scientific reports

OPEN

Emulated retinal image capture (ERICA) to test, train and validate processing of retinal images

Laura K. Young[1]✉ & Hannah E. Smithson[2]

High resolution retinal imaging systems, such as adaptive optics scanning laser ophthalmoscopes (AOSLO), are increasingly being used for clinical research and fundamental studies in neuroscience. These systems offer unprecedented spatial and temporal resolution of retinal structures in vivo. However, a major challenge is the development of robust and automated methods for processing and analysing these images. We present ERICA (Emulated Retinal Image CApture), a simulation tool that generates realistic synthetic images of the human cone mosaic, mimicking images that would be captured by an AOSLO, with specified image quality and with corresponding ground-truth data. The simulation includes a self-organising mosaic of photoreceptors, the eye movements an observer might make during image capture, and data capture through a real system incorporating diffraction, residual optical aberrations and noise. The retinal photoreceptor mosaics generated by ERICA have a similar packing geometry to human retina, as determined by expert labelling of AOSLO images of real eyes. In the current implementation ERICA outputs convincingly realistic en face images of the cone photoreceptor mosaic but extensions to other imaging modalities and structures are also discussed. These images and associated ground-truth data can be used to develop, test and validate image processing and analysis algorithms or to train and validate machine learning approaches. The use of synthetic images has the advantage that neither access to an imaging system, nor to human participants is necessary for development.

Cellular-resolution imaging of the human retina in vivo has revolutionised retinal research, providing access to structures and functional processes in the retina, and it is increasingly being used in clinical research[1–6]. Adaptive optics (AO) enabled vision research not only advances our fundamental understanding of the healthy visual system, but it also provides a route to developing valuable cellular biomarkers of disease. However, as with many biomedical imaging techniques, one challenge is the bottleneck imposed by the processing steps that are required to generate high quality images and extract information from them. In this paper we present a tool for testing, training and validating retinal image processing methods. Given the cellular-resolution that can be achieved with the use of AO, any image processing and analysis methods should operate with similar or better precision to fully utilise this capability. Furthermore, it is important to consider how the accuracy of the processing and analysis methods that are used impacts the output measures.

The approach taken here is to use synthetic data generated from a physically plausible model of image capture, which offers a number of advantages. Firstly, the properties of the simulated data capture can be systematically manipulated to investigate the effects of different image characteristics (e.g. quality) on the outcome measures, allowing us to understand the sources of error and to define tolerance limits. The parameters of the simulation (e.g. the field of view, frame rate and pixel size) can also be adjusted to simulate different instruments and investigate differences between systems. Secondly, a major advantage of using synthetic images is that they can be generated (and labeled) more rapidly and reliably than can human-derived images. Simulation therefore allows the production of the very large data sets that are typically required for machine learning approaches. Finally, synthetic data sets allow image processing and analysis methods to be developed and tested without needing access to a physical system, to human participants, or to experienced manual image graders. This more easily opens up the image processing challenge to the computer vision community. We first briefly summarise the technological advances in high resolution retinal imaging, then we introduce the data capture and image processing pipeline, including current bottlenecks, and specific use-cases.

[1]Biosciences Institute, Newcastle University, Newcastle NE2 4HH, UK. [2]Department of Experimental Psychology, University of Oxford, Oxford OX2 6GG, UK. ✉email: laura.k.young@newcastle.ac.uk

Background. AO retinal imaging provides unprecedented opportunities for fundamental vision research[5,7]. By enabling finer spatial and temporal scale measurements, AO retinal imaging also offers the potential for closer monitoring of disease[8,9]. At present the most commonly used AO retinal imaging system is the adaptive optics scanning laser ophthalmoscope (AOSLO)[5], with AO flood illumination used by many groups, and AO optical coherence tomography systems being rapidly developed and used. Pioneered by Roorda et al. in 2002[10], the AOSLO generates images by raster-scanning a point source over an area of the retina and measuring the reflected light after focusing it through a confocal pinhole. It is this type of system that we simulate in the current work.

When combined with functional methods, such as single-cell stimulation[11], blood flow measurements[12] or recording intrinsic responses of individual photoreceptors[13,14], the neurophysiology of the visual system can be studied in fine detail. Combining AO retinal imaging with psychophysical approaches provides new capabilities for studying visual perception on a cellular scale. An example relevant to the current work is the exploitation of the motion distortions introduced by scanning during retinal imaging (with and without AO correction). This allows measurement of the microscopic eye movements made during fixation to understand the role that they play in spatial vision[15–20]. Furthermore, these distortions must be removed in order to register and average multiple frames and produce a high signal-to-noise ratio image. The robust collection of good quality data from AO retinal imaging systems, as well as robust data processing automation, is clearly important for these applications.

AOSLO data analysis. There is a drive in the field to widen the adoption of AO retinal imaging instruments[5], but there are challenges that are still to be overcome. As an example, motion estimation relies on comparison to a reference frame, which should have minimal motion[21] or an additional process should remove the error introduced by motion distortions in the reference frame[22,23]. After motion extraction, registration and averaging, single images obtained at adjacent retinal locations are often stitched together to form a larger field of view montage. Errors in motion extraction will propagate through to later image processing and analysis steps. For example, Cooper et al.[24] showed that motion distortions in AOSLO images can affect the reliability and accuracy of cone mosaic metrics derived from those images, and this could have clinically-relevant implications. The image processing and data extraction steps for AOSLO images often require some manual intervention, such as manual reference frame selection or the manual labeling of cone cells. Automated and semi-automated approaches for reference frame selection[21], motion extraction and registration[25–28] and montaging[29,30] have been described. Many automated and semi-automated methods[25,29,31–35] as well as machine learning approaches[36,37] for cone localisation have been, and continue to be, explored. Assessing the accuracy of image processing and data analysis methods requires ground-truth data, which is difficult or often impossible to obtain from real images captured through the imaging system. Another difficulty for machine learning approaches is that large data sets are typically required for training and validation. Even with efficiency improvements, AOSLO data capture with human participants and the manual generation of ground-truth data is time consuming.

Accuracy estimation using ground-truth data derived from real images is possible. Motion estimation accuracy has previously been assessed by imaging a model eye and introducing known or independently-measured movements[38]. Methods of cone identification have also been validated against semi-automated or manually labeled images[33], though this is limited by the accuracy of the labels. A useful alternative approach is to use synthetic images for which ground-truth data are specified. Such an approach has been demonstrated for testing the accuracy of motion extraction from an AOSLO using synthetic data created by resampling real images[22,39]. Synthetic images have also been used to test the accuracy of cone detection algorithms[40,41] and, recently, to train deep learning models for image enhancement[42]. These approaches have generated synthetic retinal mosaics by randomly placing cells[40] or by modifying a hexagonal packing matrix[41,43,44] but stop short of simulating the full data capture pipeline.

ERICA. We present ERICA (Emulated Retinal Image CApture), the first end-to-end simulation of AOSLO data capture generating a realistic output, with specified ground-truth data, comparable to the stream of images obtained from human participants. ERICA has three main stages to generate fully synthetic images with plausible spatial and temporal structure:

1. A self-organising retinal cone mosaic generated with ground-truth cone locations, reflectances, sizes and intensity profiles.
2. A simulation of eye movements added to the raster scan to create a spatio-temporal sampling pattern.
3. Image generation that replicates data capture through an AOSLO, including diffraction, noise and residual wavefront error.

Image processing and analysis algorithms are commonly applied to images of the photoreceptor mosaic, though approaches applied to split detection imaging have also been described[30,45,46]. ERICA generates images of the cone mosaic, but we also suggest extensions to the simulation to include rods in the mosaic, to add blood vessels and to emulate non-confocal imaging. It would also be feasible to extend the model to 3D imaging, based on cellular-resolution retinal mapping in depth and an estimate of the 3D point spread function (PSF) of the eye and optical system.

