

# Peri-Saccadic Natural Vision

Michael Dorr and Peter J. Bex

Schepens Eye Research Institute, Massachusetts Eye and Ear Infirmary, Department of Ophthalmology, Harvard Medical School, Boston, Massachusetts 02114

The fundamental role of the visual system is to guide behavior in natural environments. To optimize information transmission, many animals have evolved a non-homogeneous retina and serially sample visual scenes by saccadic eye movements. Such eye movements, however, introduce high-speed retinal motion and decouple external and internal reference frames. Until now, these processes have only been studied with unnatural stimuli, eye movement behavior, and tasks. These experiments confound retinotopic and geotopic coordinate systems and may probe a non-representative functional range. Here we develop a real-time, gaze-contingent display with precise spatiotemporal control over high-definition natural movies. In an active condition, human observers freely watched nature documentaries and indicated the location of periodic narrow-band contrast increments relative to their gaze position. In a passive condition under central fixation, the same retinal input was replayed to each observer by updating the video's screen position. Comparison of visual sensitivity between conditions revealed three mechanisms that the visual system has adapted to compensate for peri-saccadic vision changes. Under natural conditions we show that reduced visual sensitivity during eye movements can be explained simply by the high retinal speed during a saccade without recourse to an extra-retinal mechanism of active suppression; we give evidence for enhanced sensitivity immediately after an eye movement indicative of visual receptive fields remapping in anticipation of forthcoming spatial structure; and we demonstrate that perceptual decisions can be made in world rather than retinal coordinates.

## Introduction

Vision has traditionally been studied with observers who maintain stable fixation and with synthetic stimuli such as narrowband gratings. Under natural conditions, however, the time-varying input to the visual system comprises a broad distribution of spatiotemporal frequencies, orientations, intensities, and contrasts (Mante et al., 2005). Furthermore, observers typically make several eye movements per second (Hering, 1879; Buswell, 1935; Stelmach et al., 1991) that dramatically change the positions of objects on the retina. Nevertheless, we experience a stable world from this sequence of discrete retinotopic snapshots and are rarely aware of eye movements or the retinal motion that they introduce.

To understand the visual processing that maintains this continuity, a large body of research has examined visual sensitivity at the time of saccades. It is now well established that sensitivity is reduced around the time of saccadic eye movements (Matin et al., 1972; Volkman et al., 1978) (for review, see Ross et al., 2001; Castet, 2010; Ibbotson and Krekelberg, 2011; Burr and Morrone, 2011).

This sensitivity loss has been attributed to several possible factors. Some authors have argued for an active, extra-retinal process of suppression (Bridgeman et al., 1975; Volkman et al., 1978) of the magnocellular pathway (Burr and Ross, 1982; Burr et al., 1994) that does not affect sensitivity to equiluminant stimuli (Burr et al., 1994). The most parsimonious explanation is based on the contrast sensitivity function (Castet and Masson, 2000; García-Pérez and Peli, 2011), which is attenuated at high spatiotemporal frequencies (Kelly, 1961). During saccades, only very low spatial frequencies fall within the detectable temporal frequency range (Burr and Ross, 1982). Additionally, the post-saccadic presence of spatial structure might introduce backward masking of the blurred peri-saccadic retinal image (Campbell and Wurtz, 1978; Diamond et al., 2000; Castet et al., 2002). However, previous experiments have generally employed impoverished stimuli such as bars, dots, and gratings and combinations of low spatial frequencies and small window sizes that affected luminance instead of pattern detection. Furthermore, these experiments often required observers to make atypical and stereotyped eye movements.

It is currently the topic of axiomatic debate as to whether results obtained with such traditional methodologies transfer to real-world behavior (Felsen and Dan, 2005; Rust and Movshon, 2005). In this article, we study vision around the time of freely made eye movements in natural dynamic scenes. In each experimental session observers watched 10 min of an uninterrupted nature documentary. A gaze-contingent display was developed to modify high-definition videos in real time and in retinal coordinates (Fig. 1). At random time intervals, local contrast increments were applied to the underlying movie in one of four

Received Sept. 11, 2012; revised Nov. 19, 2012; accepted Nov. 23, 2012.

Author contributions: M.D. and P.J.B. designed research; M.D. and P.J.B. performed research; M.D. and P.J.B. contributed unpublished reagents/analytic tools; M.D. and P.J.B. analyzed data; M.D. and P.J.B. wrote the paper.

