

# Benefits of retinal image motion at the limits of spatial vision

**Kavitha Ratnam**

School of Optometry and Vision Science Graduate Group,  
University of California, Berkeley, Berkeley, California, USA

**Niklas Domdei**

Department of Ophthalmology, University of Bonn,  
Bonn, Germany

**Wolf M. Harmening\***

Department of Ophthalmology, University of Bonn,  
Bonn, Germany

**Austin Roorda\***

School of Optometry and Vision Science Graduate Group,  
University of California, Berkeley, Berkeley, California, USA

**Even during fixation, our eyes are constantly in motion, creating an ever-changing signal in each photoreceptor. Neuronal processes can exploit such transient signals to serve spatial vision, but it is not known how our finest visual acuity—one that we use for deciphering small letters or identifying distant faces and objects—is maintained when confronted with such change. We used an adaptive optics scanning laser ophthalmoscope to precisely control the spatiotemporal input on a photoreceptor scale in human observers during a visual discrimination task under conditions with habitual, cancelled or otherwise manipulated retinal image motion. We found that when stimuli moved, acuities were about 25% better than when no motion occurred, regardless of whether that motion was self-induced, a playback of similar motion, or an external simulation. We argue that in our particular experimental condition, the visual system is able to synthesize a higher resolution percept from multiple views of a poorly resolved image, a hypothesis that might extend the current understanding of how fixational eye motion serves high acuity vision.**

## Introduction

Because our eyes are never at rest, the human visual system has to incorporate methods that transform an ever-changing retinal signal into an acute percept. Theories of visual acuity postulated that fixational eye movements (FEM), small mostly involuntary movements that occur even when we fixate, may enhance fine spatial detail by means of a dynamic sampling process (Ahissar & Arieli, 2012; Arend, 1973; Averill &

Weymouth, 1925; Marshall & Talbot, 1942). Classic experiments, limited by the technology available at the time, were unable to support these hypotheses (Kelly, 1979; Riggs, Ratliff, Cornsweet, & Cornsweet, 1953; Tulunay-Keeseey, 1960; Tulunay-Keeseey, 1982; Tulunay-Keeseey & Jones, 1976). Recent work by Rucci, Iovin, Poletti, and Santini (2007), however, demonstrated benefits of FEM for spatial frequencies as high as 10 cycles/° of visual angle in visual stimuli, attributed to a reshaping of spatiotemporal properties by equalizing or “whitening” spatial energy across the temporal domain (Kuang, Poletti, Victor, & Rucci, 2012; Rucci et al., 2007). Whereas whitening, or spectral equalization, can account for improvements in perceived contrast of retinal images that are resolvable by the cone mosaic (Rucci & Victor, 2015), its benefit cannot be readily extrapolated to predict an improvement for discrimination of high-contrast retinal images at the acuity limit, where the spacing of individual photoreceptor cells is larger than the smallest features that need to be resolved. Humans can resolve optotypes at the 20/10 acuity level (minimum gap dimension of 0.5 arcmin or 2.5 microns in the human retina; maximum spatial frequency of 60 cycles/°) and beyond (Rossi & Roorda, 2010a), suggesting that spatial resolution is not necessarily capped at the structural sampling limit of the retina (Curcio, Sloan, Kalina, & Hendrickson, 1990). It therefore remains unclear how our everyday visual abilities operate at, and even transcend, such limits especially at the fovea, and whether FEM degrade or enhance this performance.

With recent improvements in eye tracking and stimulus delivery using adaptive optics scanning laser ophthalmoscopy (AOSLO), cone-targeted stimuli can

Citation: Ratnam, K., Domdei, N., Harmening, W. M., & Roorda, A. (2017). Benefits of retinal image motion at the limits of spatial vision. *Journal of Vision*, 17(1):30, 1–11, doi:10.1167/17.1.30.



be delivered to the retina with a positional accuracy of 0.15 arcmin, less than a third the diameter of a foveal cone (Arathorn, Stevenson, Yang, Tiruveedhula, & Roorda, 2013; Harmening, Tuten, Roorda, & Sincich, 2014; Roorda et al., 2002; Yang, Arathorn, Tiruveedhula, Vogel, & Roorda, 2010). By also correcting for ocular blur, the ability to deliver near diffraction-limited, retinally stabilized stimuli facilitates testing the effects of FEM on visual tasks at the photoreceptor level. We find that FEM indeed improve discrimination tasks for poorly resolved retinal images, and discuss possible explanations.

## Methods

### Subjects

In Experiments 1 and 2 (stimuli presented with AOSLO), subjects were four adults (three males, one female; ages 30–38 years), who had no known retinal disease, had normal visual acuity, and were naïve to the purpose of the study. Subject S4 was available for Experiment 1 only and was consequently omitted from subsequent experiments. For pupil dilation and cycloplegia, a drop of 1% Tropicamide solution was instilled into the test eye 15 min prior to testing. For Experiment 3 (stimuli presented on LCD), eight additional subjects were recruited (six naïve, two of the authors; five female, three male; ages: 25–37 years). Informed consent was obtained for each subject and experimental procedures adhered to the tenets of the Declaration of Helsinki.

### AOSLO imaging and stimulation

Combined adaptive optics imaging and micro-stimulation was used to present retina-contingent visual stimuli with high contrast to targeted and spatially tracked locations in cone-resolved retinae of human observers, a method described in detail elsewhere (Arathorn et al., 2007; Rossi & Roorda, 2010b). We describe the relevant features of this procedure briefly here. The light source was a supercontinuum laser (SuperK Extreme; NKT Photonics) that provided an infrared imaging wavelength of  $842 \pm 25$  nm (luminance of  $\sim 4$  cd/m<sup>2</sup>). Retinal images were generated by raster scanning a focused spot across the retina with horizontal and vertical scan rates of 16 kHz and 30 Hz respectively. High order aberrations in the light beam exiting the eye were measured with a Shack-Hartmann wavefront sensor, and a 144-actuator, 5.5 micron stroke deformable MEMs mirror (Boston Micromachines Corp) corrected the computed wavefront error. Corrected light was continuously captured by a photomultiplier tube,

whose voltage output combined with positional signals from the scanning mirrors created  $512 \times 512$  pixel videos with a framerate of 30 Hz. A 1 s retinal video was recorded with every stimulus presentation trial (500 trials per condition per subject).