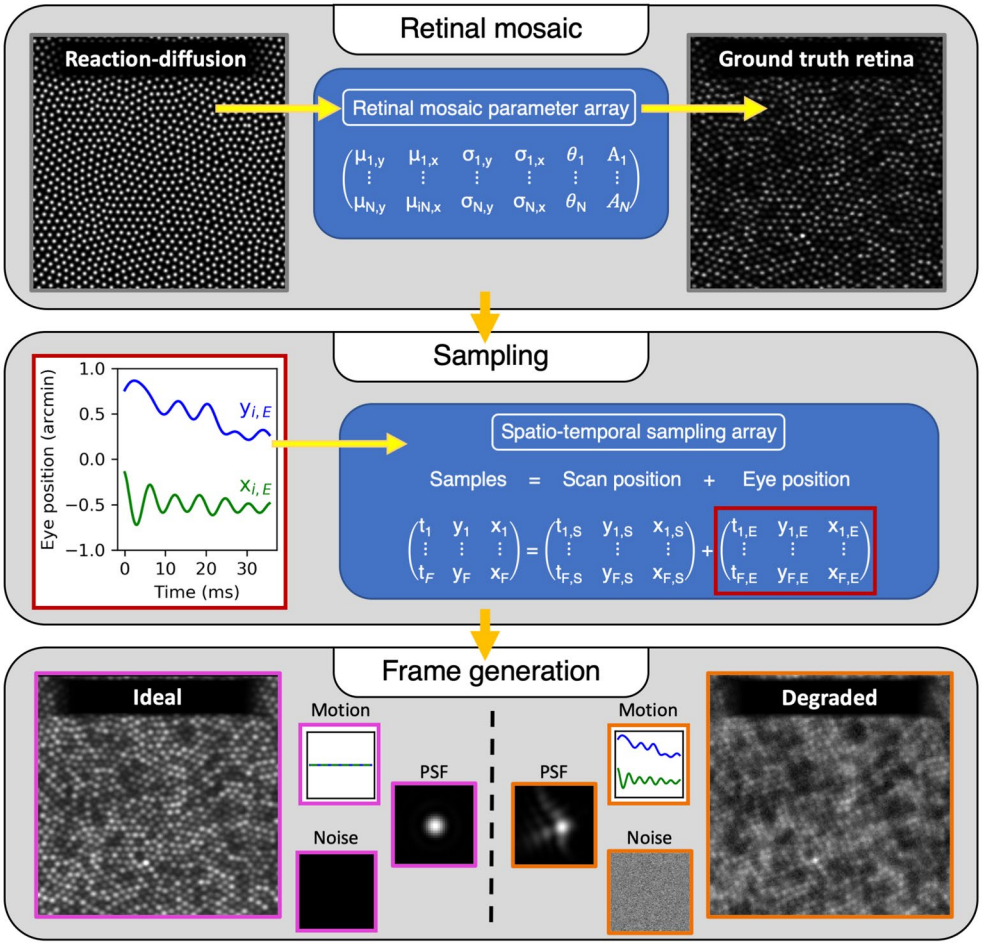

Figure 1. A schematic representation of the three main stages of ERICA. First, a retinal mosaic (left) is generated using reaction-diffusion modelling and the cone positions are extracted. Each of the N cones is defined parametrically as a Gaussian blob with parameters μ_y (vertical coordinate), μ_x (horizontal coordinate), σ_y (vertical width), σ_x (horizontal width), A (reflectance), and θ (orientation), and these are saved in the retinal mosaic parameter array. Second, eye position data are generated and added to the scan position data to define how the retina will be sampled. Third, frames are generated, as described in Fig. 4, incorporating eye motion, diffraction and residual optical aberrations (the PSF) and noise. The example on the left is a diffraction limited image with no noise or eye motion and can be used as a ground-truth frame. The example on the right includes plausible eye motion (drift and tremor, in this case), a realistic PSF based on aberration measurements[49], and noise based on measurements through a real system.

Methods and technical detail

ERICA is written in the Python programming language[47,48] and operates in three stages, outlined in detail in the following sections and summarised in Fig. 1. Firstly, a retinal mosaic is generated using a self-organising approach and this defines the locations of cone cells (and thus provides the ground truth for cone localisation). The retinal mosaic is then defined parametrically (location, size, shape and reflectance of each cone) and an image of the ground-truth retina is generated based on these parameters. The second component of the model is a simulation of the natural eye movements made during fixation. These movements are superimposed on the motion of the raster scanning system to define a spatio-temporal sampling pattern. Thirdly, for each temporal sample, the data capture through an AOSLO is modeled, including diffraction, optical aberrations and noise. Repeating this data capture simulation for each retinal position sampled produces a temporal sequence, analogous to the signals from a photodetector in an AOSLO, which can be processed as a synthetic frame. These procedures are repeated for each frame and the output is a set of images (which can be converted to a video stream, if desired) that approximate those captured from a human participant. Here we outline the current implementation of the three main modules, but alternative approaches and extensions to the simulation can be implemented within the ERICA framework.

Self-organising retinal mosaic. The first stage of the simulation is to generate a realistic mosaic of cone cells. While such a synthetic mosaic has been generated by other means[40,41,43,44,50], our approach is to use a self-

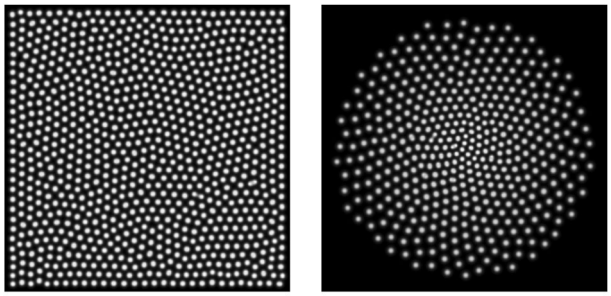

Figure 2. Self-organising mosaics generated using the Gray-Scott reaction-diffusion model. Panel (**a**) shows the output of a 300×300 element system after 20,000 time steps with coefficients: $D_u = 0.14$, $D_v = 0.06$, $f = 0.035$, $k = 0.065$. Panel (**b**) shows the output of a 300×300 element system after 60,000 time steps with coefficients: $D_u = 0.14$, $D_v = 0.4 \times 10^{-3}r + 0.04$, $f = 0.035$, $k = 0.065$, where r is the radial coordinate of the system, resulting in $D_v = 0.04$ in the centre and $D_v = 0.12$ at the furthest edge of the array. The radial variation in the spacing of the spots arises due to the radial variation in the diffusion coefficient. It should be noted that both models were seeded with the chemicals initially in the centre, and, in (**b**), the simulation has been terminated before the chemicals have diffused to the edge of the system so as to produce a circular, rather than square, arrangement of spots for display purposes.

organising method based on reaction-diffusion modelling to define the layout of the cone cells. This method generates a pseudo-hexagonal arrangement of cells without gaps or overlap and does not require a regular mosaic as a starting point. Alternatively, due to the modular nature of ERICA, the user can specify their own mosaic of cone cells simply by passing an array of horizontal and vertical positions, allowing other methods to be integrated.

Cone locations. Reaction-diffusion systems model a process by which two chemicals (U and V) each diffuse and react with one another. Such systems have been used to generate complex self-organising patterns, also known as Turing patterns[51], and have been offered as an explanation for the intricate pigmentation patterns formed in the coats of animals, such as leopards. The Gray-Scott reaction-diffusion model is one approach for generating these self-organising patterns and corresponds to the following irreversible chemical reactions that produce an inert product P:

$$U + 2V \rightarrow 3V \tag{1}$$

$$V \rightarrow P \tag{2}$$

The model computes the concentrations, u and v, of the two chemicals U and V at each pixel location, iteratively over a series of time steps using the following reaction-diffusion equations[52]:

$$\frac{\partial u}{\partial t} = D_u \nabla^2 u - uv^2 + f(1-u) \tag{3}$$

$$\frac{\partial v}{\partial t} = D_v \nabla^2 v + uv^2 - (f+k)v \tag{4}$$

where D_v is the diffusion coefficient for chemical V, D_u is the diffusion coefficient for chemical U, f is the rate at which chemical U is replenished (Eq. 1) and k is the rate at which the excess chemical V is removed (Eq. 2). To produce a pattern of spots (visible as the final concentration of V) the following parameters can be used: $D_u = 0.14$, $D_v = 0.06$, $f = 0.035$, $k = 0.065$ and an example of the output is given in Fig. 2. It is difficult to predict the pattern that will be produced for a given set of coefficients and some combinations will produce unstable outputs[52]. However, such failures are obvious in the output as lines or stripes or as a failure to produce a mosaic. Here we provide a recipe to produce cone layouts, based on extensive testing—changing D_v while keeping the other coefficients constant will produce round 'spots', so long as D_u is at least twice as large as D_v. After a stable pattern of spots is produced, a peak detection algorithm is then applied to the array (v) corresponding to the final time step, to identify the centres of each of the spots. Since the spots produced are well-defined and very high (and equal) contrast, these locations are easily and accurately extracted. It should be noted that the output of the reaction-diffusion model is only used to generate the pattern of cone locations; the widths, contrasts and intensity profiles of the spots are not used for image generation.

