

# Image Structure at the Center of Gaze during Free Viewing

Valentin Dragoi<sup>1,2</sup> and Mriganka Sur<sup>2</sup>

## Abstract

■ It is generally believed that the visual system is adapted to the statistics of the visual world. Measuring and understanding these statistics require precise knowledge of the structure of the signals reaching fovea during image scanning. However, despite the fact that eye movements cause retinal stimulation to change several times in a second, it is implicitly assumed that images are sampled uniformly during natural viewing. By analyzing the eye movements of three rhesus monkeys freely viewing

natural scenes, we report here significant anisotropy in stimulus statistics at the center of gaze. We find that fixation on an image patch is more likely to be followed by a saccade to a nearby patch of similar orientation structure or by a saccade to a more distant patch of largely dissimilar orientation structure. Furthermore, we show that orientation-selective neurons in the primary visual cortex (V1) can take advantage of eye movement statistics to selectively improve their discrimination performance. ■

## INTRODUCTION

When looking at a visual scene, primates, including humans, acquire information by selecting particular locations based on local image cues or on their behavioral relevance. This selection process typically involves saccadic eye movements that bring relevant image regions into fovea. It is well accepted that eye movements change “foveal” stimulation several times a second (Yarbus, 1967), and yet studies characterizing the statistics of stimuli on fovea have rarely included eye movements in the description of visual input.

An important characteristic of natural scenes is their correlational structure, that is, neighboring image patches are highly correlated in local features, such as line orientation or spatial frequency, whereas distant patches differ in their local structure (Dragoi, Sharma, Miller, & Sur, 2002; Reinagel & Zador, 1999; Tolhurst, Tadmor, & Chao, 1992; Field, 1987). In principle, these statistical inhomogeneities could be used by the visual system to efficiently encode incoming stimuli (Simoncelli & Olshausen, 2001; van Hateren & Ruderman, 1998; Field, 1987; Attneave, 1954) and guide subsequent eye movements. Thus, it has long been suggested that eye movements tend to cluster preferentially around image locations of high informational content, such as shape contours or sharp changes in textural gradients (Marr, 1982; Antes, 1974; Locher & Nodine, 1973). However, the manner by which local properties of image locations inspected during eye movements change during

the time course of viewing remains unknown. Answering this question could potentially clarify what local properties of the visual scene determine which areas are selected as targets for eye movements and thus contribute to our understanding of the neural principles of image encoding by the visual system. Indeed, visual processing is hypothesized to be adapted to the statistics of the visual environment and to the way we scan it (Betsch, Einhauser, Kording, & Konig, 2004; Dragoi, Sharma, Miller, et al., 2002; Dragoi, Turcu, & Sur, 2001; Simoncelli & Olshausen, 2001; van Hateren & Ruderman, 1998; Field, 1987; Attneave, 1954). Hence, understanding these statistics can give us clues about how the visual system takes advantage of rapid image sequences during eye movements to encode incoming stimuli efficiently. One example is the efficient coding hypothesis (Simoncelli, 2003; Barlow, 1990, 2001), that is, the fact that individual cortical neurons use their full range of responses to encode information. Thus, knowledge of stimulus statistics during image scanning could be used to examine whether and how the range of neuronal responses in the visual cortex has efficiently adapted to these statistics.

We addressed these issues by recording the eye position of three rhesus monkeys freely viewing natural scenes, and we asked how local image features, such as orientation, at the center of gaze change as monkeys move their eyes. We focused on the orientation of a foveal image as a key local feature that is essential for defining complex shapes in natural scenes, with a distinct representation in the visual cortex of higher mammals. We found significant anisotropy in stimulus statistics at the center of gaze; fixation to an image

---

<sup>1</sup>University of Texas-Houston Medical School, <sup>2</sup>Massachusetts Institute of Technology

patch is more likely to be followed by a saccade to a nearby patch of similar orientation structure or by a saccade to a more distant patch of largely dissimilar orientation structure. These findings are complemented by physiological recordings in V1 of alert monkey; we found that individual neurons dynamically process orientation in a way that is adapted to the statistics of natural viewing. That is, brief exposure to either nearby or orthogonal orientations can increase the dynamic range of single neurons near their peak orientation by altering the slope of orientation tuning curves. This would improve orientation discrimination performance following nearby or orthogonal adaptation, as is indeed observed psychophysically.

## METHODS

### Psychophysics

All experiments were performed under protocols that were approved by the Massachusetts Institute of Technology's Animal Care and Use Committee and conformed to National Institutes of Health guidelines. We recorded the eye positions of three adult behaving monkeys (*Macaca mulatta*) freely viewing natural scenes. During a typical free-viewing session, animals viewed a sequence of 10 natural images, each presented for 10 sec, separated by intervals of 20 sec. Images were randomized and presented for 5–7 trials each (we used a total of 200 images; each monkey viewed at least 140 images). Eye position was continuously monitored using an infrared eye tracking system (Iscan, Inc.). Eye positions were calibrated before each free-viewing session using a 5-point calibration procedure in which the animal was required to fixate on each one of the 4 corners and on the center of the monitor. During free viewing, no behavioral response was required (reward was occasionally provided to maintain alertness). We registered a fixation when the eye remained within a  $0.3^\circ \times 0.3^\circ$  window for at least 70 msec and eye velocity was less than 5% of the maximum velocity within image borders (the maximum velocity was calculated separately for each free-viewing session). Removing the eye velocity criterion underestimates the number of consecutive short-range fixations. Images subtended between  $10^\circ$  and  $11^\circ$  in width and between  $8^\circ$  and  $9^\circ$  in height (21 pixels/deg; 256 Gy) and were extracted from a high-resolution commercial photo-CD library. Images were presented both in landscape and in portrait modes and were displayed on a 60-Hz computer monitor placed 57 cm in front of the animal.