Stimuli were encoded into the scanning raster via acousto-optic modulation operating at 20 MHz that switched off the laser beam at points in the raster corresponding to the stimulus location. This produced a visible light intensity decrement stimulus (dark “E” on a red background) with high contrast. Michelson contrast between full-on and full-off stimulation was 99.9%. The beam diameter for projecting the stimulus was 5.6 mm, so diffraction reduced the actual contrast between the three bars of the projected letter “E” to between 60%–75%, depending on the specific letter size chosen for each subject (see Results). Sampling resolutions for imaging and stimulus projection ranged from 0.12 and 0.16 arcmin per pixel, depending on the exact field size used for each subject. Due to the scan rate of the vertical scanner, stimuli were presented with a background frame rate of 30 Hz, which has been shown to elicit corresponding neural signals in LGN parvocellular neurons (Sincich, Zhang, Tiruveedhula, Horton, & Roorda, 2009). Since the stimulus’s exact retinal location was embedded into the imaging video, proper stimulus encoding and stabilization were verified during postprocessing (see also Supplementary Video S1). Retinal locations  $0.8^\circ$ – $1.3^\circ$  from the fovea along the horizontal meridian (nasally in S1 and S4; temporally in S2 and S3) were selected for testing (Figure 1A through C).

### Eye movement analysis and trial rejection

Fixational eye movements were analyzed offline from recorded AOSLO videos with image based techniques with an effective temporal sampling rate of 840 Hz (Stevenson, Roorda, & Kumar, 2010). Since microsaccades rarely occurred during the 750-ms stimulus presentation interval, known to occur with fine acuity tasks (Bridgeman & Palca, 1980; Kowler & Steinman, 1979), and to eliminate any confounding effects of microsaccadic suppression on stimulus visibility, trials in which microsaccades occurred were removed from further analysis. The stimulus trajectories and eye movement characteristics analyzed in this study are hence from drift and tremor motion only. For simplicity, we refer to those motions as fixational eye motion (FEM) throughout the paper.

Proper stimulus delivery could be verified in the AOSLO videos. Trials in which the stimulus was distorted, improperly stabilized, or not presented for the full 750-ms presentation duration were removed from further analysis.

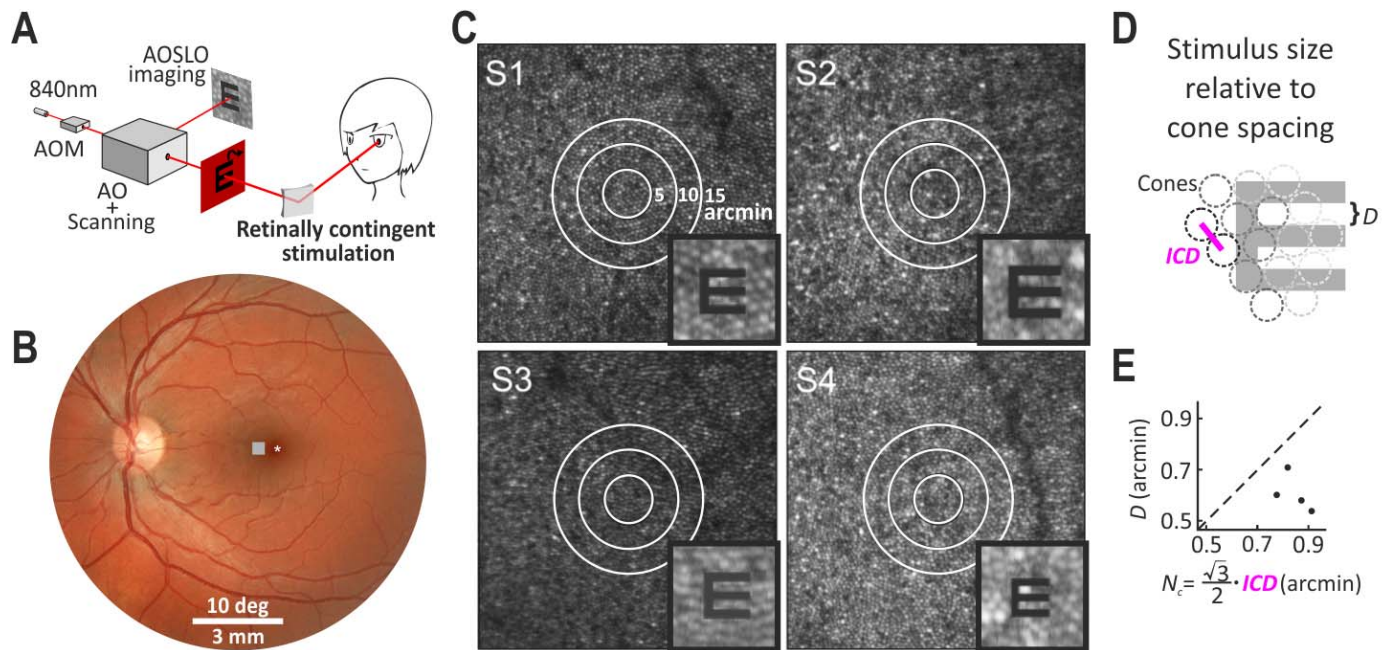


Figure 1. AOSLO microstimulation for projecting diffraction-limited stimuli to targeted retinal locations. (A) The AOSLO combines adaptive optics (AO) and high-speed scanning to record high-magnification videos of a human retina with cellular resolution. Optotypes (“E”) are projected directly onto the retina by modulating the scanning beam with a high-speed acousto-optic modulator (AOM). In this particular configuration, the subject sees a dark letter within a red square (840 nm light) that is generated by the raster scan. Real-time eye tracking is used to guide the placement of the retinal stimulus within the raster scan, enabling the delivery to targeted retinal locations (stabilized) or along any predefined path across the retina, independent of eye motion. (B) On an exemplary fundus photo, the position of test locations (gray field), placed  $\sim 1^\circ$  from the foveal center (asterisk), are shown. (C)  $1^\circ$  square AOSLO images of tested retinal locations in each subject (S1–S4). Concentric circles show retinal regions with 5 and 10 arcmin radii centered on the stimulus delivery location. Insets show  $\sim 5 \times 5$  arcmin regions overlaid with a letter stimulus shown to the scale used in the experiments. (D) Letter “E” superimposed on a hexagonal cone mosaic. The gap width between the bars of the letter “E” is indicated by  $D$ . The intercone distance is indicated by  $ICD$ . (E) At the Nyquist sampling limit, the separation between the rows of hexagonally packed cones ( $N_c$ , computed using the equation in panel E) would equal the gap width  $D$ . For all subjects in our experiments, the row to row spacing  $N_c$  was always less than the gap width,  $D$  (black dots in scatter plot).