In the human eye, the spacing of the cone mosaic varies as a function of retinal eccentricity. The density of cones, ρ, can be approximated from histological measurements[53] (raw data taken from[54]) and the spacing, s, calculated as[55],

$$s = \left(\frac{2\rho}{\sqrt{3}}\right)^{-1/2}. \tag{5}$$

In principle, the spacing between the spots in the reaction-diffusion model can be tuned through selection of the appropriate diffusion coefficients, namely D_u and D_v. Furthermore, defining a radial variation in these coefficients can produce a foveated mosaic of spots (see Fig. 2b). In theory, this radial variation could be matched to histological data to produce a full fovea by tuning the radial profile of the diffusion coefficient, although this is likely to be computationally expensive. Care must be taken when selecting the diffusion coefficients, since the outcome (and its stability) will also depend on the coefficients f and k. The size of the system must also be taken into account to ensure that the sampling is sufficiently high for the spacing desired. For those reasons, it is more efficient to alter the cone spacing by simply re-scaling a well-sampled and stable cell mosaic. This is done within ERICA by generating a retinal mosaic (see Fig. 2a) over an extended bounding box, and scaling the resulting cone positions such that the average cone spacing (determined by nearest neighbour analysis) matches the histological data.

Cone reflectance profiles. Photoreceptors act as optical waveguides[56] and the number of guided modes (e.g. linearly polarised, LP, modes) depends on the physical properties of the cell. We consider each cone to have an intensity profile that can be approximated by a Gaussian function. Small cone cells close to the centre of the fovea (approximately $< 3°$) act as single mode waveguides, and so this Gaussian model is reasonable. At larger eccentricities, cone cells exhibit biomodal (beyond 7°) or multimodal (beyond 10°) profiles and so the Gaussian approximation will be a less accurate representation as the support of higher-order modes results in more complex intensity profiles[57]. If simulation of such profiles is required ERICA could be modified to include higher-order modes in addition to the fundamental (Gaussian) mode, but clearly the modal composition should be validated against human-derived images. Furthermore, a full model of the light propagation through cone cells, including interference effects, that predicts the reflectance profiles seen in real AOSLO images has recently been developed by Meadway and Sincich[44]. Alternative models such as these could be used alongside the ERICA framework and incorporated via a user-written function that is passed as an input to ERICA's sampling function, which would then call it.

Mosaic ground-truth data. An individual retinal mosaic is specified as an array of size N x M, where N is the number of cones in the mosaic and M is the number of parameters defining an individual cone's Gaussian reflectance profile. As described above, cone locations extracted from the reaction-diffusion model are scaled so that the spacing of the cells matches histological data for a given eccentricity. The sizes of the cells are specified as the diameter of the first (Gaussian) mode of the cell, which is calculated via the Marcuse equation[58] with the cell's radius determined from histological data[59,60]. Using the retinal mosaic module within ERICA it is possible to generate these data simply by specifying the retinal eccentricity and the cone locations are scaled accordingly. ERICA specifies the following parameters for each cone in the mosaic: (1) the vertical coordinate, μ_y, (2) the horizontal coordinate, μ_x, (3) the vertical width, σ_y, (4) the horizontal width, σ_x, (5) the orientation, θ, and (6) the reflectance, A. This parameter array represents the ground truth for the make-up of the synthetic retina. All parameters can be defined by the user but by default the positions are determined by the reaction-diffusion module, the width is determined using the Marcuse equation with cells being circular by default ($\sigma_x = \sigma_y$) and the reflectance values are randomly selected from a normal distribution based on histograms computed for real images. For cone reflectance, based on observations of real images, we additionally impose a low spatial frequency variation, and this approach has been previously described by others[41]. It is also possible to specify temporal variations in the light reflected back from individual cones, such as might be seen on long time scales due to bleaching or disc shedding or on short time scales due to stimulus-evoked intrinsic changes in intensity[61]. In this case, the reflectance is taken to be that at the time point when the simulated scanned beam reaches the cone.

Using this approach it is possible to define a retinal mosaic analytically as a sum of Gaussians for N cones,

$$R(x,y) = \sum_{i=1}^{N} A_i exp\left[-\left(a(x-\mu_{x,i})^2 + 2b(x-\mu_{x,i})(y-\mu_{y,i}) + c(y-\mu_{y,i})^2\right)\right], \tag{6}$$

where

$$a = \frac{cos^2(\theta_i)}{2\sigma_x^2} + \frac{sin^2(\theta_i)}{2\sigma_y^2}, \tag{7}$$

$$b = \frac{-sin(2\theta_i)}{4\sigma_x^2} + \frac{-sin(2\theta_i)}{4\sigma_y^2}, \tag{8}$$

$$c = \frac{sin^2(\theta_i)}{2\sigma_x^2} + \frac{cos^2(\theta_i)}{2\sigma_y^2}. \tag{9}$$

An analytic representation of the retina allows the AOSLO data capture, combining the scan pattern and the motion of the eye, to be replicated without introducing sampling errors. In the case where we wish to also

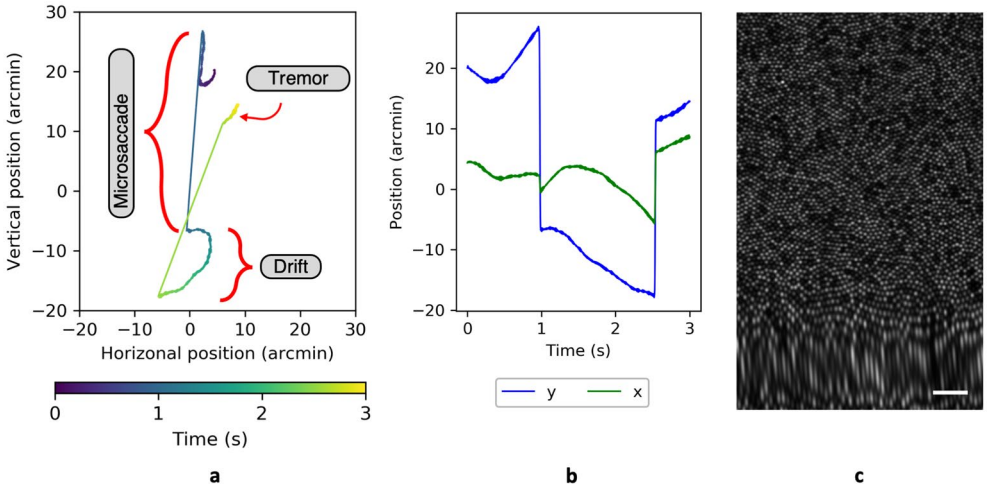

Figure 3. An example three-second eye movement trace, showing three microsaccades interspersed with periods of drift and tremor. Panel (**a**) shows the change in horizontal and vertical position over time, indicated by the colour gradient from blue ($t_{0\,s}$) to yellow ($t_{3\,s}$). Panel (**b**) shows the horizontal (x; green) and vertical (y; blue) components separately. A movie showing a synthetic retina undergoing this motion is given in the supplementary materials S1. Panel (**c**) shows frame 30 from Supplementary Video S3 (diffraction-limited resolution), corresponding to the large and mainly vertical microsaccade at approximately 1 s. This is visible as a stretch at the bottom of the frame. The scale bar represents 10 arcminutes.

replicate the effects of optical aberrations on the captured data it becomes necessary to perform additional computations, as described in a later section.

Eye movement simulation. To introduce intra- and inter-frame motion, ERICA includes a biologically plausible model of the eye movements made during fixation. Fixational eye movements are typically categorized as either tremor, drift, or microsaccades[62,63], each of which is incorporated into our model. The parameters defining each of these types of movement have default values, but these can be adjusted. Microsaccades are modeled as a ballistic motion with the peak velocity proportional to the amplitude of the displacement in log space[64]. The amplitude of each microsaccade is randomly selected from a normal distribution (default: mean of 30 arcmin and a standard deviation of 3 arcmin[64]). In the model, the occurrences of microsaccades are pre-determined by randomly selecting time intervals from a normal distribution (default: mean of 1.0 seconds and a standard deviation of 0.5 s). Between the microsaccades, eye movement data are generated by combining tremor with drift. Tremor constitutes a high temporal frequency motion and is simulated using a random motion with a temporal bandpass filter[20], which we model with a Gaussian profile (default: a centre frequency drawn from a normal distribution with a mean of 65 Hz and a standard deviation of 7.5 Hz and a bandwidth drawn from a normal distribution with a mean of 25 Hz and a standard deviation of 2.5 Hz). The amplitude of tremor is typically very small[20] and we model this by selecting it from a normal distribution with a low mean (default: a mean of 0.08 arcmin and a standard deviation of 0.03 arcmin[20]). Drift is a low temporal frequency motion modeled by a $1/f^2$ spectrum, over which tremor is superimposed[65]. The amplitude for each period of drift is randomly selected from a normal distribution (default: a mean of 6.5 arcmin and a standard deviation of 0.7 arcmin). Human participants typically maintain fixation around a central point, however there has been debate over whether microsaccades or drifts are responsible for corrective motions to bring fixation back on target[66]. The directions of microsaccades in the ERICA model are constrained to ensure that the simulated fixational eye movement maintains a locus about the centre of the synthetic retinal mosaic, simply so that blank spaces from the edge of the mosaic are not introduced into the synthetic frames. An example eye movement trace and an image containing intra-frame motion distortion are shown in Fig. 3. The parameters outlined above can be adjusted by the user to generate eye movements with particular characteristics and, for example, to test how well those characteristics are recovered by a motion extraction algorithm.

