built ultrahigh-resolution visible-light OCT (vis-OCT), which adopted the same optical design as the NIR-OCT but provided 1.1 µm axial resolution in tissue, was used to validate retinal layer delineation (Fig. 1B). As illustrated in Fig. 1, several hyperreflective layers were observed in the outer retina, including the external limiting membrane (ELM), the inner segment/outer segment junction (IS/OS), Bruch's membrane (BrM), and a thick speckling layer between the IS/OS and the BrM, which we call the *mixed layer* for convenience. As confirmed by the vis-OCT, the mixed layer contained photoreceptors' OS and RPE cells that the NIR-OCT system cannot resolve. To reveal the light-evoked dynamics in the outer retina, first we computed the changes in the optical path length (OPL) (phase difference) between the BrM and three hyperreflective bands in the outer retina, including ELM, IS/OS, and the top band of the *mixed layer* (the red color in the bottom right of Fig. 1A). To achieve high phase stability/sensitivity *in-vivo*, one needs to correct the bulk tissue motion. For this purpose, we applied our recently developed phase-restoring motion correction method to register complex-valued OCT images with subpixel precision (18). Phase traces extracted from the prolonged ORG experiments (55 s in duration) were spatially averaged across pixels within each band. Note that in this study, when converting the phase change into the OPL change, an additional negative sign was added if the target layer was anterior to the reference layer. This ensured that an increase in OPL always represented an expansion between the two layers.

As shown in Fig. 1C, distance between ELM and BrM increases after the stimulus (1 ms, 500 nm, 0.18% bleach level) at a rate of about 26 nm/s, reaching 220 nm around 20 seconds, and then slowly recovers afterwards. The IS/OS layer rapidly (∼1 s) moves away from BrM, bounces back within 2 seconds and then continues to slowly move away from the BrM. The top red band of the mixed layer (presumably OS tips) rapidly (∼1 s) moves toward the BrM by about 35 nm, then moves away from it over the next 20 seconds, followed by a slower recovery afterwards.

Conventionally, ORG monitors the movement of OS tips relative to the IS/OS layer. Due to the ambiguity of multiple tissue layers in NIR-OCT, we recorded the OPL changes from several bands in the mixed layer, indicated by various colors at the bottom right of Fig. 1A, relative to the IS/OS. As shown in Fig. 1D, the OPL changes in different bands contain a mixture of two distinct signatures: (a) rapid (∼1 s) expansion followed by a slower (∼10 s) decline, and (b) slow (∼20 s) expansion, followed by even slower recovery. The superposition of distinct signals necessitates further signal decomposition and classification to determine the tissue origin associated with each individual signal.

Unsupervised learning of spatiotemporal patterns for signal classification. The axial resolution limit of NIR-OCT leads to the mixing of light-evoked responses from both OS and RPE, thereby complicating the interpretation of the results. To identify distinct signal patterns in an unbiased manner and group them into distinct types, we projected individual

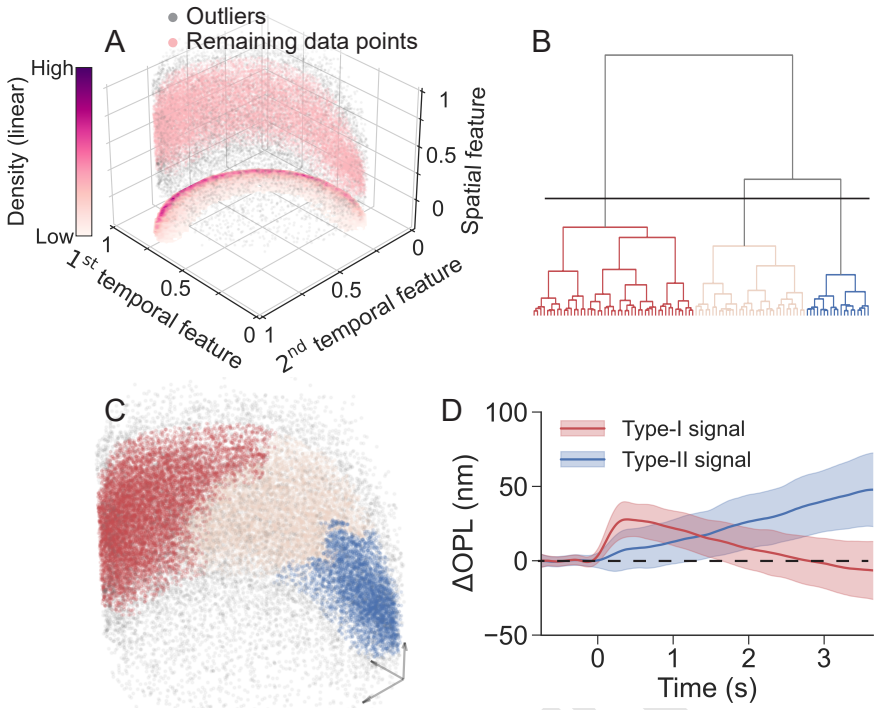

Fig. 2. Unsupervised clustering in the spatiotemporal feature space. (A) Distribution of the signals in the 3D spatiotemporal feature space. Gray dots represent the outliers identified using a distance-based detection method, while remaining data points were labeled in pink. The heatmap shows the distribution density of the remaining data points when projected onto the temporal feature plane. (B) Dendrogram showing the cluster structure of the remaining phase traces (pink dots in Fig. 2A) in the spatiotemporal feature space, with only the top 100 subclusters displayed. (C) Clustering the remaining phase traces in the spatiotemporal feature space into three clusters by thresholding the dendrogram along the solid black line. Colored dots denote different groups, and the gray dots denote outliers. (D) Corresponding representative *Type-I* and *Type-II* signals obtained by averaging the individual phase traces within each cluster. The solid lines and color bands denote the mean values and the range of standard deviations.

phase traces onto a spatiotemporal feature space and employed unsupervised learning using agglomerative hierarchical clustering based on Ward's criterion (19). We then trained a support vector machine (SVM) model on these type labels within the established feature space, enabling the extraction of each cluster's decision boundary and the classification of new phase traces using the same decision boundaries (*Materials and Methods* & *Supplementary Materials*, Fig. S1).