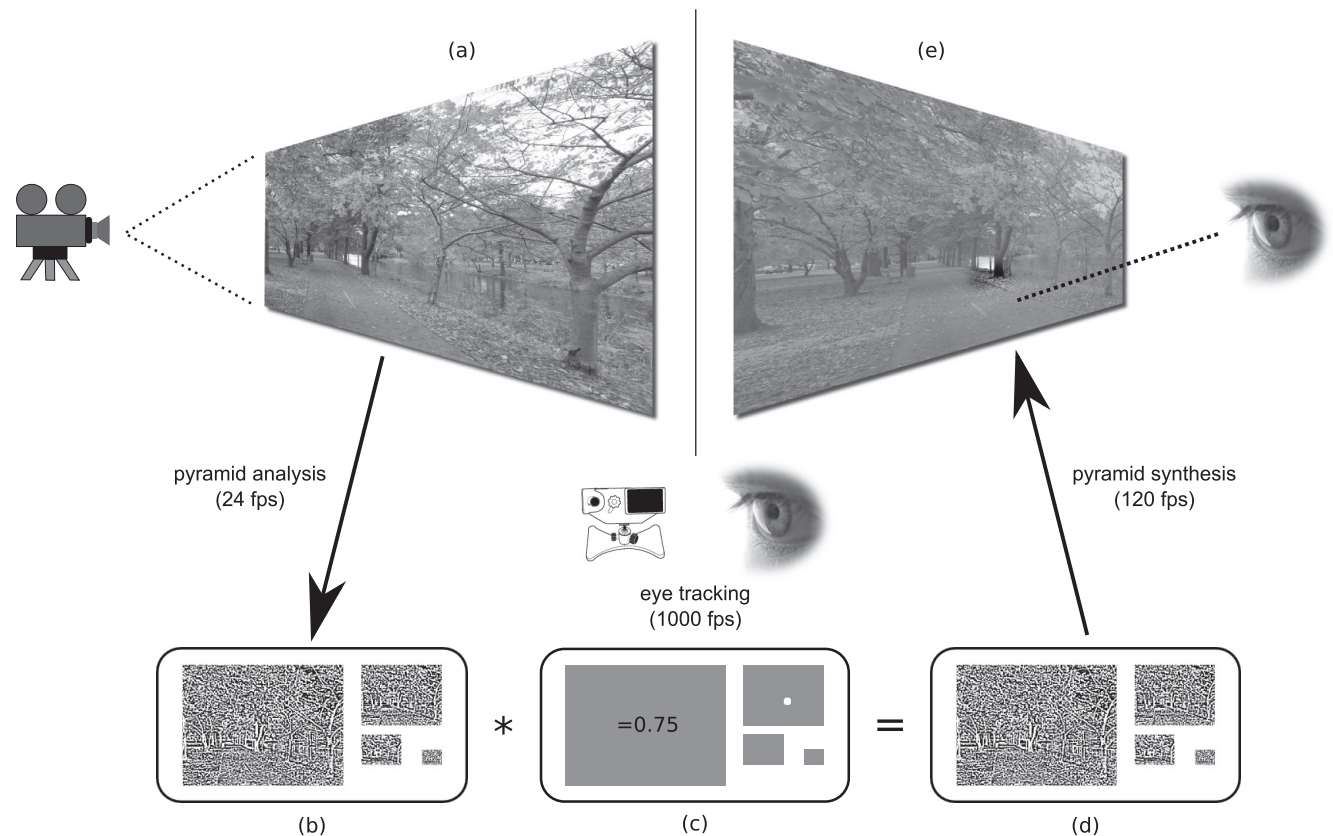
This work was supported by National Institutes of Health Grants EY018664 and EY019281 and an NVIDIA Academic Partnership.

The authors declare that they have no conflict of interest.

Correspondence should be addressed to either Michael Dorr or Peter Bex, Schepens Eye Research Institute, Massachusetts Eye and Ear Infirmary, Department of Ophthalmology, Harvard Medical School, 20 Staniford Street, Boston, MA 02114. E-mail: michael.dorr@schepens.harvard.edu or peter.bex@schepens.harvard.edu.

DOI:10.1523/JNEUROSCI.4344-12.2013

Copyright © 2013 the authors 0270-6474/13/331211-07\$15.00/0



**Figure 1.** Schematic overview of gaze-contingent display to study natural vision. *a, b*, Natural high-definition videos (*a*) are decomposed into spatiotemporal bands (four spatial bands shown in *b*) in real time. *c*, The observer's gaze point is recorded at 1000 Hz and is used to select arbitrary weights in retinal coordinates for each screen refresh (here, a small Gaussian increment next to the gaze point in one spatial band). *d, e*, During pyramid synthesis at 120 Hz, the pyramid bands are multiplied with the weights (*d*), the sum of which is a locally contrast-modulated output video (*e*). The latency of the whole system is typically <22 ms.

locations around the current gaze position. The observers' task was to indicate the location (4AFC) of the increments by pressing a corresponding button.

## Materials and Methods

Two authors and two unpaid naive subjects (males aged 29–45 years, all with normal or corrected vision) completed making observations in 20–30 sessions over a period of about a month, typically with one to four 10 min sessions in a particular day. All participants signed an informed consent form, and the study followed the guidelines of the Declaration of Helsinki and was approved by the institutional review board.

**Video processing.** Each video frame of Blu-ray nature documentaries was decomposed online at 24 frames per second (fps) into a Laplacian pyramid with seven scales (Adelson and Burt, 1981), which yields an efficient bandpass representation with individual bands ranging from 0.375 to 24 cycles per degree (cpd) in log steps.

First, frames were iteratively convoluted with a five-tap binomial lowpass filter and subsampled by a factor of 2 to create a Gaussian pyramid; such a pyramid has been used for a gaze-contingent display that simulates visual fields with a per pixel cutoff frequency (Geisler and Perry, 2002). After an upsampling and interpolation step, adjacent 24-bit pyramid levels were subtracted from each other to obtain individual spatial frequency bands at 48-bit color resolution. If these frequency bands are iteratively upsampled and added, the original image can be reconstructed; a local change to one of the bands before such synthesis leads to a localized band-specific change in the reconstructed image. Because typical videos already use the full dynamic range (pixel intensities, 0–255), contrast was reduced online (without any runtime penalty) to 75% by weighing all pyramid levels except the DC with 0.75 to create headroom for local contrast increments.

To meet real-time requirements, image processing was implemented on the graphics processing unit in the C for graphics shader language

(Mark et al., 2003); on a high-end workstation graphics board (NVIDIA Tesla C2070); pyramid analysis took about 3 ms and local target band modification and pyramid synthesis only 1.5 ms, allowing gaze-contingent screen updates every refresh cycle (120 Hz). The screen, a ViewSonic 3D VX2265wm, was carefully selected to have little output latency due to internal signal processing, which was measured with a photodiode at about 4 ms compared to a CRT. For eye tracking, a commercially available system (SR Research EyeLink 1000) was used with 1000 Hz sampling rate and a manufacturer-specified latency of 1.8 ms. Thus, the overall latency between an eye movement and its effect on the screen was 18–26 ms at the center of the screen, depending on eye movement onset relative to the screen refresh cycle. Note that this latency was accounted for in all analyses, i.e., time points refer to physical changes on the screen, not the issue time of drawing commands in the computer, which are often confounded in the literature. Our system latency is incurred largely due to limits of current display technology and not due to image processing. Shorter latencies could only be obtained with much lower spatial resolution displays, e.g., a CRT running at 200 Hz and with 800 by 600 pixels spatial resolution, which would reduce latency to a theoretical lower bound of 10–15 ms.

