Decoding network-mediated retinal response to electrical stimulation: implications for fidelity of prosthetic vision

Elton Ho[1,2], **Alex Shmakov**[3], **Daniel Palanker**[2,4]

[1]Department of Physics, Stanford University, Stanford, CA 94305, United States of America

[2]Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA 94305, United States of America

[3]Department of Computer Science, UC, Irvine, CA 92697, United States of America

[4]Department of Ophthalmology, Stanford University, Stanford, CA 94305, United States of America

Abstract

Objective.—Patients with photovoltaic subretinal implant PRIMA demonstrated letter acuity ~0.1 logMAR worse than sampling limit for 100 μm pixels (1.3 logMAR) and performed slower than healthy subjects tested with equivalently pixelated images. To explore the underlying differences between natural and prosthetic vision, we compare the fidelity of retinal response to visual and subretinal electrical stimulation through single-cell modeling and ensemble decoding.

Approach.—Responses of retinal ganglion cells (RGCs) to optical or electrical white noise stimulation in healthy and degenerate rat retinas were recorded via multi-electrode array. Each RGC was fit with linear–nonlinear and convolutional neural network models. To characterize RGC noise, we compared statistics of spike-triggered averages (STAs) in RGCs responding to electrical or visual stimulation of healthy and degenerate retinas. At the population level, we constructed a linear decoder to determine the accuracy of the ensemble of RGCs on N-way discrimination tasks.

Main results.—Although computational models can match natural visual responses well (correlation ~0.6), they fit significantly worse to spike timings elicited by electrical stimulation of the healthy retina (correlation ~0.15). In the degenerate retina, response to electrical stimulation is equally bad. The signal-to-noise ratio of electrical STAs in degenerate retinas matched that of the natural responses when $78 \pm 6.5\%$ of the spikes were replaced with random timing. However, the noise in RGC responses contributed minimally to errors in ensemble decoding. The determining factor in accuracy of decoding was the number of responding cells. To compensate for fewer responding cells under electrical stimulation than in natural vision, more presentations of the same stimulus are required to deliver sufficient information for image decoding.

eltonho@stanford.edu .

Supplementary material for this article is available online

Significance.—Slower-than-natural pattern identification by patients with the PRIMA implant may be explained by the lower number of electrically activated cells than in natural vision, which is compensated by a larger number of the stimulus presentations.

Keywords

retinal prosthesis; restoration of sight; retinal degeneration; prosthetic vision; neural computation

1. Introduction

Age-related macular degeneration (AMD) is a leading cause of untreatable blindness. Geographic atrophy (GA), the atrophic form of advanced AMD, affects around 3% of people above the age of 75, and around 25%—above90[1,2]. Due to gradual loss of photoreceptors in the central macula, GA patients experience severe deterioration in high-resolution central vision, compromising their ability to read and recognize faces. Although central vision degrades over time, patients retain their low-resolution peripheral vision, and hence typically do not lose visual acuity beyond 20/400.

One approach to restoration of sight in retinal degeneration is to replace the missing photoreceptors with photodiodes [3, 4], which convert the incident light into electric current flowing through the retina, and thus convey visual information to the secondary neurons by electrical stimulation. Massive amplification mechanisms in photoreceptors enable their operation in a very broad (about 10 orders of magnitude) range of light intensity. Photodiodes require much brighter illumination (about 1 mW mm^{-2}) in order to provide sufficient current for retinal stimulation. Therefore, photovoltaic system for restoration of sight includes augmented-reality glasses, where images captured by the camera are projected into the eye using more intense light. To avoid perception of this intense light by the remaining photoreceptors, near-infrared wavelength (880 nm) is used. To provide charge-balanced electrical stimulation, the light is pulsed, and to enable stable visual percepts, pulse repetition rate should exceed the frequency of flicker fusion (around 30 Hz). For functional restoration of sight, prosthetic vision should be encoded in a way which brain can easily decode, and, ideally, should be as close to natural, as possible.

In healthy retina, optical information (local light intensity) is converted via phototransduction into decrease of the cell potential (hyperpolarization) in photoreceptors, which, in turn, reduces the rate of release of neurotransmitter glutamate in synapses with the secondary neurons — bipolar and horizontal cells. By providing lateral inhibition that forms an antagonistic surround, horizontal cells perform the first step in spatial contrast enhancement. Bipolar cells electrically integrate inputs from multiple photoreceptors and relay these signals to tertiary retinal neurons — retinal ganglion cells (RGCs), with their inputs regulated by amacrine cells primarily via lateral inhibition. Finally, RGCs convert these signals into bursts of the action potentials ('spikes'), which propagate via optic nerve to the brain. Encoding of the visual information in different types of RGCs varies by multiple properties: ON and OFF pathways, sizes of receptive fields (RFs), transient and sustained responses, chromatic sensitivity, etc. Signals from the overlapping mosaics of the various types of RGCs are further processed in the brain before merging into a single visual

percept. In atrophic AMD, photoreceptors in the central macula slowly degenerate and disappear, while the inner retinal neurons remain largely intact, albeit with some rewiring [5–7].

Ex-vivo and *in-vivo* animal studies with photovoltaic replacement of photoreceptors demonstrated preservation of multiple features of the natural retinal signal processing: sizes of the RGC RFs, with antagonistic surround [8] and non-linear summation of subunits [3], flicker fusion [9] and adaptation to static images [10], and spatial resolution matching the pixel pitch (55 and 75 μm) [9]. Interestingly, both, *ex-vivo* and *in-vivo* studies demonstrated not only excitatory (ON) but also inhibitory (OFF) responses to electrical stimulation. The latter may be explained by stimulation of rod bipolar cells, which feed into the ON and OFF cone pathways via amacrine cells. Contrast sensitivity of prosthetic vision in rats appears to be about five times lower than natural [11], which is likely due to the fact that horizontal cells connect to the terminals of photoreceptors, in GA they become detached from the remaining neural network. Therefore, some reduction in contrast sensitivity is expected [11], and for prosthetic vision it might be partially compensated by image processing prior to projection onto the implant. Clinical trial demonstrated that patients correctly perceive various patterns of lines and letters, demonstrating monochromatic shaped vision with resolution closely matching the pixel size (100 μm in the first trial) [4]. They also report flicker fusion at frequencies exceeding 30 Hz.

One of the features of such prosthetic vision, however, appears to be the lower than normal speed of the pattern and letter recognition. In the first trial, it took about 4 s for letter identification by patients with PRIMA implants [4], while it takes less than half a second in normal subjects, when font sizes exceed the acuity limit [12]. Here, using a retrospective analysis of the previously recorded data [8], we investigate the potential retinal underpinnings of this phenomenon by comparing RGC responses to visual and electrical stimulation in healthy and degenerate rat retina recorded on a multi-electrode array (MEA). In particular, we assess the amount of noise in various cellular responses, as well as the strategies for image recognition based on ensemble encoding by a population of cells.

2. Methods

2.1. Photovoltaic implants

Photovoltaic arrays (1 mm diameter, 30 μm in thickness, with 75 μm pixels) (figure 1(a)) were manufactured from crystalline silicon, as described earlier [13], to produce anodic-first pulses. Active electrode was 20 μm in diameter, and each pixel was surrounded by a return electrode, connected into a mesh common to all pixels. Both electrodes were coated with SIROF film of about 300 nm in thickness [9].

2.2. Retinal recording

The retinal data is taken from the previously published recordings [8]. We used Long-Evans (LE, *n*=4) and Royal College of Surgeons (RCS, *n* = 4, P120–130) rats for healthy and degenerate retinal models, respectively. RCS rats were kept in a colony maintained at the Stanford Animal Facility. LE rats were purchased from Charles River (Wilmington, MA,

USC). All experimental procedures were approved by the Stanford Administrative Panel on Laboratory Animal Care, and all animals were kept in accordance with the institutional guidelines and conformed to the guidelines of the Association for Research in Vision and Ophthalmology Statement for the Use of Animals in Ophthalmic and Vision Research. Eyes were enucleated from euthanized (390 mg kg^{-1} pentobarbital sodium, 50 mg ml^{-1} phenytoin sodium) rats. A section of the retina (~3 mm × 3 mm) was dissected and placed ganglion cells side facing a 512-electrode MEA (figure 1(b)) [14]. The retina was constantly perfused with Ames' medium and bubbled with a mixture of 95% O_2 and 5% CO_2. The body temperature of rats is around 36 °C–37 °C. To allow long recordings *ex-vivo*, metabolic rate is slowed down by lowering temperature to around 30 °C. We picked 29.4 °C to be consistent with previous *ex-vivo* studies [3, 8, 10]. For electrical stimulation, an implant was placed onto the subretinal side of the tissue (figures 1(c) and (d)). A nylon mesh (~100 μm cell size) was used to press the implant and retina onto the MEA for better contact [10]. Voltage waveforms from each of the 512 electrodes on the MEA were sampled at 20 kHz frequency, amplified and digitized using custom-made readout electronics and data acquisition system [14].