To account for inter-subject differences, we combined phase traces from five rats for feature extraction and subsequent unsupervised clustering. We conducted preprocessing and principal component analysis (PCA) on the phase traces to extract their temporal features (*Materials and Methods* & *Supplementary Materials*, Fig. S2). We then constructed a three-dimensional (3D) spatiotemporal feature space using the distance to BrM (depth) as the spatial feature and the top two principal components (PCs) as the temporal features (Fig. 2A). Outliers in low-density regions were removed using a distance-based algorithm (gray dots in Fig. 2A).

To differentiate between two distinct signatures (Fig. 1D), we employed an agglomerative hierarchical clustering algorithm based on Ward's criterion (19). Thresholding the dendrogram along the solid black line in Fig. 2B grouped the remaining phase traces (pink dots in Fig. 2A) into three clusters in the spatiotemporal feature space (Fig. 2C), where a transition band (light pink dots) facilitates a better separation between two distinct dynamics signatures. Representative signals (Fig. 2D, red and blue dots in Fig. 2C) were obtained by averaging the individual phase traces within each cluster and converting them into OPL changes (ΔOPL). The first type of signal (red, *Type-I*) exhibited a rapid increase, peaking around 0.5 s, followed by a gradual decrease and a negative overshoot after 2.5 seconds (Fig. 2D). This Type-I signal matches the ORG signals from the photoreceptor outer segments in the literature, which were published at shorter time ranges (8, 10). The second type of signal (blue, *Type-II*) is characterized by a much slower rise, and its peak is not reached within the 3.5-second time range plotted in Fig. 2D.

Subsequently, we trained an SVM using the labels obtained from unsupervised clustering in Fig. 2C to set boundaries for Type-I (red surface) and Type-II (blue surface) signals in the spatiotemporal feature space (Fig. 3A). Using the pre-trained SVM, we successfully extracted Type-I and Type-II signals from new datasets. Type-I signals were located more anteriorly than Type-II signals, exhibiting a distribution close to normal (red line in Fig. 3B). Type-II signals were localized anterior to BrM, fitting a Gaussian function (blue curve in Fig. 3B). The locations of the peaks differed significantly ($^*P < 0.05$, *t-test*). The location of Type-I signals corresponded to the intensity profile of OS, while Type-II signals corresponded to location of RPE, as measured by vis-OCT (yellow line in Fig. 3B). This observation suggested that the Type-I and Type-II signals correspond to the dynamics of the OS and RPE relative to IS/OS, respectively. We interpolated and clustered the phase traces from the prolonged recording using the SVM. The Type-I signal peaked within one second, underwent a negative undershoot peaking at 10 seconds, and returned to baseline within 30 seconds (Fig. 3C). By incorporating the Type-II signal with the dynamics between IS/OS and ELM, we obtained the dynamics of SRS (from ELM to RPE). The

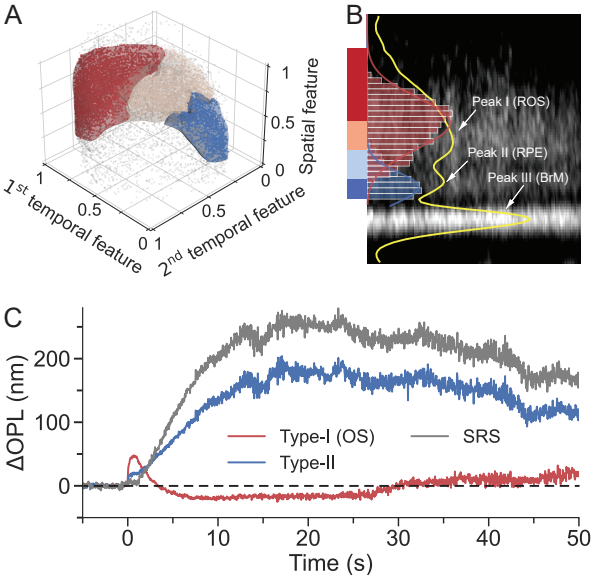

Fig. 3. Classification of new phase traces and validation of their origins. (A) The trained SVM decision boundaries for the Type-I signal (red) and the Type-II signal (blue). Dots represent the phase traces extracted from a new dataset, preprocessed and projected onto the same 3D feature space. They were classified into Type-I signals (red dots), Type-II signals (blue dots), intermediate phase traces (light pink dots) and outliers (gray dots). (B) An enlarged view of the dashed box in Fig. 1B, with contrast adjustment to enhance the RPE visibility. Histogram of Type-I (red bins) and Type-II (blue bins) signals, fitted by Gaussian functions (solid lines), shows their depth distribution on top of the averaged structural image captured by vis-OCT, with a yellow line depicting its intensity profile. The colored bars on the left correspond to the depth range from which the signals in Fig. 1D were extracted. (C) Type-I, Type-II and SRS signals from the prolonged recording.

SRS signal (gray trace in Fig. 3C) increased more slowly, peaking at about 250 nm around 20 seconds, followed by an even slower recovery. In our subsequent quantitative analyses, we focus on OS and SRS dynamics.

Signal dependence on stimulus parameters. We further quantified the signals in scotopic and photopic conditions by plotting the amplitude and latency of the peak of OS expansion and slope (expansion rate) from the SRS response (Fig. 4A). In scotopic conditions (Fig. 4B), the peak amplitude of the OS signal increased logarithmically with the stimulus strengths, from ΔOPL of about 10 (1.72) nm [mean (SD)] at a bleach level of 0.002% to about 53 (8.68) nm when the flash intensity increased 140-fold to 0.28%. The peak latencies remained within 440 – 500 ms at a bleach level < 0.1%, but increased to 877 (130) ms and 1063 (108) ms in response to stimuli at 0.19% and 0.28% bleach level, respectively. In contrast, the slope of the SRS signal increased rapidly with the stimulus intensity up to 0.06% bleach level, and stabilized at the level of about 25 nm/s after that. In photopic conditions (Fig. 4C), the retina was pre-illuminated (500 nm) for 5 minutes and the flash was at 0.28% bleach level. Background illuminance below 6×10^2 photons/(μm^2·s) did not reduce the OS response, but it decreased 5-fold at background of 6×10^4 photons/(μm^2·s). The OS peak latency was not altered with an even stronger background - up to 6×10^3 photons/(μm^2·s), but it dropped by about 30% at 6×10^4 photons/(μm^2·s). Similarly, SRS expansion rate was affected by a background illuminance only above 6×10^2 photons/(μm^2·s), which reduced the slope by about 26% to 17.8 nm/s.