**Physical setup.** Subjects were seated 75 cm away from a ViewSonic 3D VX2265wm screen in a dark room, and their heads were stabilized in a head rest. At  $47.5 \times 29.7$  cm, the screen thus covered a visual angle of about  $35^\circ \times 25^\circ$ , and the maximum resolution was 24 cycles/degree.

Eye movements were recorded monocularly in the dominant eyes of three observers and the fellow eye of one observer using the commercially available video-oculographic EyeLink 1000 eye tracker. Before every experiment session, a 13-point calibration procedure was performed followed by a 13-point validation scheme. If necessary, these steps were repeated until calibration was accepted by the manufacturer's software.

**Stimuli.** The original Blu-ray video streams (1920 × 1080 pixels, 24 fps) were cropped to the central 1680 × 1050 pixels to fit the native screen resolution.

At the beginning of a session, only the unmodified video was shown for a random initial period of 5–7 s. Then, an additional target appeared briefly that consisted of a local contrast modification in one target frequency band, with its center 2° away from the current gaze position on one of the cardinal axes, updated at 120 Hz. Smooth spatiotemporal transitions were achieved by masking the target with a spatial Gaussian (SD = 0.5° × 0.5° and support 2° by 2°) and smoothly ramping onset and offset. The temporal presentation profile followed a Gaussian (SD = 12.5 ms) and contrast was incremented in either the 0.375–0.75 cpd or the 1.5–3 cpd band. Modification strength was randomly chosen from a set of values between 1.0 and 7.0. These values were used to locally multiply contrast values in the target frequency band. Because previously all (non-DC) frequency bands had been multiplied with 0.75 to achieve a base contrast of 75%, fewer than 0.5% pixels saturated at pixel intensity 0 or 255.

Immediately after target offset, a blue cross (1° × 1°) was presented gaze-contingently at fixation, prompting the subject to respond by pressing one of four buttons of a response box. There was no time limit for response, and the movie kept playing; feedback of correct and incorrect responses was given for 400 ms. Because the response prompt followed eye movements, subjects could also determine whether calibration quality had deteriorated, e.g., due to excessive head movements. Following feedback, the response prompt disappeared and the next target was presented after a random interval of 1.5–2.5 s.

Over the course of the experiment, eight episodes of the BBC Planet Earth series of about 1 h duration each were used. Recording sessions typically consisted of around 250–300 trials and lasted about 10 min. They could be terminated early if a movie had reached its end or because of calibration shifts. The following session resumed at the previous movie position.

In the passive condition, sessions were replicated while subjects maintained stable fixation at a central blue marker. Contrast increments appeared at the same spatiotemporal (movie) locations and with the same strength. However, instead of the subject actively moving their eyes, the video was shifted on the screen and thus on the subject's retina at 120 Hz, replaying their own previously recorded eye movements. Gaze data were not preprocessed except for blinks, during which the last valid gaze position determined stimulus shift. Because the movie kept playing continuously between trials, subjects only had limited time to respond in the passive condition (up to 300 ms before the next trial would have appeared in the active condition). Trials where subjects failed to respond in a timely manner (5.9% of trials) were excluded from the analysis, for both the active and passive conditions, to maintain their pairing. Overall, about 24,700 trial pairs were collected.

The very first recording session always comprised the active condition. For subsequent sessions, there was no fixed order of active/passive conditions, but subjects were given the choice between watching novel (active) and familiar (passive) video material, with the constraint that the number of passive sessions could not exceed that of active sessions. Theoretically, in the passive condition subjects could have been influenced by knowledge about target location and timing from the identical trial in the active condition. In practice, however, the distance between paired trials typically spanned several blocks with 250–300 trials each. Most importantly, any memory effect would only serve to improve performance in the passive condition, but we here found inferior performance compared to the active condition (see Results).

**Gaze analysis.** Trials where a blink occurred in the 150 ms preceding or during stimulus presentation were discarded. Saccades were extracted with a speed-based dual-threshold algorithm (Dorr et al., 2010). For peri-saccadic sensitivity data, the saccades to be analyzed were chosen by their temporal onset proximity to stimulus midpoints. For the mislocalization analysis, saccades were categorized into four bins within 45° of the cardinal axes.

**Statistical analysis.** Raw data in Figure 2 were fitted by a least-squares solver with an asymmetric inverted Gaussian having six parameters that controlled peak location, minimum, left, and right standard deviations,

and left and right maxima. To test statistical significance, this fit was performed 5000 times in a bootstrapping procedure. Trial indices were randomly resampled (with replacement) so that each bootstrap had the same number of trials as the original dataset and that the pairing between active and passive conditions was maintained. For each bootstrap, the active–passive difference of each parameter yielded one data point; the distribution of these differences should be centered near zero for statistically indistinguishable processes. Significance was thus assessed in a two-tailed test by comparing the percentile where the distribution of the differences for each parameter reached zero. For the analysis in Figure 4 Gaussians as above were further constrained to sum to 0.75 at each time point, and 95% confidence intervals were established by 5000 bootstrap resamples.

## Results

### Basic eye movement statistics

The distribution of saccade amplitudes was heavily skewed. Maximum measured amplitude was 30°, but 75% of eye movements had an amplitude of less than 7.8° (mean amplitude, 5.6°). These numbers are comparable to previous results from free-viewing naturalistic videos if corrected for the different screen sizes (Dorr et al., 2010). Fixation durations were slightly longer, but comparable to those reported for professionally cut video material (Dorr et al., 2010) (mean in the present study, 366 vs 340 ms for Hollywood movie trailers; median, 290 vs 251 ms).