2.3. Stimulation protocol

For electrical stimulation, an 880 nm diode laser coupled via a 400 μm multimode fiber was used for illumination. The beam exiting from the fiber was collimated and homogenized using a 2° divergence microlens array. In the same optical path, we placed a yellow LED (591 nm) for visual stimulation. Both light sources were used as backlighting for an LCD screen (Holoeye HEO-0017) to generate images [3, 10]. The 8-bit LCD panel had a 60 Hz native frame rate, 1024 × 768 resolution, and a white-to-black intensity ratio of 10000:1 at 591 nm and 200:1 at 880 nm. Projected onto the retina, each screen pixel formed a 6 × 6 μm^2 square.

To characterize spatiotemporal properties of RGCs, a spatiotemporal binary white noise stimulus was used, where each pixel in each frame had a 50% chance of being bright or dark [15]. The white noise for visual stimulation was shown at 30 Hz frame rate, and made up of pixels of 60 μm in width on the retina. The white noise for electrical stimulation was displayed at 20 Hz frame rate, with the backlight laser pulsing at 4 ms, and consisted of pixels of 70 μm in width on the retina. Each white noise stimulus lasted for 30 min.

2.4. Spike sorting

Electrical stimulation generated artifacts by saturating the MEA recording amplifiers, so part of the recorded waveforms had to either be discarded or adjusted (see supplementary figure 1 (available online at https://stacks.iop.org/JNE/17/066018/mmedia)). To pre-process the data for the spike-sorting pipeline, the recording of the first 8.25 ms after the laser pulse was replaced with a randomly generated noise ('blanking') that matched the noise level of the electrode. All action potentials during this period were lost, which may lead to underestimation of the cell responsiveness. Afterwards, to remove any lingering capacitive decay outside of the blanked period, we fitted the trace with a 7th-order polynomial, and then subtracted it out from the original trace.

The artifact-subtracted raw data were then used to find and sort the action potentials ('spikes'). A negative voltage deflection exceeding three times the root-mean-squared noise on each electrode was considered a spike. Custom-made software was used to perform spike sorting, as described previously [3, 14, 16]. We applied dimensionality reduction to the detected spike waveforms using a principle component analysis, followed by expectation-maximization clustering [14]. For each putative neuron, we calculated its electrophysiological image (EI), which is the average electrical signal measured on the whole MEA when the neuron produced a spike. An EI typically shows the soma location and axonal trajectory of the RGC [17, 18]. Figures 2(a) and (b) illustrate examples of the RF mosaic of included cells for healthy and degenerate retinas, respectively. Since electrical stimulation affected only the cells directly underneath the implant, RGC RFs in the degenerate retina were observed only in that region.

All RGCs were classified into visual/electrical ON and OFF cells by inspecting the sign of the first peak preceding the spike in the time course (spike-triggered average, STA). Figures 2(c)–(e) show example RFs and time courses for natural visual response, electrical stimulation of the healthy retina, and electrical stimulation of RCS retina, respectively. Electrically stimulated cells generally have faster, but weaker STAs, similar to previous observations [8, 11]. Visual ON stimulation results in hyperpolarization of photoreceptors, and visual OFF leads to depolarization. Subretinal anodic stimulation of the LE retina will depolarize photoreceptor terminals, generating the same effect as a visual OFF stimulus. Therefore, the polarity of the peak closest to the right is inverted, and a visual LE ON cell is an electrical OFF cell. To date, there is no clear indication whether electrical ON and OFF RGCs in the degenerate retina correspond to any specific natural type of RGCs.

For our analysis, we selected cells with the following four criteria. First, for each candidate neuron, an estimate of the fraction of spikes coming from other neurons ('contaminating spikes') was obtained from the number of refractory period violations in the spike train [14]. Cells with over 10% contamination were excluded. Second, RGCs with backward propagating axonal signals were also excluded. Each of these criteria removed less than 10% of cells. Third, RGCs with the time course signal-to-noise ratio (SNR) below 3 were excluded from the analysis. For the SNR calculation, the peak value of the time course was used as a signal, and the root-mean-square value of the 10 time course values farthest from the time of the action potential was used as noise [8, 11]. This criterion excluded a higher ratio of RCS RGCs than LE RGCs, exemplified by the SNR difference between the time courses in figures 2(c) and (e). Depending on the preparation, this criterion removed around 24%–40% of the initially identified LE visual cells, and 92%–94% of the RCS cells. Fourth, for LE retinas, we only included cells that responded to both visual and electrical stimulation, and have their somas under the implant. The restricted area of interest allows for a fairer comparison between the number of responsive cells across retina types. This procedure retained around ~19%–24% of LE cells. Within the implant region, approximately 23%–32% of the visually responsive cells were also electrically responsive. Overall, compared to the initially identified cells, around 13%–18% of the LE RGCs and 6%–8% of RCS RGCs were included in our analysis.

2.5. Modeling RGC responses

Experimental data for each cell was fitted to a linear–nonlinear (LN) model [15] and a convolutional neural network (CNN) model [19], as illustrated in figure 3. Due to drifting in recording data, we used a train–test–discard split of 20/20/60 (see supplementary figure 2). Only 20% was chosen for training, because we fitted our models to segments of data which were consistent with respect to data drifting, while keeping enough data to avoid model overfitting. The model was tested on another 20% of the test data in the unused 80%. Since there were no repeated stimuli, we compared the model predictions to experimental data by applying Gaussian broadening to each spike with $\sigma = 2$ white noise frames for smoothening (figure 5(a)). This was applied to both the spikes measured in experiments and model-predicted spikes. We then computed the Pearson correlation coefficient for the resulting traces. In addition, all model fits were five-fold cross-validated, where we equally sized the sections dividing the entire recorded data. Neither model overfitted to any particular segment of the data.

2.5.1. LN model—Mathematically, the LN model (figure 3(a)) is set as follows:

$$R(s) = N(w \cdot s) \tag{1}$$

where R = response

s = stimulus

w = linear weights/filter

N = static nonlinearity

The linear filter w can be computed through STA response to the white noise stimulus. The static nonlinearity can be extracted by mapping the empirical cell activity to stimulus convolved with the linear weights ($w \cdot s$) [15].

2.5.2. CNN model—An implementation of a CNN model has been proven useful for modeling the healthy salamander retina [19], and here we used a similar architecture (figure 3(b)) with two convolution blocks followed by a dense layer. Each convolution block consisted of a 2D convolution (weights), a parametric rectifier linear unit, batch normalization (norm), and a dropout layer. These last two components sped up the training while also regularizing the network to prevent overfitting. The number of filters and their dimensions for each convolution block were picked to optimize model performance on test data while avoiding overfitting and are listed in table 1. The output of the second convolution layer was flattened into a 1D vector before being fed into the final dense linear layer, which had a number of units matching the number of recruited RGCs in the retina.

The network was trained using the gradient-descent ADAM optimizer [20] and a Poisson log-likelihood. L2 weight regularization was employed on the convolution and linear layers, while L1 regularization was used on the output of the network. Especially for cells with lower firing rates, L1 can efficiently zero-out many weights. The complete loss function was defined as follows:

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^{N} R_i - \widehat{R}_i \log R_i + \frac{\alpha}{2} \sum_{j=1}^{M} \left| W_j \right|_F^2 + \frac{\beta}{N} \sum_{i=1}^{N} \left| R_i \right| \tag{2}$$

where R_i = model response i

$\widehat{R}_i$ = target response i

$\left| W_j \right|_F$ = Frobenius norm of the jth convolution or linear layer weight matrix

N = total number of training samples M = total number of weight matrices

α, β = L2 and L1 regularization coefficients, respectively.