## Experiment 1: Natural versus manipulated retinal motion

In Experiment 1, we compared visual discrimination under two manipulated retinal image motion conditions (stabilized and incongruent) with natural viewing. In natural viewing, retinal image motion due only to FEM was allowed to occur (Figure 2A, Figure 3A). In stabilized viewing, the stimulus was locked onto a targeted set of cones (Figure 2B, Figure 3B, and Supplementary Video S1). Incongruent motion was defined as retinal motion that was manipulated by the AOSLO stimulation procedure in the following way: Habitual FEM were detected and compensated for in real-time. Additionally, the stimulus followed a trajectory that was extracted from the subject’s earlier eye motion traces. Since subjects showed idiosyncratic eye movements, stimulus trajectories were randomly selected from a pool of the subject’s own eye movements. The resulting net movement of the stimulus relative to

the retina was thus a path similar to typical eye motion, but one that was incongruent to the eye motion happening at the time of presentation (Figure 3C).

In a four-alternative-forced-choice task, subjects reported the orientation of an “E” optotype of maximum negative contrast. As in conventional Snellen eye charts, the height and width of the letter were five times the line thickness. Subjects were instructed to gaze at a fixation laser target while attending to the acuity task. The parafoveal location was chosen for two reasons: (a) to prevent the subjects from trying to follow the stabilized stimulus, which would appear to move relative to the scanning raster, and (b) because image stabilization performance is more robust for retinal locations just off the fovea where cone photoreceptors are easier to resolve. The exact region of stimulated cones and statistics of how many cones were involved in the task was determined from acquired AOSLO videos for each subject (Figure 1C). The stimulus size was deliberately set to be undersampled



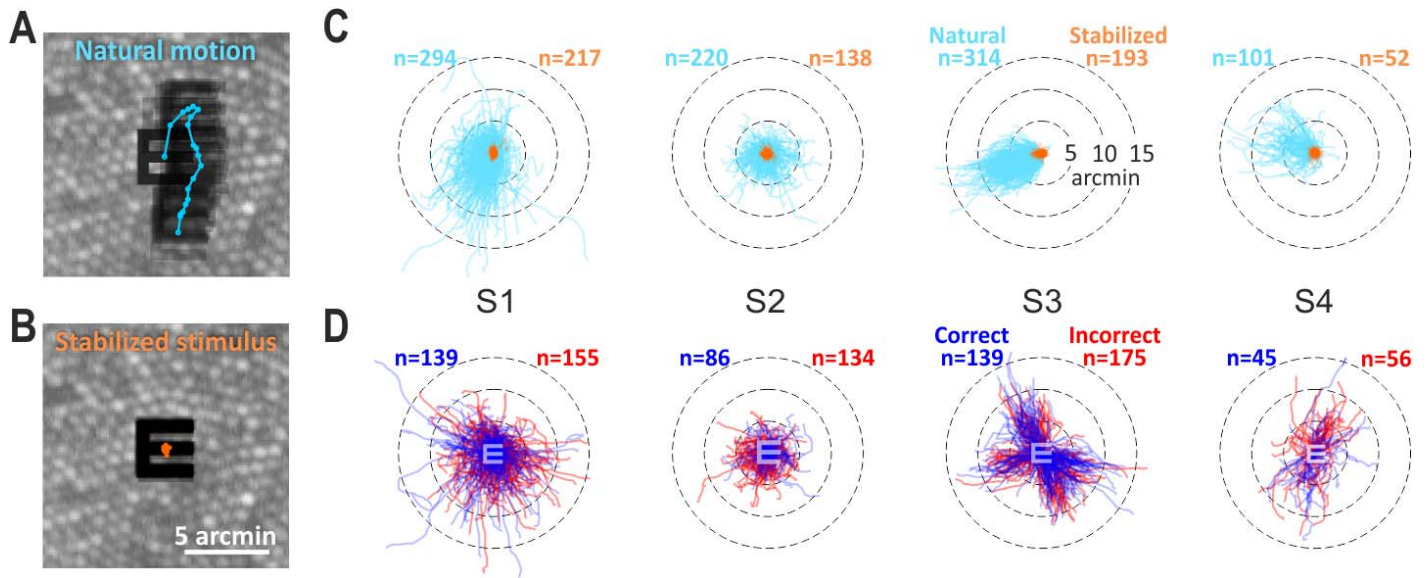


Figure 2. Retinal image motion due to FEM and motion manipulation. (A) Projected stimuli are directly encoded into the AOSLO video, allowing for an unambiguous record of the relative locations of the retina and the stimulus over the course of each trial. Here, the path of the stimulus over the course of one trial (duration: 750 ms, colored dots denote stimulus location in each of 23 video frames) of a naturally moving eye is shown. Due to fixational eye motion, the “E” moves over many photoreceptors. (B) When stimuli were presented stabilized, residual stimulus movement was smaller than the diameter of single cones. (C) Retinotopic stimulus trajectories for natural (blue) and stabilized (orange) conditions are shown across S1–S4; subjects exhibit idiosyncratic differences in FEM, sometimes with microneystagmus type orientation preferences (e.g., S3). Concentric circles represent 5, 10, and 15 arcmin radii of visual angle around the retinal location of stimulus starting location (compare Figure 1C). (D) Trajectories from the natural condition corresponding with correct (blue) and incorrect (red) psychophysical responses are replotted relative to stimulus orientation. There is no clear relation between how the stimulus is sampled and discrimination performance. The size of the letter for each subject is superimposed for reference.

by the cone mosaic, with an average number of cones sampling the image (number of cone centers within a convex hull surrounding the letter) at any given time being  $\sim 21$ . By comparison, the number of hexagonally packed cones required to sample the three bars of the letter “E” at the Nyquist limit is 23. Illustrations and the equation describing this relationship are in Figure 1D and E. Subjects were able to correctly discriminate about 40%–60% of letters presented at this stimulus size in an earlier acuity experiment. This performance range was selected to ensure subjects performed well above the 25% guessing rate but also below the performance plateau for the task. The stimulus gap size we used at this retinal location was, on average, about 0.6 arcmin, corresponding to a Snellen optotype of 20/12 which, for an “E,” has a dominant spatial frequency of 50 cycles/°. The stimulus was presented for 500 trials in each of the viewing conditions, pseudorandomly interleaved and split into ten experimental blocks over a two hour time span. Comparison of stabilized and incongruent motion with natural viewing was performed in two successive experimental sessions. After trial rejection (see previous section), 400–450 trials per condition remained for further analyses.