### Image Processing

The mean orientation and orientation selectivity index (OSI) of image patches were calculated as follows: Fourier components were extracted from the orienta-

tion magnitude histogram for each patch:  $a = \sum_{i=0}^{N-1} M(\theta_i) \cos(2\theta_i)$ ;  $b = \sum_{i=0}^{N-1} M(\theta_i) \sin(2\theta_i)$ , where  $M(\theta_i)$  represents the orientation magnitude obtained for the set of  $N = 36$  orientations ( $\theta_i, i = 0, 1, \dots, N - 1$ ), which are uniformly distributed over  $0$ – $180^\circ$ . The patch mean orientation  $\theta$  was calculated as  $\theta = 0.5 \arctan(b/a)$  if  $a > 0$  or as  $\theta = 90 + 0.5 \arctan(b/a)$  if  $a < 0$ . If  $a > 0$  and  $b < 0$ ,  $\theta = 180 + 0.5 \arctan(b/a)$ . Patch OSI was given by  $OSI = c/M$ , where  $c = \sqrt{(a^2 + b^2)}$  and  $M$  is the mean orientation magnitude.

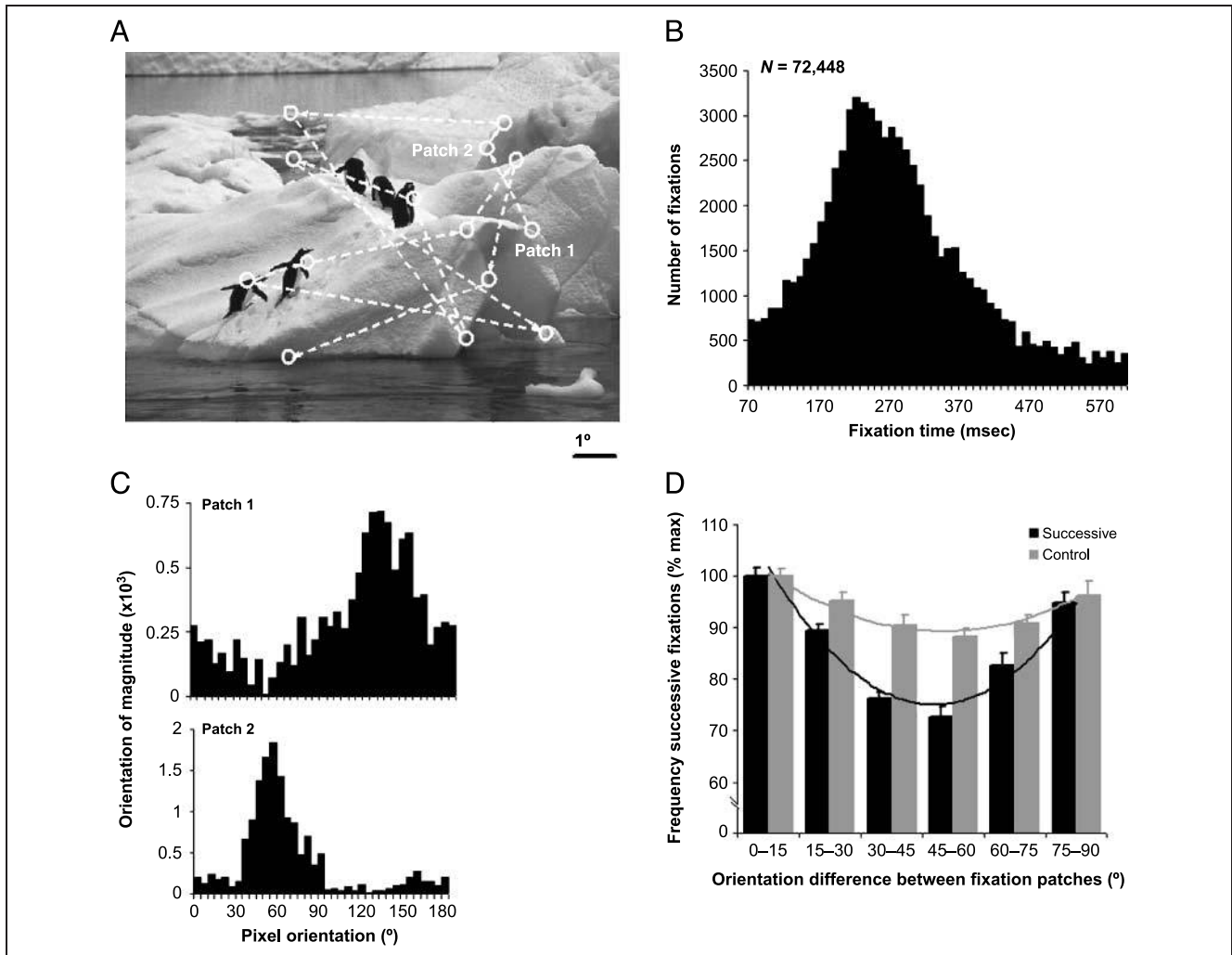
## Electrophysiology

Single neuron recordings were made from two alert macaque monkeys with extracellular electrodes lowered transdurally. Details of recordings are given elsewhere (Dragoi, Sharma, Miller, et al., 2002; Dragoi, Turcu, et al., 2001). Eye position was continuously monitored using an infrared eye tracking system (Iscan, Inc.). Monkeys were trained to fixate on a small spot ( $0.1^\circ$ ) presented on a video monitor placed at a distance of 57 cm. Once the animal had achieved stable fixation for 100 msec, a visual stimulus was presented within neurons' receptive field (receptive fields were located within  $2.5^\circ$  of the center of fixation). The monkeys were required to maintain fixation throughout the stimulus presentation to earn a juice reward; the trial was automatically aborted if fixation instability exceeded  $0.25^\circ$  at any time during stimulus presentation. Stimuli were movie strips in which each frame consisted of a  $5^\circ \times 5^\circ$  sine-wave grating of 2 cycles/deg spatial frequency and 75% contrast level presented using a binocular. Each movie strip was presented for approximately 1.86 sec (16 orientations  $\times$  7 repeats  $\times$  60 Hz). Stimulus spatial phase was randomized across frames for each orientation. Control and adaptation trials were grouped in blocks of 400–500 trials, with 2 repeats each. Reverse correlation calculation was averaged over the entire movie duration of 1.86 sec during control, adaptation, and recovery conditions (eliminating onset response to movie sequence; i.e., the first 200–400 msec did not alter our results significantly). To increase the effectiveness of adaptation, we used drifting sine-wave gratings (temporal frequency = 3 Hz; similar spatial characteristics on flashed gratings) as adapting stimuli because we found that drifting gratings evoke stronger responses than flashed stimuli. Preferred orientation was calculated as described previously (Dragoi, Sharma, Miller, et al., 2002; Dragoi, Sharma, & Sur, 2000). To determine orientation discrimination performance, we calculated  $d'$  before and after adaptation for each time delay between spike train and movie sequence as  $d' = (R_1 - R_2) / \sqrt{([SD_1]^2 + [SD_2]^2)/2}$ , that is, the difference between mean spike rates at two adjacent orientations divided by root mean square standard deviation (Green & Swets, 1966).

