A paradoxical misperception of relative motion

Josephine C. D'Angelo[a] (ID), Pavan Tiruveedhula[a] (ID), Raymond J. Weber[b] (ID), David W. Arathorn[b] (ID), and Austin Roorda[a,1] (ID)

Detecting the motion of an object relative to a world-fixed frame of reference is an exquisite human capability [G. E. Legge, F. Campbell, *Vis. Res.* **21**, 205–213 (1981)]. However, there is a special condition where humans are unable to accurately detect relative motion: Images moving in a direction consistent with retinal slip where the motion is unnaturally amplified can, under some conditions, appear stable [D. W. Arathorn, S. B. Stevenson, Q. Yang, P. Tiruveedhula, A. Roorda, *J. Vis.* **13**, 22 (2013)]. We asked: Is world-fixed retinal image background content necessary for the visual system to compute the direction of eye motion, and consequently generate stable percepts of images moving with amplified slip? Or, are nonvisual cues sufficient? Subjects adjusted the parameters of a stimulus moving in a random trajectory to match the perceived motion of images moving contingent to the retina. Experiments were done with and without retinal image background content. The perceived motion of stimuli moving with amplified retinal slip was suppressed in the presence of a visible background; however, higher magnitudes of motion were perceived under conditions when there was none. Our results demonstrate that the presence of retinal image background content is essential for the visual system to compute its direction of motion. The visual content that might be thought to provide a strong frame of reference to detect amplified retinal slips, instead paradoxically drives the misperception of relative motion.

motion perception | eye movements | adaptive optics

The most sensitive judgments of motion require the presence of a world-fixed reference object near the moving target, like detecting a moving satellite among the stars or a bug crawling across a tree trunk. In a study exploring this discriminative ability, Legge and Campbell (1) found that in conditions with a frame of reference, displacement thresholds were as low as 0.3 arcminutes. By comparison, the center-to-center spacing of foveal cones is approximately 0.5 arcminutes (2). Detecting relative motion is therefore considered a hyperacuity, which is a class of visual tasks where the motion or displacement detection thresholds are smaller than the sampling limits of the photoreceptors (3). In the absence of a frame of reference, motion judgments are impaired (1, 4).

Yet, there are at least two reports of a special condition under which humans fail to accurately detect relative motion. Riggs et al. (5) used an optical lever technique to either stabilize the retinal image or to amplify the retinal slip. They reported that under the amplified conditions, stimuli appeared to be "locked in place." The authors never quantified the perceived motion under these amplified conditions nor did they ever do follow-up studies on this observation. Years later, Arathorn et al. (6) found something similar when they used a system (an earlier version of the system used in this report) to directly project images onto the retina in a retina-contingent manner. They found that under conditions where the retinal slip was unnaturally amplified, the retinal image appeared stable, despite the presence of unavoidable, world-fixed retinal background content that might have served as a frame of reference.

This so-called "illusion of relative stability" for all images moving in the direction of retinal slip, regardless of amplitude, suggests that the visual system knows its direction of motion and that anything moving in a direction consistent with the direction of retinal slip is rendered in the percept to be stable. This is an important extension of the well-known fact that, despite ever-present fixational eye movements which are large enough to make details of the scene sweep across the retina by a detectable amount, the world appears stable. The mechanisms for perceptual visual stabilization are unknown. How is the direction of eye motion determined? Specifically, is world-fixed retinal image background content needed to compute the direction of eye motion or are nonvisual cues (e.g. efferent copy) sufficient? Arathorn et al. (6) were unable to answer these questions because they could not adequately control or remove world-fixed retinal content from the visual scene. We have addressed that limitation using an updated system wherein we can fully control or remove retinal image content. We designed a method-of-adjustment

Author affiliations: [a]Herbert Wertheim School of Optometry and Vision Science, University of California, Berkeley, CA 94720; and [b]Electrical and Computer Engineering Department, Montana State University, Bozeman, MT 59717-3780

[1]To whom correspondence may be addressed. Email: aroorda@berkeley.edu.

experiment to quantify the perceived motion of stimuli moving in different directions with respect to eye motion in conditions with and without retinal image background content.

Our findings indicate that the sensory signals that inform the visual system about its direction of motion are retinal-image based. In conditions with retinal image background content, the perceived magnitude of motion of images moving with amplified retinal slip was significantly lower than that of images moving in the same direction as eye motion. When we performed the same experiments with no visible background—an unusual circumstance in the real world—we found that these perceptions reversed; images with increased retinal slip were perceived to have a high magnitude of motion while images moving in the same direction were perceived as having little to no motion. The retinal background content—that would normally be considered to provide a strong frame of reference for detecting motion—paradoxically drives the misperception of relative motion.

Results

We measured the perceived motion of stimuli moving with increased retinal slip (Gain −1.5) and moving in the same direction as retinal motion (Gain +1.5). The definition of the Gains are shown in Fig. 1B. Because the motion of these stimuli is dependent on retinal motion, we call these "retina-contingent stimuli." The specific Gains were chosen for two reasons: First, because they both generate the same amount of world motion for any given retinal motion, and second, because if we chose Gain +1 and −1, the Gain +1 condition, being stabilized on

the retina, would quickly fade from view and preclude reliable matching. We additionally tested world-fixed stimuli (Gain 0) as a control. Motion perceptions were recorded under two conditions: in the presence of rich retinal image background content (Fig. 1D) and with no visual cues (Fig. 1E). In this study, we refer to these conditions as "background-present" and "background-absent," respectively.

The experiment ran as follows (full details in *Materials and Methods*). Subjects simultaneously viewed a world-fixed projector display and a second display, delivered through an adaptive optics scanning light ophthalmoscope or AOSLO (Fig. 1A). For each presentation, the subject held their gaze on a projector-generated fixation cross while attending to the motion of an AOSLO-generated circular stimulus presented 2° in the horizontal periphery. The retina-contingent stimulus was presented in the first time interval. Following a 500-ms interstimulus interval, the stimulus in the second interval moved with preprogrammed random walk offsets, independent of eye motion, with a controllable magnitude quantified by its diffusion constant; see *Materials and Methods* and Eq. 1.