## Experiment 2: Contrast matching and discrimination performance

This experiment consisted of two parts, a contrast matching task and a discrimination task. For the contrast matching task, subjects matched perceptual contrast for stimuli presented under stabilized and incongruent motion conditions. The aim of this task was to quantify the extent of perceptual fading of the stabilized stimulus compared to the moving stimulus. Stimulus duration, size, and retinal location were identical to those of Experiment 1 with the exception that squares rather than letters were employed (Figure 4A). The use of a square allowed subjects to focus on stimulus contrast, the relevant attribute for this experiment, without unnecessarily attending to orientation. Two vertically offset squares were simultaneously presented, one retinally stabilized and one moving identical to the incongruent viewing condition. Simultaneous stimulus presentation ensured that both stabilized and incongruently moving squares would appear to move relative to the AOSLO scanning raster (Arathorn et al., 2013), so the stimuli would not be easily differentiated other than by their relative contrast. For each trial, subjects indicated which stimulus appeared

darker (i.e. had more contrast). The contrast of the stabilized square was held constant while the incongruently moving square's contrast was adjusted over repeated staircases to converge onto the value for which it appeared similar to the stabilized square.

A sequence of seven one-down-one-up staircases was employed; staircases terminated after seven reversals, and the threshold was calculated as the mean value from the last four reversals. For the first three staircases, both squares started at maximum physical contrast. The mean threshold value from these staircases was doubled and then used as the starting value for the moving square in the following three staircases. For the final staircase, the starting value was the mean of the previous six staircase thresholds, and the contrast step sizes were made smaller as to provide finer resolution when determining the final contrast value.

This reduced contrast value was then incorporated into the second part of Experiment 2, a discrimination task. The protocol for Experiment 1 was repeated except that discrimination of naturally moving, maximum contrast and naturally moving, reduced contrast "E" optotypes were compared. The maximum and reduced contrast conditions were pseudorandomly interleaved with 250 trials each, and subjects reported letter orientation.

### Experiment 3: External computer-based simulation

To better gauge the amount of visual information needed to benefit from image motion, a simulation of cone activation patterns derived from the AOSLO experiments was constructed and presented to an independent subject group in a separate discrimination experiment. Simulation stimuli were computed with custom written Matlab scripts (Figure 5A and B). First, a spatial representation of cone apertures was constructed using a randomly jittered hexagonal array with center-to-center distances equaling those for cone outer segment distances (Curcio et al., 1990). The light acceptance profile of each cone aperture was represented by a two-dimensional Gaussian whose full-width at half-maximum was 48% of the inner segment diameter for the mean eccentricity used in AOSLO experiments— $1^\circ$  (MacLeod, Williams, & Makous, 1992). A bitmap image of the "E" stimulus at threshold size was spatially convolved with a two-dimensional Gaussian to represent residual blur that is due to diffraction in an AOSLO system. This stimulus representation was then filtered by the cone array and summed across each cone aperture to find a model activation value (ranging from 0 to 1) for each cone. Next, the model cone array was replaced by a Voronoi diagram representing cone locations, i.e., finding

boundaries between each cone that minimize the distance from cone centers to any point of the boundary. Each Voronoi cell was assigned a gray value equal to the cone model activation value. The physical size of the simulated field was magnified on the computer screen such that visual acuity did not limit the task to identify the correct orientation of the "E" (stimulus gap size: 5 arcmin of visual angle). Eight subjects (six naïve, two of the authors) discriminated the orientation of such simulated "E" stimuli via a standard LC-display at 2 m viewing distance. A head and chinrest stabilized subject positioning. Subject responses were recorded via button presses on a computer keyboard. 150 trials for two viewing conditions, static and dynamic, were presented pseudorandomly interleaved. In static viewing, the location of cones and the stimulus was held constant. In dynamic viewing, the position of each cone relative to the stimulus was changed from frame to frame (30 frames/s) based on a motion path drawn from all paths recorded with the AOSLO in one of the subjects in Experiment 1. All subjects were presented with the same set of motion paths but in randomly permuted sequences. This motion produced the rendition of a static stimulus viewed through a moving set of Voronoi patches (see Supplementary Video S2). Intertrial progression was self-paced and stimulus presentation time was 750 ms.

## Results

### Nature of FEM during the acuity task and quality of stabilization

AOSLO imaging and micro-stimulation allowed us to study the exact nature of FEM during the specific task (Figure 2A) as well as to provide an unambiguous record of tracking performance (Figure 2B). It is important to note that the FEM behavior shown here represent FEM (a) that occur during 1-s epochs around the time that the stimulus is presented in a self-paced task, (b) that occur when the eye is fixating on a target while attending to the task in the near periphery, and (c) that do not contain microsaccades and poorly tracked trials, which comprised between 10% and 20% of the trials.

Observed in Experiment 1, of the total 929 analyzed trials with a naturally moving stimulus, retinal motion due to FEM exhibited idiosyncratic differences across subjects. Some subjects showed relatively random FEM directions between each trial, resulting in a more circular overall shape when all FEM trajectories are plotted superimposed (Figure 2C, subjects S1 and S2). FEM trajectories from S3 and S4 revealed a clear

preference for specific directions of motion during the task (Figure 2C, S3 and S4). Across subjects, absolute trajectory length averaged across all individual trials was similar. Relative to the underlying mosaic of photoreceptors, the stimulus traversed a retinal distance equaling about 10.5 unique cones during each 750-ms presentation during natural viewing (an example stimulus trajectory close to this average is shown in Figure 2A). In 600 analyzed trials under the stabilized condition, the residual stimulus motion that occurred due to imperfections of the tracking and stabilization techniques was small. Here, the stimulus traversed 0.4 cones on average across subjects. Expressed differently, stimulus trajectory amplitudes under stabilization were about 25 times less than in natural viewing (Figure 2C). This analysis confirmed that the exact same set of cones was stimulated during the stabilized condition, whereas many more cones were stimulated during natural viewing.