## RESULTS

We used 200 images depicting foods, landscapes, or animals in their natural environment, which were presented for 10 sec each, whereas a pupil tracking system recorded eye position. Figure 1A shows an example of an eye scan path (white line) and fixation location centers (circles) during one free-viewing episode. A cumulative fixation time histogram (Figure 1B) confirms previous

reports (Mazer & Gallant, 2003; Gallant, Connor, & Van Essen, 1998) that eye movements typically consist of 100–400 msec of fixations separated by 20–50 msec of saccadic transitions (only fixations within image borders were considered). For each fixation, we calculated the orientation at each pixel location for all pixels within a  $1^\circ$  square array at the center of gaze; from pixel orientation magnitude histograms, we determined the local orientation and strength of orientation “tuning” for each



**Figure 1.** Stimulus statistics at the center of gaze during free viewing of natural scenes. (A) Example of an eye scan path (white line) and fixation location centers (circles) during one free-viewing episode. (B) Cumulative fixation time histogram from the three monkeys. (C) Examples of orientation magnitude histograms for selected  $1^\circ \times 1^\circ$  fixation patches. We determined the orientation of each pixel based on the direction of the local gray scale gradient in a standard 25-pixel array from the arc tangent of the partial derivative of brightness in a  $5 \times 5$  kernel in the vertical direction, divided by this value in the horizontal direction. For each fixation, we calculated the orientation of each pixel at the center of fixation; from pixel orientation magnitude histograms, we determined the orientation and orientation magnitude for each fixation patch. (D) Relative frequency of successive and control fixations. We considered all fixations, irrespective of the OSI of the image patch being fixated. Successive fixations (black bars) represent fixation vectors recorded during free viewing. Control fixations (gray bars) represent the average of shuffled and random fixations. Shuffled fixations were obtained by randomizing the order in which fixations were generated while maintaining the mean saccade length (the mean amplitude of shuffled fixations was within  $0.3^\circ$  of the mean amplitude before shuffling; only sequences containing at least 4 consecutive fixations were shuffled). Shuffled fixation numbers were corrected by the ratio of the total unshuffled fixations to shuffled fixations. Random fixations were generated by choosing random image locations within image borders such that the mean random saccade length in each session matched the mean saccade length of successive fixations throughout the session. Each bin was normalized to the maximum value of successive fixations or control histograms. Trend lines are shown as second-order polynomial fits for each histogram. Error bars represent *SEM* calculated across sessions.

fixation patch (Figure 1C). In principle, depending on viewing distance, a 1° square patch could contain complex objects, such as faces, that are composed of signals tuned to different orientations. We prevented this situation as much as possible by selecting only those images that contained complex familiar stimuli in a detailed, rather than perspective, view.

### Local Properties of Visual Input during Free Viewing

We investigated how image orientation at the center of gaze changes during image scanning by calculating the difference between the orientation of successive fixation patches for each scanning session. If images were sampled uniformly, then the full range of orientations composing a scene would ensure that the distribution of orientation differences between successive fixation patches would be isotropic. However, Figure 1D indicates significant anisotropy in the distribution of eye movements during free viewing (i.e., images are sampled such that consecutive fixations are more likely to occur between patches of similar or largely dissimilar orientation structures;  $p < .01$ , Wilcoxon sign test, comparing the frequency of successive fixations for orientation differences  $\Delta\theta < 15^\circ$  and  $\Delta\theta > 75^\circ$ , with the frequency of successive fixations for  $\Delta\theta$  between  $30^\circ$  and  $60^\circ$ ). Indeed, Figure 1D shows that consecutive fixations to image patches of similar ( $\Delta\theta < 15^\circ$ ) and largely dissimilar ( $\Delta\theta > 75^\circ$ ) orientations occur about 25% more often than consecutive fixations to patches of intermediate orientation difference ( $\Delta\theta$  between  $30^\circ$  and  $60^\circ$ ).

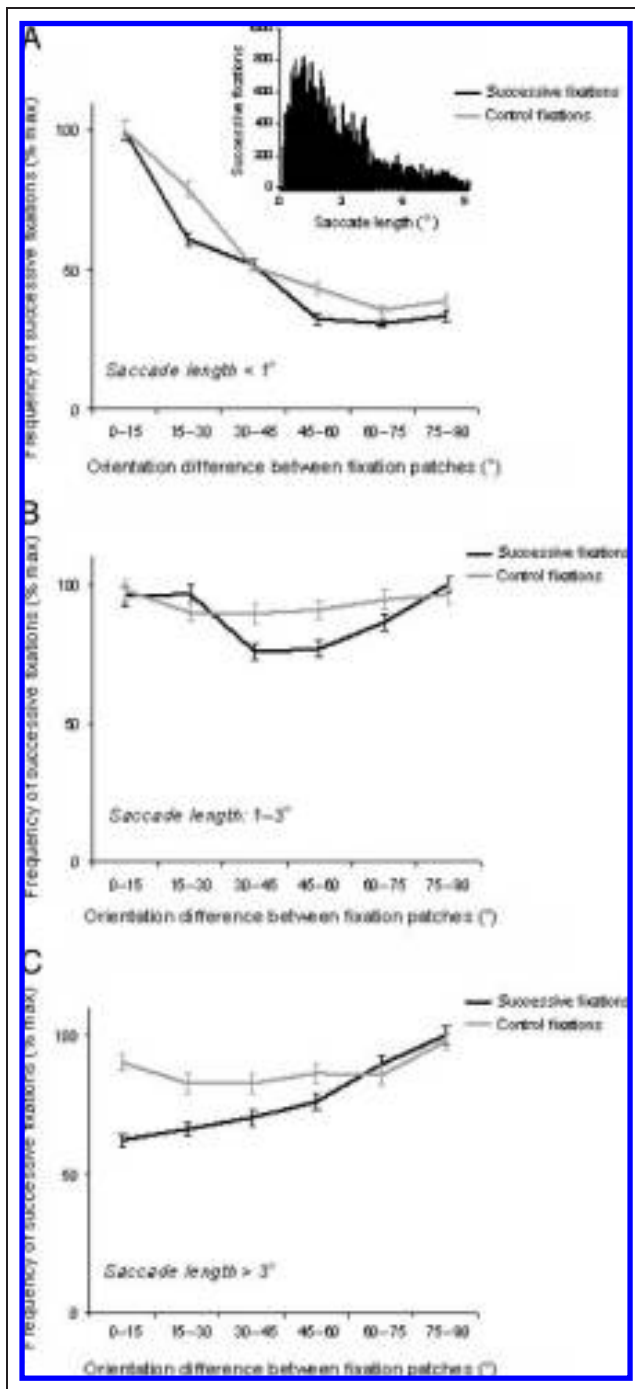
The correlational structure of natural scenes (Tolhurst et al., 1992; Attneave, 1954) could, in principle, explain the anisotropy reported in Figure 1D. Indeed, because neighboring image patches tend to be highly correlated in spatial structure (e.g., orientation) whereas distant patches tend to be only weakly correlated (Dragoi, Sharma, Miller, et al., 2002), it is conceivable that more frequent saccades to nearby patches or to distant image patches might account for the U-shaped curve shown in Figure 1D. To ensure that the orientation distribution of fixations is not influenced by the particular orientation structure of the images in our set or by the statistics of saccade distances during free viewing, we generated control fixations by: (1) shuffling the recorded saccade vector such that the order in which saccades were generated was randomized (the mean saccade length was fixed), and (2) choosing random image locations such that the mean random saccade length in each session matched the mean saccade length of successive fixations throughout the session. Figure 1D shows that the orientation distribution of control saccades exhibits only a weak bias for small and large orientation differences—a bias caused by known overrepresentation of cardinal orientations in natural scenes (Dragoi, Turcu, et al., 2001; Coppola, White, Fitzpatrick, & Purves, 1998; be-