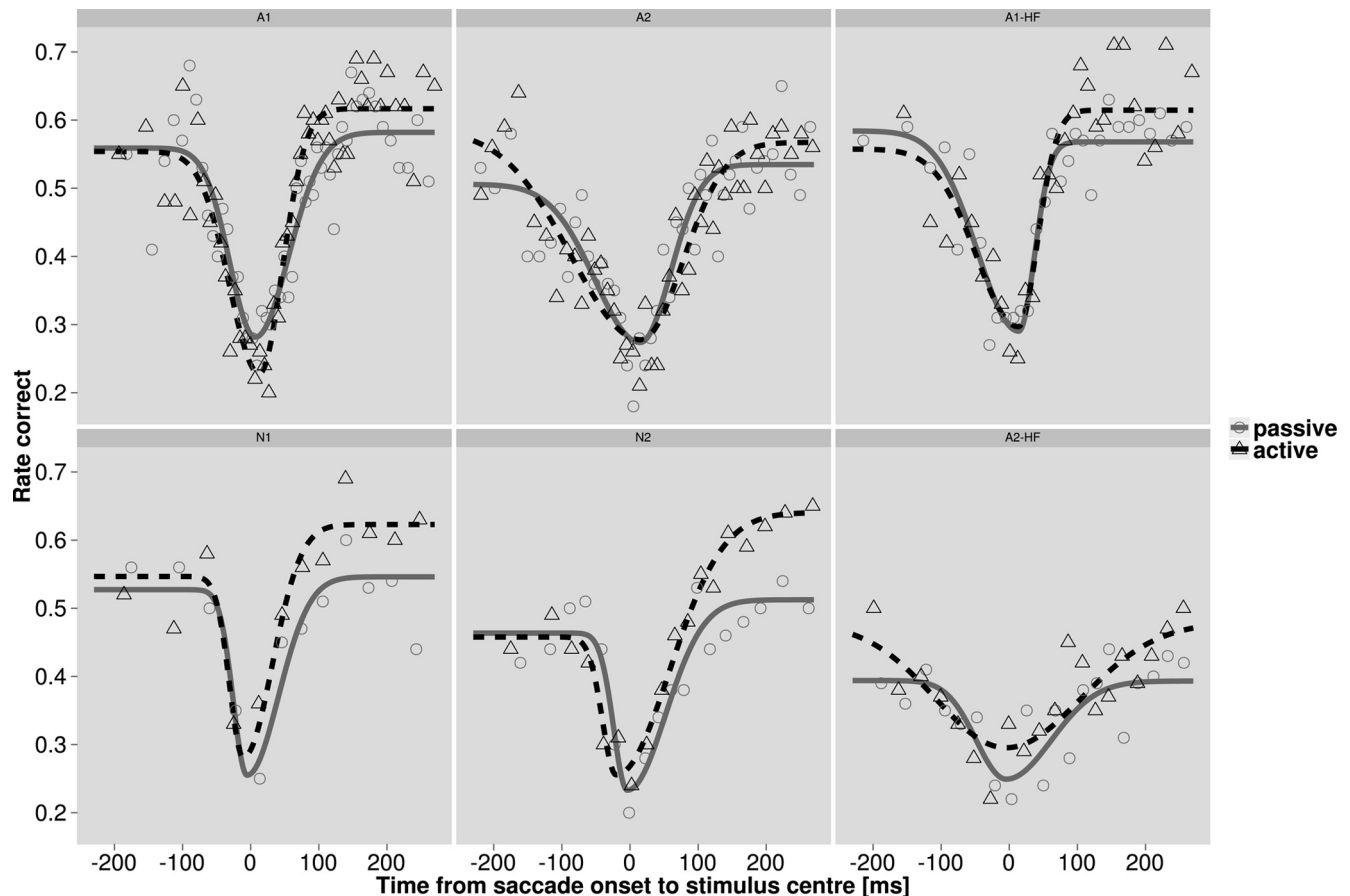
### Peri-saccadic sensitivity loss

We used our gaze-contingent display to analyze visual performance of the authors and two naive observers during uninterrupted viewing of natural dynamic scenes. In a first, active condition, viewing behavior was unconstrained. Therefore, stimuli could be presented at any time relative to freely made eye movements and fixations. Figure 2 shows target localization rates as a function of relative timing of stimulus presentation and saccade onset pooled across contrast levels for each observer.

The dashed black data are for the active condition where subjects freely made eye movements. Data were fit with an asymmetric inverted Gaussian (see Materials and Methods). For both low (0.375–0.75 cpd; Fig. 2, left four panels) and high (1.5–3 cpd; Fig. 2, right two panels) spatial frequency target bands, performance exhibits a clear dip at around the time of saccade. The data are consistent with previous reports (Volkman et al., 1978; Burr et al., 1994) obtained with traditional psychophysical stimuli.

To examine whether this effect reflects a process of extra-retinal visual suppression or is related to the higher retinal speed during fast eye movements, we extended our system to generate simulated, i.e., passive saccades. Video processing and stimulus generation were the same as in the active condition above. However, observers maintained steady fixation at the screen center while the whole processed video frame was shifted on the screen to reproduce the approximate retinal input (discretized at 120 Hz) of previous active recordings. Under both active and passive conditions, therefore, backward masking caused by post-saccadic spatial structure was also constant. The solid gray curves in Figure 2 show performance in this passive condition. These data quantify the change in sensitivity that can be attributed to the spatiotemporal characteristics of the retinal image during saccades. There was a loss in sensitivity that was not significantly different between active and passive conditions around the time of a saccade. This result indicates that retinal image speed rather than an extra-retinal suppression is the main cause of sensitivity loss during saccadic eye movements under natural viewing of complex scenes.





**Figure 2.** Perisaccadic sensitivity. Left four panels show data for contrast modifications in the 0.375–0.75 cpd band; right two panels show data for the 1.5–3 cpd band. Symbols denote bins of 100 trials; solid lines are the best-fitting asymmetric Gaussians. In the active condition (dashed line, triangles), subjects freely made eye movements; in the passive condition (solid line, circles), the same retinal input was replayed by shifting the video while subjects maintained central fixation. Sensitivity was impaired during both active and passive saccades. When contrast increments were presented immediately after saccades ( $t > 0$ ), sensitivity was enhanced compared to the passive condition.