Given the nature of our orientation discrimination acuity task (four main orientations of the Snellen E), we wondered if the eye can adjust FEM relative to the orientation of the optotype to maximize transient information content (e.g. motion preferably perpendicular to the bar orientation), and whether specific motion traces offer advantages for the task compared to others. In Figure 2D the same motion paths as in Figure 2C are plotted, but now rotated relative to the orientation of the optotype orientation during presentation, and with indication of correct and incorrect psychophysical responses. We observed no clear trends in this analysis. In this short period of time the eye does not seem to adjust its FEM behavior according to the orientation of the letter, and certain directions of eye motion do not appear to confer clear benefits.

### Experiment 1: Discrimination benefits from FEM at the resolution limit

Discrimination performance with retinal image stabilization dropped on average by 23% across subjects (Figure 3D;  $p < 0.05$ , two-tailed binomial  $z$  test). Thus, fine spatial resolution was impaired in the absence of retinal image motion due to FEM, or visual acuity was enhanced by FEM. In fact, the visual resolution achieved in our experimental setup was higher than what simple spatial sampling models of the cone mosaic would predict. For each subject, the stimulus gap, or distance between adjacent bars of the “E,” was compared to the Nyquist limit ( $N_C$ ) of the tested retinal location (Figure 1E). The stimulus gap constitutes the primary image detail subjects use to discriminate orientation (Rossi & Roorda, 2010b). For each subject, the gap size was smaller than  $N_C$  (gap

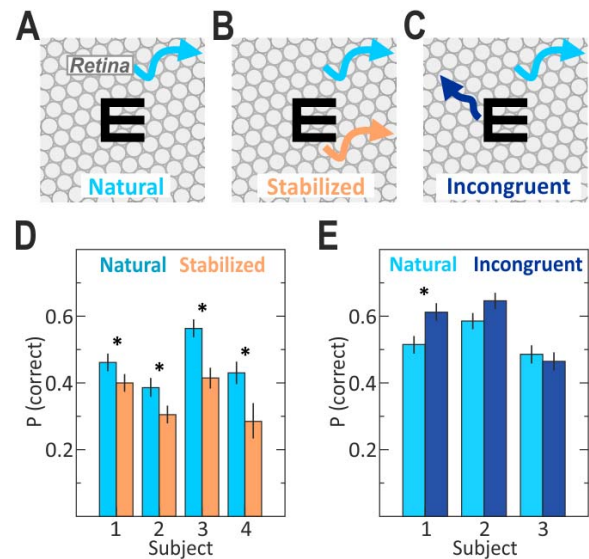


Figure 3. Stimulus motion improves acuity at the resolution limit. (A) In natural viewing, the stimulus (“E”) is fixed in space and the retinal cone mosaic (circles) moves due to fixational eye motion (FEM, light blue arrow). (B) In stabilized viewing, the stimulus moves with the retina (orange arrow), such that it stays locked on the same cones during presentation. (C) In the incongruent motion condition, the stimulus moves - while the eye performs its habitual FEM - in a path according to a previously recorded FEM trace. (D) Stimulus stabilization reduced discrimination performance in all subjects by an average of 23%. (E) Relative to the natural viewing condition, subjects performed equally well or better when incongruent motion was employed. Asterisk (\*) denotes  $p$  value  $< 0.05$ . Error bars are standard error of the mean.

size/ $N_C = 0.61/0.90$ ,  $0.74/0.85$ ,  $0.63/0.80$ ,  $0.57/0.94$  arcmin for S1 through S4, respectively).

Subjects performed similarly or better under the incongruent than under the natural condition (Figure 3E; S1,  $p < 0.01$ ; S2 and S3,  $p > 0.05$ ; two-tailed binomial  $z$  test,  $n = \sim 450$ ). These findings demonstrate that the visual system can benefit from retinal image motion even when the activity is independent of FEM at the time of stimulus presentation.

### Experiment 2: Contrast reduction during stabilization is not critical

To determine whether contrast was reduced under stabilization and how performance may have been affected, we devised a pair of experiments. The perceived contrast of stabilized versus moving stimuli was indeed reduced by about 20%, but performance was similar ( $p > 0.05$ , two-tailed binomial  $z$  test,  $n = \sim 250$ ) when subjects discriminated naturally moving stimuli presented at full and reduced (80%) contrast (Figure 4). These results suggest reduced contrast was



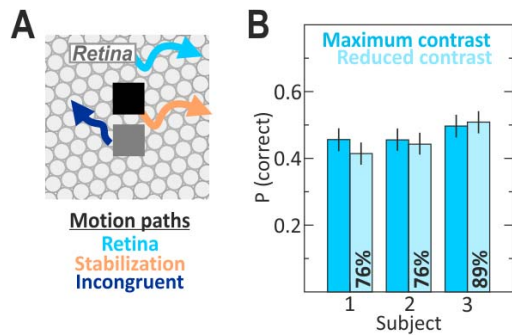


Figure 4. Contrast matching and discrimination at reduced contrast. (A) Two squares with identical dimensions to the “E” stimuli in Experiment 1 were simultaneously presented retinally stabilized and in an incongruent motion similar to subjects’ own eye movements. Over multiple staircases, the contrast of the moving square was updated until both squares appeared perceptually similar to the subject. These reduced contrast values—percentages indicated in (B)—were used in the second part of the experiment. (B) Discrimination performance for naturally moving, maximum contrast and naturally moving, reduced contrast “E”s were compared. Reduced contrast values, indicated as a percentage of maximum contrast, are shown for each subject. Subjects performed similarly for both conditions. Error bars are standard error of the mean.

not responsible for decreased performance under stabilized conditions.

### Experiment 3: Dynamic cone activation patterns suffice for discrimination benefit

We tested whether the dynamic information present at the photoreceptor level, effectively a series of poorly sampled “snapshots” of the stimulus, is sufficient for enhancing discrimination under natural motion conditions. Using an external monitor, subjects viewed simulations of cone excitation patterns resulting from moving and stabilized “E” optotypes (Figure 5A and B and Supplementary Video S2).

Subjects performed significantly worse, by an average of 27%, in the static simulation condition (Figure 5C; S1 and S3,  $p < 0.01$ ; remaining subjects,  $p < 0.001$ ; two-tailed binomial  $z$  test,  $n = 150$ ). This degradation was similar to the 23% performance drop observed in Experiment 1; performance ratios of the natural versus stabilized AOSLO experiment and dynamic versus static simulation experiment were not significantly different (mean ratio: 1.30 and 1.38, respectively;  $p = 0.57$ ; Wilcoxon rank sum test). These results reinforce the fact that the retinal output via dynamic cone excitation patterns, regardless of ongoing FEM, contains sufficient information to improve discrimination of images at the sampling resolution limit.