The subject's task was to observe the motion in each interval and then adjust the diffusion constant of the stimulus in the second interval (random walk stimulus) until its motion looked equivalent to that of the stimulus in the first interval (retina-contingent stimulus). No specific criteria were given for making a match and subjects expressed no difficulty or frustration to achieve one. Subjects could view as many self-initiated presentations as they wished for each Gain condition while adjusting the match. They were instructed to view their final match at
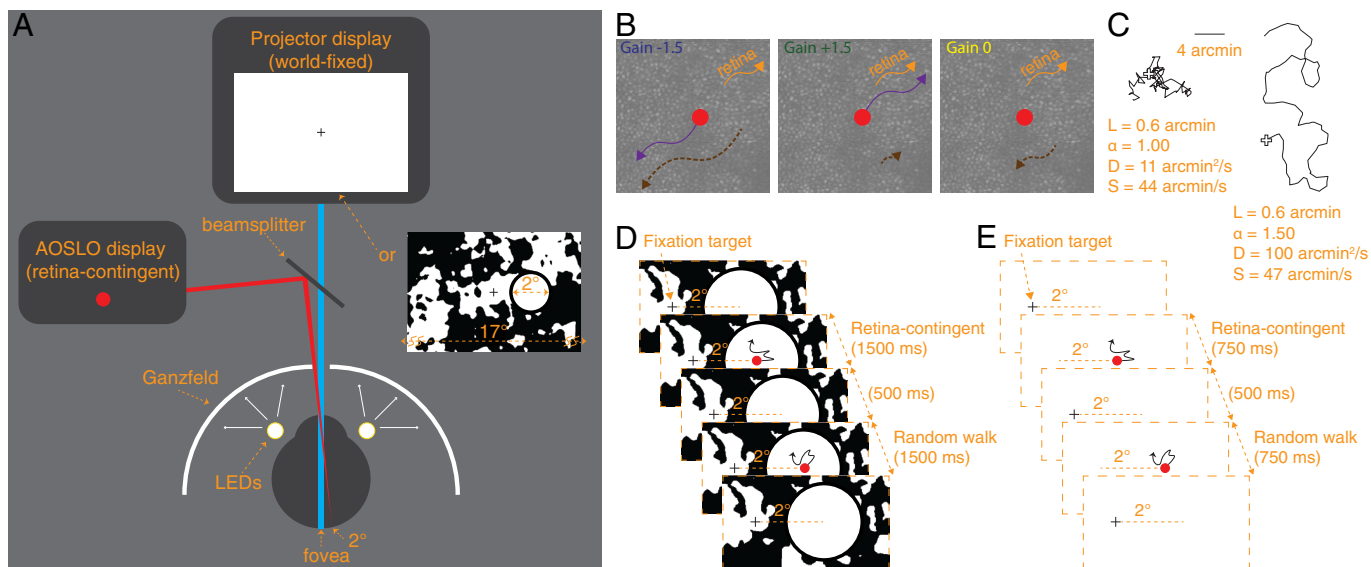
Fig. 1. Display configuration. (A) The projector drew a fixation target and the surrounding 17° field displayed either white light or a binarized noise pattern. The AOSLO generated the stimuli positioned 2° temporal to the fovea (nasal field). These images were simultaneously projected onto the retina through a beamsplitter. (B) Rules for the three retina-contingent stimuli. Each panel shows an identical retinal trajectory indicated by the orange arrows. The purple arrows indicate the stimulus' trajectory in the world. A Gain −1.5 stimulus moves with increased retinal slip, a Gain +1.5 stimulus moves in advance of retinal motion, and a Gain 0 stimulus is world-fixed. The dashed-brown arrows indicate the stimulus' trajectory across the retina. A Gain −1.5 stimulus moves 2.5× more on the retina than a Gain 0 stimulus, and a Gain 0 stimulus moves 2× more on the retina than a Gain +1.5 stimulus. (C) Simulated trajectories showing uncorrelated, Brownian motion (*Left*) and positively correlated, persistent motion (*Right*). The white cross indicates the starting position for each trace. Both paths were generated from the same average step length, L = 0.6 arcminutes, and have similar speeds (S), yet have different diffusion constants (D) due to different degrees of persistence, indicated by the value α. Full descriptions of these parameters are in *Materials and Methods*. (D and E) Experiment sequence. Subjects fixated on the target and attended to the stimuli positioned 2° temporally. The retina-contingent stimulus moved contingent to each subject's idiosyncratic fixational eye motion. The random walk stimulus moved independent of eye motion, with pregenerated random offsets. Under background-present conditions (D), the projector field displayed a binarized noise pattern which changed after each presentation and the fixation target remained on for the entire duration. Under background-absent conditions (E), a "Ganzfeld effect" was achieved by setting up a white paper with an aperture in front of the display permitting only the AOSLO and projector beams to enter the eye, shown in (A). LEDs were taped around the eye to illuminate the paper and the luminance was adjusted so that the subject saw only the stimuli in a white full-field surround. Owing to its proximity to the eye, the natural blur of the aperture rendered the transition between the display and luminance-matched paper invisible. The fixation target was timed to turn off when the stimuli were presented.
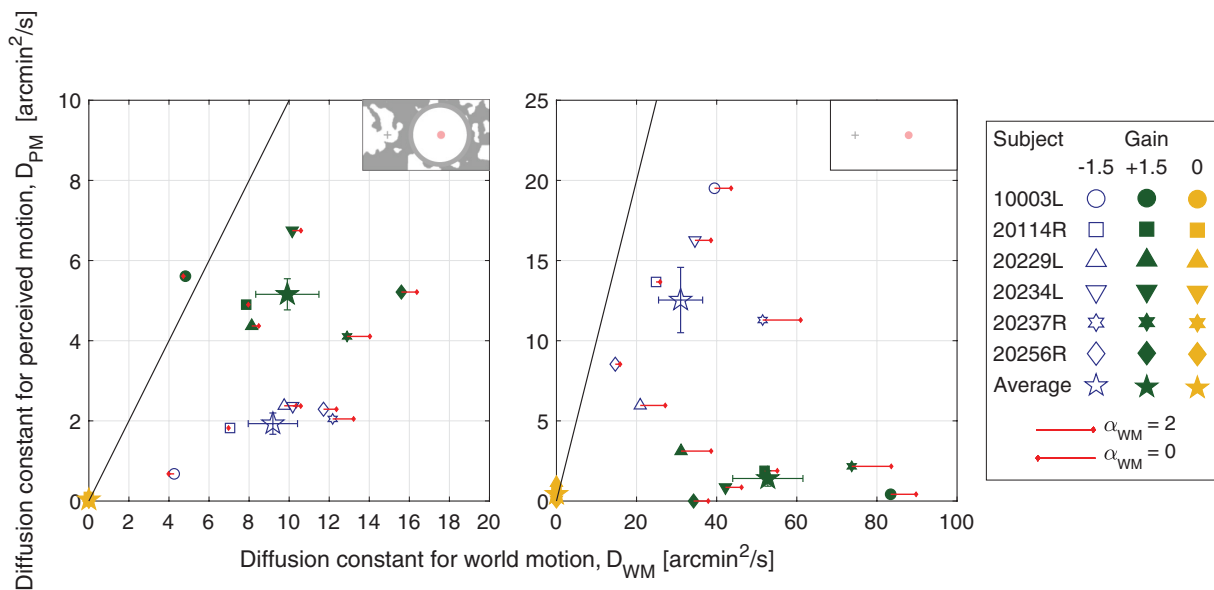
Fig. 2. Average diffusion constants for perceived motion (D_{PM}) versus diffusion constants for world motion (D_{WM}) from six subjects. Experiments were tested under background-present and background-absent conditions, indicated by labels on the *Top Right* corner of each graph. The small symbols represent each subject's average perceptual match for Gain -1.5 stimuli (blue open symbols), Gain $+1.5$ stimuli (green filled symbols), and Gain 0 stimuli (yellow filled symbols). The large stars are the group averages with SE of the mean bars. The red arrows show the extent to which the eye motion, and consequent retina-contingent stimulus' world motion (α_{WM}), deviated from Brownian. Arrows pointing *Right* indicate persistence ($\alpha_{WM} > 1$), arrows pointing *Left* indicate antipersistence ($\alpha_{WM} < 1$), and no arrow means that the motion was Brownian ($\alpha_{WM} = 1 +/- 0.02$). Longer arrows correspond to higher deviations from Brownian motion. The arrow length in the legend indicates pure persistence ($\alpha_{WM} = 2$, straight line trajectory at constant velocity) if pointing *Right* or pure antipersistence ($\alpha_{WM} = 0$, oscillatory motion) if pointing *Left*.

least three times before submitting. The diffusion constant of the random walk stimulus which the subject submitted as a match in perceived motion (PM) will be called D_{PM}. All the Gain conditions were presented in a pseudorandom order until six matches were submitted for each.

The AOSLO is a custom-built device that can image and track the retina in real time and deliver the retina-contingent stimuli with subarcminute accuracy (7). An AOSLO video of the retina containing a version of the stimulus embedded as a decrement on each video frame was recorded for each presentation. The videos were analyzed offline to compute the following: a continuous eye motion trace; the trajectory of each retina-contingent stimulus; and the accuracy of the retina-contingent stimulus. We computed the diffusion constant for the eye motion (EM) (D_{EM}) and the diffusion constant for the world motion (WM) of the stimulus (D_{WM}). Also, since the eye does not always exhibit true Brownian motion, we computed a second parameter, α, which quantified the extent to which the eye motion (and consequently the retina-contingent stimulus motion) was either persistent ($\alpha > 1$) or antipersistent ($\alpha < 1$) (8). Across all traces for each trial, we computed a single D_{WM} and α_{WM} per match, from the average of all valid retina-contingent presentations. Last, we computed the average drift speed (S_{EM}). Full descriptions of the motion parameters are in *Materials and Methods*.

Fig. 2 plots the average responses from experiments tested under background-present (*Left*) and background-absent (*Right*) conditions. The retina-contingent stimulus' diffusion constant for perceived motion (D_{PM}) is plotted as a function of its diffusion constant for world motion (D_{WM}). Each small data point represents one subject and is the average of six trials. Refer to *SI Appendix*, Fig. S1 for individual subject distributions of the six trials. The large stars represent the group averages with SE of the mean bars. The data points that lie along the diagonal 1:1 line represent trials where the perceived motion of the

retina-contingent stimulus was equivalent to the stimulus' actual motion occurring in the world. The red arrows depict the extent to which the α for the world motion (α_{WM}) deviated from Brownian motion. Arrows pointing right represent persistence ($\alpha_{WM} > 1$), arrows pointing left represent antipersistence ($\alpha_{WM} < 1$), and the absence of an arrow indicates Brownian motion ($\alpha_{WM} = 1 +/- 0.02$). Each red arrow represents the average α_{WM} from six trials, individual trial distributions are shown in *SI Appendix*, Fig. S1. We did not show red arrows for the Gain 0 stimuli because the stimuli are world-fixed and motion statistics do not apply.