cause horizontal and vertical orientations outnumber obliques, random saccades would have a higher probability of connecting cardinal targets). However, importantly, successive and control saccades have different orientation statistics. That is, there is a larger relative difference between successive fixations to targets of small and large orientation differences and successive fixations to targets of intermediate orientation difference ( $p < .01$ , Wilcoxon sign test) than would be expected from a random sampling of the same scenes. We performed additional controls by shuffling the images in our set by assigning each fixation vector to a random image. In this case, we found that the orientation distribution of nonshuffled fixations was significantly different from that of shuffled fixations. (Additional control experiments are described in Supplementary Figure 1.)

One variable that could influence the results shown in Figure 1D is saccade length (Figure 2A, inset). Indeed, the correlational structure of natural scenes ensures that short saccades are more likely to reach nearby image patches of similar orientation, whereas long saccades are more likely to reach more distant image patches of dissimilar orientation (Dragoi, Sharma, Miller, et al., 2002). To determine whether and how saccade amplitude influences the orientation statistics of successive fixations, we analyzed our eye movement data as a function of the distance between successive fixation patches (Figure 2) and we normalized the results to determine whether the profile of successive fixation histogram differs from that of control (shuffled and random) fixation histograms. Confirming our expectation, we found that short saccades ( $< 1^\circ$ ) were almost indistinguishable in terms of orientation statistics from shuffled and random saccades of similar length (Figure 2A;  $p > .1$ , Wilcoxon sign test). In contrast, as saccades increase in size, a higher fraction of eye movements tends to connect image patches of different orientation structures relative to patches of similar orientation structures. Importantly, saccades larger than  $3^\circ$  tend to preferentially reach image regions that are largely dissimilar, whereas image regions of similar orientation tend to be avoided. This bias was almost abolished when saccades were shuffled or randomized. Indeed, control saccades do not show any significant bias for small or large orientation differences relative to intermediate differences ( $p > .1$ , Wilcoxon sign test).

### Low-level and High-level Factors in the Selection of Visual Targets

The orientation asymmetry of successive fixations during image scanning may reflect the interaction of low-level (stimulus-dependent) and high-level (cognitive) variables in the selection of visual targets (Delorme, Rousselet, Mace, & Fabre-Thorpe, 2004; Mazer & Gallant, 2003; Koch & Ullman, 1985). Although the exact mechanism involved in selecting particular spatial locations



**Figure 2.** (A–C) Orientation structure of successive fixations for different saccade amplitudes. Inset: cumulative saccade amplitude histogram from the three monkeys. The histogram represents only saccades between 0° and 9°. Black lines: successive fixations; gray lines: control fixations (average of shuffled and random fixations). Control fixations are defined as in Figure 1D. Error bars represent SEM calculated across sessions.

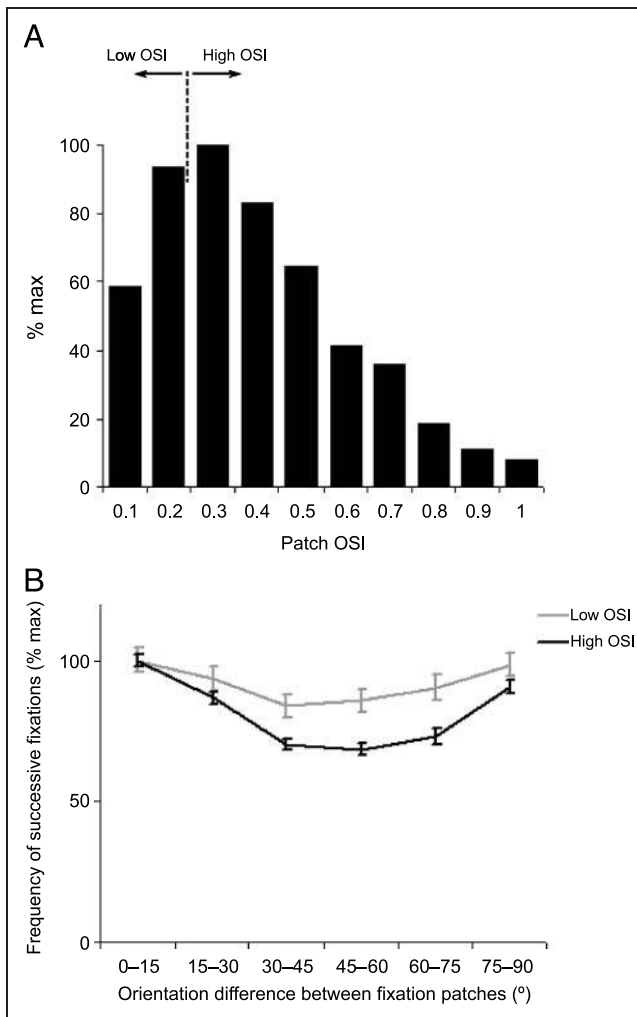
during image scanning remains unknown, it has been hypothesized that both bottom-up and top-down processes are involved in the rapid visual analysis of natural scenes. Hence, we proceeded to a more in-depth anal-

ysis of how low-level (such as local orientation) and high-level factors (such as viewing time, image familiarity, and image cognitive content) contribute to the orientation specificity of consecutive fixations during free viewing. These control experiments can potentially clarify the source of the orientation anisotropy of successive fixations during free viewing.