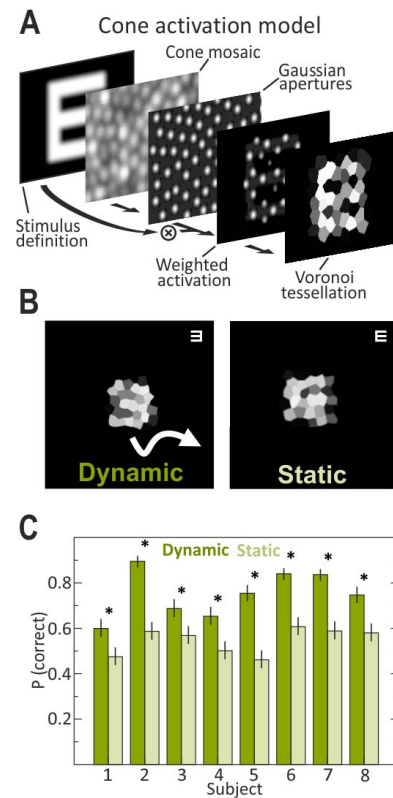


Figure 5. Modelled dynamic cone activation produces a similar benefit to feature discrimination as actual retinal motion. (A) A model of cone activation was derived by convolution of size-matched stimuli with a Voronoi patch of cone photoreceptor positions (see Methods for details). (B) Presented on a standard computer display, stimuli were either computed on a nonmoving model mosaic (Static), or on one that moved based on fixational eye movements from the AOSLO experiments (Dynamic); see also Supplementary Video S2. (C) Similar as in natural versus stabilized viewing, discrimination performance of all subjects dropped when stimuli were presented statically. Asterisk (\*) denotes  $p$  value  $< 0.05$ . Error bars are standard error of the mean.

## Discussion

Our results demonstrate that discrimination of high contrast optotypes at the eye’s resolution limit benefit from image motion that is similar to that caused by FEM, regardless of whether the motion was self-induced, a playback of earlier motion, or presented in an external simulation of cone activation. It should be reiterated that the benefits of eye motion observed in this study are restricted to those caused by ocular drift only. Subjects rarely exhibited microsaccades during the stimulus presentation interval, and any trials with microsaccades were removed to eliminate any confounding effects of microsaccadic suppression on stimulus visibility. In light of the current understanding

of the functional impact of FEM on vision, our findings offer cause to extend such theories.

Theories of spatial whitening postulate that the temporal modulations induced by FEM equalize the spectral power of natural images across spatial frequencies, effectively filtering out low-frequency image correlations and enhancing higher spatial frequencies. The temporal modulations induced by typical FEM amplitudes have been shown convincingly to improve contrast thresholds for stimuli up to 10 cycles/° in the presence of lower frequency noise or natural image statistics (Kuang et al., 2012; Rucci et al., 2007). Also, since stabilized stimuli fade due to neural adaptation (Ditchburn & Ginsborg, 1952; Riggs et al., 1953; Riggs & Ratliff, 1952), we first needed to explore the extent to which the reduced performance observed for the stabilized condition could have been due to a simple reduction in perceptual contrast of the stimulus. It is important to note here that stimuli delivered via the AOSLO raster scanner are, inherent to its mode of operation, continuously modulated at 30 Hz, corresponding to the system's frame rate (see Methods). It is known that such temporal modulation is preserved in visual signals up to postretinal stages, as those 30 Hz signals in neural activity, including those measured under stabilized stimulus conditions, have been observed in LGN parvocellular neurons (Sincich et al., 2009). Whereas this raster refresh rate may have mitigated the fading to some extent, perceptual fading of relatively stable, but flickering, stimuli is still known to occur (Schieting & Spillmann, 1987). The small degree of fading that we measured for the stabilized condition could not explain the overall reduction in performance under the stabilized condition, and we generally observed that in our specific task, contrast did not limit discrimination performance (Experiment 2). Moreover, our visual stimulus was undersampled by the photoreceptor array, a situation that is not explicitly considered in previous studies.

As such, and ruling out contrast as a contributing factor, it is not clear that whitening theories can readily explain the results of our study, and we suggest that alternative explanations or additional factors for how FEM enhance acuity might be needed. One plausible approach to the problem can be found in multiframe superresolution algorithms in the field of computer vision, in which a high-resolution image is reconstructed from a series of lower resolution frames, enabling the synthesis of images surpassing the spatial resolution of the original camera (Ben-Ezra, Zomet, & Nayar, 2005; Farsiu, Robinson, Elad, & Milanfar, 2004). These superresolution techniques include multi-exposure noise reduction, in which image signal-to-noise ratio is improved by averaging multiple exposures together, and subpixel image location, in which the centroid of a light distribution, blurred due to under-

sampling, can be computed with subpixel accuracy. Both mechanisms are feasible within the visual system. Shifter circuits (Anderson & Van Essen, 1987), interpolation circuits (Barlow, 1979; Crick, Marr, & Poggio, 1981), neural networks (Pitkow, Sompolinsky, & Meister, 2007), and neuronal phase locked loops (Ahissar & Arieli, 2012) have all been proposed as mechanisms by which the signals from a moving retinal image can be correctly synthesized. Compensation of the retinal image motion due to FEM may help provide a stable percept of the external world and also serve as a mechanism by which multiple-exposure noise reduction occurs in the visual system. It is also well documented that the visual system is capable of a form of subpixel resolution in a phenomenon termed hyperacuity (Westheimer, 1987), in which relative stimulus positioning can be judged at a resolution three to five times higher than the cone sampling limit (Klein & Levi, 1985), a feat which is also robust against retinal image motion (Westheimer & McKee, 1977).