**Comparing Perceived Motion for Gain $+1.5$ and -1.5.** Under background-present conditions (Fig. 2, *Left*), images moving in the same direction as eye motion (Gain $+1.5$) were perceived to have more motion than images moving with increased retinal slip (Gain -1.5). This means that despite the Gain -1.5 stimuli having similar world motion, as well as approximately five times more retinal motion than the Gain $+1.5$ stimuli, all subjects reported seeing it as having significantly less motion ($P = 0.0029$, post hoc Tukey–Kramer). This is consistent with the findings of Arathorn et al. (6) for which some world-fixed background content could not be avoided.

Under background-absent conditions (Fig. 2, *Right*), the motion perceptions reversed. Images moving with increased retinal slip (Gain -1.5) were now perceived to be moving with a high magnitude of motion whereas subjects reported perceiving little to no motion when images moved in the same direction as eye motion (Gain $+1.5$).

The striking reversal in perceptions shows how profoundly the presence of retinal image background content impacts perceived motion. For Gain -1.5 stimuli, all subjects reported perceiving a higher magnitude of motion under background-absent conditions (average $D_{PM} = 12.54$ arcmin^2/s, sem $+/$

$- 2.04$ arcmin^2/s) compared to background-present conditions (average $D_{PM} = 1.93$ arcmin^2/s, sem $+/-$ 0.27 arcmin^2/s). Motion perceptions between these conditions were significantly different ($P = 0.0051$, post hoc Tukey–Kramer). For Gain $+1.5$ stimuli, all subjects reported perceiving a lower magnitude of motion under background-absent conditions (average $D_{PM} = 1.41$ arcmin^2/s, sem $+/-$ 0.48 arcmin^2/s) compared to background-present conditions (average $D_{PM} = 5.16$ arcmin^2/s, sem $+/-$ 0.39 arcmin^2/s). Motion perceptions between these conditions were also significantly different ($P = 0.0052$, post hoc Tukey–Kramer).

Ratios Analysis. To compare the diffusion constant for perceived motion (D_{PM}) and diffusion constant for world motion (D_{WM}) between experiments tested under background-present and background-absent conditions, we computed the ratio $[D_{PM}/D_{WM}]$ for each retina-contingent condition. If subjects always perceived the stimulus' actual motion in the world, then D_{PM}/D_{WM} would be equal to one regardless of background condition. Fig. 3 shows the ratios in conditions with background-present plotted on the y-axis and the ratios in conditions with background-absent plotted on the x-axis. The results do not lie on or even close to the 1:1 line—instead the connected data points for each subject show a negative slope, indicative of the reversal of motion perception with background condition.
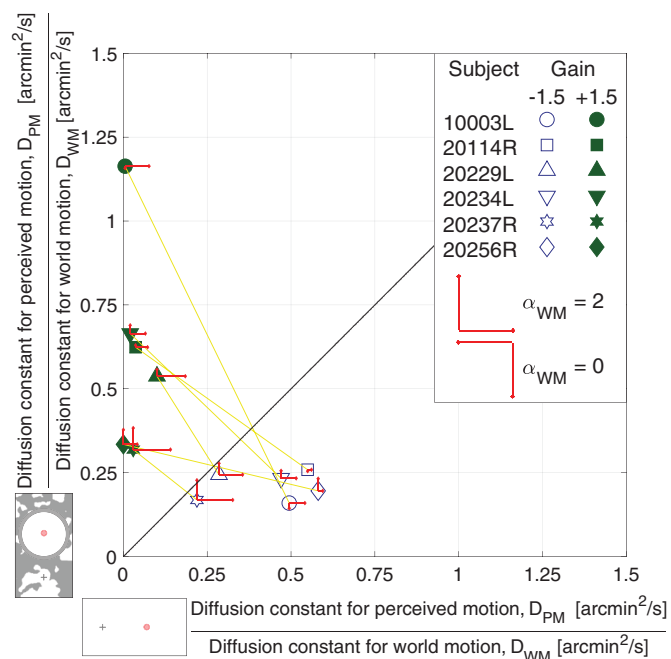
Fig. 3. Computed ratios from six subjects tested under background-present (y-axis) and background-absent (x-axis) conditions. The symbols represent each subject's [average diffusion constant for perceived motion]/[average diffusion constant for world motion] for two Gain conditions: Gain -1.5 stimuli (blue open symbols) and Gain $+1.5$ stimuli (green filled symbols). The red arrows show the extent to which the eye motion, and consequent retina-contingent stimulus' world motion (α_{WM}), deviated from Brownian. Arrows pointing up and *Right* indicate persistence ($\alpha_{WM} > 1$) under background-present and background-absent conditions, respectively. Arrows pointing down and *Left* indicate antipersistence ($\alpha_{WM} < 1$) under background-present and background-absent conditions, respectively. No arrow means that the motion was Brownian ($\alpha_{WM} = 1 +/- 0.02$). Longer arrows correspond to higher deviations from Brownian motion. The arrow length in the legend indicates pure persistence ($\alpha_{WM} = 2$, straight line trajectory at constant velocity) if pointing *Up/Right* or pure antipersistence ($\alpha_{WM} = 0$, oscillatory motion) if pointing *Down/Left*.

Perceived Motion for Gain 0. Under both background-present and background-absent conditions, all subjects reported perceiving little to no motion for world-fixed images (Gain 0) (Fig. 2). Under background-present conditions, the average $D_{PM} = 0.033$ arcmin^2/s, sem $+/-$ 0.020 and under background-absent conditions, the average $D_{PM} = 0.43$ arcmin^2/s, sem $+/-$ 0.18. Motion perceptions between these conditions were not significantly different ($P = 0.064$, post hoc Tukey–Kramer). The average perceived motion for Gain 0 stimuli was lower than the average perceived motion of Gain -1.5 and Gain $+1.5$ stimuli, with the exception of 20256R, where the subject perceived no motion for both Gain 0 and Gain $+1.5$ under background-absent conditions (Fig. 2, *Right*).

Comparing Perceived Motion for Gain +1.5 and 0 under Background-Absent Conditions. Subjects reported seeing little to no motion when viewing the Gains 0 and $+1.5$ stimuli under background-absent conditions. Motion perceptions were not significantly different between the Gains 0 and $+1.5$ stimuli ($P = 0.057$, post hoc Tukey–Kramer). The eye motion was not significantly different between Gain $+1.5$ and Gain 0 presentations, shown in Fig. 4A. However, the Gain 0 stimuli moved across the retina with greater retinal motion (RM) than Gain $+1.5$ stimuli: The average diffusion constant for retinal motion (D_{RM}) of a Gain 0 stimulus was 15.55 arcmin^2/s, while the average D_{RM} of a Gain $+1.5$ stimulus was 5.86 arcmin^2/s (*Materials and Methods*).

Eye Motion Depends on Background and Stimulus Condition. The diffusion constant for eye motion (D_{EM}), α for eye motion (α_{EM}), and speed for eye motion (S_{EM}) for all Gains and background conditions are plotted on Fig. 4.
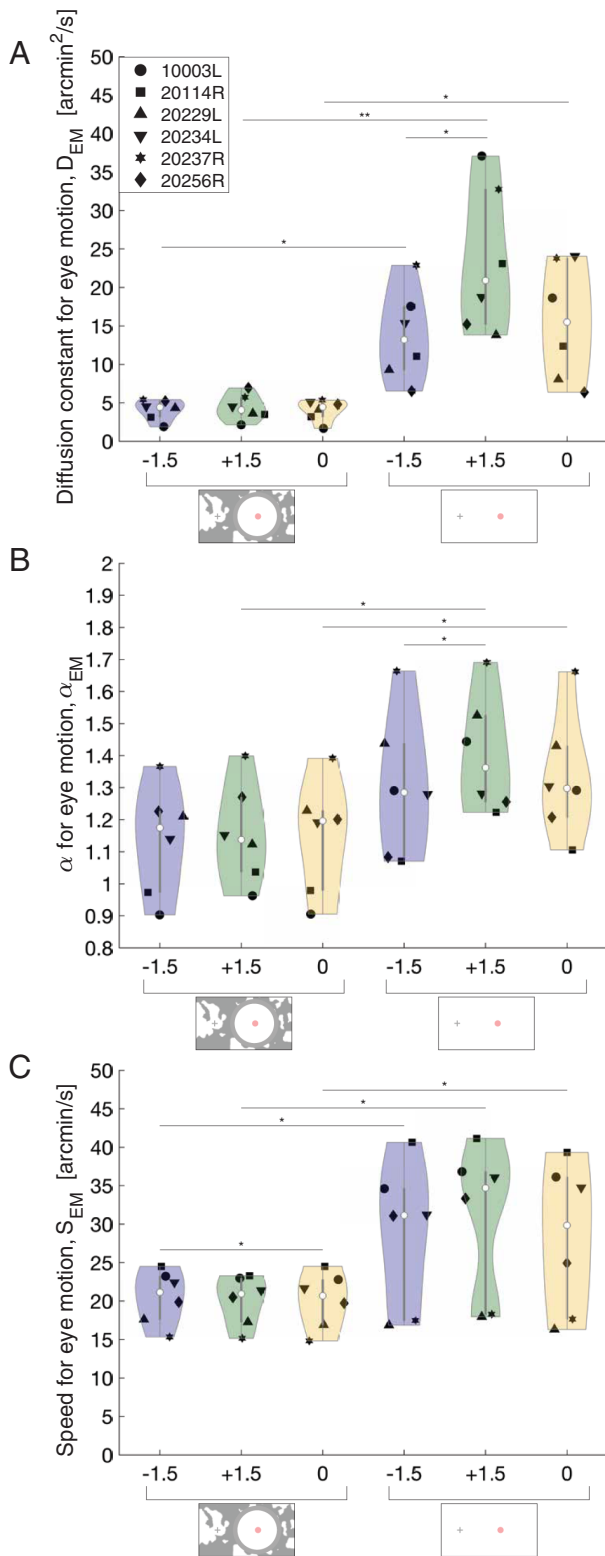
Under background-present conditions, D_{EM} was not significantly different between Gain conditions (Fig. 4A). Most, but not all, eyes exhibited a small degree of persistence ($\alpha_{EM} > 1$) but the α_{EM} did not significantly differ between Gain conditions (Fig. 4B). The S_{EM} during Gain -1.5 stimuli presentation was slightly higher than that of the Gain 0 stimuli presentation (Fig. 4C).