One important issue is whether and how the particular orientation structure at the center of gaze is able to influence the orientation anisotropy of successive fixations. Because visual targets of high orientation magnitude may be easier to discriminate, it is, in principle, possible that selecting those fixation patches of higher orientation magnitude could yield a stronger orientation anisotropy. To address this issue, we measured the strength of the orientation signals at the center of gaze from the orientation magnitude histogram of each fixation patch by calculating the OSI, which is a measure that indicates how well the patch is “tuned” to the mean orientation within the 1° × 1° foveal area. Figure 1C, for instance, shows that Patch 1 is more broadly tuned to its dominant orientation than is Patch 2 (i.e., the orientation magnitude histogram of Patch 1 is wider than that of Patch 2). The OSI histogram in Figure 3A indicates that patch OSI is normally distributed with a peak around 0.3. We thus divided fixation patches into two groups depending on whether patch OSI values were smaller (low OSI) or greater (high OSI) than the mode, and then we calculated the frequency of successive fixations for visual targets of low and high orientation tuning strength. Figure 3B shows that when visual targets are more discriminable (higher OSI values), subjects tend to fixate more frequently on similar and largely dissimilar successive patches relative to fixations on successive patches of lower orientation selectivity ( $p < .05$ , Wilcoxon sign test). This implies that the degree of the orientation anisotropy of successive fixations is closely linked to “low-level” factors, such as the strength of orientation signals of visual targets.

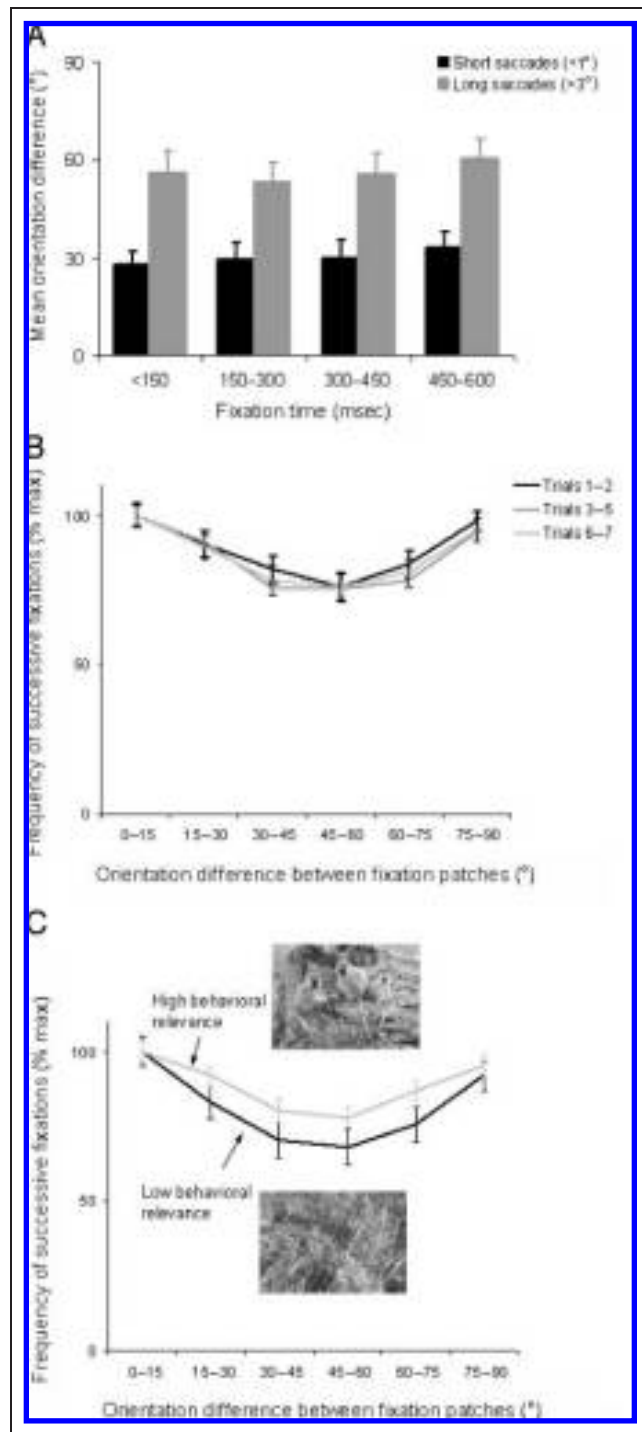
How do possible “high-level” or top-down factors influence target selection? We examined the effect of three variables: fixation period (longer fixations may allow more access to information), “early” vs. late trials (a correlate of stimulus familiarity), and behavioral relevance of images, on the orientation structure of patches at successive fixations.