Image generation. The PSF of a double-pass confocal microscope, such as the AOSLO, is defined as the autocorrelation of the single pass PSF[67–69]. However, the retina is not a uniform diffuse reflector. To replicate the process of data capture from the retinal mosaic a series of computations must be performed as described in Fig. 4. Briefly: (1) the input PSF should be multiplied by the synthetic retinal mosaic, giving the reflectance profile of the illuminated patch of retina; (2) the illuminated patch should be convolved with the output PSF to replicate reimaging it on to the pinhole; and (3) the image formed at the pinhole should be multiplied by the transmission function defining the pinhole. The sum of this final image is the signal that would be received at the detector at a particular moment in time. Within the image generation module of ERICA, the pupil diameter, imaging wavelength and pinhole size can be defined by the user to match the configuration of their system. Pupil apodization can also be applied to simulate the optical Stiles Crawford effect by specifying a Gaussian

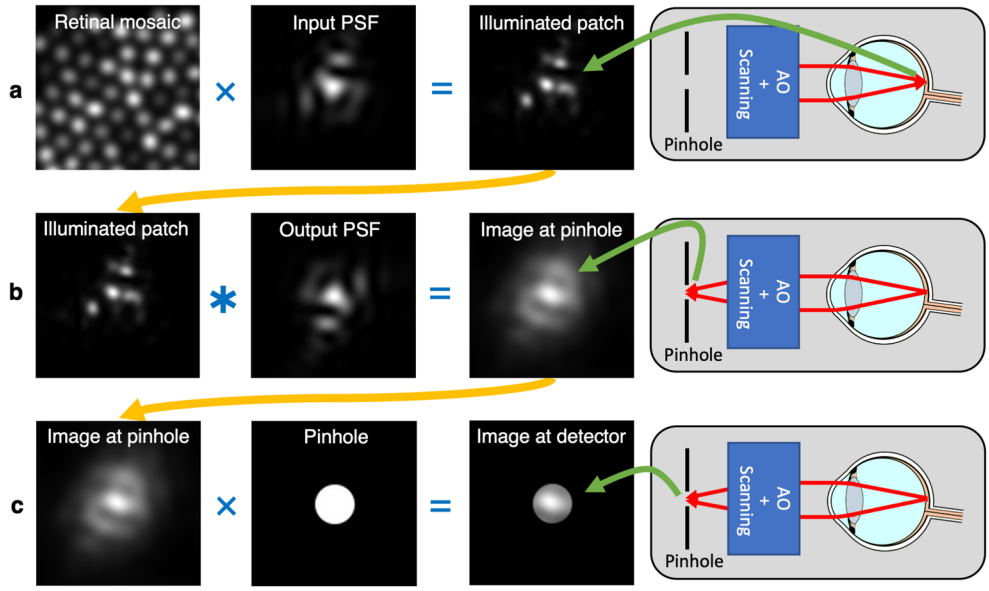

Figure 4. The steps to replicate data capture through a scanned confocal retinal imaging system, such as an AOSLO, which are shown schematically on the right. Step (**a**) shows the illumination of a patch of retina by a point source directed through the system and the eye's optics (Input PSF), which may be degraded by the eyes optics. This step is represented by the multiplication of the reflectance profile of the retinal patch and the input PSF. Step (**b**) shows the process of re-imaging the illuminated patch of the retina through the system and the eye's optics (Output PSF) to the plane where the pinhole is located. This step is computed as a convolution of the illuminated patch of retina with the output PSF, noting that the wavefront is inverted in the outward direction (this operation is equivalent to the cross-correlation of the illuminated patch with the input PSF). Step (**c**) shows the transmission of the final image through the pinhole to give the image that would be seen by the detector. This step is computed as the multiplication of the image at the pinhole by the transmission function of the pinhole, which we consider to be 1 inside and 0 outside a uniform disc. The final step is summation of the image formed behind the pinhole, to replicate integration by the detector (neglecting noise), giving the signal at the current position in the scan, $I(x,y)$. Steps (**a–c**) are repeated for every sample location as the beam is raster scanned over a region of the retina. Note that while this model allows for a different input and output PSF, in practice we assume that the input and output PSFs are the same (the PSF of the eye, which includes diffraction effects).

pupil transmission function. It is possible to perform these computations by adding terms and operations to Eq. (6), such as by using the extended Nijboer–Zernike approach to approximate the analytic PSF[70,71]. However, this adds complexity and that method also assumes some wavefront symmetry[70]. It is therefore simpler to use a numerical approach and use the discrete Fourier transform to render images with residual aberration (which can vary from frame-to-frame), taking care to oversample appropriately to avoid sampling artefacts. Finer sampling resolutions produce images closer to an analytic solution but with added computational expense. To test this, a set of retinal mosaics was generated with a range of sampling resolutions and this set was used to generate synthethic frames with a realistic residual wavefront error (based on a publicly available data set[49]). A sampling resolution at least two times higher than the desired resolution of the resulting frame should be used and the test showed that a factor increase in sampling resolution of ≥ 4 produces almost identical outputs, defined by the sum squared difference from the a high ($10\times$) sampling resolution image.

Following the steps in Fig. 4, ERICA generates a data stream similar to that which would be obtained from an AOSLO (or any scanned confocal system, such as a standard scanning laser ophthalmoscope). A raw image frame is then generated from the data stream using knowledge of the positions on the scanning mirrors at each time point. To generate a motion free reference image, a data stream is generated by replicating the spatio-temporal sampling of the raster scanning system—a sinuoidal horizontal fast scan and an approximately linear vertical slow scan. To generate a realistic sampling pattern during fixation, eye movement would be added to the raster scan pattern, as described above. The scan pattern used in the simulation can be defined by the user and we use a model of the scan pattern in our AOSLO[72]. This is a sinusoidal fit to the fast scanner motion and a sawtooth function for the slow scanner, which includes the flyback region. Custom specification of the scan pattern allows the use of ERICA to investigate the effects of desinusoiding errors or instability in the scanning system. This data stream is then typically processed using the same desinusoiding algorithm that is applied to real, human-derived intensity data. All images presented in this paper have been desinusoided for clarity. Noise can be added to images after following the steps in Fig. 4 and we assume that this is an additive process. By default, noise is generated by selecting values from a normal distribution, the mean and standard deviation of which are matched to the noise

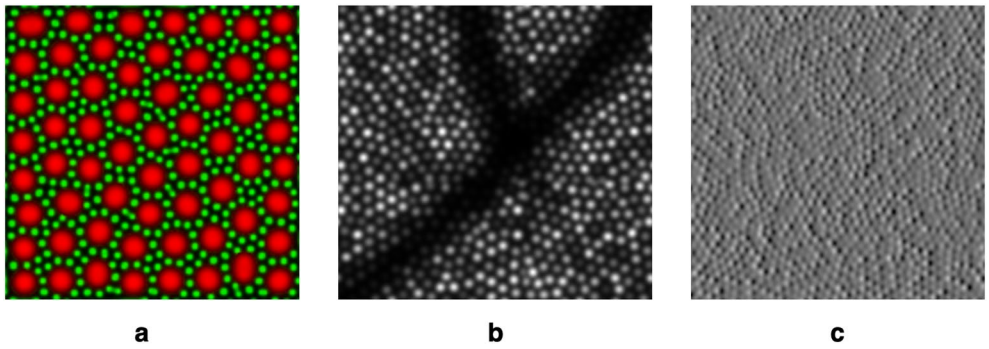

Figure 5. Examples to show possible extensions to ERICA. Panel (**a**) shows the output of a reaction-diffusion model that includes a rod mosaic (shown in green), self-organising itself between cone cells (shown in red). Panel (**b**) shows simulation of the shadow caused by a retinal blood vessel overlaid on a synthetic cone mosaic image. Panel (**c**) shows the approximation of a split-detection image, formed by reproducing the intensity profile of a photoreceptor inner segment as it typically appears in such images. This is assumed to be a differential image of a mosaic of Gaussian profiles.

characteristics of our AOSLO. This array of noise samples must be either added before desinusoiding, or must must undergo the same desinusioding procedure before being added to the image.