The D_{EM} and S_{EM} were uniformly greater for all Gains under background-absent conditions, meaning there was more eye motion. While α_{EM} values were greater for all Gains under background-absent conditions, only Gain $+1.5$ and 0 conditions showed a significant increase.

Under background-absent conditions, the D_{EM} of the Gain $+1.5$ stimuli was significantly greater than that of the Gain -1.5. The α_{EM} values were greater than one for all Gain conditions but were greatest for the Gain $+1.5$ condition.

To offer a physical sense of the differences in α_{EM} values between background conditions, all the gaze trajectory traces during the Gain -1.5 retina-contingent stimulus presentations are plotted for one of the subjects in Fig. 5 *A* and *B* and for all subjects and all conditions in *SI Appendix*, Fig. S2.

The Presence of Persistent Eye Motion ($\alpha_{EM} > 1$) Affects Interpretation of the Results. Subjects adjusted the diffusion constant of the random walk stimulus until its motion looked perceptually equivalent to the motion of the respective retina-contingent stimulus. Although Fig. 4 shows that α for eye motion (α_{EM}) changed between conditions and was mostly persistent, we chose to constrain the motion of the random walk matching stimulus to be Brownian motion with α, on average, equal to one.

Fig. 4. (A–C) Violin plots (9) showing the distribution of the (A) diffusion constant for eye motion, D_{EM}, (B) α for eye motion, α_{EM}, and (C) speed for eye motion, S_{EM}, for three Gains (-1.5, $+1.5$, and 0) under two conditions (background-present and background-absent). In the center of the violins, the white circles and dark bars represent the median and interquartile range, respectively; these are surrounded by the density trace (10). The black points are the (A) average D_{EM}, (B) average α_{EM}, and (C) average S_{EM} for each subject across their six respective trials for each Gain. The asterisks indicate statistical significance ($P < 0.05$ and $P < 0.01$, respectively) from a post hoc Tukey–Kramer test following a two-factor repeated-measures ANOVA.

We constrained the motion of the random walk matching stimulus to be Brownian because we found that subjects perform poorly at discriminating a highly persistent motion (e.g. $1.3 < \alpha < 1.7$) with a high diffusion constant from a less persistent motion (e.g. $1 < \alpha < 1.3$) with a low diffusion constant. This was found when we ran a control experiment with the same method-of-adjustment procedure, but with controlled random walk stimuli in both intervals. The stimulus in the first interval was a preprogrammed random walk stimulus with a selected α and diffusion constant. The subject's task was to adjust the parameters of the random walk stimulus in the second interval until its motion looked equivalent to that of the random walk stimulus in the first interval. Subjects could adjust two parameters—the diffusion constant as well as the α. We repeated each match two times. We found that for the same test α and diffusion constant, sometimes the subject would correctly match the α and diffusion constant, while for other matches the α and diffusion constant were both much higher or both much lower (*SI Appendix*, Fig. S3). What could contribute to this mismatch, is that random walks with high α and diffusion constant values can be generated from the same average step length as random walks with low α and diffusion constant values, examples are shown in Fig. 1C. We concluded that introducing a second adjustable parameter, α, complicated the interpretation of results and also made the experimental procedure more difficult for the subject. We therefore ran experiments with one adjustable parameter, the diffusion constant, and set the α to be on average equal to one.

We could not, however, control the α for the world motion (α_{WM}) of the retina-contingent stimuli because the subjects' fixational eye motion governed this parameter.

When the α_{WM} is different from the α_{PM}, comparing only their diffusion constants is too simplistic. This is because the α and diffusion constant are dependent on each other. A subject with persistent eye motion, would have a high α_{WM} value and thus a higher diffusion constant, compared to a trajectory with the same average step length but with an α equal to one. This could explain why, in Fig. 3, the subjects with high α_{WM} did not report D_{PM} equal to the D_{WM}. The subjects with more persistent motion ($\alpha_{WM} > 1$, larger red arrows) tended to have lower ratio values. These higher α_{WM} values correspond to higher D_{WM} which increase the value in the denominator.

Fig. 6 plots the ratios from Fig. 3 as a function of each subject's α_{WM}, for two of the conditions: Gain -1.5 stimuli tested under background-absent conditions and Gain $+1.5$ stimuli tested under background-present conditions. To predict how the presence of persistent eye motion could impact the interpretation of the results, we modeled how the ratio of the diffusion constants would vary with α_{WM} under a specific condition where the subject equated the mean square displacement (MSD) between the motion of the retina-contingent stimulus and the motion of the random walk stimulus over a time interval ($\triangle$T) of 2 frames. In the model, the random walk stimulus' α, which represents the α for perceived motion (α_{PM}), is equal to one (Brownian motion) and the retina-contingent stimulus' α_{WM} ranges from antipersistent ($0.9 <= \alpha_{WM} < 1$), to Brownian ($\alpha_{WM} = 1$), to persistent ($1 < \alpha_{WM} <= 1.8$). The model is represented by the solid curve in Fig. 6. The derivation and full descriptions of these parameters are in *Materials and Methods*.

As the retina-contingent stimulus' α_{WM} becomes more persistent, it on average traverses further than a Brownian stimulus and therefore the mean square displacement of the trajectory increases faster over a time interval compared to that of the Brownian stimulus.
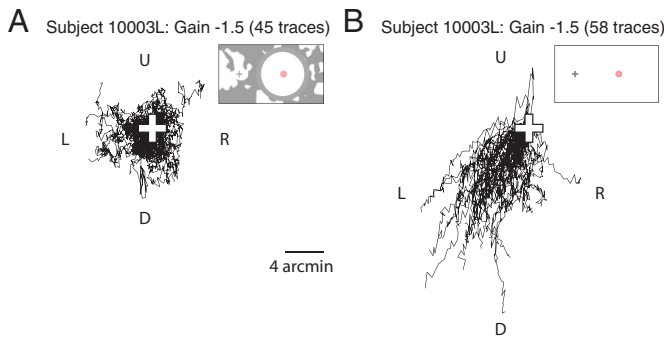
Fig. 5. Example gaze traces for subject 10003L during Gain −1.5 retina-contingent presentations under (A) background-present (1,500-ms duration and $\alpha_{EM} = 0.90$) and (B) background-absent (750-ms duration and $\alpha_{EM} = 1.29$) conditions. The white cross indicates the starting position for each trace. The gaze directions are labeled: *Left* (L), *Right* (R), *Up* (U), and *Down* (D).

The model shows that there is an exponential decay relationship between D_{PM}/D_{WM} and α_{WM}—as the eye motion becomes more persistent ($\alpha_{WM} > 1$), the corresponding D_{WM} increases and subsequently the ratio D_{PM}/D_{WM} decreases. The human subject data are largely predicted by this model.

Discussion

When an object moves in the world, the image that is cast onto the retina has retinal motion both due to the object's motion as well as the fixational eye motion. To properly perceive an object's motion in the world, the visual system is tasked with disentangling the two (11, 12). Normally, the visual system performs this task exceptionally well; humans are able to reliably perceive world-fixed objects as stable and can identify moving objects within it with hyperacuity (1, 3). Yet, in conditions where an image is programmed to move with amplified retinal slip, humans exhibit a paradoxical inability to accurately perceive the motion of the image relative to a high-contrast, world-fixed background.