***En-face* functional maps.** Using repeated volumetric scans, we mapped the SRS and OS dynamics, analogous to multifocal ERG, but with a single flash. Figure 5A displays an OCT volume with a 12° FOV, with the structural contrast in gray and angiographic contrast highlighted in red. The spatiotemporal distribution of SRS and OS signals was obtained at a 0.10% bleach level. The temporal resolution was 8 Hz, and the total recording time was 5 seconds with 1 second baseline. Figure 5B shows the OS and SRS signals at specific time points. Fig. 5C maps the spatial distribution of the phase

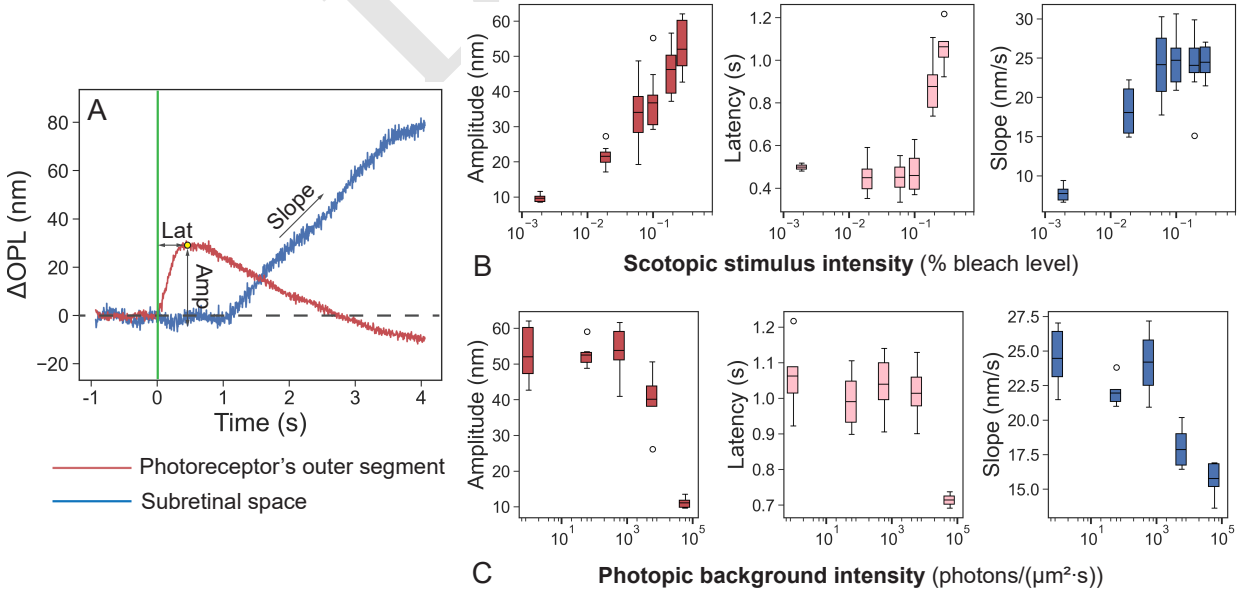

Fig. 4. Responses of the outer segment (OS) and subretinal space (SRS) in different conditions. (A) Representative traces of the light-evoked responses in photoreceptor OS and in SRS. (B) Amplitude and latency of the OS response, and slope of the SRS expansion as a function of stimulus intensity on scotopic background (C) Same as a function of the background illuminance. For box plots, horizontal bar: mean value, box edges: 25 and 75 percentiles, whiskers: 1.5 × standard deviations (SDs).

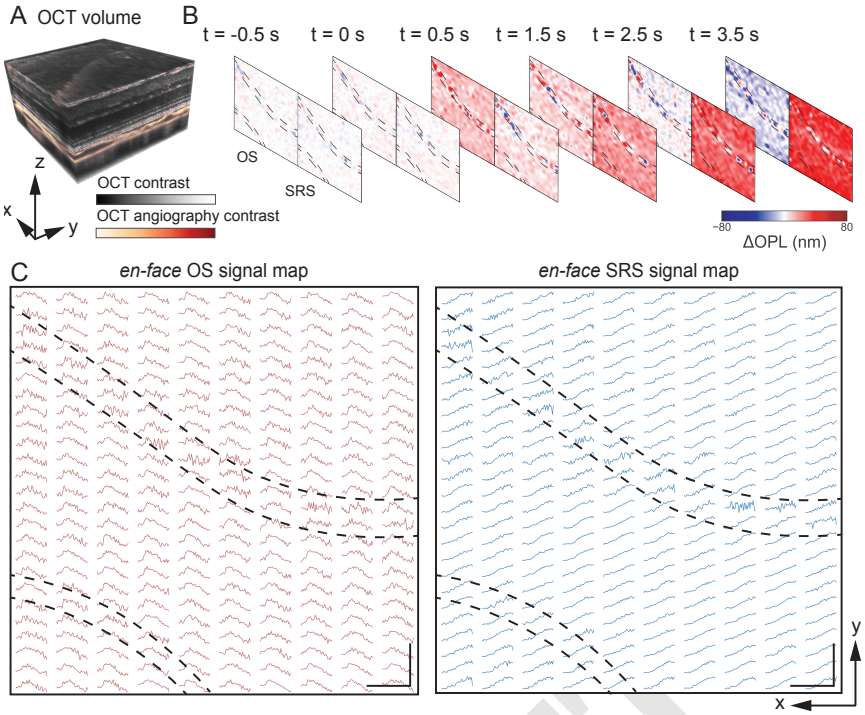

Fig. 5. Representative *en-face* maps of outer segment (OS) and subretinal space (SRS) signals. (A) A volumetric scan covered a 12° field of view, with the structural contrast in gray and angiographic contrast highlighted in red. (B) The OS and SRS signals at selected time points. Locations blocked by large blood vessels are outlined by dashed lines. (C) Spatiotemporal evolution of the OS and SRS signals over the entire FOV. Each curve presents the average response from a 1.2° × 0.48° ($x \times y$) area. Scale bar: 200 µm.

traces, with each grid representing the average response from a 1.2° × 0.48° area. In general, the OS and SRS signal maps show high-fidelity detection over the entire FOV, and high signal variance was observed underneath two large blood vessels, outlined by dashed lines in Figs. 5B-C. A video illustrating the spatiotemporal evolution of the OS and SRS signals can be found in *Supplementary Materials (Movie S1)*.