The input stimulus was similar to that used in computing STAs, while the response now included all activity and inactivity. For visual stimulation, 20 consecutive movie frames (spanning 600 ms) were considered one stimulus, and the spike rate during 33 ms following the stimulus was taken as the target response. Similarly, for electrical stimulation, five movie frames (250 ms) and the following 50 ms of activity was considered a stimulus–response pair. To improve precision of spike timing while increasing the training sample count, the electrical stimulus was up-sampled with linear interpolation to 250 Hz, and the corresponding RGC spiking activity was binned to match the stimulus frame rate. During validation, the predicted activations were down-sampled back to the original frame rate before the correlation was computed.

The CNN is parameterized by 13 different hyperparameters, including filter count, size, stride, and nonlinearity for each of the two convolution blocks. In addition, we also explored different values for the learning rate, L1 and L2 coefficients, batch size, and dropout probability. We performed 100 trials for each dataset with randomized values for all parameters using the SHERPA hyperparameter optimization library [21], and the best-performing hyperparameters were picked. Networks were trained for 50 epochs on electrical datasets and 50 epochs on visual datasets. The total training time on a single NVIDIA Titan X GPU was 30 min and 2 h, respectively.

2.6. RGC noise estimation

We estimated the noise level in RGC firing under electrical stimulation using the algorithm illustrated in figure 4(a). First, we passed the white noise stimulus through an LN model for an LE RGC under visual stimulation, generating a simulated spike train. For the noise-injection step, we then removed spikes randomly at some predefined ratio, which we denote noise ratio (NR). Spontaneous spikes were added into the spike train to match the original average spike rate via a Poisson process. With the new noise-injected spike train, we re-computed the STA. The boxed regions in figure 4(b) with dimensions (length, width, time) = (3 px, 3 px, 4 frames) were used for the subsequent analyses.

For each cell and its STA, we computed a characteristic correlation curve (CCC) and its corresponding area-under-curve (AUC), as shown in figure 6(a). Let $F(n)$ be the 24 STA

frames preceding the nth spike, N be the total number of spikes, and $\text{STA}\left(\frac{n}{N}\right)$ be the STA computed with only a fraction n/N of the spike train. We have

$$\text{STA}\left(\frac{n}{N}\right) = \frac{1}{n}\sum_{i=1}^{n} F(i) \tag{3}$$

$$\begin{aligned}
\text{CCC}\left(\frac{n}{N}\right) &= \text{Corr}\left(\text{STA}(1), \text{STA}\left(\frac{n}{N}\right)\right) \\
&= \frac{\sum\left((\text{STA}(1) - \overline{\text{STA}(1)}) \cdot \left(\text{STA}\left(\frac{n}{N}\right) - \overline{\text{STA}\left(\frac{n}{N}\right)}\right)\right)}{\sqrt{\sum\left(\text{STA}(1) - \overline{\text{STA}(1)}\right)^2} \cdot \sqrt{\sum\left(\text{STA}\left(\frac{n}{N}\right) - \overline{\text{STA}\left(\frac{n}{N}\right)}\right)^2}}
\end{aligned} \tag{4}$$

$$\text{AUC} = \int_{0}^{1} \text{CCC}\left(\frac{n}{N}\right) d\left(\frac{n}{N}\right) \tag{5}$$

Note that STA(1) is when all the spikes are included in computing the STA.

If a cell responds perfectly only to one single type of stimulus with no spontaneous firing, then $\text{STA}\left(\frac{1}{N}\right) = \text{STA}\left(\frac{N}{N}\right)$, and AUC= 1. Figure 6(a) shows the CCC's of three example cells, and figure 6(b) shows the distribution of cell count for various levels of AUC.

Out of all LE retinas, we selected the cell that had the median AUC under visual stimulation as a reference. Noise-injection into the STA of this cell yielded a family of CCCs shown in figure 6(c). To characterize the noise of each cell under electrical stimulation, we matched its CCC to the curve in the family that has the most similar AUC, and the resulting matching NR characterizes the cell (figure 6(d)).

2.7. Ensemble encoding

To evaluate how much information is encoded by the ensemble of cells for the pattern recognition task, we simulated projection of pixelated Landolt-C onto a piece of retina (figure 7(a)). Each presentation of the C lasted for five movie frames, and was spatially pixelized into either 70 μm (for electrical stimulus) or 60 μm (for visual stimulus) pixels. The brightness of each pixel was then rounded to the nearest one of eight evenly spaced greyscale levels. The resulting simulated stimulus ($\hat{s}$) had dimensions (length, width, time) = (64, 32, 24 frames) for visual stimuli or (20, 20, 24 frames) for electrical stimuli. The first 19 frames were all dark, and the remaining frames were bright, where the Landolt-C was displayed. The simulated stimulus was then convolved with the STAs (w) of each cell to produce an input strength ($w \cdot \hat{s}$), similar to equation (1). We then inspected the white noise stimulus for blocks of consecutive frames (s, same dimensions to $\hat{s}$) that share a similar input strength, mathematically defined as

$$s = \mathrm{argmin}_s \left| w \cdot s - w \cdot \hat{s} \right| \qquad (6)$$

The 500 blocks that best satisfy the above criterion are chosen, and the average RGC activity 30 ms following each block was considered a response to the Landolt-C. Afterwards, responses of all cells were concatenated into a template with time bins of 5 ms. Four different templates were created for four orientations (up, down, left, right) of the C using the same procedure. We then simulated 10000 trials with random orientations that will be decoded and identified for its ground truth orientation. For each trial, the number of spikes in each time bin was simulated as a Poisson process with its mean matching the spike rate in the bin in the corresponding template. The generated spiking pattern for each trial was correlated to all templates, and one with the highest Pearson's r was considered the decoded orientation. Decoding accuracy was taken as the ratio of correctly decoded trials to total. For the four LE retinas, the cell counts were 49, 49, 21, and 20; for the four RCS retinas, the cell counts were 19, 14, 13, and 9. To study the effect of number of cells on decoding accuracy, we fixed the size of the C at 14 pixels. To study the effect of C size and number of flashes, we included all cells on each retina into the decoder.

3. Results

3.1. Single-cell response modelling

Each selected RGC recorded on an MEA were fitted with both an LN model [15] and a CNN model (mcintosh2018). The 30 min long recordings were split into a train–test–discard ratio of 20/20/60. After model training, both spikes in the test data and the model predictions were broadened with a Gaussian filter of $\sigma = 2$ white noise frames. The filtered test data was then correlated with the filtered model predictions using Pearson's correlation coefficient.

For the natural response of healthy retina, the LN model fitted to levels similar to previous reports in salamander and rat retinas (correlation in the range of 0.3) [19, 22]. The CNN model fitted much better to the spike trains elicited by visual stimulation in ON and OFF cells (correlation of about 0.6, figure 5(a) top), agreeing with earlier studies in the salamander retina [23]. However, both models predictions correlated with retinal responses to electrical stimulation of the healthy or degenerate retina significantly worse (figure 5(a), center and bottom). Across a population of cells and multiple retinas ($n = 4$ each), CNN fits to electrical data reached a correlation of only ~0.15, significantly lower than 0.6 for the LE visual response (figure 5(b)). The LN model fitted distinctly better to electrical OFF cells than ON cells in LE retinas ($p < 10^{-7}$, two-sample t-test), while the CNN model fitted with less discrepancy between the two cell types ($p = 0.013$). For electrical ON cells, the CNN model fitted significantly better than the LN model ($p < 10^{-9}$), but the same cannot be said for electrical OFF cells. In RCS retinas, correlation with the CNN model was similar to the LE retina under electrical stimulation (figure 5(b)). However, correlations with the LN model were far worse for RCS OFF cells than for LE electrical OFF cell, while that for electrical ON cells remained similar.

3.2. Noise estimation in RGC firing

To characterize the noise of each RGC, we computed a CCC and its corresponding AUC for each cell and its STA according to equations (3)–(5) (see section 2). If a cell responds perfectly only to one single type of stimulus with no spontaneous firing, the AUC would be 1 and the CCC curve would be a flat line at correlation equal to 1. If the cell is firing strictly spontaneously, the CCC would resemble the noise curve in figure 6(a).