Our findings support those of Arathorn et al. (6) and demonstrate that the visual system suppresses the perceived motion of anything moving in a direction consistent with retinal slip (Gains less than 1), despite their magnitude. The neural mechanism for doing this has yet to be described, but from a vision standpoint, it is functionally sufficient, since the likelihood that any real-world object would be moving in the same direction of retinal slip for any appreciable duration is effectively zero. All other directions, such as our Gain +1.5 and all other Gain conditions described in Arathorn et al. (6), are perceived as moving. As a control, we tested the same experiment for one subject, presenting a stimulus that moved with a Gain of +1.5, but programmed to move orthogonal to eye motion. There was no significant difference in motion perception between this stimulus and the stimulus that moved in advance of eye motion, Gain +1.5 (*SI Appendix*, Fig. S4).

This paper shows that the direction of eye motion must arise from the retinal image content itself (inflow) and is not relayed to the visual system by any other nonvisual means, such as efference copy.

To further understand the process, we first review what is known functionally and structurally about perception of motion.

Perceiving the World As Stable. Previous studies showed that the retinal input informs the visual system about image motion on the retina due to drift. Poletti et al. (13) presented gaze-contingent

stimuli using a dual Purkinje eye tracker in conditions with and without frames of references. Consistent with our findings on Fig. 2, stimuli moving in the same direction as gaze (theirs was a Gain +1 condition, which meant the image was stabilized) were perceived to move when frames of references were present, and perceived as stable when all visual content was removed. They concluded that the retinal input drives the compensation of retinal motion.

A possible origin of the retinal signal is directionally sensitive retinal ganglion cells (DSRGCs) for which evidence of their existence in primate and human retina continues to mount (14–16). The subclass of ON–OFF DSRGCs, which project to the lateral geniculate nucleus and subsequently the visual areas of the cortex, are the most likely candidates. Combined signals from a sufficient number of these cells to eliminate ambiguity could be used to compute the direction of retinal slip, which may provide partial information to the downstream circuitry to assist in the stabilization process. Exactly how and where the stabilization occurs, however, is not known and evidence is mixed (17–20).

Detecting Moving Objects. Humans have a hyperacute ability to detect (21) and resolve (1) the motion of objects within a visual scene. The neural underpinning for detection of these stimuli in the presence of incessant eye movements may lie in the Object Motion Sensing ganglion cells. Yet to be found in primates, this class of retinal ganglion cells is very effective at identifying an object that is moving differently than the surround (12).

But despite increasing knowledge of the neural systems that underlie our ability to perceive a stable and moving world (22), what remains unclear is how two objects that move in a direction
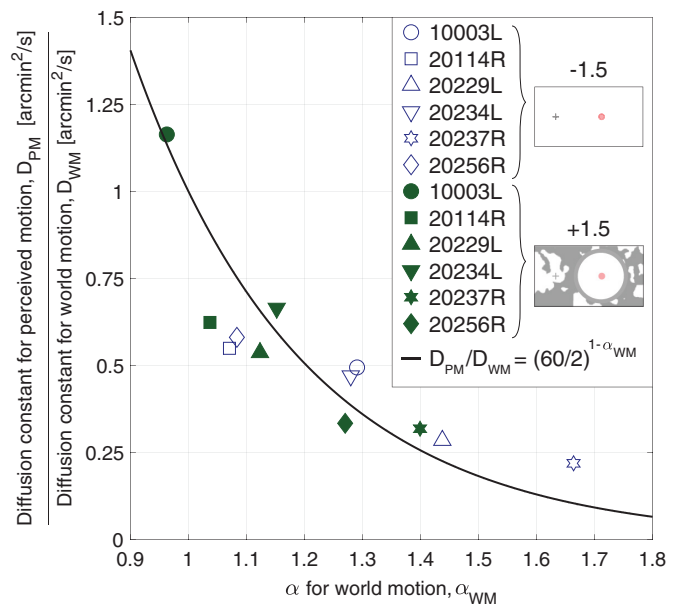
Fig. 6. Computed ratios from six subjects plotted as a function of each subject's α for world motion, α_{WM}. The symbols represent each subject's [average diffusion constant for perceived motion]/[average diffusion constant for world motion]. Data from two experiments are shown: Gain −1.5 stimuli tested under background-absent conditions (blue open symbols) and Gain +1.5 stimuli tested under background-present conditions (green filled symbols). The black curve predicts the D_{PM}/D_{WM} versus α_{WM}, considering that the subject matched the mean square displacements of the random walk stimulus with the retina-contingent stimulus over a time interval of 2 frames. Because the random walk stimulus' motion is Brownian ($\alpha_{PM} = 1$) and the retina-contingent stimulus' motion ranges from antipersistent ($0.9 \leq \alpha_{WM} < 1$), to Brownian ($\alpha_{WM} = 1$), to persistent ($1 < \alpha_{WM} \leq 1.8$), the curve shows an exponential decay, which largely matches the data.

consistent with retinal slip but with different velocities can be rendered in the percept to be fixed relative to each other.

The Role of Efference Copy. Studies have shown that efference copy or other nonvisual cues can inform the visual system about its gaze direction during drift (23, 24), albeit not with great accuracy. In our experiments, however, it is very clear from the background-absent condition that neither efference copy nor any other nonvisual signal is being used effectively to determine the direction of motion in a manner that guides perception of moving objects. That being said, the perception of motion in the background-absent condition cannot be explained as being due simply to retinal motion either. For example, the Gain 0 condition, which actually moved more across the retina than the Gain +1.5 condition, was perceived to be moving either the same (subject 20256R), or less than the Gain +1.5 condition. Further investigations of this condition are warranted.

Precedence in Prior Literature. The level of stimulus control in this study is unprecedented so, aside from an earlier study by coauthors of the current paper (6) and a brief mention in the early stabilization literature (5), there is no other study that is directly comparable. However, some level of control of the retinal slip can be accomplished during a smooth pursuit task. In the most relevant study, Turano and Heidenreich (25) explored the detectability of changes in the motion of sinusoidal grating while the subject pursued a simultaneously presented smooth pursuit stimulus that moved faster or slower than the grating. When the smooth pursuit stimulus moved slower than the grating, the subject exhibited an expected sensitivity to differences in grating motion. However, when the smooth pursuit target moved faster than the grating, subjects exhibited an unexpected difficulty detecting changes in grating motion. This second condition is similar in a way to our own in that their grating, on average, slipped in a direction consistent with retinal motion, but with a different magnitude than the surround. So, while the authors were unable to explain these phenomena, they are not surprising or unexpected in the context of our results.

Another field of study involves the perception of motion using widely studied stimuli like those that give rise to the Ouchi illusion (26) and the related jitter aftereffect (27). This class of experiment provides strong evidence that retinal image content provides important cues to eye motion. Interestingly, our study explores the inability to perceive relative motion where it does exist, while this other class of studies explore the opposite: perceived relative motion where it does not exist. Nevertheless, these are complementary to each other.

Control Experiments to Validate Results. In this study, we presented stimuli for 1,500 ms under background-present conditions and for 750 ms under background-absent conditions. These duration choices were motivated by a trade-off to achieve robust motion judgments while also optimizing retinal tracking. A longer interval duration gives the subject sufficient time to discern the motion magnitude of the stimuli but can also lead to tracking failures if the eye drifts too far away from fixation. We performed control experiments on subjects 10003L and 20114R to verify that different stimulus durations did not contribute to the results. We presented the stimuli for 750-ms intervals under background-present conditions and found similar trends, although less robust, as experiments with 1,500-ms intervals (*SI Appendix*, Fig. S5). Under background-absent conditions, the eye had faster drift and larger corrective microsaccades which led to

tracking failures and as a result, we were unable to perform the experiments with 1,500-ms stimulus intervals.

In the experiments tested under background-absent conditions, we ensured no frames of reference were visible to the subject, even removing the fixation target during the retina-contingent interval and the random walk interval. Under these conditions, we could not truly assess the amount of motion perceived. Specifically, we could not conclude that the subject did not perceive motion when they matched a Gain 0 condition to a nonmoving stimulus in the second interval. As a check, we performed a control experiment for one subject with the same experimental protocol, except that the fixation target remained on to serve as a frame of reference during the random walk interval. We found no difference in matches to the Gain 0 stimuli whether the fixation target was on or off during the random walk interval (*SI Appendix*, Fig. S6).