Signals from SRS and OS, recorded over a wide field in response to a single flash, offer a convenient approach to diagnostic mapping, in contrast to the slow and low-resolution multifocal ERG. This new technique expands the practical applicability of optoretinography to studies of not only photoreceptors but also the RPE's control of water dynamics in SRS in health and disease, thus becoming a more complete replacement for ERG.

Discussion

After disambiguation of various signals in the outer retina, dynamics of its cellular layers can be described independently. Time derivative of the tissue deformation, i.e., its expansion rate, is very informative since it may reveal the water influx rate. As shown in Fig. 6, expansion rate of OS reaches its maximum of about 165 nm/s within 0.1 s after the stimulus and drops back to zero during the next 1 s. RPE, on the other hand, reaches the maximum contraction rate of about -9 nm/s at 2 s after the stimulus, leading to the maximum contraction of approximately 20 nm at 3 s, and slowly recovering afterwards. Expansion of SRS begins about 1.5 s after the stimulus, reaching a rate of 15 nm/s in the following 10 s, after which the expansion rate gradually decreases. SRS expansion may originate from water transport across the RPE in response to a decrease in K^+ and an increase in Na^+ concentrations due to phototransduction in photoreceptors, as observed earlier using other methods (20, 21). The inner segment is compressed by about 10 nm during 1.5 s, after which it follows the dynamics of SRS, albeit at a lower amplitude (100 nm expansion at maximum).

It is interesting to relate the ORG signals to ERG previously recorded in the same species – pigmented rats, in response to a white flash (22), shown in Fig. 6E. The a-wave in ERG – the photoreceptors' response to light, begins right after the flash (within 10 ms), corresponding to the beginning of the OS expansion in ORG. Continuation of the a-wave is obscured by the b-wave, which starts a few tens of ms later and corresponds to the electric current generated by the second-order retinal neurons. C-wave in ERG takes over after about 0.5 seconds and reaches its maximum at around 1.5 seconds. It corresponds to a potassium current through RPE, induced by the photoreceptors' response to light. The RPE contraction in ORG, which reaches its maximum rate between 1 and 2 s, overlaps with the peak of c-wave (Figs. 6D-E). This novel optical signature opens the window into physiology of the photoreceptor-RPE interactions and the interphotoreceptor matrix. Water transport is limited by the membrane permeability and hence its dynamics is much slower than that of electric current. Using a model of the OS elongation due to osmotic imbalance during phototransduction (6), we estimated the membrane permeability coefficient based on the OS expansion rate. The water permeability coefficient of the rod OS membrane was found to be 5.9×10^{-3} cm /s, not too far from 2.6×10^{-3} cm /s measured earlier in vitro – see *Supplementary Materials* and Table S1 for details.

Intriguingly, the hyperpolarized RPE layer is getting compressed in this process, unlike swelling of the hyperpolarized

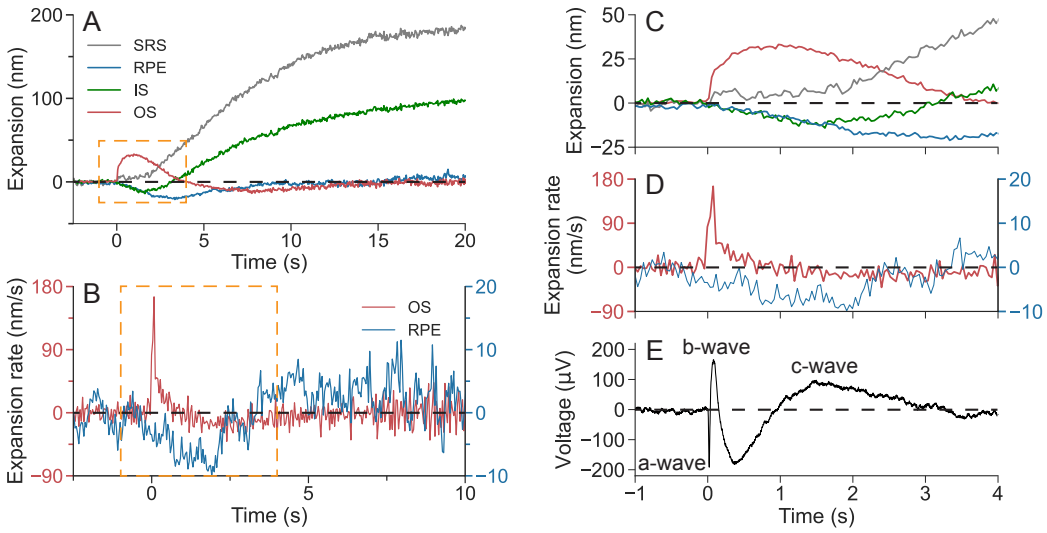

Fig. 6. Deformation of cellular layers over time and comparison of ORG with ERG. (A) Representative dynamics of the deformations of the rod outer segment (OS), inner segment (IS), retinal pigment epithelium (RPE), and subretinal space (SRS) after a 1 ms green stimulus at 0.26% bleach level. (B) The expansion rate of the OS (red curve) and RPE (blue curve) averaged across 5 measurements. (C) and (D) Zoom-in view of the orange boxed data in (A) and (B), respectively. (E) An example electroretinography (ERG) trace in response to a white flash, where the a, b and c wave are labeled. Modified from Tan *et al.* (22). OPL change was converted into physical deformation with a refractive index of 1.41.

outer segments driven by osmotic pressure changes in response to released osmolytes during phototransduction.