Comparison to human cone mosaics. To verify that the packing geometry of the synthetic mosaics is representative of the cone mosaic of typical human participants, we made comparisons at a set of matched retinal eccentricities. Images were collected from seven human participants at three eccentricities (1.5°, 3.0° and 6.0°) as part of another study, which received approval from the University of Oxford Research Ethics Committee (reference: R57315/RE001). The study was conducted in accordance with the Declaration of Helsinki and informed consent was obtained from all participants. We did not attempt imaging in the centre of the fovea as without dilation the cone cells would not be well-resolved and the estimates of cone packing would likely be unreliable. Cone positions were manually labeled in both the human and synthetic images by five experienced graders (one of whom was the first author) and metric data were averaged across graders. Manual grading of the synthetic images, rather than using the ground-truth data, ensured that apparent changes in the packing geometry caused by weakly reflecting cones, which could be missed by a human grader, were captured for a fair comparison to the human images. Presentations of synthetic and human images at different eccentricities were randomly interleaved and the grader was blind to the type of image displayed on any given trial. Using synthetic images, comparisons can also be made between the manual labels from human observers with the ground-truth cone locations to understand how the accuracy of human observers impacts cone density measurements. This will be presented in a future publication.

Extensions. The current implementation of ERICA simulates the cone mosaic as visualised through a scanning confocal imaging system. This model is limited to a 2D representation of the retina and residual wavefront aberrations are added based on a 2D model of the PSF. Modelling the effects of changing the axial resolution would be possible with an extension of the simulation to perform image generation based on the 3D PSF and with a cellular-resolution model of the retina in depth. A model of the light propagation through the cone cells, such as that developed by Meadway and Sincich could also be incorporated into the ERICA framework. ERICA is designed in a modular manner with example Jupyter notebooks demonstrating its use. It is anticipated that users who wish to develop additional features will create their own functions to be called by existing functions within ERICA (e.g. a function defining the reflectance profile of a cone cell based on ray tracing would be called by ERICA's sampling function) or to create modified versions of the existing inputs to ERICA (e.g. a different method for generating cone locations would add those locations to a retinal mosaic parameter array to be read by the sampling function).

A simpler approach to generating images of other retinal structures, either in depth or within the same retinal layer, is to estimate the intensity profile of those structures. Examples are given in Fig. 5. Figure 5a shows that rod cells could be added into the images. This is possible by modification of the reaction-diffusion model to self-organise smaller spots in amongst a mosaic of larger spots. This will be presented in more detail in a future publication. Another structure that is likely to be of interest is the retinal vasculature, particularly as it causes shadows in the images of the photoreceptor mosaic that may affect image registration. A basic approximation, as shown in Fig. 5, would be to introduce a darkened region, either during or after image generation. Non-confocal imaging geometries detect multiply-scattered light returning from structures other than the photoreceptors. Accurately replicating this process would require a ray-tracing model, which is not currently included within ERICA. However, it is possible to generate approximate non-confocal synthetic images using an estimate of the intensity profile of the structures observed in these images. As an example, split-detection imaging is a useful tool for detecting photoreceptor inner segments even when the reflectance in the confocal image is low. Methods for automatic detection of cells in these images and for montaging have been developed[30,45,46]. An example of a synthetic split-detection image is given in Fig. 5c, which has been generated by approximating the expected

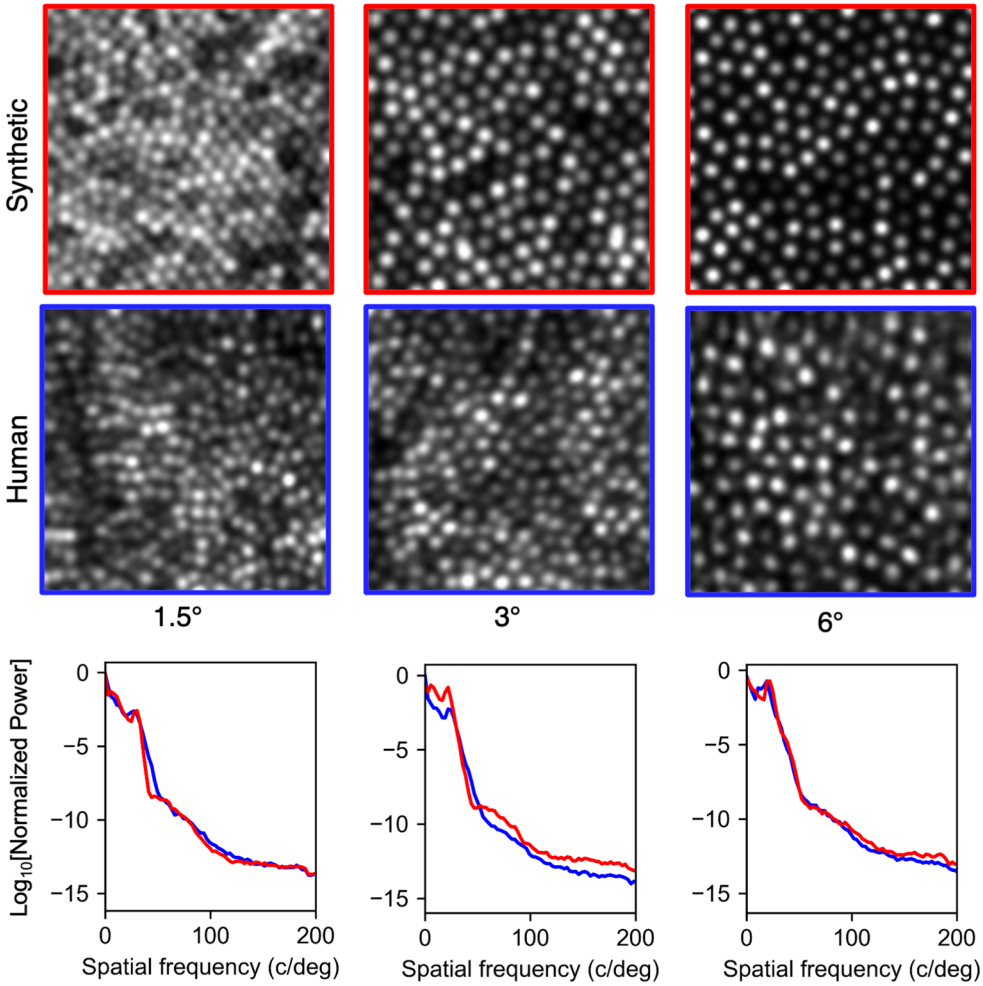

Figure 6. Example images from a human participant collected from our AOSLO[72] and synthetic images generated for the same eccentricities—1.5°, 3° and 6°. The synthetic images are generated with no noise, residual aberration or eye motion and so have better image quality than the human images. The pupil diameter for the simulation was set to 5 mm, which is approximately the same as the diameter of the human participant's pupil during image capture. The synthetic images were spatially scaled to account for the human participant's axial length being different to that of the model. The synthetic images presented here have been intensity-scaled linearly and then clipped such that the number of pixels that are saturated is the same as in the human images (approximately 10% for the detector gain that was used during imaging). The lower panels show the power spectra of the synthetic images in red and the human images in blue. In each spectrum, the peak below 50 c/deg corresponds to the modal spacing of the cones (Yellot's ring).

intensity profile as the differential of a Gaussian distribution. The example given shows differentiation in the horizontal direction, replicating a left-right split-detection geometry. Similarly, a model image of the retinal pigment epithelium (RPE), is theoretically possible, such as by using the reaction-diffusion model to generate cell locations and using an estimated dark-field RPE cell intensity profile. By updating the ground-truth retina, ERICA allows emulation of the AOSLO image capture from such structures, including realistic eye motion, aberrations and noise.

Results

Examples of synthetic images produced by ERICA are given in Fig. 6 and example movies showing temporal variation in eye motion and wavefront error, based on real aberration measurements[49], are additionally given in the supplementary materials S1 (scale bar represents 10 arcminutes). The appearances of the synthetic and human images are convincingly similar. The lower panels in Fig. 6 show the normalised power spectra of the images in the top panels. The spacing of the cells is matched within ERICA based on the retinal eccentricity and the axial length of the participant's eye and so the modal spacing (the peak in the power spectrum at < 50 c/deg, which corresponds to Yellott's ring) is also matched. Therefore, the scale of the spatial frequency axis is matched between human and synthetic images, but these spectra also show similar profiles.