Summary. This study employs a powerful combination of high-resolution retina tracking with stimulus delivery to reveal striking properties of absolute and relative motion perception that would otherwise not have been found in natural viewing conditions. The retinal-image-based signals that the visual system uses to generate a stable percept of the world are so potent that they can disrupt our fundamental skill of detecting relative motion. This dichotomy suggests that evolution settled on a mechanism that optimized behavior in the majority and most critical visual circumstances and tolerated anomalous behavior in outlier observing conditions. The knowledge provided has crucial implications for those who aim to elucidate the neural underpinnings of this important property of human vision.

Materials and Methods

Subjects. Six subjects (4 experienced and 2 naive) were recruited. The experiment was approved by the Institutional Review Board at the University of California, Berkeley. Prior to the experiment, subjects provided informed consent to participate. We applied topical eye drops of 1.0% tropicamide and 2.5% phenylephrine hydrochloride to dilate and cycloplege subjects prior to each experiment.

Adaptive Optics Scanning Light Ophthalmoscopy. Experiments were performed using a multiwavelength AOSLO (28, 29). A 940-nanometer (nm) laser beam measured the eye's wavefront and a deformable mirror corrected for the optical imperfections of the eye to confine the laser beams to a small spot. A focused 840-nm laser beam scanned sinusoidally across the retina by a 16 kHz horizontal scanner and a 60 Hz vertical scanner. The field size was set to image a 1.71° raster square of the retina. An acousto-optic modulator (AOM) modulated a 680-nm laser beam to deliver a circular 12 arcminute diameter increment stimulus onto targeted retinal locations. Concurrently, the AOM modulated the 840-nm imaging beam to turn off at the same targeted retinal locations. This rendered a 3×3 arcminute decrement in the image which enabled unambiguous tracking of the 680-nm increment stimuli. We used custom software to move the stimuli contingent to the retinal motion (7). The AOSLO records real-time, high-resolution 512 by 256 pixel videos of the retina. Each pixel subtends 0.2 by 0.4 min of visual angle. The average power of the 940-nm, 840-nm, and 680-nm laser beams were 58.6 µW, 111.4 µW, and 11.6 µW giving rise to equivalent luminances of 0.0036 cd/m^2, 0.84 cd/m^2, and 3,960 cd/m^2, respectively using methods described by Domdei et al. (30). Pupil sizes for subjects ranged from 5.9 to 7.2 mm.

Projector Display. A digital light processing LightCrafter projector (DLP4500; Texas Instruments, Dallas, TX, USA) displayed a 17° background which projected over the 840-nm imaging raster. The mean luminance of the display, which was approximately 540 cd/m^2, effectively canceled perception of the 840- and

940-nm rasters. A fixation target and patterns were drawn over the background and timed with the stimulus delivery dependent on the experiment conditions (Fig. 1 *D* and *E*). The procedure was programmed with MATLAB (MathWorks, Natick, MA, USA) using the Psychophysics Toolbox (31–33).

Real-Time Eye Tracking. Fixational eye movements cause distortions in the raw AOSLO videos and these distortions encode the eye motions that occur during image acquisition (34). We extracted the eye motion traces from these videos by sectioning each frame into 32 strips, selecting a reference frame, and then cross-correlating the strips of every subsequent frame to the reference frame (7). This real-time eye tracking enabled targeted stimulus delivery; in each frame at a critical strip, the AOM would modulate the 680-nm and 840-nm laser beams to place increment and decrement stimuli at the targeted retinal location. The critical strip occurred 2 ms before stimulus delivery and this reduced lag enabled subarcminute accuracy (35). Full details on the computation of the target position for the retina-contingent stimuli can be found in *SI Appendix*, section 1.

Retina-Contingent Conditions. The high-resolution, real-time eye tracking of the AOSLO allowed us to deliver stimuli that moved contingent to fixational eye motion transformed by a "Gain." The sign of the Gain indicates the direction the image moved with respect to the eye, and the stimulus' world displacement was equal to the Gain magnitude times the eye displacement. We tested three Gains: −1.5, +1.5, and 0. The Gain −1.5 stimulus moved directly opposite the direction of eye motion with a magnitude that was 1.5 times that of the eye motion. For example, if the eye moved 2′ right, then the stimulus moved in the world 3′ left, producing a total retinal displacement of 5′. Therefore, in this condition, the stimulus moved on the retina with 2.5 times more motion than a natural, world-fixed object. The Gain +1.5 stimulus moved in the same direction as the eye with a magnitude equal to 1.5 times that of the eye motion. For example, if the eye moved 2′ right, then the stimulus moved in the world 3′ right. The stimulus moved in advance of eye motion with retinal image motion that was half the retinal motion of a natural, world-fixed object. The Gain 0 stimulus did not move in the world, which means that this was a natural, world-fixed stimulus and served as control.

Random Walks. Prior to the experiment, we preprogrammed 100 s of 750-ms and 1,500-ms duration random walks with varying diffusion constants for the subjects to use in the matching task. These paths were generated by computing randomized step lengths in x and y drawn from a normal distribution with a range of SDs from 0.05 to 1.6 arcminutes per step (or AOSLO frame), in increments of 0.05 arcminutes. For the sake of brevity, we will refer to the SD of the randomized step lengths as simply step lengths throughout the manuscript. This process effectively randomized the length and direction of each step and, because all directions were equally probable, the α of the paths was approximately equal to one. However, owing to the random nature of generating paths in this method, the same randomized step length could generate paths with a range of possible diffusion constants. We computed the average diffusion constant for each step length, by inputting the step length as the △X and △Y in the mean square displacement formula in Eq. **2**. After solving for the mean square displacement, we input this value into Eq. **1** to solve for the diffusion constant. For example, a step length of 1.6 arcminutes corresponds to a diffusion constant of 76.8 arcmin^2/s. Therefore, for each of the 32 step lengths, we generated 100 paths and from these, we selected the 10 paths with diffusion constants closest to the respective average diffusion constant.

Experimental Design.

Set up. The subject's pupil was aligned with the AOSLO beam and their head was immobilized by the use of a dental bite bar.

Conditions. We tested two conditions. Under the background-present condition, the 17° projector background was filled with rich retinal image background content: a blurred and binarized 1/f noise pattern that changed after every presentation (Fig. 1*D*). The fixation target was present for the entire duration and each stimulus was presented for 1,500 ms. Under the background-absent condition, all visual cues were removed (Fig. 1*E*). This was achieved by attaching a white paper with a small central aperture directly in front of the subject. The

aperture permitted only the raster and projector light to pass into the subject's eye. The proximity of the aperture to the subject's eye (about 5 cm) ensured that its edges appeared out-of-focus and blurred and therefore blended with the projector view. We then taped light-emitting diodes (LEDs) around the subject's eye and the subject adjusted the power of the LEDs until the white projector light was indistinguishable from the paper. Additionally, the fixation target disappeared during the intervals with stimulus presentation. Without a fixation target, eyes generally exhibit faster drift and larger corrective microsaccades (36). To mitigate the effects of this increased retinal motion on the tracking accuracy, each stimulus was presented for 750 ms.

Task. The projector display drew a fixation target positioned 2° nasally away from the AOSLO raster, which remained stationary Fig. 1 *D* and *E*. The 680-nm light from the AOSLO delivered increment circular images onto the retina, which subtended 12 min of visual angle. We tested one retina-contingent condition per trial, either Gain −1.5, +1.5, or 0. In each trial, the subject could initiate as many presentations as necessary. In a single presentation, the subject attended to the stimulus in the first time interval (retina-contingent), followed by a 500-ms break, followed by the stimulus in the second time interval (random walk). The subject's task was to adjust the diffusion constant of the random walk stimulus until its motion looked perceptually equivalent to the respective retina-contingent stimulus. When the subject found a match, they initiated a minimum of three final presentations and then submitted their response. In one block, there were six trials and the trial order was randomized. Subjects completed three blocks per background condition. Therefore, for each background condition, there were six trials tested for each Gain. All perceptual responses were made on a gamepad.

Quality Control. During the experiment, if the subject looked away from the fixation target, the tracking would fail and the stimulus would not be delivered. This ensured that the subject maintained fixation when making motion judgments. Blinks and large saccades also led to tracking failures and the stimulus misdelivery was usually recognizable to the subjects, who were informed ahead of the experiment to disregard presentations with poor stimulus delivery.

After the experiments, we used custom software (37) to extract a continuous eye motion trace from each recorded video. We then used a video-analysis script to determine the retinal positions of the 3 × 3 arcminute stimulus-tracking decrement in each frame. We compared the eye motion trace to the stimulus motion trace to verify that the stimulus moved appropriately contingent to the retina, dependent on the Gain. In this way, videos were evaluated for tracking errors and stimulus misdelivery.