Figure 6(a) illustrates the characteristic correlation curves (CCCs) of three example cells, as well as the CCC for an STA generated from randomly sampled white noise frames, which is described by a square root dependence on the normalized number of spikes. The CCC for RCS is the closest to the noise curve and has the lowest AUC, followed by LE electrical. Distribution of AUCs across the cell population in retinas, shown in figure 6(b), confirms that RCS responses were the noisiest, followed by the LE electrical responses. By replacing spikes in LE visual responses with randomly timed spikes (see section 2), we can generate a family of CCCs with various NRs (figure 6(c)). At a certain NR, the noise-injected visual CCC matches that from electrical responses. Compared to the LE visual response with median AUC, the NR was $55.3 \pm 22.5\%$ and $78.2 \pm 6.5\%$ for LE electrical and RCS, respectively (figure 6(d)). All RCS responses were at least 55% noisier than natural.

3.3. Ensemble encoding

To evaluate the efficacy of ensemble encoding of visual information, we simulated the flashing of Landolt-C patterns, and the resultant spiking patterns of RGCs on a retina. The spiking patterns of all included cells on the same piece of retina were used to train a decoder (see section 2). Afterwards, with test trials, we compute the accuracy of decoding the orientation of a displayed Landolt-C.

Accuracy of decoding the orientation of Landolt-C rises with increasing number of recruited cells (figure 7(b)). Since the cells were ranked by their independent decoding accuracies, the first few cells contributed to the faster rise in accuracy. In addition, beyond the first few, recruited cells started carrying redundant information, which improved accuracy with diminishing returns. Such trend is generally observed in decoding the ensemble of neural signals for applications in brain–machine interfaces [24]. Notably, neither presence of photoreceptors nor stimulation type affected the decoding accuracy significantly, despite the spiking being much more stochastic under electrical stimulation, as discussed previously. Decoding accuracy also rises steeply with increasing C size until it reaches 4–5 pixels (figure 7(b)), where the accuracy flattens out because the gap in C now exceeds one pixel, and hence it is fully resolved. The asymptotic level of accuracy was determined primarily by the number of recorded cells in the retina. For example, retinas with 49 recorded cells can reach 60%–80% accuracy, while only 50% can be achieved with 19–21 cells (figure 7(c)).

Increasing the number of stimulus presentations also increased the decoding accuracy (figure 7(d)). Therefore, to compensate for fewer cells responding in RCS retinas, more presentations are required to accumulate the same amount of information for the image decoding. To achieve 75% accuracy in decoding the orientation of letter C, LE retinas required 2–3 flashes of the image, while the RCS retinas needed 13 presentations.

4. Discussion

The fact that predictive retinal models perform worse for RCS under electrical stimulation than for LE retina under visual stimulation is not surprising and can be explained by several factors, including the increased spontaneous firing rate in the degenerate retina [25, 26], likely due to higher uptake of retinoic acid [27]. There is evidence that even in healthy mice retinas, intrinsic variability of the RGCs response to network-mediated electrical activation is higher than with natural visual stimulation [28]. Replicability and reliability of the network-mediated stimulation also decreases with retinal rewiring during degeneration [29]. In addition, fewer and weaker responding cells in the RCS retina compared to normal could be expected both *in-vivo* and *ex-vivo* due to limited penetration depth of electric field into the inner nuclear layer (INL), especially with residual debris separating the degenerate retina from the stimulating array.

Our model predictions for the network-mediated stimulation of RCS retina are worse than in other studies that directly stimulate RGCs, either electrically [22,30] or optogenetically [31]. Direct stimulation of RGCs bypasses many noise sources, such as bipolar and amacrine cells, connecting synapses, and any pathological rewiring of the retina, and therefore each RGC can be more reliably modeled. The tradeoff for this reliability is the loss of inherently preserved retinal signal processing, such as antagonistic center-surround organization [8] and subunit summation [3]. As demonstrated in our ensemble encoding simulation and evidenced in clinical studies [4], it appears that for pattern perception, the ensemble encoding is more relevant than reproducibility of the individual cell responses.

On the other hand, we found that response of the healthy retina to electrical stimulation is much noisier than natural. Also, the number of electrically activated RGCs is lower than natural by a factor of ~3–4. This might be related to the difference in mechanisms of natural and electrical activation of the photoreceptors. In natural vision, due to the rather slow phototransduction cascade, a millisecond flash causes photoreceptor hyperpolarization for tens of milliseconds [32]. Under electrical stimulation, however, membrane potential is affected directly, and therefore it closely follows the electrical pulse duration (<10 ms), much shorter than the natural response. Another factor might be related to the fact that our experiments were performed in the dark. Since the dark-adapted photoreceptors are depolarized, further depolarization of the terminals by electric field is quite limited, effectively restricting the dynamic range. Both factors likely contribute to the lower-than-natural SNR and fewer activated cells when the healthy retina is stimulated electrically.

The cell count is further reduced when we compare across healthy and degenerate retinas, even when both are under electrical stimulation. Within the implant boundary, smaller number of RCS RGCs and their corresponding spotted distribution of RFs may be due to multiple reasons. First, limited penetration depth of electric field has stronger effect on the INL than photoreceptors. Photoreceptor outer segments all reached the implant, allowing electrical access to all cells. However, cell bodies in the INL are staggered in depth, so the deeper bipolar cells might not reach the stimulation threshold. Second, since RCS rat models the retinitis pigmentosa (RP), significant retinal rewiring could have resulted in less- or even non-responsive regions. Retina in AMD patients is unlikely to rewire to the same

extent as in the end-stage RP, and hence we can expect a higher cell count and a more complete mosaic in clinical testing. Third, the thinner than normal RCS retina is more prone to mechanical damage during tissue preparation and mounting. Fourth, to avoid mechanical damage, we elected to remove vitreous less vigorously, which may have left some residual vitreous and debris that impeded electrical coupling to the MEA. These last two reasons would cause a lower cell count *ex-vivo* than in clinical practice. With the SNR >3 selection criterion, it is likely that a fraction of ON–OFF cells was excluded because the ON and OFF responses counteracted and canceled or weaken each other in STA computation. In addition, since we did not further sub-classify cells beyond ON and OFF, it is likely that some direction-selective cells are included, although f or electrical responses, it is unclear whether direction selectivity is preserved.

By construction, under a radially symmetric stimulus, the STA of an RGC is the first-order term in the Wiener kernel series expansion of the cell's response function [33]. Therefore, the LN model can be considered a single-filter approximation, while the CNN model can fit better due to inclusion of multiple linear filters and a better approximation of nonlinearities [34]. Indeed, previous studies have shown that CNN models fit markedly better than the LN model in the salamander retina [19], which we also observed here in the healthy rat retina under visual stimulation. Surprisingly, there is little difference between the two models fitted to LE retina OFF cells stimulated electrically, indicating that the responses were predominantly single-filter (figure 5(b)). An interpretation is that OFF cell responses can be described nearly completely using only a single RF, which means these cells only respond to a limited subspace of stimuli. Consequentially, these RGCs may fail to respond to certain classes of spatial patterns that require subunit computation, such as null stimuli [35], where the linear filter of a cell is scaled and subtracted away from an otherwise response-inducing white noise stimulus. Computationally, LN is a simpler model that captures lower-order computations, while CNN can learn a better higher-order representation, if it exists. The disparity between LN and CNN models for RCS data suggests that even the degenerate retina retains some degree of higher-order computation. For example, from the response to alternating grating (both *ex-vivo* and *in-vivo*), we know that nonlinear summation of subunits occurs also in the degenerate retina [3], but whether the number of computational subunits for each RGC matches that of the healthy retina remains unknown.

It is of note that the correlation we perform here is not conventional. Commonly, when comparing between a model prediction and experimental data, some form of repeated stimulus is required. For example, it could be a 30 s long white noise movie repeated for ten times. The average response is then correlated to the model prediction. However, in our case, we did not have repeated stimuli. As a workaround, we applied a Gaussian filter to broaden the spikes recorded over 6 min of non-repeating white noise. If cells spike with perfect timing and repeatability, the correlation should resemble the conventional method. However, if spike rates and spike times are highly variable, such as RGCs in a degenerate retina, the correlation should be lower than in conventional methods. In addition, without repeated stimuli, we cannot select cells using reliability-based cell selection criteria [22]. Instead, we selected cells with SNR above 3 in the STA time course, which may include noisier cells. Consequently, the absolute correlation value reported here is not a one-to-one comparison with previous studies [22, 23].