One variable that might have influenced our results is the duration of the fixation period (i.e., longer fixations could allow more access to visual information in order to minimize the noise in target selection induced by spurious shorter fixations and to possibly enhance the orientation anisotropy reported here). However, Figure 4A shows that the orientation specificity of both short and long saccades is not affected by fixation time ( $p > .05$ , Wilcoxon sign test). The mean orientation difference between successive fixations is roughly constant for both short saccades (orientation difference



**Figure 3.** Target fixation and orientation selectivity. (A) Histogram of fixation patch OSI. Target patches were binned in 0.1 OSI groups and then were represented as percentage patches relative to the maximum patch count (corresponding to the 0.3 OSI bin). (B) Frequency of successive fixations for visual targets of low and high orientation tuning strength (% max). Fixation patches were divided into two groups: low OSI (<0.3, gray line) and high OSI (>0.3, black line). Error bars represent *SEM* calculated across sessions.

around 30°) and long saccades (orientation difference around 57°) and is not correlated with fixation time. Another variable that could have influenced our results is stimulus familiarity. It is, in principle, possible that monkeys may employ different image scanning strategies depending on the degree of familiarity with the image. That is, novel images presented during early trials in each session may be rapidly analyzed to extract orientation nonspecific high-level cues, whereas more familiar images presented late in the session may be analyzed to extract orientation-specific low-level features. Hence, the orientation anisotropy of successive fixations may be stronger for familiar images presented within the last few trials in the session. Our analysis, however, indicates an opposite conclusion (Figure 4B): The degree



**Figure 4.** Top-down factors involved in target selection. (A) Fixation duration – Mean orientation difference between successive fixations connected by short (<1°, black bars) and long (>3°, gray bars) saccades as a function of fixation time. Orientation differences are binned in 150-msec fixation time intervals. We did not consider fixations beyond 600 msec. (B) Image familiarity – Orientation histogram of successive fixations for early trials (Trials 1–2), middle trials (Trials 3–5), and late trials (Trials 6–7). (C) Behavioral relevance – Orientation histogram of successive fixations for scenes of high behavioral relevance (gray lines) and scenes of low behavioral relevance (black lines). Both distributions are significantly different from random saccade distributions. Error bars represent *SEM* calculated across sessions.

of orientation anisotropy is not significantly different ( $p > .05$ , Wilcoxon sign test) for successive fixations recorded during Trials 1–2, 3–5, and 6–7 (we have obtained similar results by dividing the 7 trials into two categories; for instance, Trials 1–3 and 4–7). Because absolute fixation locations change across trials, this argues that images can be scanned in a variety of ways, yet local principles of scanning (i.e., the orientation statistics of successive fixations) are preserved.

One concern about using natural images in psychophysical studies is that real-world scenes vary widely with respect to their cognitive content or behavioral relevance. It is, therefore, possible that our subjects may employ different scanning strategies depending on the informational content of the scene. To test whether the orientation statistics of successive fixations are influenced by the cognitive content of the scene being inspected, we divided the 200 images in our set into two groups—“high behavioral relevance” (e.g., animals in their natural habitat or foods) and “low behavioral relevance” (e.g., landscape or outdoors pictures)—and rejected 43 images that did not fit in either group (we found that, on average, monkeys scanned highly relevant images approximately 15% more frequently than images of lower behavioral relevance). Figure 4C shows that the orientation inhomogeneity of successive fixations tends to be accentuated when images of low behavioral content are used. The orientation statistics of successive fixations depend on the informational content of the scene in a significant manner only for successive fixations within the  $45\text{--}75^\circ \Delta\theta$  range ( $p < .05$ , Wilcoxon sign test), whereas orientation anisotropy is not significant for small or large  $\Delta\theta$  values. This suggests a stronger role of top-down influences on eye movements when viewing behaviorally relevant scenes, despite a highly significant orientation anisotropy of successive fixations. This demonstrates the limitations of our proposal, that is, the fact that saccades are not generated solely based on the orientation structure of a stimulus, but based on the interplay between bottom-up (e.g., image orientation and contrast) and top-down variables (such as behavioral relevance). However, it is remarkable that images of highly heterogeneous orientation and luminance distribution and random behavioral relevance follow the same scanning pattern.

### **Could Visual Signals at Peripheral Locations Contribute to Target Selection?**

Our results indicate that the visual system uses fovea to examine details about images at small spatial scales, and then larger eye movements move fovea to more distant image regions of significantly different structure. This raises the question of how the visual system could guide saccades to image regions of largely different orientation structure despite the fact that these regions are located

away from the fovea (Figure 2C), where visual sensitivity is diminished. Could the drop in the sensitivity of the visual system away from the point of fixation alter the pooling of orientation signals from peripheral locations? To answer this question, we used a Gaussian blur filter to model the reduction in spatial resolution outside the fovea (Peli, 1996). We found (Supplementary Figure 2) that, although the local orientation of fixation image patches changed slightly after blurring, the magnitude of orientation signals actually increased. This suggests that, despite the reduction in spatial resolution outside fovea, the visual system has access to reliable information about the orientation of peripheral image regions that may be used for saccade planning.

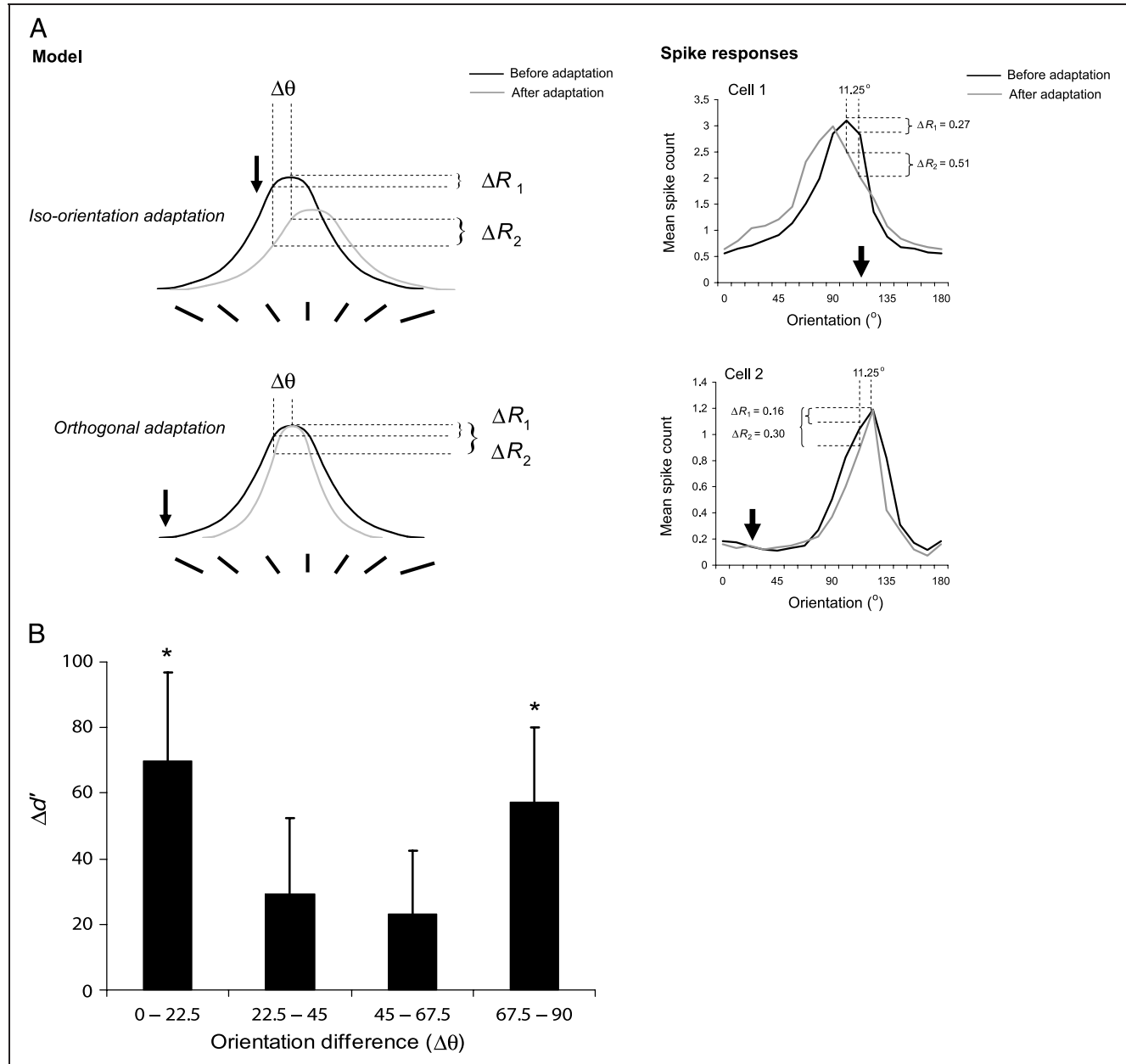
### **Implications for the Neural Processing of Visual Features**