We determined the number of videos where the stimulus delivery was poor but might not have been recognizable to the subject. This was achieved by comparing the eye position to the stimulus-tracking decrement position in frames with retina-contingent presentation. We computed the SD of the stimulus delivery accuracy for each video across all the frames that had retina-contingent stimulus presentation. We considered a video to have poor stimulus delivery when the SD of the stimulus misdelivery from the targeted location was greater than 0.9 arcminutes. If the majority of the videos used by the subject had SDs greater than this threshold, we removed the trial. This strict criterion ensured that the motion judgments were not contaminated by presentations with poor tracking. This was especially important for the Gain −1.5 stimuli because tracking errors could cause stimuli motions that no longer slipped in directions consistent with retinal motion, thus disrupting the "illusion of relative stability."

Because we are interested in motion perception during periods of fixational drift, we used a final filter that removed traces with microsaccades. Microsaccades are linear and ballistic and this would have biased the random walk analysis.

Analysis of Eye Motion and Retina-Contingent Stimulus Motion.

Random walk analysis. Fixational drift, the eye motion that occurs in between microsaccades, resembles a random trajectory similar to Brownian motion (8, 38–42). That is, the motion is two-dimensional with a velocity that varies according to a normal distribution. Studies have shown that drift deviates from uncorrelated random motion, instead having correlated or anticorrelated properties depending on the timescale (8, 39). To quantify the statistical properties of drift, we computed the diffusion constant (D) which quantifies the amount the eye moves from its starting position, and we computed the α

which measures the extent to which the steps in the trajectory are uncorrelated, positively correlated, or negatively correlated. Eq. **1** shows a method to solve for the diffusion constant and the scaling exponent α (43, 44).

$$D = \frac{MSD}{2d\triangle T^\alpha} \qquad [1]$$

From Eq. **1**, the diffusion constant measures the temporal changes of the MSD. The mean square displacement is the squared Euclidean distance of the horizontal and vertical positions of the eye between two time points specified by the time interval ($\triangle T$) which, in our case, is the time between frames (Eq. **2**). The dimension (d) is equal to two since the eye moves in both x and y directions. As expressed in Eq. **2**, we computed the MSD over nonoverlapping $\triangle T$'s. Overlapping pairs are very correlated, and therefore computing over nonoverlapping pairs aimed to mitigate accidentally biasing the α to higher values. This was primarily important when computing over a low number of traces. As suggested by Saxton (45), the number of $\triangle T$'s was determined by one-quarter times the total time steps (in our case, total number of frames). We did not calculate beyond one-quarter of the time steps because higher $\triangle T$'s have fewer samples and are thus more vulnerable to noise (45). In Eq. **2**, N represents the total time steps and $\lfloor\ \rfloor$ represents the greatest integer function.

$$MSD = \frac{1}{\lfloor N/\triangle T\rfloor} \sum_{j=1}^{\lfloor N/\triangle T\rfloor} ((x_{j\triangle T+1} - x_{j\triangle T-\triangle T+1})^2$$
$$+ (y_{j\triangle T+1} - y_{j\triangle T-\triangle T+1})^2). \qquad [2]$$

From Eq. **1**, if $\alpha = 1$, then the MSD increases linearly with $\triangle T$. This is uncorrelated random motion because the steps in the trajectory are independent from previous steps. This is referred to as Brownian motion. When $\alpha > 1$, the MSD increases faster than linearly with $\triangle T$. The steps in the trajectory are positively correlated; each step has a tendency to continue moving in the same direction as the previous step. This is referred to as superdiffusion (46), persistence (8, 43, 47), or diffusion with flow (44). When $\alpha < 1$, the MSD increases slower than linearly with $\triangle T$. The steps in the trajectory are negatively correlated; each step has a tendency to move in the opposite direction as the previous step. This is referred to as subdiffusion (46), antipersistence (8, 43, 47), or caged motion (44). In this study, we use the terms Brownian motion, persistence, and antipersistence.

We compute the diffusion constant and α by plotting the $\log_{10}$(MSD) as a function of the $\log_{10}(\triangle T)$. The slope of the line across these data points is α and the y-intercept is $\log_{10}$(diffusion constant). Therefore, across all of the traces from each analysis, we quantified the amount of motion (diffusion constant) and its deviation from Brownian motion (α).

We performed this random walk analysis on the eye's motion (D_{EM} and α_{EM}), the retina-contingent stimuli's motion (D_{WM} and α_{WM}), and the random walk stimuli's motion at the setting for a perceptual match (D_{PM} and α_{PM}).

Speed analysis. Across all traces with retina-contingent stimulus presentation, we computed the average displacement of the eye over time to compute the speed (S_{EM}). Eye positions were sampled at 60 Hz.

Statistics. We performed a two-factor repeated-measures ANOVA of the alpha for eye motion (α_{EM}), diffusion constant for eye motion (D_{EM}), diffusion constant for perceived motion (D_{PM}), and speed for eye motion (S_{EM}). There were two background conditions (background-present and background-absent) and three retina-contingent stimuli (Gain -1.5, Gain $+1.5$, and Gain 0), which were within-subject factors. Following the repeated-measures ANOVA, we performed a post hoc Tukey–Kramer test to measure the significance of the pairwise comparisons. We compared between Gain categories within the same background condition and for each Gain we compared differences between background conditions.

Comparing the Eye's Motion and the Retina-Contingent Stimuli's World Motion. The diffusion constant for world motion (D_{WM}) and α for world motion (α_{WM}) describe how the retina-contingent stimuli move in the world. For the Gain 0 stimulus, D_{WM} is equal to zero because the stimulus is not moving in the world and α_{WM} does not apply. The world motion of the Gains -1.5 and $+1.5$ stimuli are dependent on the parameters of the eye's motion: D_{EM} and α_{EM}. How the eye moves (α_{EM}) is equal to how the nonzero retina-contingent stimuli move (α_{WM}) in the world. However, the eye's magnitude of motion (D_{EM}) is different from the retina-contingent stimulus' magnitude of motion (D_{WM}) because the retina-contingent stimulus' motion is transformed by a Gain. There is a squared relationship between the diffusion constant and mean square displacement, therefore, for the Gains -1.5 and $+1.5$ stimuli, the $[D_{WM}] = [D_{EM}] * (1.5)^2$.

Comparing the Eye's Motion and the Retina-Contingent Stimuli's Retinal Motion. The diffusion constant for retinal motion (D_{RM}) and α for retinal motion (α_{RM}) describe how the retina-contingent stimuli move across the retina. For the Gain 0 stimulus, D_{RM} and α_{RM} are equal to D_{EM} and α_{EM}, respectively, because all retinal motion is the result of the stimulus slipping consistent to the fixational eye motion. The retinal motion of the Gains $+1.5$ and -1.5 stimuli are dependent on the parameters of the eye's motion: D_{EM} and α_{EM}. How the eye moves (α_{EM}) is equal to how the nonzero retina-contingent stimuli move (α_{RM}) on the retina. However, similar to the above paragraph the D_{EM} is not equal to the D_{RM}. The Gain -1.5 stimulus slips with 2.5 times more motion than the eye. The Gain $+1.5$ stimulus moves ahead of eye motion and moves with half the motion of the eye. Therefore, for the Gain -1.5 stimuli, the $[D_{RM}] = [D_{EM}] * (2.5)^2$; and for the Gain $+1.5$ stimuli, the $[D_{RM}] = [D_{EM}] * (0.5)^2$.

Simulated Ratios. The black curve in Fig. 6 models the ratio values versus α for world motion (α_{WM}), where the random walk stimulus' α, which represents the α for perceived motion (α_{PM}), is equal to one (Brownian motion) while the retina-contingent stimulus' α_{WM} ranges from antipersistent ($0.9 <= \alpha_{WM} < 1$), to Brownian ($\alpha_{WM} = 1$), to persistent ($1 < \alpha_{WM} <= 1.8$). This model shows how the ratio of the diffusion constants would vary with α_{WM} under a specific condition where the subject equated the MSD between the motion of the retina-contingent stimulus and the motion of the random walk stimulus with a specific time interval ($\triangle T$). We tested a range of time intervals and determined by a least-squares procedure that a time interval of 2 frames best predicted the measured data. The units of the time interval are converted to seconds because the AOSLO has a 60 Hz frame rate. Eqs. **3** and **4** show the derivation, where the random walk stimulus' motion is represented as the "PM" and the retina-contingent stimulus' motion is represented as the "WM". As the retina contingent stimulus' motion becomes more persistent ($\alpha_{WM} > 1$), the $\triangle T$ over which the stimulus reaches higher MSDs becomes smaller compared to the Brownian random walk stimulus. The model assumes that there are no directional biases.