### Post-saccadic sensitivity enhancement

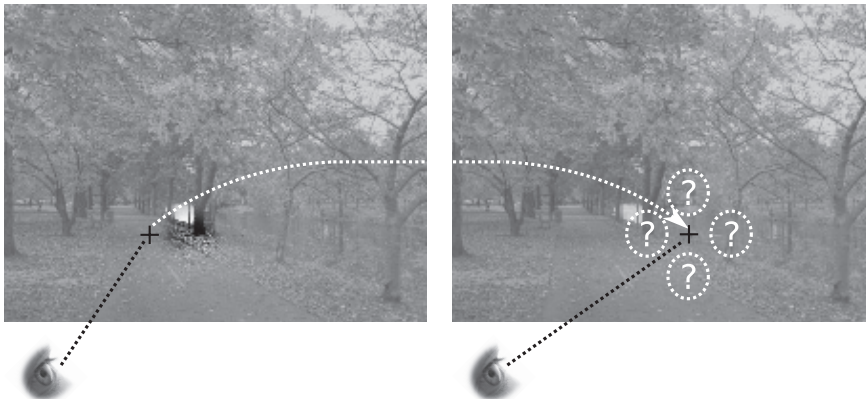
Importantly, when saccades were initiated before stimulus onset (positive time points in Fig. 2), visual sensitivity was significantly higher in the active than in the passive condition ( $p < 0.01$  across all observers and, individually,  $p < 0.05$  for five of six cases). In principle, such a sensitivity difference could possibly be explained by a difference in fixational eye movements between active and passive conditions; in the latter, observers maintained steady fixation on the central fixation cross, and indeed fixational eye movements were suppressed (data not shown). However, this lower rate of fixational eye movements did not differ between pre-saccadic and post-saccadic episodes except for a very small time window ( $< 50$  ms) around saccades, which cannot account for the active–passive difference at later points in time.

An equivalent comparison between pre-saccadic and post-saccadic performance for the active-only condition is frustrated by: (1) the presence of the response prompt at fixation after stimulus offset; (2) image differences that affect saccade rates; and (3) the lack of perfect pairing of stimulus strengths and pedestal contrasts that is possible for the active–passive comparison. Nevertheless, within the active, but not passive, condition, post-saccadic performance was significantly higher than the pre-saccadic performance in two of six cases. These results suggest an active role of anticipation of impending retinal image displacement when volitional eye movements are about

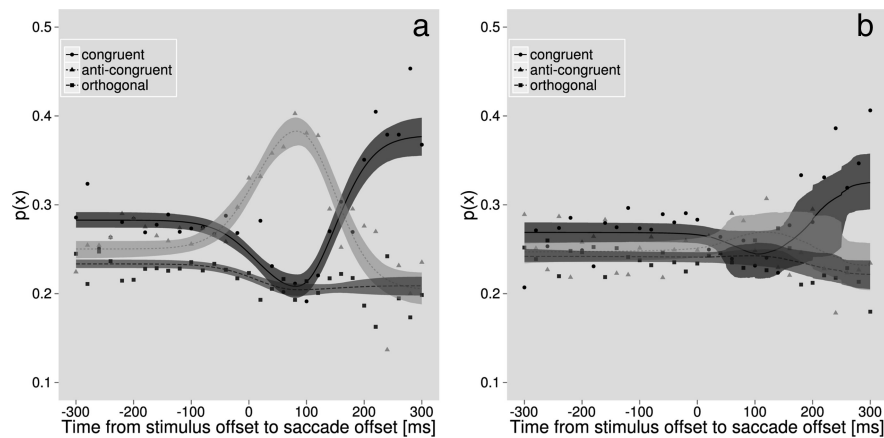
to be made, which causes an increase of visual sensitivity at the saccade endpoint.

### Pre-saccadic geotopic mislocalization

Introspectively, when saccades were made immediately after the stimulus (negative time points in Fig. 2), observers reported that they frequently detected the contrast increment, but failed to localize it correctly (Fig. 3). This effect could not be explained by peri-saccadic compression of space (Ross et al., 1997), which would not alter the relative location of the target. We therefore examined the interaction between saccade direction and perceived target position relative to gaze. Figure 4*a* shows an analysis of the active condition expressed as the probability of response directions relative to eye movement directions. For a given eye movement, a response could have been made in one of three directions relative to the eye movement: the same direction (“congruent”; black circles, solid), the opposite direction (“anti-congruent”; light gray triangles, dashed), or “orthogonal” (medium gray squares) to the eye movement. Note that we here averaged the two possible orthogonal directions so that the probabilities at every time point in Figure 4*a* sum to 0.75. There are three time periods of interest in Figure 4*a*. When eye movements were executed earlier than 50 ms before the stimulus was presented, as expected the subsequent response was mostly unrelated to the eye movement direction (all three curves close to 25%). Eye movements executed later than 150 ms after the stim-



**Figure 3.** Dissociation between retinotopic and geotopic coordinate systems. In rich natural scenes, target locations frequently coincide with identifiable features. In this example, a contrast increment is presented retinotopically to the right of the fovea and geotopically appears on a tree trunk. After a large saccade near stimulus offset, retinotopic and geotopic positions can be reversed. Here, the response in retinal coordinates remains “right” after a large right-ward saccade, whereas the geotopic target location (tree trunk) is now on the left of the fovea. Subjects frequently erroneously report the geotopic location of the target (see Fig. 4*a*).



**Figure 4.** Geotopic mislocalization. Symbols denote means in 20-ms-wide temporal bins; shaded areas indicate 95% confidence intervals of the probability of eye movement directions relative to response directions. Congruent eye movements (solid black) are in the same direction as the subsequent response, whereas anti-congruent eye movements (dashed light gray line) are in the opposite direction. For example, a “left” response (to the tree trunk) in Figure 3 would be incongruent with the rightward saccade. **a**, Data for the active condition. Briefly after stimulus offset, responses were most likely to be in the opposite direction of the eye movement (light gray triangles, dashed line). Later, after stimulus offset, responses were most likely to be in the same direction (solid black curve) as the eye movement. **b**, Data for the passive condition.

ulus was presented were presumably often triggered by detection of the stimulus and, consequently, there was a higher probability of a response in the same direction as the eye movement (rise of the black solid curve on the right of Fig. 4*a*). Eye movements executed between 50 ms before and 150 ms after the stimulus were planned before the stimulus was presented and therefore could not have been influenced by the stimulus. Surprisingly, there was a high probability of response in the opposite direction to the eye movement (peak of the light gray dashed curve ~80 ms after stimulus offset). This result shows that observers frequently made the geotopic localization judgment illustrated in Figure 3. To disambiguate effects of the retinal input and volitional eye movement control, we performed the same analysis for the passive condition (Fig. 4*b*). When eye movements were only simulated, the relationship between response direction and retinal image shift (simulated eye movement) direction was the same as that in the active condition for orthogonal (medium gray) and congruent (solid black) cases (Fig. 4*b*). Importantly, however,

there was no indication of a geotopic mislocalization for stimuli presented just before an image shift (corresponding to an eye movement from the active condition).