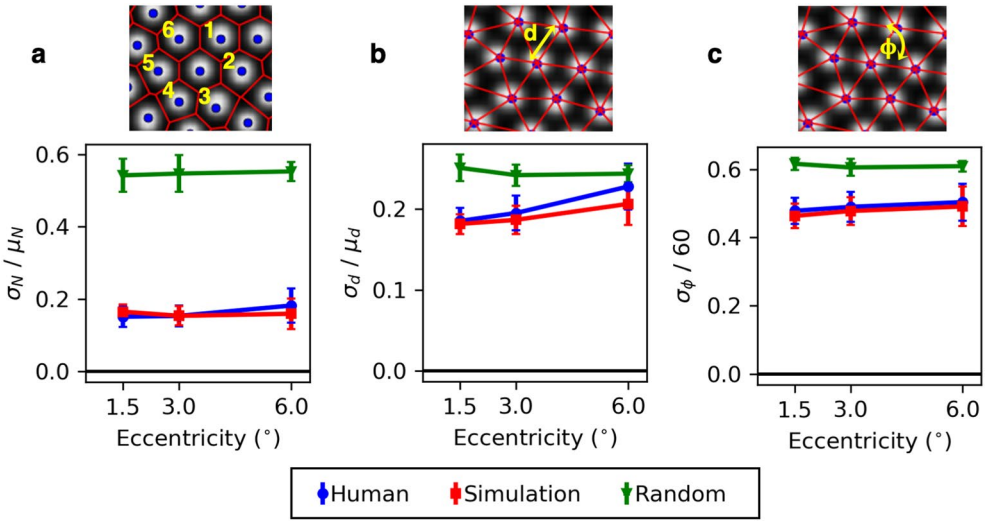

Figure 7. Three metrics of packing regularity (eccentricities the same as in Fig. 6) are compared for human-derived images and for synthetic images generated either using reaction-diffusion modelling or by randomly placing cells (i.e completely irregular packing with the same number of cones). For each metric the standard deviation is used as a measure of variability, normalised by the mean. A value of zero is equivalent to a completely regular mosaic. Each data point represents the mean, and each error bar the standard deviation, calculated over 27 human or 12 synthetic images. Images were manually labeled by five experienced markers and the data plotted here represent the average across markers. The panels show the variability in (**a**) the number of nearest neighbours, (**b**) the nearest neighbour distance and (**c**) the vertex angle between nearest neighbours.

We analysed the packing geometry of the synthetic and human cone mosaics using three summary statistics (see Fig. 7): (a) the number of nearest neighbours to a cone (i.e. the number of Voronoi regions associated with an individual cone), (b) the first nearest neighbour distance, and (c) the vertex angle between the nearest neighbour and that neighbour's closest common neighbour. For each image and at each eccentricity, the three parameters were calculated for 100 randomly selected cones. The metric value for an individual image was defined as the standard deviation of the calculated parameter, normalised by the mean (or 60° in the case of the vertex angle measurement). This gives a measure of variance that would be zero for a perfectly regular hexagonal matrix. Data were then averaged across images (across participants) for each eccentricity. To help interpret the scale of these values, a cone mosaic was generated with a random packing arrangement (but the same number of cells) and without any constraint on the minimum spacing between cells. This provides an estimate of the metric value for a completely irregular mosaic. Fig. 7 shows a comparison between the packing regularity of synthetic and human images. These results show that ERICA produces cone packing geometries close to those that we observe in the human retina through our AOSLO[72]. It is notable that agreement between the human images and the simulation is reduced at the largest eccentricity when considering the variation in nearest neighbour distance. We suggest this is likely due to the presence of rods in the human images that are absent in the simulation, which may have affected the manual labels[73].

Discussion

Adaptive optics retinal imaging is rapidly emerging as an important tool for studying the eye[4,5,74]. A significant bottleneck in the wider use of these systems is the need for robust, accurate and automated image processing and analysis methods to generate good quality images and extract reliable data. The use of a generative model for algorithm testing is not new, but its implementation in the AO retinal imaging field has so far been limited[40–43]. ERICA uses a self-organising approach that produces mosaics with similar packing geometries to human-derived images. An advantage to using reaction-diffusion modelling is that cells naturally arrange themselves without gaps. Self-organising approaches has been suggested previously as a model for retinal sampling[50,75], using different methods. However, the reaction-diffusion approach does offer another advantage, as demonstrated in Fig. 5a, since multiple interleaved mosaics can be generated that both self-avoid and avoid each other.

Previous approaches for generating synthetic images have performed image generation in a similar way to the method we have developed. These approaches model cone cells as Gaussian intensity profiles and convolve the retinal image with the PSF[41,42]. ERICA fully replicates the image generation process, as shown in Fig. 4, and we include a step to separately specify the input and output PSFs to include, for example, differences in the input and output pupil functions.

Beddgood and Metha developed a model to add motion distortions to simulated images using a different approach to ours[22]. They used a flood-illumination image as the ground truth and the rows of this were re-indexed based on an eye movement trace that had been captured from a human participant. By contrast, ERICA includes a model of biologically plausible eye motion allowing motion extraction accuracy to be characterised relative to the specific measures of interest, such as the amplitude or velocity of microsaccades or drifts. These

movements are introduced during a replication of the non-linear scan pattern and this can additionally be performed analytically to avoid sampling artefacts. De-sinusoiding is a necessary step for processing images from an AOSLO (and from a scanning laser ophthalmoscope without AO correction). Dependent on the voltage ramp of the galvo, it is also possible that the slow scanner is not perfectly linear. By including a model for the scan pattern we can test the effect that non-linearities have on motion extraction, in terms of both the remapping of spatial information and the temporal sampling of eye motion, which ideally would be linear.

In this paper we have presented a complete end-to-end simulation of AOSLO data capture that allows generation of large data sets of retinal images with ground-truth data. The properties of these synthetic images, such as residual wavefront error, noise and motion distortion, as well as the content (e.g. by varying retinal eccentricity or changing cone reflectances) can be specified. ERICA has been designed to be flexible so that users can adjust any parameters in order to replicate their own system. Any part of the simulation could be extended or modified by interacting with the Python code, and we envisage new additions for public release would be developed via a branch, which would then be moderated before being merged with the master branch. Images can be generated over a user-defined field of view (noting the limitations of anisoplanatism beyond 2°) and eccentricity (noting the lack of rod cells in the current implementation and the limitations of the single-mode Gaussian model). Finally, we openly share ERICA with the community and make a number of synthetic data sets available. We hope that this will be a useful contribution to the field to further the design, development and testing of reliable automated image processing and analysis methods for adaptive optics retinal imaging.

Data availability

ERICA is available to the community via GitHub (https://github.com/LauraKateYoung/ERICA) under the GNU GPL3 license. Example data sets are also available for use (https://data.ncl.ac.uk/authors/Laura_Young/9553463). The reference implementation is available via Zenodo (https://doi.org/10.5281/zenodo.4767154).