In ensemble encoding, we made two important assumptions: First, the input strength ($w\,s$) was calculated with linear weights extracted from the binary white noise. Since pixels in the stimulus are spatiotemporally independent, the resulting trained weights are generally biased against spatiotemporally correlated stimuli, such as long straight edges and bars, drifting objects, and even natural scenes [36]. Therefore, the current method leads to underestimating the accuracies in the LE retina. It is unknown whether directional sensitivity remains intact in the degenerate retina or how many nonlinear subunits exist under electrical stimulation, so the accuracy curves for the RCS retina in figure 7(b) may or may not be underestimated. Second, the letter C was placed at the same location over the five frames it was displayed. Normally with microsaccades, visual pattern can be displaced by ~70 μm between the frames presented at 20–33 Hz [37]. Since RFs form a tightly packed mosaic, we assumed translational symmetry in response, i.e. no matter where the stimulus is displayed on the retina, the retinal output will carry equivalent amount of information. This assumption cannot be made if the retina is foveated, as the density of photoreceptors varies with eccentricity. However, since the rat retina is not foveated, we decided to follow this assumption. As a result, the current analysis assumes that the effect of the eye movements on amount of information for pattern identification is well-approximated even without moving the letter C.

Single-cell SNR played little role in ensemble encoding of the visual information. As demonstrated in figure 7(b), all retinas had similar accuracies, even though LE RGCs under visual stimulation had far better SNR. A reason might be that better SNR as calculated might not limit the amount of information propagating downstream, if the encoded visual signals were orthogonal to the major noise eigen-modes [38]. Since visual information is distributed across the retina, the more cells recruited for decoding, the higher is the accuracy. Unlike natural visual response, electrical stimulation affects the bipolar cells stronger if they reside near the electrode surface, and hence fewer RGCs were responding than in natural stimulation. To compensate for the reduced amount of visual information transmitted, the stimulus needs to be replayed multiple times. This may explain the longer time patients require to recognize letters and other patterns in clinical trials [4].

As in Sloan font, the gap in a Landolt-C is 1/5 of the letter size [39]. For letter sizes smaller than 4 pixels, the gap is not fully resolved but encoded in some shade of grey different from the rest of the ring, which led to lower accuracy in identification. Once the gap is fully resolved, i.e. letter size greater than 5 pixels, decoding accuracy remains relatively stable. This signifies that with the pixel size used in these studies (70 μm), the limiting factor in resolution is strictly at the implant level (pixel size), but not biological (subunit size). Essentially, prosthetic vision can resolve spatial features down to the pixel size with stable accuracy, which matches our previous *in-vivo* measurements [3].

In conclusion, we found that LN and CNN models matched the RGC activity elicited by subretinal electrical stimulation less accurately than that for natural responses, likely due to the weaker than natural response and higher spontaneous firing in the degenerate retina. Despite the noisier signal, visual information is still encoded across the ensemble of cells in the retina, which allows patients to perform visual discrimination tasks, albeit slower due to the reduced number of responding RGCs, compared to natural vision.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

[1]. Wong WL, Su X, Li X, Cheung CM, Klein R, Cheng CY and Wong TY 2014 Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis Lancet Glob. Health 2 e106–16 [PubMed: 25104651]

[2]. Friedman DS, Tomany SC, McCarty C and De Jong P 2004 Prevalence of age-related macular degeneration in the United States Arch. Ophthalmol. 122 564–72 [PubMed: 15078675]

[3]. Lorach H et al. 2015 Photovoltaic restoration of sight with high visual acuity Nat. Med. 21 476–82 [PubMed: 25915832]

[4]. Palanker D, Le Mer Y, Mohand-Said S, Muqit M and Sahel J A 2020 Photovoltaic restoration of central vision in atrophic age-related macular degeneration Ophthalmology 127 1097–1104 [PubMed: 32249038]

[5]. Mazzoni F, Novelli E and Strettoi E 2008 Retinal ganglion cells survive and maintain normal dendritic morphology in a mouse model of inherited photoreceptor degeneration J. Neurosci. 28 14282–92 [PubMed: 19109509]

[6]. Humayun MS et al. 1999 Morphometric analysis of the extramacular retina from postmortem eyes with retinitis pigmentosa *Invest*. Ophthalmol. Vis. Sci. 40 143–8 https://iovs.arvojournals.org/article.aspx?articleid=2162046

[7]. Kim SY, Sadda S, Pearlman J, Humayun MS, de Juan E Jr., Melia BM and Green WR 2002 Morphometric analysis of the macula in eyes with disciform age-related macular degeneration Retina 22 471–7 [PubMed: 12172115]

[8]. Ho E, Smith R, Goetz G, Lei X, Galambos L, Kamins TI, Harris J, Mathieson K, Palanker D and Sher A 2017 Spatiotemporal characteristics of retinal response to network-mediated photovoltaic stimulation J. Neurophysiol. 119 389–400 [PubMed: 29046428]

[9]. Ho E, Lei X, Flores T, Lorach H, Huang T, Galambos L, Kamins T, Harris J, Mathieson K and Palanker D 2019 Characteristics of prosthetic vision in rats with subretinal flat and pillar electrode arrays J. Neural. Eng. 16 066027 [PubMed: 31341094]

[10]. Goetz G, Smith R, Lei X, Galambos L, Kamins T, Mathieson K, Sher A and Palanker D 2015 Contrast sensitivity with a subretinal prosthesis and implications for efficient delivery of visual information Invest. Ophthalmol. Vis. Sci. 56 7186–94 [PubMed: 26540657]

[11]. Ho E, Lorach H, Goetz G, Laszlo F, Lei X, Kamins T, Mariani J-C, Sher A and Palanker D 2018 Temporal structure in spiking patterns of ganglion cells defines perceptual thresholds in rodents with subretinal prosthesis Sci. Rep. 8 3145 [PubMed: 29453455]

[12]. Chung ST, Mansfield JS and Legge GE 1998 Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision Vision Res. 38 2949–62 [PubMed: 9797990]

[13]. Wang L et al. 2012 Photovoltaic retinal prosthesis: implant fabrication and performance J. Neural. Eng. 9 046014 [PubMed: 22791690]

[14]. Litke AM et al. 2004 What does the eye tell the brain?: development of a system for the large-scale recording of retinal output activity IEEE Trans. Nucl. Sci. 51 1434–40

[15]. Chichilnisky EJ 2001 A simple white noise analysis of neuronal light responses Network 12 199–213 [PubMed: 11405422]

[16]. Field GD et al. 2010 Functional connectivity in the retina at the resolution of photoreceptors Nature 467 673–7 [PubMed: 20930838]

[17]. Li PH, Gauthier JL, Schiff M, Sher A, Ahn D, Field GD, Greschner M, Callaway EM, Litke AM and Chichilnisky EJ 2015 Anatomical identification of extracellularly recorded cells in large-scale multielectrode recordings J. Neurosci. 35 4663–75 [PubMed: 25788683]

[18]. Petrusca D, Grivich MI, Sher A, Field GD, Gauthier JL, Greschner M, Shlens J, Chichilnisky EJ and Litke AM 2007 Identification and characterization of a Y-like primate retinal ganglion cell type J. Neurosci. 27 11019–27 [PubMed: 17928443]

[19]. McIntosh L, Maheswaranathan N, Nayebi A, Ganguli S and Baccus S 2016 Deep learning models of the retinal response to natural scenes Adv. Neural Inf. Process. 29 1369–1377

[20]. Kingma DP and Ba J 2014 Adam: a method for stochastic optimization (arXiv preprint arXiv: 14126980)

[21]. Hertel L, Collado J, Sadowski P, Ott J and Baldi P 2020 Sherpa: robust hyperparameter optimization for machine learning (arXiv preprint arXiv: 200504048)