Does the cortical processing of visual features take advantage of orientation anisotropy in the selection of visual targets during eye movements? We propose a relationship between the anisotropy of stimulus statistics at the center of gaze and specific adaptive schemes that visual cortical neurons use to enhance their capacity to identify local features during viewing. Specifically, it is known that following exposure to spatially correlated image patches, cortical neurons can adapt at scales similar to those of visual fixation (Dragoi & Sur, 2003). Brief adaptation to spatially correlated image patches is known to induce short-term changes in the response of V1 neurons. In the orientation domain, there are two important aspects of adaptation: iso-orientation adaptation (Dragoi, Sharma, Miller, et al., 2002; Dragoi, Sharma, & Sur, 2000; Muller, Metha, Krauskopf, & Lennie, 1999) (exposure to a certain image pattern affects subsequent viewing and discrimination of similar patterns) and orthogonal adaptation (Dragoi, Sharma, Miller, et al., 2002; Felsen et al., 2002) (exposure to a certain image pattern affects subsequent viewing and discrimination of largely dissimilar patterns).

We propose that brief adaptation in V1 neurons is implicit during visual fixation. The overrepresentation of successive fixations to images of similar and largely dissimilar orientation structures (Figure 1) is thus equivalent to a preponderance of iso-orientation and orthogonal adaptation episodes during natural viewing (relative to intermediate orientation adaptation episodes). From a functional standpoint, brief adaptation has been shown to affect the signaling capabilities of V1 neurons: Brief adaptation to an oriented grating induces a repulsive shift in the preferred orientation of individual V1 neurons and broadens orientation selectivity when an adapting grating is similar to cells' optimal orientation, whereas it maintains orientation preference and sharpens orientation tuning when an adapting grating is largely dissimilar to cells' optimal orientation (Dragoi, Sharma, Miller, et al., 2002; Felsen et al., 2002).

What could be the functional significance of rapid adaptation? An ecological view of adaptation might suggest that rapid adaptation improves orientation discrimination of V1 neurons, and we propose that such improvements are possible both in iso-orientation adap-

tation and orthogonal adaptation. In contrast, we predict that V1 neurons would exhibit only a mild improvement in orientation discrimination performance after adaptation at intermediate orientations (this type of adaptation is less frequent during natural viewing;



**Figure 5.** Effects of orientation adaptation on discrimination performance by V1 neurons. (A) Model: iso-orientation adaptation. Brief adaptation to an oriented grating (arrow) increases the slope of the tuning curve in the vicinity of preadaptation preferred orientation to amplify the neuronal response difference ( $\Delta R_2 > \Delta R_1$ ) elicited by nearby stimuli of orientation difference  $\Delta\theta$ . Orthogonal adaptation. Brief adaptation to an oriented grating (arrow) orthogonal to cells' preferred orientation increases the slope of the tuning curve in the vicinity of preadaptation preferred orientation to amplify the neuronal response difference ( $\Delta R_2 > \Delta R_1$ ) elicited by nearby stimuli of orientation difference  $\Delta\theta$ . Spike responses: changes in orientation discrimination performance of V1 neurons of behaving monkeys after brief adaptation. We show orientation tuning curves of two representative V1 neurons before (black line) and after 400 msec of adaptation (gray line). For each cell, adaptation improves orientation discrimination near the peak of the tuning curve ( $\Delta R_2 > \Delta R_1$ ). (B) Change in the discriminability ( $d'$ ) of two gratings differing in orientation by 11.25°, after adaptation to a range of orientations ( $\Delta\theta$ ). Each bar represents the mean change in discriminability for neurons pooled in different  $\Delta\theta$  intervals shown on the x-axis. Asterisks denote statistical significance ( $p < .05$ ) from preadaptation values. Error bars represent SEM.