Received: 8 February 2021; Accepted: 4 May 2021
Published online: 27 May 2021

References

1. Godara, P., Dubis, A. M., Roorda, A., Duncan, J. L. & Carroll, J. Adaptive optics retinal imaging: Emerging clinical applications. *Optom. Vis. Sci.* **87**, 930–941. https://doi.org/10.1097/OPX.0b013e3181ff9a8b (2010).
2. Carroll, J., Kay, D. B., Scoles, D., Dubra, A. & Lombardo, M. Adaptive optics retinal imaging: Clinical opportunities and challenges. *Curr. Eye Res.* **38**, 709–721. https://doi.org/10.3109/02713683.2013.784792 (2013).
3. Roorda, A. & Duncan, J. L. Adaptive optics ophthalmoscopy. *Annu. Rev. Vis. Sci.* **1**, 19–50. https://doi.org/10.1146/annurev-vision-082114-035357 (2015).
4. Morgan, J. I. W. The fundus photo has met its match: Optical coherence tomography and adaptive optics ophthalmoscopy are here to stay. *Ophthal. Physiol. Opt.* **36**, 218–239. https://doi.org/10.1111/opo.12289 (2016).
5. Marcos, S. *et al.* Vision science and adaptive optics, the state of the field. *Vis. Res.* **132**, 3–33. https://doi.org/10.1016/j.visres.2017.01.006 (2017).
6. Georgiou, M. *et al.* Adaptive optics imaging of inherited retinal diseases. *Br. J. Ophthalmol.* **102**, 1028–1035. https://doi.org/10.1136/bjophthalmol-2017-311328 (2018).
7. Roorda, A. Adaptive optics for studying visual function: A comprehensive review. *J. Vis.* **11**, 6–6. https://doi.org/10.1167/11.5.6 (2011).
8. Sun, L. W. *et al.* Assessing photoreceptor structure in retinitis pigmentosa and usher syndrome. *Invest. Opthal. Vis. Sci.* **57**, 2428. https://doi.org/10.1167/iovs.15-18246 (2016).
9. Sun, L. W. *et al.* Multimodal imaging of photoreceptor structure in choroideremia. *PLoS ONE* **11**, e0167526. https://doi.org/10.1371/journal.pone.0167526 (2016).
10. Roorda, A. *et al.* Adaptive optics scanning laser ophthalmoscopy. *Opt. Express* **10**, 405. https://doi.org/10.1097/OPX.0b013e3181ff9a8b0 (2002).
11. Yang, Q., Arathorn, D. W., Tiruveedhula, P., Vogel, C. R. & Roorda, A. Design of an integrated hardware interface for AOSLO image capture and cone-targeted stimulus delivery. *Opt. Express* **18**, 17841. https://doi.org/10.1097/OPX.0b013e3181ff9a8b1 (2010).
12. Zhong, Z., Petrig, B. L., Qi, X. & Burns, S. A. In vivo measurement of erythrocyte velocity and retinal blood flow using adaptive optics scanning laser ophthalmoscopy. *Opt. Express* **16**, 12746. https://doi.org/10.1364/OE.16.012746 (2008).
13. Jonnal, R. S. *et al.* In vivo functional imaging of human cone photoreceptors. *Opt. Express* **15**, 16141. https://doi.org/10.1097/OPX.0b013e3181ff9a8b3 (2007).
14. Grieve, K. & Roorda, A. Intrinsic signals from human cone photoreceptors. *Invest. Opthalmol. Vis. Sci.* **49**, 713. https://doi.org/10.1167/iovs.07-0837 (2008).
15. Raghunandan, A., Frasier, J., Poonja, S., Roorda, A. & Stevenson, S. B. Psychophysical measurements of referenced and unreferenced motion processing using high-resolution retinal imaging. *J. Vis.* **8**, 14–14. https://doi.org/10.1097/OPX.0b013e3181ff9a8b5 (2008).
16. Song, S. *et al.* Fixational eye movements for normal and strabismic amblyopic observers. *J. Vis.* **10**, 456–456. https://doi.org/10.1097/OPX.0b013e3181ff9a8b6 (2010).
17. Kumar, G. & Chung, S. T. L. Characteristics of fixational eye movements in people with macular disease. *Invest. Opthalmol. Vis. Sci.* **55**, 5125. https://doi.org/10.1167/iovs.14-14608 (2014).
18. Ratnam, K., Domdei, N., Harmening, W. M. & Roorda, A. Benefits of retinal image motion at the limits of spatial vision. *J. Vis.* **17**, 30. https://doi.org/10.1097/OPX.0b013e3181ff9a8b8 (2017).
19. Agaoglu, M. N., Sheehy, C. K., Tiruveedhula, P., Roorda, A. & Chung, S. T. L. Suboptimal eye movements for seeing fine details. *J. Vis.* **18**, 8. https://doi.org/10.1097/OPX.0b013e3181ff9a8b9 (2018).
20. Bowers, N. R., Boehm, A. E. & Roorda, A. The effects of fixational tremor on the retinal image. *J. Vis.* **19**, 8. https://doi.org/10.3109/02713683.2013.7847920 (2019).
21. Salmon, A. E. *et al.* An automated reference frame selection (ARFS) algorithm for cone imaging with adaptive optics scanning light ophthalmoscopy. *Transl. Vis. Sci. Technol.* **6**, 9. https://doi.org/10.1167/tvst.6.2.9 (2017).
22. Bedggood, P. & Metha, A. De-warping of images and improved eye tracking for the scanning laser ophthalmoscope. *PLoS ONE* **12**, e0174617. https://doi.org/10.3109/02713683.2013.7847922 (2017).
23. Zhang, M. *et al.* Strip-based digital image registration for distortion minimization and robust eye motion measurement from scanned ophthalmic imaging systems. *Biomed. Opt. Express* **12**, 2353. https://doi.org/10.1364/BOE.418070 (2021).

24. Cooper, R. F., Lombardo, M., Carroll, J., Sloan, K. R. & Lombardo, G. Methods for investigating the local spatial anisotropy and the preferred orientation of cones in adaptive optics retinal images. *Vis. Neurosci.* **33**, E005. https://doi.org/10.3109/02713683.2013.7847924 (2016).
25. Mujat, M., Ferguson, R. D., Iftimia, N. & Hammer, D. X. Compact adaptive optics line scanning ophthalmoscope. *Opt. Express* **17**, 10242. https://doi.org/10.3109/02713683.2013.7847925 (2009).
26. Dubra, A. & Harvey, Z. Registration of 2D Images from Fast Scanning Ophthalmic Instruments. In *Biomedical Image Registration* Vol. 6204 (eds Hutchison, D. *et al.*) 60–71 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2010). https://doi.org/10.1007/978-3-642-14366-3_6.
27. Ramaswamy, G. & Devaney, N. Pre-processing, registration and selection of adaptive optics corrected retinal images. *Ophthal. Physiol. Opt.* **33**, 527–539. https://doi.org/10.1111/opo.12068 (2013).
28. Chen, H., He, Y., Wei, L., Li, X. & Zhang, Y. Automatic dewarping of retina images in adaptive optics confocal scanning laser ophthalmoscope. *IEEE Access* **7**, 59585–59599. https://doi.org/10.1109/ACCESS.2019.2914463 (2019).
29. Xue, B., Choi, S. S., Doble, N. & Werner, J. S. Photoreceptor counting and montaging of en-face retinal images from an adaptive optics fundus camera. *J. Opt. Soc. Am. A* **24**, 1364. https://doi.org/10.3109/02713683.2013.7847929 (2007).
30. Chen, M. *et al.* Multi-modal automatic montaging of adaptive optics retinal images. *Biomed. Opti. Express* **7**, 4899. https://doi.org/10.1364/BOE.7.004899 (2016).
31. Li, K. Y. & Roorda, A. Automated identification of cone photoreceptors in adaptive optics retinal images. *J. Opt. Soc. Am. A* **24**, 1358. https://doi.org/10.1146/annurev-vision-082114-0353571 (2007).
32. Chui, T. Y., Song, H. & Burns, S. A. Adaptive-optics imaging of human cone photoreceptor distribution. *J. Opt. Soc. Am. A* **25**, 3021. https://doi.org/10.1146/annurev-vision-082114-0353572 (2008).
33. Garrioch, R. *et al.* Repeatability of in vivo parafoveal cone density and spacing measurements. *Optom. Vis. Sci.* **89**, 632–643. https://doi.org/10.1097/OPX.0b013e3182540562 (2012).
34. Mohammad, F., Ansari, R., Wanek, J. & Shahidi, M. Frequency-based local content adaptive filtering algorithm for automated photoreceptor cell density quantification. In *19th IEEE International Conference on Image Processing* 2325–2328 (IEEE, Orlando, FL, USA, 2012). https://doi.org/10.1109/ICIP.2012.6467362.
35. Chiu, S. J. *et al.* Automatic cone photoreceptor segmentation using graph theory and dynamic programming. *Biomed. Opt. Exp.* **4**, 924. https://doi.org/10.1364/BOE.4.000924 (2013).
36. Cunefare, D. *et al.* Open source software for automatic detection of cone photoreceptors in adaptive optics ophthalmoscopy using convolutional neural networks. *Sci. Rep.* **7**, 6620. https://doi.org/10.1038/s41598-017-07103-0 (2017).
37. Cunefare, D. *et al.* Deep learning based detection of cone photoreceptors with multimodal adaptive optics scanning light ophthalmoscope images of achromatopsia. *Biomed. Opt. Express* **9**, 3740. https://doi.org/10.1364/BOE.9.003740 (2018).
38. Sheehy, C. K., Tiruveedhula, P., Sabesan, R. & Roorda, A. Active eye-tracking for an adaptive optics scanning laser ophthalmoscope. *Biomed. Opt. Express* **6**, 2412. https://doi.org/10.1364/BOE.6.002412 (2015).
39. Tam, J. Speed quantification and tracking of moving objects in adaptive optics scanning laser ophthalmoscopy. *J. Biomed. Opt.* **16**, 036002. https://doi.org/10.1146/annurev-vision-082114-0353578 (2011).
40. Turpin, A., Morrow, P., Scotney, B., Anderson, R. & Wolsley, C. Automated identification of photoreceptor cones using multi-scale modelling and normalized cross-correlation. In *Image analysis and processing: ICIAP 2011* Vol. 6978 (eds Maino, G. & Foresti, G. L.) 494–503 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2011). https://doi.org/10.1007/978-3-642-24085-0_51.
41. Mariotti, L. & Devaney, N. Performance analysis of cone detection algorithms. *J. Opt. Soc. Am. A* **32**, 497. https://doi.org/10.1111/opo.122890 (2015).
42. Li, W. *et al.* Quality improvement of adaptive optics retinal images using conditional adversarial networks. *Biomed. Opt. Express* **11**, 831. https://doi.org/10.1364/BOE.380224 (2020).
43. Cooper, R. F. *et al.* Effects of intraframe distortion on measures of cone mosaic geometry from adaptive optics scanning light ophthalmoscopy. *Transl. Vis. Sci. Technol.* **5**, 10. https://doi.org/10.1167/tvst.5.1.10 (2016).
44. Meadway, A. & Sincich, L. C. Light reflectivity and interference in cone photoreceptors. *Biomed. Opt. Express* **10**, 6531–6554 (2019).
45. Cunefare, D. *et al.* Automatic detection of cone photoreceptors in split detector adaptive optics scanning light ophthalmoscope images. *Biomed. Opt. Express* **7**, 2036. https://doi.org/10.1364/BOE.7.002036 (2016).
46. Davidson, B. *et al.* Automatic cone photoreceptor localisation in healthy and Stargardt afflicted retinas using deep learning. *Sci. Rep.* **8**, 7911. https://doi.org/10.1038/s41598-018-26350-3 (2018).
47. van der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy array: A structure for efficient numerical computation. *Comput. Sci. Eng.* **13**, 22–30. https://doi.org/10.1109/MCSE.2011.37 (2011).
48. Oliphant, T. E. *Guide to NumPy* (Continuum Press, Austin, Tex., 2015).
49. Jarosz, J. *et al.* High temporal resolution aberrometry in a 50-eye population and implications for adaptive optics error budget. *Biomed. Opt. Exp.* **8**, 2088. https://doi.org/10.1364/BOE.8.002088 (2017).
50. Cottaris, N. P., Jiang, H., Ding, X., Wandell, B. A. & Brainard, D. H. A computational-observer model of spatial contrast sensitivity: effects of wave-front-based optics, cone-mosaic structure, and inference engine. *J. Vis.* **19**, 8. https://doi.org/10.1167/19.4.8 (2019).
51. Turing, A. The chemical basis of morphogenesis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **237**, 37–72. https://doi.org/10.1111/opo.122898 (1952).
52. Pearson, J. E. Complex patterns in a simple system. *Science* **261**, 189–192. https://doi.org/10.1126/science.261.5118.189 (1993).
53. Curcio, C. A., Sloan, K. R., Kalina, R. E. & Hendrickson, A. E. Human photoreceptor topography. *J. Comp. Neurol.* **292**, 497–523. https://doi.org/10.1016/j.visres.2017.01.0060 (1990).
54. Watson, A. B. A formula for human retinal ganglion cell receptive field density as a function of visual field location. *J. Vis.* **14**, 15. https://doi.org/10.1016/j.visres.2017.01.0061 (2014).
55. Williams, D. R. & Coletta, N. J. Cone spacing and the visual resolution limit. *J. Opt. Soc. Am. A* **4**, 1514. https://doi.org/10.1016/j.visres.2017.01.0062 (1987).
56. Enoch, J. M. Optical properties of the retinal receptors. *J. Opt. Soc. Am.* **53**, 71. https://doi.org/10.1364/JOSA.53.000071 (1963).
57. Liu, Z., Kocaoglu, O. P., Turner, T. L. & Miller, D. T. Modal content of living human cone photoreceptors. *Biomed. Opt. Express* **6**, 3378. https://doi.org/10.1364/BOE.6.003378 (2015).
58. Marcuse, D. Loss analysis of single-mode fiber splices. *Bell Syst. Tech. J.* **56**, 703–718. https://doi.org/10.1002/j.1538-7305.1977.tb00534.x (1977).
59. Polyak, S. *The Retina* (University of Chicago Press, Chicago, Illinois, 1941).
60. Tyler, C. W. Analysis of visual modulation sensitivity II Peripheral retina and the role of photoreceptor dimensions. *J. Opt. Soc. Am. A* **2**, 393. https://doi.org/10.1016/j.visres.2017.01.0066 (1985).
61. Cooper, R. F., Brainard, D. H. & Morgan, J. I. W. Optoretinography of individual human cone photoreceptors. *Opt. Express* **28**, 39326. https://doi.org/10.1016/j.visres.2017.01.0067 (2020).
62. Martinez-Conde, S., Macknik, S. L. & Hubel, D. H. The role of fixational eye movements in visual perception. *Nat. Rev. Neurosci.* **5**, 229–240. https://doi.org/10.1016/j.visres.2017.01.0068 (2004).
63. Rucci, M. & Poletti, M. Control and functions of fixational eye movements. *Annu. Rev. Vis. Sci.* **1**, 499–518. https://doi.org/10.1146/annurev-vision-082114-035742 (2015).
64. Hauperich, A.-K., Young, L. K. & Smithson, H. E. What makes a microsaccade? A review of 70 years research prompts a new detection method. *J. Eye Movement Res.*https://doi.org/10.16910/JEMR.12.6.13 *(2020).*