[22]. Höfling L, Oesterle J, Berens P and Zeck G 2020 Probing and predicting ganglion cell responses to smooth electrical stimulation in healthy and blind mouse retina Sci. Rep. 10 1–20 [PubMed: 31913322]

[23]. Maheswaranathan N et al. 2018 Deep learning models reveal internal structure and diverse computations in the retina under natural scenes (BioRxiv: 340943)

[24]. Carmena JM, Lebedev MA, Crist RE, O'Doherty JE, Santucci DM, Dimitrov DF, Patil PG, Henriquez CS and Nicolelis MAL 2003 Learning to control a brain–machine interface for reaching and grasping by primates PLoS Biol. 1 e42 [PubMed: 14624244]

[25]. Sekirnjak C, Jepson LH, Hottowy P, Sher A, Dabrowski W, Litke AM and Chichilnisky EJ 2011 Changes in physiological properties of rat ganglion cells during retinal degeneration J. Neurophysiol. 105 2560–71 [PubMed: 21389304]

[26]. Stasheff SF 2008 Emergence of sustained spontaneous hyperactivity and temporary preservation of OFF responses in ganglion cells of the retinal degeneration (rd1) mouse J. Neurophysiol. 99 1408–21 [PubMed: 18216234]

[27]. Denlinger B, Helft Z, Telias M, Lorach H, Palanker D and Kramer RH 2020 Local photoreceptor degeneration causes local pathophysiological remodeling of retinal neurons JCI Insight 5 2

[28]. Sekhar S, Ramesh P, Bassetto G, Zrenner E, Macke JH, Rathbun DL 2020 Characterizing Retinal Ganglion Cell Responses to Electrical Stimulation Using Generalized Linear Models Front. Neurosci. 14 [PubMed: 32047422]

[29]. Yoon YJ, Lee J-I, Jang YJ, An S, Kim JH, Fried SI and Im M 2020 Retinal degeneration reduces consistency of network-mediated responses arising in ganglion cells to electric stimulation IEEE Trans. Neural. Syst. Rehabil. Eng. 28 1921–1930 [PubMed: 32746297]

[30]. Maturana MI, Apollo NV, Hadjinicolaou AE, Garrett DJ, Cloherty SL, Kameneva T, Grayden DB, Ibbotson MR, Meffin H and Pillow JW 2016 A Simple and Accurate Model to Predict Responses to Multi-electrode Stimulation in the Retina PLoS Comput. Biol. 12 e1004849 [PubMed: 27035143]

[31]. Ferrari U, Deny S, Sengupta A, Caplette R, Trapani F, Sahel J-A, Dalkara D, Picaud S, Duebel J and Marre O 2020 Towards optogenetic vision restoration with high resolution PLoS Comput. Biol. 16 e1007857 [PubMed: 32667921]

[32]. Bortoff A 1964 Localization of slow potential responses in the Necturus retina Vision Res. 4 627–IN9 [PubMed: 4170193]

[33]. Schwartz O, Pillow JW, Rust NC and Simoncelli EP 2006 Spike-triggered neural characterization J. Vis. 6 13

[34]. Moskovitz TH, Roy NA and Pillow JW 2018 A comparison of deep learning and linear-nonlinear cascade approaches to neural encoding (BioRxiv: 463422)

[35]. Shah NP et al. 2019 (BioRxiv: 496422) Inference of nonlinear receptive field subunits with spike-triggered clustering eLife 9 e45743

[36]. Sharpee T, Rust NC and Bialek W 2004 Analyzing neural responses to natural signals: maximally informative dimensions Neural Comput. 16 223–50 [PubMed: 15006095]

[37]. Møller F, Laursen M, Tygesen J and Sjølie A 2002 Binocular quantification and characterization of microsaccades Graefes Arch. Clin. Exp. Ophthalmol. 240 765–70 [PubMed: 12271375]

[38]. Rumyantsev OI, Lecoq JA, Hernandez O, Zhang Y, Savall J, Chrapkiewicz R, Li J, Zeng H, Ganguli S and Schnitzer MJ 2020 Fundamental bounds on the fidelity of sensory cortical coding Nature 580 100–5 [PubMed: 32238928]

[39]. Sloan LL 1959 New test charts for the measurement of visual acuity at far and near distances Am. J. Ophthalmol. 48 807–13 [PubMed: 13831682]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript
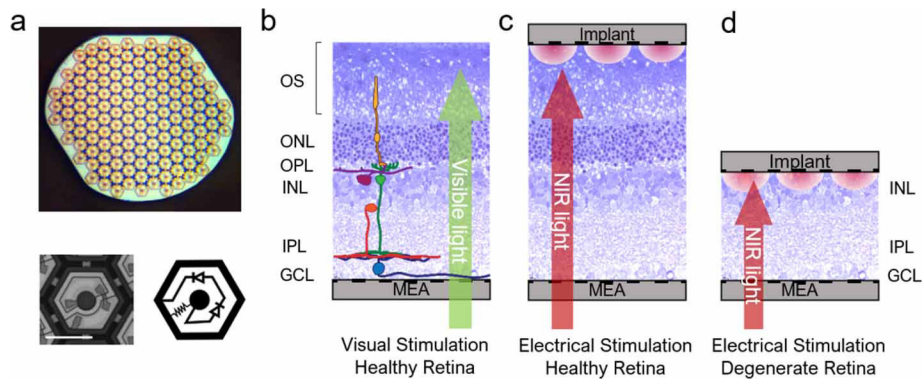
Figure 1.