The responses of rods and cones to light vary in many aspects, including sensitivity, time constant and light adaptation (23). These differences may be attributed to different expressions of isoforms involved in the phototransduction cascade and distinct structural organizations of plasma membranes in rods and cones (24, 25). Until now, ORG was observed mainly in cone photoreceptors, since detecting rod signals in human subjects is more challenging because rods are generally smaller and densely packed around cones (26–28). In rats, however, 97% of photoreceptors are rods (29), and hence our OS signals are likely dominated by the rod outer segments. Adaptive optics (AO) enables imaging of single rods in peripheral human retina, and one AO-OCT study reported that a flash bleaching 0.05% rhodopsin resulted in rod OS elongation of about 60 nm, while using the same flash bleaching 0.2% opsin did not result in a detectable cone OS elongation (10). These results agree with our observations: 0.06% bleach caused OS elongation by 20-30 nm in rats.

In a previous study using conventional intensity-based OCT (6), Zhang *et al.* observed a slow (peak latency of 10-100 seconds, depending on the bleach level) increase in distance between IS/OS and BrM in response to light, and attributed it to elongation of the OS. Notably, this slow signal resembles our SRS expansion, while the actual OS signal is much faster to rise and fall in our observation and in previous reports (10).

Another interesting finding is the undershoot of the OS signal, with its tens of seconds-long recovery (Fig. 6A), which was not reported in previous studies due to shorter observation time. It may be related to water transport from the OS to SRS when the osmolytes ($G\alpha_t$, $G\beta_1\gamma_1$) rebind to the cell membrane during deactivation of phototransduction. Restoration of the OS osmotic pressure during expansion of the SRS may be accompanied by its slight (20 nm) compression, which recovers later along with the other cellular structures in the outer retina.

Expansion of the SRS, closely related to phototransduction and water transport via RPE, is larger than that of the OS. Such signature may serve as a more sensitive measure of retinal physiology, providing additional diagnostic insights for various diseases involving the outer retina. ORG conveniently combines structural and functional retinal imaging in the same machine.

Materials and Methods

System setup. All the ORG imaging experiments in this study were performed using a custom-built spectral-domain OCT (Fig. 7A). This NIR-OCT system operated with a broadband superluminescent diode (cBLMD-T-850-HP-I, λ_c = 840 nm, $\Delta\lambda$ = 146 nm, Superlum, Ireland), providing an axial resolution of 2.0 μm in tissue. A spectrometer interfaced with a line-scan camera (Cobra-800, Octoplus, E2V) acquired the spectral interference fringes at an A-scan speed of 250 kHz, corresponding to an image depth of 1.07 mm in air. For posterior segment imaging in rodent eyes, a lens-based afocal telescope conjugated the axis of a galvo scanner to the pupil. To reduce the incident beam size and increase the scan angle, the telescope's magnification was 0.17 (scan lens: 80 mm focal length; ocular lens: 30 mm + 25 mm focal length). The theoretical diffraction-limited lateral resolution was 7.2 μm with a standard rat eye model (30). A function generator (PCIe-6363, National Instruments, USA) synchronized the camera acquisition (both A-scan and B-scan acquisitions), galvanometer scanning, and visual stimulation.

For visual stimulation, a LED (MBB1L3, Thorlabs, USA) was collimated by an aspheric condenser lens, and a narrow bandpass filter (500 ± 5 nm, #65-694, Edmund Optics, Singapore) was used to reshape the spectrum to optimize the sensitivity to rhodopsin (31). The LED's response time was ∼300 μs, equivalent to 6% of the B-scan frame acquisition time, and LED's response waveform to trigger is shown in *Supplementary Materials*, Fig. S3. A 43.4° Maxwellian illumination was projected to the posterior eye to cover an area of approximately 6.75 mm^2. The power and duration of the light stimulus were converted to the bleach percentages of rhodopsin using a published method (32, 33), detailed in *Supplementary Materials*.

A vis-OCT at higher axial resolution of 1.1 μm in tissue was constructed to resolve the ultra-fine laminated structure of the outer