Although fixational eye movements are large enough to be perceptually visible, our world appears stable (Murakami, 2003). On the other hand, if stimuli are presented with similar amplitudes of retinal image motion but incongruent with actual FEM, they are perceived as moving (Arathorn et al., 2013). In order to correct for ocular jitter and provide a stable percept of the external world, it has been suggested that the visual system decodes retinal signals relative to FEM (Burak, Rokni, Meister, & Sompolinsky, 2010; Coakley, 1983; Eizenman, Hallett, & Frecker, 1985; Shakhnovich, 1977). If spatiotemporal cone signals are also synthesized relative to ongoing FEM, then the benefits of retinal image motion may be restricted only to that induced by natural eye movements. In Experiment 1, we devised an incongruent viewing condition in which the stimulus moved in a retinal trajectory recorded from subjects' previous FEM. Given that discrimination performance was similar under natural and incongruent motion conditions, efference-based processes are unlikely to contribute to the integration of dynamic cone signals as occurred in the incongruent viewing condition. In any case, efference copies, which are generated at central motor stations, are not expected to have sufficient resolution for resolving spatial ambiguities of the fine details ( $< 1$  arcmin) presented in the current study (Havermann, Cherici, Rucci, & Lappe, 2014). Additionally, even if their resolution was sufficient, their involvement would not explain why similar performance was achieved in the natural and incongruent conditions. Therefore, afferent retina-based mechanisms for encoding motion of images on the retina, whether they arise from FEM or by some other means, such as proposed models involving elongated arrays of retinal ganglion cells (Ahissar, Ozana, & Arieli, 2015), would better



explain the preserved performance observed in the incongruent viewing condition.

We are currently developing a photoreceptor-based model of visual acuity that can support the empirical results demonstrated here, which should also help to better identify receptor and postreceptor mechanisms involved in dynamic acuity tasks. To test the extent to which the same benefits might be realized for challenging visual tasks under more natural viewing conditions, experiments using natural optics and stabilizing the image with a tracking scanning laser ophthalmoscope (no adaptive optics; Sheehy et al., 2012) are underway.

Our results, demonstrating benefits from eye motion at the visual acuity limits, may have some practical implications. It may explain certain visual behaviors in patients with retinal disease, such as how patients with retinal degenerative diseases maintain excellent visual acuity despite massive reductions in foveal cone density (Ratnam, Carroll, Porco, Duncan, & Roorda, 2013) or how the increased FEM in patients with central vision loss could be an adaptive mechanism to reap the same dynamic benefits for the larger receptive fields outside the fovea (Hennig & Worgotter, 2004). The extent of cone activation during FEM may also serve as a biomimetic principle for the refinement of image processing algorithms in computer vision and the design of retinal prosthetics (Dagnelie, 2012). Finally, it can explain why it takes time to recognize the finest features in a visual scene or to reach maximum performance on a visual acuity task (Baron & Westheimer, 1973).

## Supplemental material

Supplementary Video S1 is an example of AOSLO videos showing natural (left), stabilized (center), and incongruent (right) conditions for Experiments 1 and 2. The videos show how the stimulus is directly encoded into the videos, allowing an unambiguous record of eye motion and stimulus position for each trial. The white crosshairs were encoded in the videos during recording and were not visible to the subjects. Stimuli were presented at 30 Hz.

Supplementary Video S2 is an example of the dynamic simulation condition from Experiment 3 for each letter orientation. Stimuli were displayed for 750 ms during the experiment and are shown here for a longer duration for ease of viewing. The static condition would have consisted of a single frame of these videos and can be experienced by pausing the video at any time.

*Keywords:* adaptive optics, micro-stimulation, fixational eye motion, photoreceptor sampling

## Acknowledgments

This research was supported by EY023591 (AR); Foundation Fighting Blindness (AR); Deutsche Forschungsgemeinschaft DFG: Ha5323/3-1 (WMH), Ha5323/4-1 (WMH), Ha5323/5-1 (WMH); and Carl Zeiss Förderfonds (WMH).

\*WMH and AR contributed equally to this article. Commercial relationships: A.R. has a patent (USPTO #7118216) licensed to Canon USA Inc. and Boston Micromachines Corp. Both he and the companies stand to benefit financially from publication of these results. Corresponding author: Wolf M. Harmening. Email: wolf.harmening@ukb.uni-bonn.de. Address: Department of Ophthalmology; University of Bonn, Bonn, Germany.

## References

- Ahissar, E., & Arieli, A. (2012). Seeing via miniature eye movements: A dynamic hypothesis for vision. *Frontiers in Computational Neuroscience*, 6, 89.
- Ahissar, E., Ozana, S., & Arieli, A. (2015). 1-D vision: Encoding of eye movements by simple receptive fields. *Perception*, 44, 986–994.
- Anderson, C. H., & Van Essen, D. C. (1987). Shifter circuits: A computational strategy for dynamic aspects of visual processing. *Proceedings of the National Academy of Sciences, USA*, 84, 6297–6301.
- Arathorn, D. W., Stevenson, S. B., Yang, Q., Tiruveedhula, P., & Roorda, A. (2013). How the eye sees a stable and moving world. *Journal of Vision*, 13(10):22, 1–19, doi:10.1167/13.10.22. [PubMed] [Article]
- Arathorn, D. W., Yang, Q., Vogel, C. R., Zhang, Y., Tiruveedhula, P., & Roorda, A. (2007). Retinally stabilized cone-targeted stimulus delivery. *Optics Express*, 15, 13731–13744.
- Arend, L. E. (1973). Spatial differential and integral operations in human vision: Implications of stabilized retinal image fading. *Psychological Review*, 80, 374–395.
- Averill, H. I., & Weymouth, F. W. (1925). Visual perception and the retinal mosaic, II. The influence of eye movements on the displacement threshold. *Journal of Comparative Psychology*, 5, 147–176.
- Barlow, H. B. (1979). Reconstructing the visual image in space and time. *Nature*, 279, 189–190.
- Baron, W. S., & Westheimer, G. (1973). Visual acuity