Photovoltaic subretinal implant and stimulation scheme. (a) Top: implant of 1 mm in width with 70 μm pixels. Bottom left: a single photovoltaic pixel. Bottom right: circuit diagram of a pixel. (b) Visual stimulation of the healthy rat retina placed on top of a transparent MEA. (c) Electrical stimulation of the healthy retina by an implant placed on top of photoreceptors. (d) Electrical stimulation of the degenerate retina by an implant placed on top of the inner nuclear layer (INL).
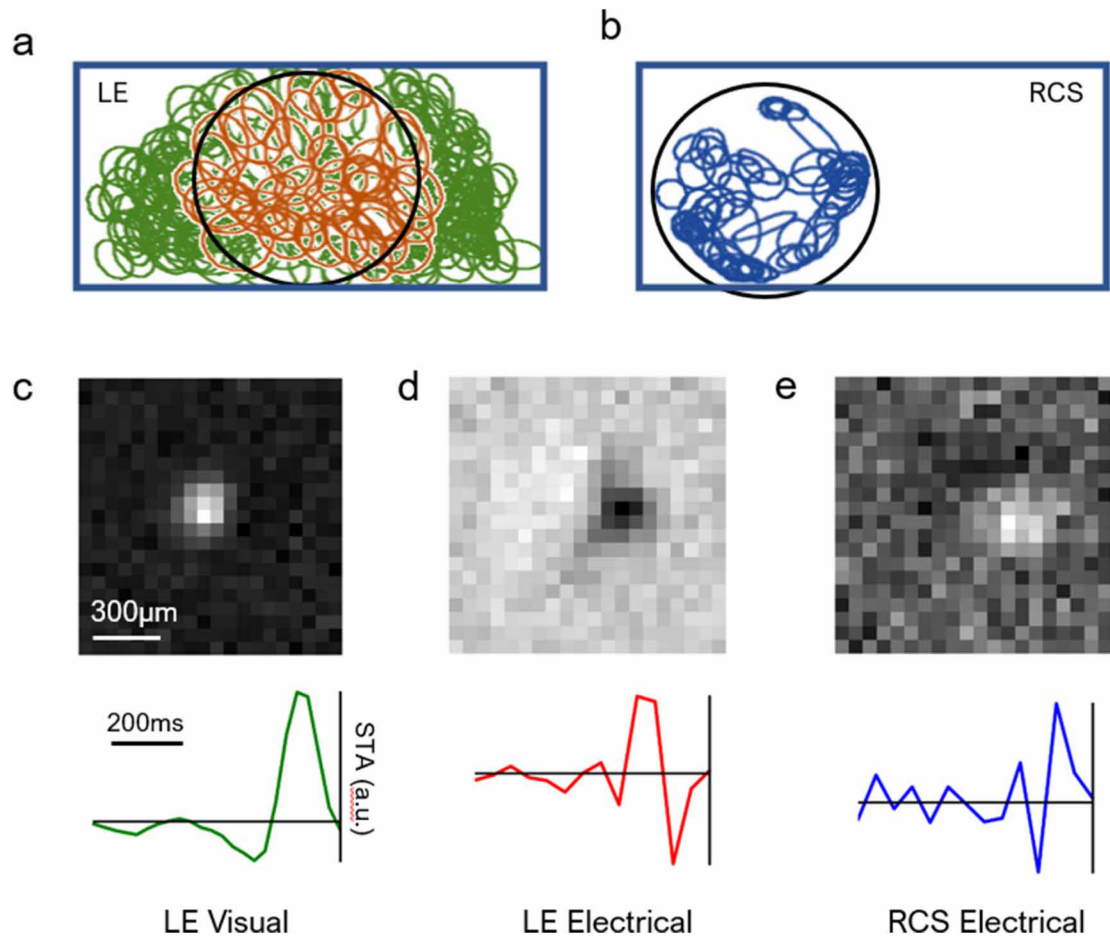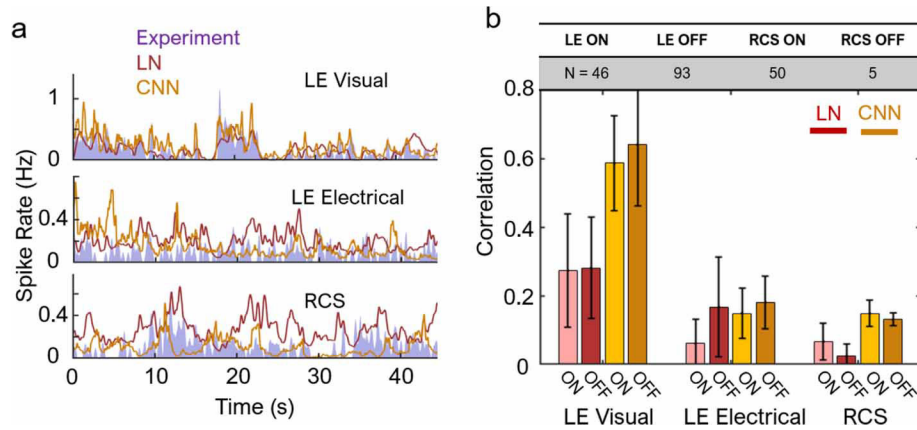
Figure 2.
MEA recording of the RGC responses to stimulation. (a) Mosaic of the natural receptive fields (RF) in LE retina. The 1 mm × 2 mm rectangle marks the boundary of the MEA, and the 1 mm circle indicates the edge of the implant. Due to limited precision of the implant mounting procedure, the photovoltaic array may not be centered on the MEA. Orange ellipses correspond to cells responding to both visual and electrical stimulation, and green ellipses correspond to cells responding only to visual stimulation. Only the orange cells were included in our subsequent analyses for LE retinas. (b) Mosaic of RFs in RCS retina upon electrical stimulation. RGCs outside the implant were insensitive to electrical stimulation, and hence are not presented here. (c) RF and the time course (STA) of a cell in the LE retina responding to visual white noise. (d) Same cell responding to electrical stimulation. (e) RF and the time course of an electrical ON cell in the RCS retina.
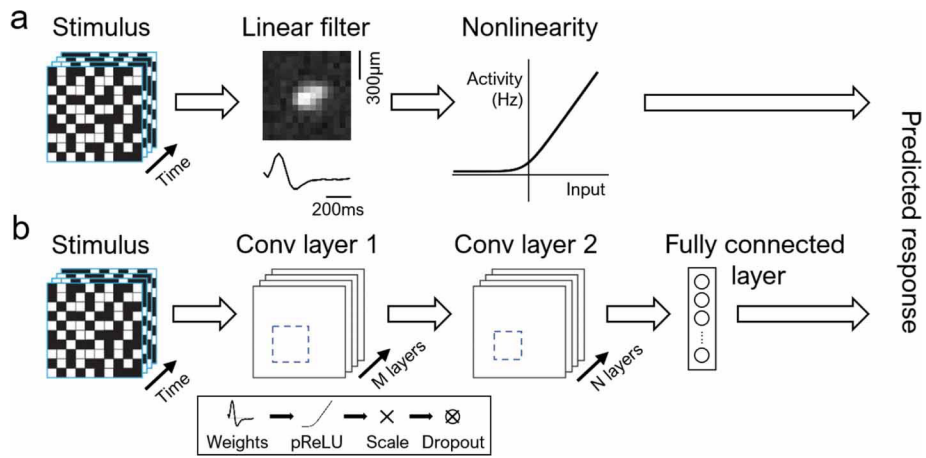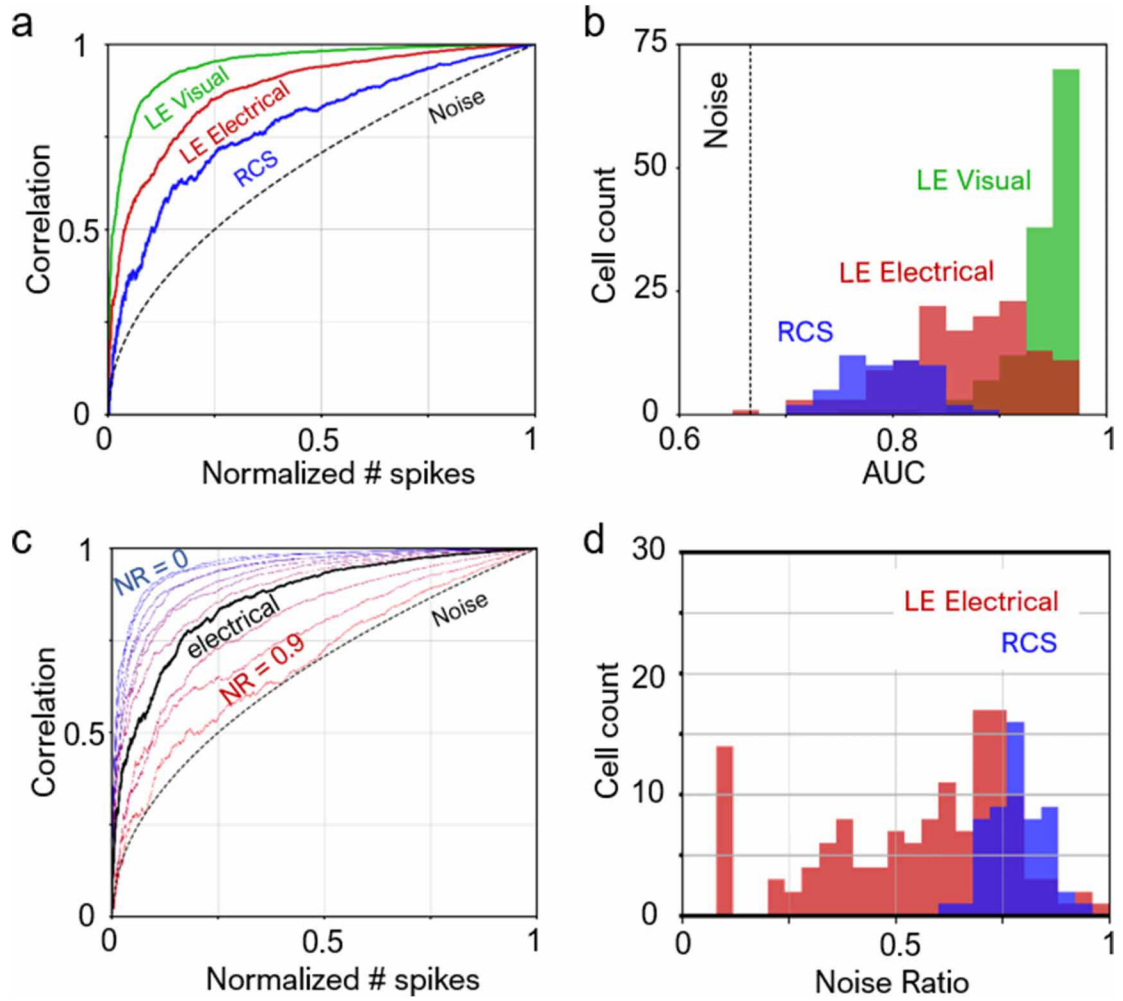
Figure 3.
Computational models for individual RGCs. (a) Linear–nonlinear model. (b) Convolutional neural network. Each convolutional layer consists of a sequence of linear filters (weights), parametric rectifying linear unit (pReLU), batch renormalization (scaling), and a dropout layer.
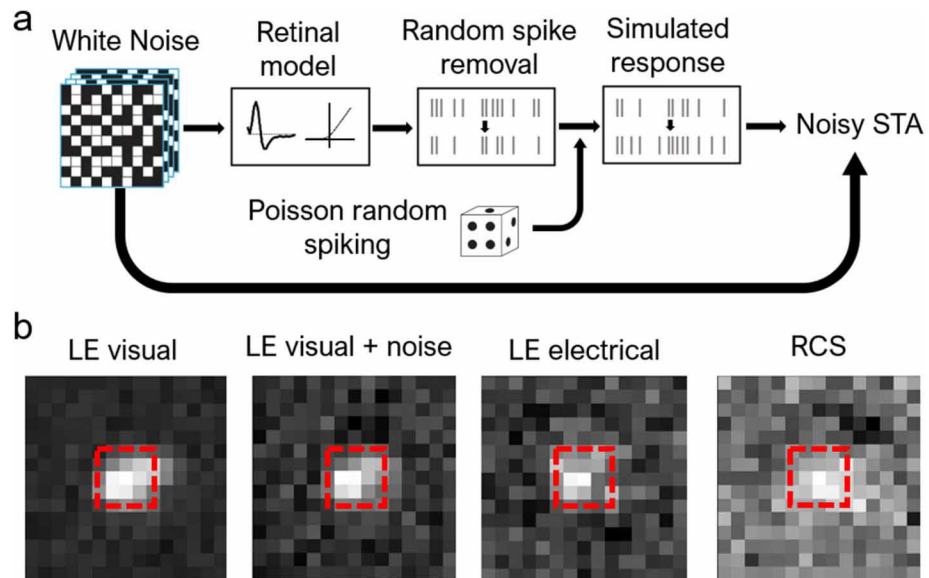
Figure 4.
Adding noise to STAs. (a) Algorithm used for adding noise to STAs. Spikes were first generated by a model of the visual response, and then partially removed and replaced with spikes generated through a Poisson process. The resulting simulated response was then used to re-compute the STA. (b) Examples of STAs under different stimulation and retina types. With properly chosen noise ratio, the noise-injected visual STA resembles that of the LE retina under electrical stimulation. The region bound by the red dash line was used for subsequent analyses.
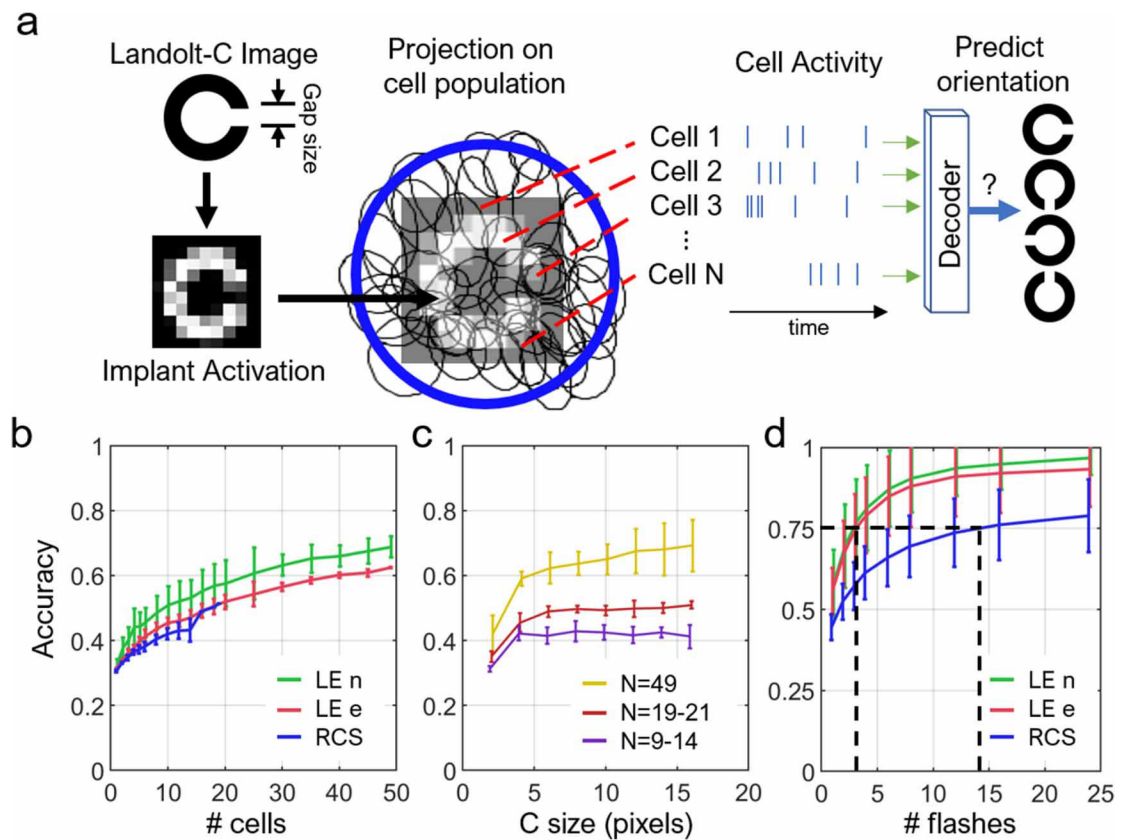
Figure 5.
Single-cell model predictions and its accuracy. (a) Examples of LN and CNN model predictions for a single cell, alongside the experimental data. Both models fit well the LE visual responses, but not the electrical stimulation of either LE or RCS retinas. (b) Correlation with the test set data, averaged over a population of cells (table). Model fit to electrically stimulated cells (~0.15 for CNN with LE and RCS data) is significantly worse than to the visual response (~0.6) ($p < 10^{-4}$ for both LN and CNN models). For all pairs of LN versus CNN for the same stimulus and cell type (e.g. LE visual LN vs LE visual CNN), $p < 10^{-4}$ (two-sample t-test), except for LE electrical, which was statistically insignificant ($p = 0.58$). Error bars indicate the standard deviation (s.d.) over the population of cells.