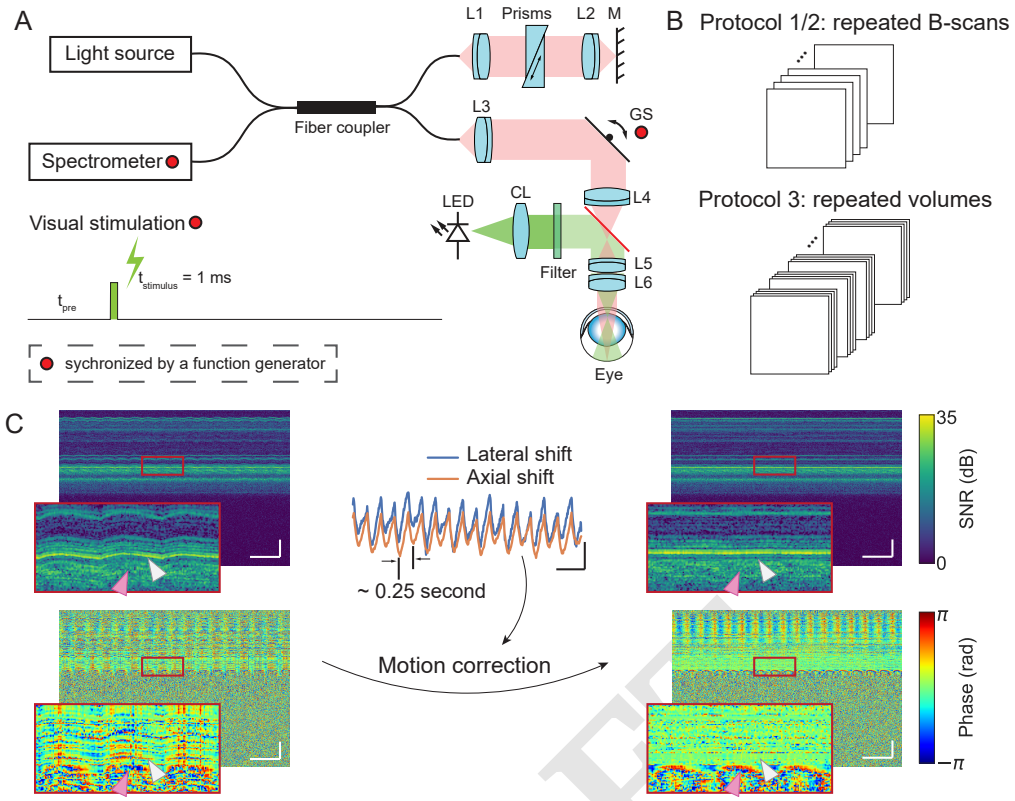

Fig. 7. System setup, stimulus scheme, acquisition protocols and image registration. (A) A spectral-domain OCT was used to image the posterior segment of rat eyes. A line-scan camera interfaced with the spectrometer was used to acquire the interference fringes. A function generator synchronized the line-scan camera acquisition, galvo scanner rotation, and flash timing. L1-L6: doublet lenses. CL: condenser lens. GS: galvo scanner. Filter spectral window: 500 ± 5 nm. (B) Three acquisition protocols were used: repeated B-scans (Protocol 1: 1000 A-scans per B-scan, 200 B-scans per second, Protocol 2: 1000 A-scans per B-scan, 25 B-scans per second) and repeated volumes (Protocol 3: 1000 A-scans per B-scan, 25 B-scans per volume, 8 volumes per second). The details of acquisition and stimulation protocols were listed in *Supplementary Materials*, Table S3. (C) Left column: time-elapsed intensity (upper) and phase (lower) M-scans without correction, when no light stimulus was delivered to the retina. Scale bar: 500 ms (horizontal), 100 µm (vertical). Center column: the subpixel-level bulk motion estimated by locating the peak of the upsampled cross-correlation map between repeated B-scans. Scale bar: 500 ms (horizontal), 5 µm (vertical). Right column: corresponding time-elapsed intensity (upper) and phase (lower) M-scans after the phase-restoring subpixel motion correction. Scale bar: 500 ms (horizontal), 100 µm (vertical). Gray and pink arrows label loci of the outer retina and the choroid, respectively.

retina. Briefly, the system used a supercontinuum laser (SuperK Extreme, NKT Photonics, Denmark) with a spectrum truncated between 435 and 650 nm. Vis-OCT adopted a similar optical design as the NIR system, albeit all the components work in the visible spectral range. The details of the vis-OCT setup are in *Supplementary Materials*. A spectrometer (Cobra VIS, Wasatch, USA) interfaced with a line-scan camera acquired the spectra of the interference fringes at an A-scan speed of 50 kHz. Volumetric raster scans (500 A-scans × 500 B-scans) were collected within a 12° × 1.2° (1.14 mm × 0.11 mm) rectangle field of view. Sequential B-scans were aligned and then averaged into a single frame.

Animal experiment protocol. All experiments were conducted in accordance with guidelines and approvals from Institutional Animal Care and Use Committee (IACUC), SingHealth (2020/SHS/1574). Brown Norway rats (N = 43) were used with details listed in Supplementary Materials, Table S2). Animals were sedated using a ketamine/xylazine combination to better maintain retinal functional responses compared to other commonly used anesthetics, such as isoflurane and urethane, while minimizing eye motion (34, 35). Vital signals, including heart and respiration rates, were monitored throughout the imaging sessions. Animals were placed in a prone position with their head restrained stereotaxically. Two mydriasis drops, 1% Tropicamide (Alcon, Geneva, Switzerland) and 2.5% Phenylephrine (Alcon, Geneva, Switzerland), were administrated onto the cornea before the imaging. The cornea was moisturized by a balanced salt solution frequently throughout the imaging sessions.

M-opsin and rhodopsin have similar sensitivity spectra (peaking at approximately 500 nm), while S-opsin's sensitivity peaks at 350 nm, with minimal overlap with rhodopsin sensitivity spectrum. Since M-opsin co-expression ratio decreases from ventral to dorsal (31), the scanning region was limited to the dorsal area to minimize the influence of M-opsin in cones.

Details of the acquisition and stimulation protocols are listed in *Supplementary Materials*, Table S3. Three scanning protocols were used in this study. In the first protocol, we acquired 1000 A-scans per B-scan and 200 B-scans per second, with a total acquisition time of 5 seconds. In the second protocol, the B-scan time interval was increased to 40 ms and 1375 B-scans were acquired over a period of 55 seconds. The third protocol consisted of 40 repeated volumes (25 B-scans per volume) recorded at 8 volumetric scans per second for a period of 5 seconds. Flash intensity, dark/light adaptation and inter-flash time interval varied in different experiments.

Automated extraction of the outer retina dynamics from speckle patterns.

Preprocessing of OCT images and phase traces. The raw interference fringe first underwent standard OCT post-processing steps, including spectral calibration, k-linearity, dispersion compensation, and discrete Fourier transform (DFT), which resulted in complex-valued OCT images.

Phase-resolved OCT is very susceptible to bulk tissue motion, which results in apparent image distortions in the time-elapsed intensity M-scan and degrade the phase stability (see Fig. 7C, left column). To correct the motion-induced phase error, we estimated subpixel-level translational displacements between repeated B-scans using the single-step DFT algorithm (see Fig. 7C, center column) (36). We registered the complex-valued OCT images using our

recently developed phase-restoring subpixel motion correction algorithm (18), where the lateral and axial displacements were corrected by multiplying the corresponding exponential terms in the spatial frequency domain and the spectrum domain, respectively. After image registration (see Fig. 7C, right column), we achieved excellent motion stability at photoreceptor layers (gray arrows), while periodical oscillations from the vascular pulsation (pink arrows) could be observed in the choroid.

Light-evoked dynamics of the outer retina was then extracted by computing the temporal phase difference between the pixel pairs from the outer retinal bands. Several hyperreflective bands, including ELM, IS/OS, BrM, and a thick speckling layer that contains photoreceptors' OS and RPE cells, were automatically segmented using graph theory and dynamic programming (*Supplementary Materials*, Fig. S1A) (37). In self-referencing measurements, for each pixel in a target layer, we computed its temporal phase change with respect to a reference layer, detailed in *Supplementary Materials*.