## Discussion

Traditional experimental paradigms with simple stimuli and stereotypical eye behaviors have provided us with widely accepted models of visual function in mammalian brains. However, the ultimate purpose of the visual system is to guide behavior in natural environments that fundamentally differ from laboratory settings. The translation from laboratory to natural yet controlled conditions and tasks has until now been frustrated by technical limitations.

We here implemented a novel methodology to study naturalistic vision, supporting real-time, gaze-contingent modifications of high-resolution natural videos that were freely watched by observers. This technology allowed us to investigate some classic phenomena under more natural conditions and to make observations that would be difficult or impossible to make with simple stimuli and controlled eye movements. Specifically, we here make three main observations.

Firstly, when observers freely made eye movements, sensitivity to localized transient contrast increments was impaired around the time of eye movements. This confirms previous reports (for review, see Ross et al., 2001; Castet, 2010; Ibbotson and Krekelberg, 2011) with traditional stimuli and unnaturally large (Burr et al., 1994) saccade amplitude distributions or microsaccades (Ditchburn, 1955) and extends them to natural stimuli and saccades (Dorr et al., 2010).

It has previously been claimed that the sensitivity loss is confined to very low spatial frequencies below 0.5 cycles per degree (Diamond et al., 2000). However, in the published data (Volkman et al., 1978; Burr et al., 1994) sensitivity is reduced by about 30–50% even for higher spatial frequencies. In our experiment, we used two different frequency bands. One band (0.375–0.75 cpd) corresponded to the lowest frequencies recorded in neurons of the primate visual system (De Valois et al., 1982), and the other band (1.5–3 cpd) corresponded roughly to the peak of the human contrast sensitivity function (Campbell and Robson, 1968). In both cases, there was a comparable loss of sensitivity, challenging the claim that peri-saccadic suppression is limited to the magnocellular pathway.

Furthermore, peri-saccadic suppression phenomena have previously been attributed to a top-down, extra-retinal efference copy that serves to minimize the perception of sudden global retinal image displacements (Burr et al., 1994). Other authors have argued that sensitivity losses are a bottom-up consequence of retinal image motion (Castet, 2010; García-Pérez and Peli,

2011) or masking during the saccade (Campbell and Wurtz, 1978; Diamond et al., 2000; Castet et al., 2002).

We attempt to settle this controversy for naturalistic viewing conditions by comparing sensitivity for real and simulated eye movements that produce approximately the same retinal input and differ only in the intentional control of the former. Previous work (Diamond et al., 2000) has used mirror-simulated eye movements to examine this question. Phase discrimination of an isolated target on a blank background was impaired by real but not simulated eye movements, which is consistent with both positional uncertainty and contrast sensitivity loss. In the presence of a structured background, however, contrast sensitivity was lost for both active and passive conditions.

Naturalistic stimuli contain chromatic variation that should not be actively suppressed and may even be enhanced during saccades (Burr et al., 1994; Diamond et al., 2000). With our apparatus, we replayed the same naturalistic video input with targets of both low and high spatial frequencies to dissociate active from passive retinal image translations. This matched the spatio-temporal spectrum of the retinal image in both conditions and thus directly quantified the contrast sensitivity account of peri-saccadic sensitivity loss. Sensitivity was equally impaired during active and passive saccades, and there was no evidence for enhancement of chromatic sensitivity. These data do not necessarily rule out evidence for saccadic suppression under more constrained conditions. However, for natural vision, we conclude that the most parsimonious explanation for peri-saccadic sensitivity loss is the known poor sensitivity to high-speed luminance (Kelly, 1979) and chromatic (Kelly and van Norren, 1977) retinal images. This effect of peri-saccadic vision change is fully characterized by the passive condition (gray curves in Fig. 2).

Our second main observation concerns the remarkable stability of visual perception across eye movements. This problem is not solved by a loss in sensitivity during saccades alone. To achieve this stability, the visual system must also integrate pre- and post-saccadic information. Electrophysiological studies have shown that receptive fields of single units in the parietal cortex remap to new locations in anticipation of eye movements (Duhamel et al., 1992; Duhamel et al., 1997). A recent behavioral study (Rolfes et al., 2011) showed pre-saccadic perceptual enhancement for briefly flashed Gabors at the future retinotopic locations of attended locations, which was attributed to an updating of attention before an eye movement.

Based on these data, it has therefore been proposed that remapping may be the mechanism by which trans-saccadic stability is achieved. Our methodology allows a direct test of this hypothesis, because predictive remapping is only possible under active but not passive conditions. For stimuli that were presented immediately after active saccades (positive time points in Fig. 2), our results show that sensitivity was enhanced. Previous work has reported sensitivity changes related to eye movement behavior. Stabilizing the retinal input by cancelling microsaccadic eye movements reduces sensitivity to high spatial frequencies (Kelly, 1979; Kelly, 1981; Rucci et al., 2007), but this equally affects both pre- and post-saccadic fixations, unlike the differences observed here. Similarly, attention can increase sensitivity during smooth pursuit eye movements (Schütz et al., 2008). Here, we observe an elevation in sensitivity after an active saccade and therefore provide direct behavioral evidence for sensitivity changes under natural conditions. This result is consistent with receptive field remapping in anticipation of changes in the retinal input caused by the observer's own saccades.