- as a function of exposure duration. *Journal of the Optical Society of America* 63, 212–219.
- Ben-Ezra, M., Zomet, A., & Nayar, S. K. (2005). Video super-resolution using controlled subpixel detector shifts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27, 977–987.
- Bridgeman, B., & Palca, J. (1980). The role of microsaccades in high acuity observational tasks. *Vision Research*, 20, 813–817.
- Burak, Y., Rokni, U., Meister, M., & Sompolinsky, H. (2010). Bayesian model of dynamic image stabilization in the visual system. *Proceedings of the National Academy of Sciences, USA*, 107, 19525–19530.
- Coakley, D. (1983). *Minute eye movement and brain stem function*. Boca Raton, FL: CRC Press.
- Crick, F. H. C., Marr, D. C., & Poggio, T. (1981). An information-processing approach to understanding the visual cortex. In F. O. Schmitt (Ed.), *The organization of the cerebral cortex* (pp. 505–533). Cambridge, MA: MIT Press.
- Curcio, C. A., Sloan, K. R., Kalina, R. E., & Hendrickson, A. E. (1990). Human photoreceptor topography. *Journal of Comparative Neurology*, 292, 497–523.
- Dagnelie, G. (2012). Retinal implants: Emergence of a multidisciplinary field. *Current Opinion in Neurology*, 25, 67–75.
- Ditchburn, R. W., & Ginsborg, B. L. (1952). Vision with a stabilized retinal image. *Nature*, 170, 36–37.
- Eizenman, M., Hallett, P. E., & Frecker, R. C. (1985). Power spectra for ocular drift and tremor. *Vision Research*, 25, 1635–1640.
- Farsiu, S., Robinson, D., Elad, M., & Milanfar, P. (2004). Fast and robust multi-frame super-resolution. *IEEE Transactions on Image Processing*, 13, 1327–1344.
- Harmening, W. M., Tuten, W. S., Roorda, A., & Sincich, L. C. (2014). Mapping the perceptual grain of the human retina. *Journal of Neuroscience*, 34, 5667–5677.
- Havermann, K., Cherici, C., Rucci, M., & Lappe, M. (2014). Fine-scale plasticity of microscopic saccades. *Journal of Neuroscience*, 34, 11665–11672.
- Hennig, M. H., & Worgotter, F. (2004). Eye micro-movements improve stimulus detection beyond the Nyquist limit in the peripheral retina. In S. Thrun, L. K. Saul, and B. Scholkopf (Eds.), *Advances in neural information processing systems* 16, (pp. 1475–1482). Cambridge, MA: MIT Press.
- Kelly, D. H. (1979). Motion and vision. I. Stabilized images of stationary gratings. *Journal of the Optical Society of America*, 69, 1266–1274.
- Klein, S. A., & Levi, D. M. (1985). Hyperacuity thresholds of 1 sec: Theoretical predictions and empirical validation. *Journal of the Optical Society of America*, 2, 1170–1190.
- Kowler, E., & Steinman, R. M. (1979). Miniature saccades: Eye movements that do not count. *Vision Research*, 19, 105–108.
- Kuang, X., Poletti, M., Victor, J. D., & Rucci, M. (2012). Temporal encoding of spatial information during active visual fixation. *Current Biology*, 22, 510–514.
- MacLeod, D. I., Williams, D. R., & Makous, W. (1992). A visual nonlinearity fed by single cones. *Vision Research*, 32, 347–363.
- Marshall, W. H., & Talbot, S. A. (1942). Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. *Biological Symposia—Visual Mechanisms*, 7, 117–164.
- Murakami, I. (2003). Illusory jitter in a static stimulus surrounded by a synchronously flickering pattern. *Vision Research*, 43, 957–969.
- Pitkow, X., Sompolinsky, H., & Meister, M. (2007). A neural computation for visual acuity in the presence of eye movements. *PLoS Biology*, 5, e331.
- Ratnam, K., Carroll, J., Porco, T. C., Duncan, J. L., & Roorda, A. (2013). Relationship between foveal cone structure and clinical measures of visual function in patients with inherited retinal degenerations. *Investigative Ophthalmology & Visual Science*, 54, 5836–5847. [PubMed] [Article]
- Riggs, L. A., & Ratliff, F. (1952). The effects of counteracting the normal movements of the eye. *Journal of the Optical Society of America*, 42, 872–873.
- Riggs, L. A., Ratliff, F., Cornsweet, J. C., & Cornsweet, T. N. (1953). The disappearance of steadily fixated visual test objects. *Journal of the Optical Society of America*, 43, 495–501.
- Roorda, A., Romero-Borja, F., Donnelly, W., Queener, H., Hebert, T., & Campbell, M. (2002). Adaptive optics scanning laser ophthalmoscopy. *Optics Express*, 10, 405–412.
- Rossi, E. A., & Roorda, A. (2010a). Is visual resolution after adaptive optics correction susceptible to perceptual learning? *Journal of Vision*, 10(12):11, 1–14, doi:10.1167/10.12.11. [PubMed] [Article]
- Rossi, E. A., & Roorda, A. (2010b). The relationship between visual resolution and cone spacing in the human fovea. *Nature Neuroscience*, 13, 156–157.
- Rucci, M., Iovin, R., Poletti, M., & Santini, F. (2007).

- Miniature eye movements enhance fine spatial detail. *Nature*, *447*, 851–854.
- Rucci, M., & Victor, J. (2015). The unsteady eye: An information processing stage, not a bug. *Trends in Neuroscience*, *38*, 195–206.
- Schietering, S., & Spillmann, L. (1987). Flicker adaptation in the peripheral retina. *Vision Research*, *27*, 277–284.
- Shakhnovich, A. R. (1977). *The brain and regulation of eye movement*. New York, NY: Plenum.
- Sheehy, C. K., Yang, Q., Arathorn, D. W., Tiruveedhula, P., de Boer, J. F., & Roorda, A. (2012). High-speed, image-based eye tracking with a scanning laser ophthalmoscope. *Biomedical Optics Express*, *3*, 2612–2622.
- Sincich, L. C., Zhang, Y., Tiruveedhula, P., Horton, J. C., & Roorda, A. (2009). Resolving single cone inputs to visual receptive fields. *Nature Neuroscience*, *12*, 967–969.
- Stevenson, S. B., Roorda, A., & Kumar, G. (2010). Eye tracking with the adaptive optics scanning laser ophthalmoscope. In: S.N. Spencer (Ed.), *Proceedings of the 2010 symposium on eye-tracking research and applications* (pp. 195–198). New York, NY: Association for Computed Machinery.
- Tulunay-Keesey, U. (1960). Effects of involuntary eye movements on visual acuity. *Journal of the Optical Society of America*, *50*, 769–774.
- Tulunay-Keesey, U. (1982). Fading of stabilized retinal images. *Journal of the Optical Society of America*, *72*, 440–447.
- Tulunay-Keesey, U., & Jones, R. M. (1976). The effect of micromovements of the eye and exposure duration on contrast sensitivity. *Vision Research*, *16*, 481–488.
- Westheimer, G. (1987). Visual acuity and hyperacuity: Resolution, localization, and form. *Journal of the Optical Society of America*, *64*, 567–574.
- Westheimer, G., & McKee, S. P. (1977). Integration regions for visual hyperacuity. *Vision Research*, *17*, 89–93.
- Yang, Q., Arathorn, D. W., Tiruveedhula, P., Vogel, C. R., & Roorda, A. (2010). Design of an integrated hardware interface for AOSLO image capture and cone-targeted stimulus delivery. *Optics Express*, *18*, 17841–17858.