65. Ezenman, M., Hallett, P. & Frecker, R. Power spectra for ocular drift and tremor. *Vis. Res.* **25**, 1635–1640. https://doi.org/10.1016/0042-6989(85)90134-8 (1985).
66. Cherici, C., Kuang, X., Poletti, M. & Rucci, M. Precision of sustained fixation in trained and untrained observers. *J. Vis.* **12**, 31–31. https://doi.org/10.1136/bjophthalmol-2017-3113282 (2012).
67. Artal, P., Marcos, S., Williams, D. R. & Navarro, R. Odd aberrations and double-pass measurements of retinal image quality. *J. Opt. Soc. Am. A* **12**, 195. https://doi.org/10.1136/bjophthalmol-2017-3113283 (1995).
68. Venkateswaran, K., Roorda, A. & Romero-Borja, F. Theoretical modeling and evaluation of the axial resolution of the adaptive optics scanning laser ophthalmoscope. *J. Biomed. Opt.* **9**, 132. https://doi.org/10.1136/bjophthalmol-2017-3113284 (2004).
69. Webb, R. H. Confocal optical microscopy. *Rep. Prog. Phys.* **59**, 427–471. https://doi.org/10.1136/bjophthalmol-2017-3113285 (1996).
70. Janssen, A. J. E. M. Extended Nijboer–Zernike approach for the computation of optical point-spread functions. *J. Opt. Soc. Am. A* **19**, 849. https://doi.org/10.1364/JOSAA.19.000849 (2002).
71. Braat, J., Dirksen, P. & Janssen, A. J. E. M. Assessment of an extended Nijboer–Zernike approach for the computation of optical point-spread functions. *J. Opt. Soc. Am. A* **19**, 858. https://doi.org/10.1364/JOSAA.19.000858 (2002).
72. Young, L. K., Morris, T. J., Saunter, C. D. & Smithson, H. E. Compact, modular and in-plane AOSLO for high-resolution retinal imaging. *Biomed. Opt. Express* **9**, 4275. https://doi.org/10.1364/BOE.9.004275 (2018).
73. Morgan, J. I. W., Vergilio, G. K., Hsu, J., Dubra, A. & Cooper, R. F. The reliability of cone density measurements in the presence of rods. *Transl. Vis. Sci. Technol.* **7**, 21. https://doi.org/10.1167/tvst.7.3.21 (2018).
74. Burns, S. A., Elsner, A. E., Sapoznik, K. A., Warner, R. L. & Gast, T. J. Adaptive optics imaging of the human retina. *Prog. Retin. Eye Res.* **68**, 1–30. https://doi.org/10.1167/11.5.60 (2019).
75. Balasuriya, S. & Siebert, P. A biologically inspired computational vision front-end based on a self-organised pseudorandomly tessellated artificial retina. *Proc. IEEE Int. Jt. Conf. Neural. Netw.* **5**, 3069–3074. https://doi.org/10.1109/IJCNN.2005.1556415 (2005).

Acknowledgements
The authors would like to thank Sarah Regan and Mital Shah for collecting the human images and the five expert raters for producing manual labels for them. The authors would like to acknowledge funding from Fight For Sight (1467/8), the University of Oxford Wellcome Trust Institutional Strategic Support Fund (105605/Z/14/Z), the University of Oxford Medical Research Fund (MRF/LSV2015/2161), the EPA Cephalosporin Fund (CF 277), the John Fell Oxford University Press (OUP) Research Fund (103/786 and 151/139). Laura Young is supported by a UKRI Future Leaders Fellowship (MR/T042192/1) and is also supported by a Reece Foundation Fellowship in Translational Systems Neuroscience (Newcastle University).

Author contributions
L.Y. designed and wrote the simulation, performed the analyses and wrote the main manuscript text. Both authors contributed to the conceptualisation of the approach and to writing and editing the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-90389-y.

Correspondence and requests for materials should be addressed to L.K.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021