Thirdly, it has been proposed that a functional purpose of remapping mechanisms is to support a stable perception of the visual world. However, remapping is only a necessary but not sufficient process for visual stability, and recent reviews (Wurtz, 2008; Ibbotson and Krekelberg, 2011) also conclude that at some stage there must exist a spatiotopic (i.e., geotopic) representation. Despite the importance of such geotopic coding, behavioral evidence for its existence remains controversial. Several adaptation phenomena are stronger when retinotopic and geotopic coordinates are aligned (Nishida et al., 2003). Furthermore, adaptation (Melcher, 2005; Melcher, 2007) and sensory thresholds sum across saccades at the same geotopic location (Melcher and Morrone, 2003), and memory for target position is stronger in geotopic than in retinotopic coordinates (Lin and Gorea, 2011). While these findings argue in favor of geotopic encoding, others have argued that this transfer is global but not geotopic (Knapen et al., 2009) and that spatial information decays faster in geotopic than retinotopic coordinates (Golomb and Kanwisher, 2012).

Traditional stimulus methodologies typically have a limited number of isolated stimuli, explicit spatial references (such as compulsory fixation markers), and stereotyped eye movements. For example, the conflicting results for faster decay rates of retinotopic (Lin and Gorea, 2011) or geotopic (Golomb and Kanwisher, 2012) memory thus may be attributed to the quantity of landmarks in the scene (Golomb and Kanwisher, 2012). Therefore, geotopic and retinotopic position are often confounded, which complicates direct behavioral tests of geotopic representation.

Our methodology allowed a more sensitive analysis of the coordinate system within which spatial location is encoded. There was a systematic relationship between eye movement direction and the perceived location of the contrast increment. When stimuli were presented immediately before an eye movement was executed, the subsequent response had a high probability of being in the opposite direction of the eye movement (dashed light gray curve in Fig. 4*a*). One possible explanation for this effect is that observers may shift their attention toward the landing point of a saccade before it is actually executed (Deubel and Schneider, 1996). Consequently, the perceived fixation location may precede the physical location by up to 40 ms (Hunt and Cavanagh, 2009). In principle, this could place the stimulus location on the opposite side of the perceived fixation location and give rise to a response in the opposite direction to the eye movement. This predictive phenomenon means that the peak of the dashed light gray curve in Figure 4*a*, which occurs at  $\approx 80$  ms, should not occur later than  $t = 0$ , when stimulus and saccade offsets coincide.

Alternatively, targets may not be perceived in retinal coordinates, but in coordinates defined by physical landmarks in the scene. In this case, the geotopic position of the stimulus will also be shifted in the opposite direction of the eye movement, as illustrated in Figure 3. This hypothesis requires only that an eye movement is made after stimulus offset. The data, with a peak at about 80 ms after stimulus offset, therefore provide evidence that the visual system can encode spatial position in natural scenes in geotopic rather than retinotopic coordinates.

In summary, in this manuscript we developed a technological framework to conduct controlled research with ecologically relevant stimuli and tasks. With continuous gaze-contingent displays and with unconstrained eye movements, we demonstrate several key mechanisms used by the visual system to maintain a stable representation of the visual world across abrupt changes in



the retinal image caused by eye movements. Sensitivity to the intrusion of retinal motion caused by the saccade is attenuated by the spatiotemporal tuning of the visual system. Additionally, encoding of the post-saccadic input is facilitated between receptive fields by remapping of information across saccades. Lastly, under natural conditions, pre- and post-saccadic retinal images must be registered against a geotopic representation of a complex world that is not well characterized by traditional stimuli.

## Notes

Supplemental material for this article is available at <http://michaeldorr.de/perisaccadicvision/>. The website includes example videos visualizing the experimental procedure. This material has not been peer reviewed.