$$D_{PM} = \frac{MSD}{2d\triangle T^{\alpha_{PM}}} \qquad D_{WM} = \frac{MSD}{2d\triangle T^{\alpha_{WM}}} \qquad [3]$$

$$\frac{D_{PM}}{D_{WM}} = \frac{(2/60)^{\alpha_{WM}}}{(2/60)^1} = (60/2)^{1-\alpha_{WM}}. \qquad [4]$$

1. G. E. Legge, F. Campbell, Displacement detection in human vision. *Vis. Res.* **21**, 205–213 (1981).
2. Y. Wang *et al.*, Human foveal cone photoreceptor topography and its dependence on eye length. *eLife* **8**, e47148 (2019).
3. G. Westheimer, Visual acuity and hyperacuity. *Invest. Ophthalmol. Vis. Sci.* **14**, 570–572 (1975).
4. S. P. McKee, L. Welch, D. G. Taylor, S. F. Bowne, Finding the common bond: Stereoacuity and the other hyperacuities. *Vis. Res.* **30**, 879–891 (1990).
5. L. A. Riggs, F. Ratliff, J. C. Cornsweet, T. N. Cornsweet, The disappearance of steadily fixated visual test objects. *J. Opt. Soc. Am.* **43**, 495–501 (1953).
6. D. W. Arathorn, S. B. Stevenson, Q. Yang, P. Tiruveedhula, A. Roorda, How the unstable eye sees a stable and moving world. *J. Vis.* **13**, 22 (2013).
7. D. W. Arathorn *et al.*, Retinally stabilized cone-targeted stimulus delivery. *Opt. Express* **15**, 13731–13744 (2007).
8. R. Engbert, R. Kliegl, Microsaccades keep the eyes' balance during fixation. *Psychol. Sci.* **15**, 431–436 (2004).
9. B. Bechtold, Violin plots for Matlab. Github. https://github.com/bastibe/Violinplot-Matlab. Accessed 22 July 2024.
10. J. L. Hintze, R. D. Nelson, Violin plots: A box plot-density trace synergism. *Am. Stat.* **52**, 181–184 (1998).
11. J. J. Gibson, *The Perception of the Visual World* (Houghton Mifflin Co., 1950).
12. B. P. Ölveczky, S. A. Baccus, M. Meister, Segregation of object and background motion in the retina. *Nature* **423**, 401–408 (2003).
13. M. Poletti, C. Listorti, M. Rucci, Stability of the visual world during eye drift. *J. Neurosci.* **30**, 11143–11150 (2010).
14. A. Y. Wang *et al.*, An on-type direction-selective ganglion cell in primate retina. *Nature* **623**, 381–386 (2023).
15. S. S. Patterson *et al.*, Conserved circuits for direction selectivity in the primate retina. *Curr. Biol.* **32**, 2529–2538 (2022).
16. Y. J. Kim *et al.*, Origins of direction selectivity in the primate retina. *Nat. Commun.* **13**, 2862 (2022).
17. C. H. Anderson, D. C. Van Essen, Shifter circuits: A computational strategy for dynamic aspects of visual processing. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6297–6301 (1987).
18. M. Gur, D. M. Snodderly, Visual receptive fields of neurons in primary visual cortex (v1) move in space with the eye movements of fixation. *Vis. Res.* **37**, 257–265 (1997).
19. B. Motter, G. Poggio, Dynamic stabilization of receptive fields of cortical neurons (VI) during fixation of gaze in the macaque. *Exp. Brain Res.* **83**, 37–43 (1990).
20. J. M. McFarland, A. G. Bondy, B. G. Cumming, D. A. Butts, High-resolution eye tracking using v1 neuron activity. *Nat. Commun.* **5**, 4605 (2014).
21. S. P. McKee, K. Nakayama, The detection of motion in the peripheral visual field. *Vis. Res.* **24**, 25–32 (1984).
22. D. Kerschensteiner, Feature detection by retinal ganglion cells. *Annu. Rev. Vis. Sci.* **8**, 135–169 (2022).
23. A. Raghunandan, J. Frasier, S. Poonja, A. Roorda, S. B. Stevenson, Psychophysical measurements of referenced and unreferenced motion processing using high-resolution retinal imaging. *J. Vis.* **8**, 14 (2008).
24. Z. Zhao, E. Ahissar, J. D. Victor, M. Rucci, Inferring visual space from ultra-fine extra-retinal knowledge of gaze position. *Nat. Commun.* **14**, 269 (2023).
25. K. A. Turano, S. M. HEeidenreich, Speed discrimination of distal stimuli during smooth pursuit eye motion. *Vis. Res.* **36**, 3507–3517 (1996).
26. H. Ouchi, *Japanese Optical and Geometrical Art* (Courier Corporation, 2013).
27. I. Murakami, P. Cavanagh, A jitter after-effect reveals motion-based stabilization of vision. *Nature* **395**, 798–801 (1998).
28. A. Roorda *et al.*, Adaptive optics scanning laser ophthalmoscopy. *Opt. Express* **10**, 405–412 (2002).
29. S. Mozaffari, F. LaRocca, V. Jaedicke, P. Tiruveedhula, A. Roorda, Wide-vergence, multi-spectral adaptive optics scanning laser ophthalmoscope with diffraction-limited illumination and collection. *Biomed. Opt. Express* **11**, 1617–1632 (2020).
30. N. Domdei *et al.*, Ultra-high contrast retinal display system for single photoreceptor psychophysics. *Biomed. Opt. Express* **9**, 157–172 (2018).
31. D. H. Brainard, S. Vision, The psychophysics toolbox. *Spat. Vis.* **10**, 433–436 (1997).
32. M. Kleiner *et al.*, What's new in psychtoolbox-3. *Perception* **36**, 1–16 (2007).
33. D. G. Pelli, The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spat. Vis.* **10**, 437–442 (1997).
34. S. B. Stevenson, A. Roorda, "Correcting for miniature eye movements in high-resolution scanning laser ophthalmoscopy" in *Ophthalmic Technologies XV*, F. Manns, P. Soderberg, A. Ho, Eds. (SPIE, 2005), vol. 5688, pp. 145–151.
35. Q. Yang, D. W. Arathorn, P. Tiruveedhula, C. R. Vogel, A. Roorda, Design of an integrated hardware interface for AOSLO image capture and cone-targeted stimulus delivery. *Opt. Express* **18**, 17841–17858 (2010).
36. C. Cherici, X. Kuang, M. Poletti, M. Rucci, Precision of sustained fixation in trained and untrained observers. *J. Vis.* **12**, 31 (2012).
37. M. N. Agaoglu, M. Sit, D. Wan, S. T. Chung, Revas: An open-source tool for eye motion extraction from retinal videos obtained with scanning laser ophthalmoscopy. *Invest. Ophthalmol. Vis. Sci.* **59**, 2161–2161 (2018).
38. A. M. Clark, J. Intoy, M. Rucci, M. Poletti, Eye drift during fixation predicts visual acuity. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2200256119 (2022).
39. R. Engbert, K. Mergenthaler, P. Sinn, A. Pikovsky, An integrated model of fixational eye movements and microsaccades. *Proc. Natl. Acad. Sci. U.S.A.* **108**, E765–E770 (2011).
40. X. Kuang, M. Poletti, J. D. Victor, M. Rucci, Temporal encoding of spatial information during active visual fixation. *Curr. Biol.* **22**, 510–514 (2012).
41. K. Mergenthaler, R. Engbert, Modeling the control of fixational eye movements with neurophysiological delays. *Phys. Rev. Lett.* **98**, 138104 (2007).
42. J. A. Roberts, G. Wallis, M. Breakspear, Fixational eye movements during viewing of dynamic natural scenes. *Front. Psychol.* **4**, 60833 (2013).
43. J. J. Collins, C. J. De Luca, Open-loop and closed-loop control of posture: A random-walk analysis of center-of-pressure trajectories. *Exp. Brain Res.* **95**, 308–318 (1993).
44. H. Qian, M. P. Sheetz, E. L. Elson, Single particle tracking. Analysis of diffusion and flow in two-dimensional systems. *Biophys. J.* **60**, 910–921 (1991).
45. M. J. Saxton, Single-particle tracking: The distribution of diffusion coefficients. *Biophys. J.* **72**, 1744–1753 (1997).
46. C. J. Herrmann, R. Metzler, R. Engbert, A self-avoiding walk with neural delays as a model of fixational eye movements. *Sci. Rep.* **7**, 12958 (2017).
47. B. B. Mandelbrot, J. W. Van Ness, Fractional brownian motions, fractional noises and applications. *SIAM Rev.* **10**, 422–437 (1968).
48. J.C. D'Angelo, A. Roorda, Danglojc/paradoxicalmisperception. v1.1. Zenodo. https://zenodo.org/records/13351937. Deposited 20 August 2024.