Figure 6.
Noise in retinal responses. (a) Characteristic correlation curves. A cell with little noise would have a curve farther away from the pure noise curve. As a corollary, the greater the area-under-curve (AUC), the less noisy is the cell. Each cell had a different total number of spikes used for computing their respective STAs. Normalized number of spikes refers to the ratio of included spikes to the total number of spikes (see *n/N* in equations (3)–(5)). (b) AUC for a population of cells. Almost all LE cells responding to visual stimulus exhibit less noise than all RCS cells under electrical stimulation. (c) Example correlation curves for LE visual with different added noise ratios (NR). At NR ~0.65, the noise-injected correlation curve matches that for LE electrical. (d) Noise ratios over the cell population. For LE electrical, the average noise ratio is $55.3 \pm 22.5\%$; for RCS, the noise ratio is $78.2 \pm 6.5\%$.

Figure 7.
Ensemble encoding of the visual information. (a) Algorithm for evaluation of the ensemble encoding accuracy. Using the cell activity under a certain projection of a Landolt-C, decoder discerns its orientation. Error bars represent the standard deviation of accuracies across retinas. (b) Accuracy versus number of cells included in the decoder, with a C size of 14 pixels. (c) Accuracy as a function of the letter C size. (d) Accuracy versus number of presentations (flashes) of the letter. Since electrical stimulation activates fewer cells in the RCS retina than visual stimulation in healthy retina, ~5 times more flashes are required to achieve the same decoding accuracy as natural. (four LE retinas, N cells = 20, 21, 49, 49; four RCS retinas, N cells = 9, 10, 13, 14).

Table 1.

Number of filters and their dimensions for each convolution block and for each stimulation type.

		Stimulation	
		Visual	**Electrical**
Layer	Convolution block 1	8 filters, 13×13 each	16 filters, 5×5 each
	Convolution block 2	16 filters, 9×9 each	32 filters, 5×5 each