## References

- Adelson EH, Burt PJ (1981) Image data compression with the Laplacian pyramid. In: *Proceeding of the conference on pattern recognition and image processing*, 218–223. Los Angeles: IEEE Computer Society.
- Bridgeman B, Hendry D, Stark L (1975) Failure to detect displacement of the visual world during saccadic eye movements. *Vis Res* 15:719–722. [CrossRef Medline](#)
- Burr DC, Morrone MC (2011) Spatiotopic coding and remapping in humans. *Philos Trans R Soc Lond B Biol Sci* 366:504–515. [CrossRef Medline](#)
- Burr DC, Ross J (1982) Contrast sensitivity at high velocities. *Vis Res* 22:479–484. [CrossRef Medline](#)
- Burr DC, Morrone MC, Ross J (1994) Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature* 371:511–513. [CrossRef Medline](#)
- Buswell GT (1935) *How people look at pictures: A study of the psychology of perception in art*. Chicago: University of Chicago.
- Campbell FW, Robson JG (1968) Application of Fourier analysis to the visibility of gratings. *J Physiol* 197:551–566. [Medline](#)
- Campbell FW, Wurtz RH (1978) Saccadic omission: why we do not see a grey-out during a saccadic eye movement. *Vis Res* 18:1297–1303. [CrossRef Medline](#)
- Castet E (2010) Perception of intra-saccadic motion. In: *Dynamics of visual motion processing: neuronal, behavioral, and computational approaches*, Chap 10 (Ilg UJ, Masson GS, eds), pp 213–238. New York: Springer.
- Castet E, Masson GS (2000) Motion perception during saccadic eye movements. *Nat Neurosci* 3:177–183. [CrossRef Medline](#)
- Castet E, Jeanjean S, Masson GS (2002) Motion perception of saccade-induced retinal translation. *Proc Natl Acad Sci U S A* 99:15159–15163. [CrossRef Medline](#)
- Deubel H, Schneider WX (1996) Saccade target selection and object recognition: evidence for a common attentional mechanism. *Vis Res* 36:1827–1837. [CrossRef Medline](#)
- De Valois RL, Albrecht DG, Thorell LG (1982) Spatial frequency selectivity of cells in macaque visual cortex. *Vis Res* 22:545–559. [CrossRef Medline](#)
- Diamond MR, Ross J, Morrone MC (2000) Extraretinal control of saccadic suppression. *J Neurosci* 20:3449–3455. [Medline](#)
- Ditchburn RW (1955) Eye-movements in relation to retinal action. *Optica Acta* 1:171–176. [CrossRef](#)
- Dorr M, Martinetz T, Gegenfurtner K, Barth E (2010) Variability of eye movements when viewing dynamic natural scenes. *J Vis* 10(10):28 1–17. [CrossRef Medline](#)
- Duhamel JR, Colby CL, Goldberg ME (1992) The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* 255:90–92. [CrossRef Medline](#)
- Duhamel JR, Bremmer F, Ben Hamed S, Graf W (1997) Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature* 389:845–848. [CrossRef Medline](#)
- Felsen G, Dan Y (2005) A natural approach to studying vision. *Nat Neurosci* 8:1643–1646. [CrossRef Medline](#)
- García-Pérez MA, Peli E (2011) Visual contrast processing is largely unaltered during saccades. *Front Psychol* 2:247. [CrossRef Medline](#)
- Geisler WS, Perry JS (2002) Real-time simulation of arbitrary visual fields. In: *ETRA '02: Proceedings of the 2002 symposium on eye tracking research and applications*, pp 83–87. New York: ACM.
- Golomb JD, Kanwisher N (2012) Retinotopic memory is more precise than spatiotopic memory. *Proc Natl Acad Sci U S A* 109:1796–1801. [CrossRef Medline](#)
- Hering E (1879) Über die Muskelgeräusche des Auges [On muscle sounds of the eye] (in German). *Sitzberichte der kaiserlichen Akademie der Wissenschaften in Wien. Mathematisch-naturwissenschaftliche Klasse* 79:137–154.
- Hunt AR, Cavanagh P (2009) Looking ahead: the perceived direction of gaze shifts before the eyes move. *J Vis* 9(9):1 1–7. [CrossRef Medline](#)
- Ibbotson M, Kregelberg B (2011) Visual perception and saccadic eye movements. *Curr Opin Neurobiol* 21:553–558. [CrossRef Medline](#)
- Kelly DH (1961) Visual response to time-dependent stimuli. I. Amplitude sensitivity measurements. *J Opt Soc Am* 51:422–429. [CrossRef Medline](#)
- Kelly DH (1979) Motion and vision. II. Stabilized spatiotemporal threshold surface. *J Opt Soc Am* 69:1340–1349. [CrossRef Medline](#)
- Kelly DH (1981) Disappearance of stabilized chromatic gratings. *Science* 214:1257–1258. [CrossRef Medline](#)
- Kelly DH, van Norren D (1977) Two-band model of heterochromatic flicker. *J Opt Soc Am* 67:1081–1091. [CrossRef Medline](#)
- Knapen T, Rolfs M, Cavanagh P (2009) The reference frame of the motion aftereffect is retinotopic. *J Vis* 9(5):16 1–7. [CrossRef Medline](#)
- Lin IF, Gorea A (2011) Location and identity memory of saccade targets. *Vis Res* 51:323–332. [CrossRef Medline](#)
- Mante V, Frazor RA, Bonin V, Geisler WS, Carandini M (2005) Independence of luminance and contrast in natural scenes and in the early visual system. *Nat Neurosci* 8:1690–1697. [CrossRef Medline](#)
- Mark WR, Glanville RS, Akeley K, Kilgard MJ (2003) Cg: a system for programming graphics hardware in a C-like language. In: *SIGGRAPH '03: ACM SIGGRAPH 2003 papers*, 896–907, New York: ACM.
- Matin E, Clymer AB, Matin L (1972) Metacontrast and saccadic suppression. *Science* 178:179–182. [CrossRef Medline](#)
- Melcher D (2005) Spatiotopic transfer of visual-form adaptation across saccadic eye movements. *Curr Biol* 15:1745–1748. [CrossRef Medline](#)
- Melcher D (2007) Predictive remapping of visual features precedes saccadic eye movements. *Nat Neurosci* 10:903–907. [CrossRef Medline](#)
- Melcher D, Morrone MC (2003) Spatiotopic temporal integration of visual motion across saccadic eye movements. *Nat Neurosci* 6:877–881. [CrossRef Medline](#)
- Nishida S, Motoyoshi I, Andersen RA, Shimojo S (2003) Gaze modulation of visual aftereffects. *Vis Res* 43:639–649. [CrossRef Medline](#)
- Rolfs M, Jonikaitis D, Deubel H, Cavanagh P (2011) Predictive remapping of attention across eye movements. *Nat Neurosci* 14:252–256. [CrossRef Medline](#)
- Ross J, Morrone MC, Burr DC (1997) Compression of visual space before saccades. *Nature* 386:598–601. [CrossRef Medline](#)
- Ross J, Morrone MC, Goldberg ME, Burr DC (2001) Changes in visual perception at the time of saccades. *Trends Neurosci* 24:113–121. [CrossRef Medline](#)
- Rucci M, Iovin R, Poletti M, Santini F (2007) Miniature eye movements enhance fine spatial detail. *Nature* 447:851–854. [CrossRef Medline](#)
- Rust NC, Movshon JA (2005) In praise of artifice. *Nat Neurosci* 8:1647–1650. [CrossRef](#)
- Schütz AC, Braun DI, Kerzel D, Gegenfurtner KR (2008) Improved visual sensitivity during smooth pursuit eye movements. *Nat Neurosci* 11:1211–1216. [CrossRef Medline](#)
- Stelmach LB, Tam WJ, Hearty PJ (1991) Static and dynamic spatial resolution in image coding: An investigation of eye movements In: *Human vision, visual processing and digital display II, Proceedings of the SPIE*, Vol. 1453, pp 147–152. Los Angeles: IEEE Computer Society.
- Volkman FC, Riggs LA, White KD, Moore RK (1978) Contrast sensitivity during saccadic eye movements. *Vis Res* 18:1193–1199. [CrossRef Medline](#)
- Wurtz RH (2008) Neuronal mechanisms of visual stability. *Vis Res* 48:2070–2089. [CrossRef Medline](#)