We also conducted temporal filtering to remove unwanted signal frequencies. The extracted signals were first processed by bandstop filters to minimize residual artifacts induced by heartbeat and breathing. A low-pass filter with a cut-off frequency of 10 Hz was subsequently used to filter out high-frequency oscillations. After filtering, both ends of the phase traces exhibited high variance due to edge effects, so they were removed from the data analysis afterwards (*Supplementary Materials*, Fig. S1C).

Construction of spatiotemporal feature space. We used the principal component analysis (PCA) to compress high-dimensional phase traces into a lower-dimensional feature space to facilitate the analysis (*Supplementary Materials*, Fig. S2C). Traces with a pre-stimulus SD larger than 60 mrad were excluded, mostly from pixels with low SNR or underneath blood vessels. Then each trace was normalized by subtracting its mean value and dividing by its SD (*Supplementary Materials*, Fig. S2). To account for inter-subject differences, we combined normalized traces extracted from five rats to calculate PC coefficients. The top two PCs capturing a total variance of 73.3% were selected as temporal features and its axial distance to the BrM was selected as the spatial feature. We normalized both the top two PCs and the depth information to construct a three-dimensional spatiotemporal feature space.

To avoid undesirable elongated clusters in subsequent unsupervised clustering analysis, we used a distance-based outlier detection method to preclude outliers distributed in low-density regions in the spatiotemporal feature space (38, 39). For each data point in the feature space, we calculated the minimum radius of a sphere that is centered at that point and can cover 2% of the remaining data points. This radius reflected the local distribution density around each data point in the feature space. Then, a given point would be labeled as an outlier if its corresponding radius was larger than $Q_3 + (Q_3 - Q_1)/5$, where Q_1 and Q_3 are the first and third quartiles of the calculated radii.

Unsupervised clustering with a hierarchical clustering algorithm. We used an agglomerative hierarchical clustering algorithm under the Ward criterion to group individual points in the established spatiotemporal feature space (19). The Euclidean distance of each pair of points was computed and similar subclusters were then iteratively merged into larger clusters. Under the guidance of the Ward method or minimum variance method, each merger guaranteed a minimum increase of total within-cluster variance.

Training a support vector model in the established feature space. We trained a SVM in the spatiotemporal feature space with the previously obtained labels, including outlier, intermediate phase trace, Type-I signal, and Type-II signal, to obtain their decision boundaries. The input features were standardized before training, the Gaussian kernel was selected, and automatic hyperparameters optimization was turned on. Evaluated with the 10-fold cross-validation strategy, the trained SVM achieved a classification accuracy of 99.2%. The SVM is useful for processing a new dataset using the same criterion. Phase traces extracted from the new dataset will be preprocessed and projected onto the same feature space. Then Type-I and Type-II signals can be automatically extracted based on the classification results from the SVM.

Processing of datasets with lower temporal resolution. Prolonged recording (protocol 2) and volumetric scans (protocol 3) had an 8-fold and 25-fold lower temporal sampling rate than repeated B-scans (protocol 1), respectively. To cluster these phase traces in the same spatiotemporal feature space, they were first linearly interpolated into 5 ms time intervals and smoothed using a Gaussian filter. Other procedures were similar to those used for processing the repeated B-scans datasets.

ACKNOWLEDGMENTS. We thank Drs. R. Sabesan, A. Roorda, E. Pugh and M. Ismail for fruitful discussions. T. Ling acknowledges the support of the National Research Foundation, Singapore for the NRF Fellowship Award (NRF-NRFF14-2022-0005), the Startup Grant (SUG) from Nanyang Technological University, the seed funding programme under NMRC Centre Grant – Singapore Imaging Eye Network (SIENA) (NMRC/CG/C010A/2017), and the Ministry of Education, Singapore under its AcRF Tier 1 Grant (RS19/20; RG28/21). L. Schmetterer acknowledges the funding support for the National Medical Research Council (CG/C010A/2017_SERI; OFIRG/0048/2017; OFLCG/004c/2018; TA/MOH-000249-00/2018 and MOH-OFIRG20nov-0014), National Research Foundation Singapore (NRF2019-THE002-0006 and NRF-CRP24-2020-0001), A*STAR (A20H4b0141), the Singapore Eye Research Institute & Nanyang Technological University (SERI-NTU Advanced Ocular Engineering (STANCE) Program), the Duke-NUS Medical School (Duke-NUS-KP(Coll)/2018/0009A), and the SERI-Lee Foundation (LF1019-1) Singapore. V. A. Barathi acknowledges the support for Singapore NMRC Centre Grant (NMRC/CG/M010/2017/Pre-Clinical). D. Palanker and Y. Zhuo acknowledge the NIH grant U01 EY025501 and AFOSR grant FA9550-20-1-0186.

References

1. O Strauss, The retinal pigment epithelium in visual function. *Physiol Rev* **85**, 845–81 (2005).
2. JR Sparrow, D Hicks, CP Hamel, The retinal pigment epithelium in health and disease. *Curr Mol Med* **10**, 802–23 (2010).
3. A Lakkaraju, et al., The cell biology of the retinal pigment epithelium. *Prog. retinal eye research* **78**, 100846 (2020).
4. n B. Oakley, DG Green, Correlation of light-induced changes in retinal extracellular potassium concentration with c-wave of the electroretinogram. *J. Neurophysiol.* **39**, 1117–1133 (1976).
5. DP Joseph, SS Miller, Apical and basal membrane ion transport mechanisms in bovine retinal pigment epithelium. *J Physiol* **435**, 439–63 (1991).
6. P Zhang, et al., In vivo optophysiology reveals that g-protein activation triggers osmotic swelling and increased light scattering of rod photoreceptors. *Proc. Natl. Acad. Sci.* **114**, E2937–E2946 (2017).
7. VP Pandiyan, et al., The optoretinogram reveals the primary steps of phototransduction in the living human eye. *Sci. Adv.* **6**, eabc1124 (2020).
8. D Hillmann, et al., In vivo optical imaging of physiological responses to photostimulation in human photoreceptors. *Proc. Natl. Acad. Sci.* **113**, 13138–13143 (2016).
9. VP Pandiyan, PT Nguyen, J Pugh, E. N., R Sabesan, Human cone elongation responses can be explained by photoactivated cone opsin and membrane swelling and osmotic response to phosphate produced by rgs9-catalyzed gtpase. *Proc Natl Acad Sci U S A* **119**, e2202485119 (2022).
10. M Azimipour, et al., Optoretinogram: optical measurement of human cone and rod photoreceptor responses to light. *Opt. Lett.* **45**, 4658–4661 (2020).
11. S Tomczewski, et al., Light-adapted flicker optoretinograms captured with a spatio-temporal optical coherence-tomography (stoc-t) system. *Biomed. Opt. Express* **13**, 2186–2201 (2022).
12. JW Goodman, *Statistical optics.* (John Wiley & Sons), (2015).
13. F Zhang, K Kurokawa, A Lassoued, JA Crowell, DT Miller, Cone photoreceptor classification in the living human eye from photostimulation-induced phase dynamics. *Proc. Natl. Acad. Sci.* **116**, 7951–7956 (2019).
14. A Lassoued, et al., Cone photoreceptor dysfunction in retinitis pigmentosa revealed by optoretinography. *Proc Natl Acad Sci U S A* **118** (2021).
15. TH Kim, B Wang, Y Lu, T Son, X Yao, Functional optical coherence tomography enables in vivo optoretinography of photoreceptor dysfunction due to retinal degeneration. *Biomed Opt Express* **11**, 5306–5320 (2020).
16. B Chang, et al., Cone photoreceptor function loss-3, a novel mouse model of achromatopsia due to a mutation in gnat2. *Invest Ophthalmol Vis Sci* **47**, 5017–21 (2006).
17. P Calvert, et al., Phototransduction in transgenic mice after targeted deletion of the rod transducin β-subunit. *Proc. Natl. Acad. Sci.* **97**, 13913–13918 (2000).
18. H Li, et al., Shot-noise limited phase-sensitive imaging of moving samples by phase-restoring subpixel motion correction in fourier-domain optical coherence tomography. *bioRxiv* (2022).
19. JH Ward Jr, Hierarchical grouping to optimize an objective function. *J. Am. statistical association* **58**, 236–244 (1963).
20. B Huang, C Karwoski, Light-evoked expansion of subretinal space volume in the retina of the frog. *J. Neurosci.* **12**, 4243–4252 (1992).

21. AV Dmitriev, VI Govardovskii, HN Schwahn, RH Steinberg, Light-induced changes of extracellular ions and volume in the isolated chick retina-pigment epithelium preparation. *Vis Neurosci* **16**, 1157–67 (1999).
22. B Tan, E Mason, B MacLellan, KK Bizheva, Correlation of visually evoked functional and blood flow changes in the rat retina measured with a combined oct+ erg system. *Investig. Ophthalmol. Vis. Sci.* **58**, 1673 (2017).
23. JI Korenbrot, Speed, sensitivity, and stability of the light response in rod and cone photoreceptors: Facts and models. *Prog. Retin. Eye Res.* **31**, 442–466 (2012).
24. NT Ingram, AP Sampath, GL Fain, Why are rods more sensitive than cones? *The J. Physiol.* **594**, 5415–5426 (2016).
25. D Mustafi, AH Engel, K Palczewski, Structure of cone photoreceptors. *Prog Retin. Eye Res* **28**, 289–302 (2009).
26. A Dubra, et al., Noninvasive imaging of the human rod photoreceptor mosaic using a confocal adaptive optics scanning ophthalmoscope. *Biomed. optics express* **2**, 1864–1876 (2011).
27. N Doble, et al., In vivo imaging of the human rod photoreceptor mosaic. *Opt. letters* **36**, 31–33 (2011).
28. Z Liu, F Zhang, K Zucca, A Agrawal, DX Hammer, Ultrahigh-speed multimodal adaptive optics system for microscopic structural and functional imaging of the human retina. *Biomed Opt Express* **13**, 5860–5878 (2022).
29. Y Fu, KW Yau, Phototransduction in mouse rods and cones. *Pflugers Arch. - Eur. J. Physiol.* **454**, 805–819 (2007).
30. M Campbell, A Hughes, An analytic, gradient index schematic lens and eye for the rat which predicts aberrations for finite pupils. *Vis. Res.* **21** (1981).
31. YV Wang, M Weick, JB Demb, Spectral and temporal sensitivity of cone-mediated responses in mouse retinal ganglion cells. *J Neurosci* **31**, 7670–81 (2011).
32. I Perlman, Kinetics of bleaching and regeneration of rhodopsin in abnormal (rcs) and normal albino rats in vivo. *The J. Physiol.* **278**, 141–159 (1978).
33. P Zhang, M Goswami, RJ Zawadzki, J Pugh, Edward N., The photosensitivity of rhodopsin bleaching and light-induced increases of fundus reflectance in mice measured in vivo with scanning laser ophthalmoscopy. *Investig. Ophthalmol. & Vis. Sci.* **57**, 3650–3664 (2016).
34. V Choh, et al., Isoflurane and ketamine:xylazine differentially affect intraocular pressure-associated scotopic threshold responses in sprague-dawley rats. *Documenta Ophthalmol.* **135**, 121–132 (2017).
35. G Nair, et al., Effects of common anesthetics on eye movement and electroretinogram. *Documenta Ophthalmol.* **122**, 163–76 (2011).
36. M Guizar-Sicairos, ST Thurman, JR Fienup, Efficient subpixel image registration algorithms. *Opt. letters* **33**, 156–158 (2008).
37. F LaRocca, et al., Robust automatic segmentation of corneal layer boundaries in sdoct images using graph theory and dynamic programming. *Biomed. Opt. Express* **2** (2011).
38. F Ros, S Guillaume, A hierarchical clustering algorithm and an improvement of the single linkage criterion to deal with noise. *Expert. Syst. with Appl.* **128**, 96–108 (2019).
39. EM Knox, RT Ng, Algorithms for mining distancebased outliers in large datasets in *Proceedings of the international conference on very large data bases*. (Citeseer), pp. 392–403 (1998).