Figure 1D). How does rapid adaptation influence the dynamic coding of visual stimuli? In the case of iso-orientation adaptation during the time course of fixation, there is a suppression of responses on the flank of the V1 tuning curve toward the adapting stimulus and a facilitation of responses on the opposite flank, which together constitute a repulsive shift in preferred orientation (Figure 5A, left). We suggest that brief adaptation (the adapting orientation is marked by the arrow) improves the discrimination of orientations close to neurons' optimal orientation ( $\Delta\theta$ ) by increasing the slope of the tuning curve in the vicinity of preadaptation preferred orientation (Simoncelli & Olshausen, 2001; Muller et al., 1999). This leads to an amplification of the neuronal response difference ( $\Delta R_2 > \Delta R_1$ ) elicited by nearby stimuli of orientation difference  $\Delta\theta$ . This improvement in discrimination performance after iso-orientation adaptation has psychophysical support (Clifford, Wyatt, Arnold, Smith, & Wenderoth, 2001; Regan & Beverley, 1985) and could be a mechanism that might be used by the visual system to resolve fine orientation details during visual fixation. For large saccades, subsequent fixation is likely to be made on image patches of largely dissimilar orientation (Figure 2C). In this case, orthogonal adaptation has been shown to improve neuronal performance in V1 by sharpening orientation tuning curves. This sharpening leads to an increase in the slope of the tuning curve in the vicinity of preadaptation preferred orientation (Figure 5A, left), thus an amplification of the response difference ( $R_2 > R_1$ ) elicited by nearby stimuli of orientation difference  $\Delta\theta$ . There is a corresponding improvement of psychophysical performance on orientation discrimination (Dragoi, Sharma, Miller, et al., 2002; Clifford et al., 2001), suggesting that orthogonal adaptation could be a mechanism that might be used by the visual system in order to rapidly resolve fine orientation details after a large saccade.

We have examined these hypotheses by electrophysiologically recording the responses of neurons in V1 of behaving monkeys using the reverse correlation method (Dragoi, Sharma, Miller, et al., 2002) and by focusing on changes in orientation discriminability of responses before and after brief adaptations (400 msec duration). Figure 5A (right) shows that orientation tuning curves of two representative V1 neurons are altered by adaptation. First, brief exposure to a grating oriented close to cells' preferred orientation (Cell 1, Figure 5, right) induces a repulsive shift in preferred orientation and an increase in the slope of the tuning curve in the vicinity of preadaptation preferred orientation. A measure of neurons' capacity to discriminate orientations is the response difference ( $\Delta R$ ) elicited by nearby stimuli presented close to the peak of the tuning curve. The increase in the slope of the neuronal tuning curve after adaptation leads to an amplification of the response difference  $\Delta R_1$  elicited by nearby stimuli of an orientation difference

of  $11.25^\circ$  (this is the orientation resolution of our V1 recordings). Thus, the response difference after adaptation  $\Delta R_2$  is almost twice the response difference before adaptation. By this measure, Cell 1 shows an improvement in orientation discrimination performance after iso-orientation adaptation. Second, brief exposure to a grating oriented orthogonally with respect to cells' preferred orientation (Cell 2, Figure 5A) sharpens orientation tuning and increases the slope of the tuning curve in the vicinity of preadaptation preferred orientation. Hence, the response difference  $\Delta R_1$  elicited by nearby stimuli of an orientation difference of  $11.25^\circ$  relative to the peak of the tuning curve is amplified ( $\Delta R_2 > \Delta R_1$ ) to improve orientation discrimination performance.

However, amplification in the slope of the tuning curve after adaptation does not ensure an improvement in orientation discrimination performance. Discrimination performance also depends on the variability of responses. For instance, if adaptation increases variability, then orientation discrimination could be impaired despite the increase in the tuning curve slope near the peak. To control for this possibility, we calculated the neurons' ability to discriminate gratings near its preferred orientation by calculating the mean sensitivity  $d'$  (Green & Swets, 1966) to discriminate orientations of  $11.25^\circ$  relative to the peak presented on the flank of the tuning curve on which the adapting stimulus was located. This is a useful measure because it takes into account both the slope of the tuning curve and the response variance. To obtain reliable measures of response variance, we recorded orientation-selective responses in 86 neurons for at least 400 trials in each condition (control and adaptation) using the reverse correlation method.

Figure 5B shows the changes in orientation discrimination performance after adaptation ( $\Delta d'$ ) for a range of differences between the adapting stimulus and cells' preferred orientation ( $\Delta\theta$ ). Importantly, adaptation reliably improves neuronal orientation discrimination performance ( $p < .05$ ) only when the adapting stimulus is similar ( $\Delta\theta < 22.5^\circ$ ) or orthogonal ( $\Delta\theta > 67.5^\circ$ ) with respect to cells' preferred orientation. Adaptation at intermediate orientations ( $\Delta\theta$  between  $22.5^\circ$  and  $67.5^\circ$ ) is known to induce a smaller repulsive shift in preferred orientation (Dragoi, Sharma, Miller, et al., 2002) compared to iso-orientation and orthogonal adaptation. This causes a smaller increase in the slope of the postadaptation tuning curve in the vicinity of preadaptation preferred orientation and thus a reduction in postadaptation enhancement in orientation discrimination performance relative to iso-orientation adaptation. Indeed, Figure 5B shows that the population of neurons shows only a mild improvement in orientation discrimination performance after adaptation at intermediate orientations. Because a large fraction of cells actually exhibits impairment in orientation discrimination after intermediate orientation

adaptation, the overall population effect does not reach the significance threshold ( $p > .05$ ).

## DISCUSSION

The relationship between environmental statistics and the structure of the visual system is a persistent theme in systems neuroscience (Simoncelli & Olshausen, 2001; van Hateren & Ruderman, 1998; Atick & Redlich, 1990; Field, 1987; Marr, 1982; Attneave, 1954). However, despite considerable work that has generated important results at the experimental and computational levels, establishing a firm quantitative link between properties of natural stimuli and efficient information processing by individual neurons and networks has been a difficult task. Our results reveal novel relationships between statistical properties of the environment and processing of visual inputs: (1) eye movements dynamically shape visual input reaching the fovea to induce a significant anisotropy in stimulus statistics at the center of gaze, and (2) information processing at the early stages of cortical processing (e.g., V1) is adapted to eye movement statistics to improve the discrimination of elementary features, such as orientation. We interpret the relationship between stimulus statistics during free viewing and orientation processing in V1 functionally, not mechanistically. That is, the reason why successive fixations tend to land on image regions of similar or largely dissimilar orientations is not because the V1 neurons involved in stimulus processing during fixation exhibit orientation adaptation effects. Indeed, the improvement in orientation discrimination after adaptation occurs only after a saccade has been made and a new visual target has been chosen. Hence, adaptation in V1 is unlikely to contribute to target selection. Instead, we propose that, during evolution or development, orientation-selective neurons in V1 have evolved to take advantage of eye movement statistics. Specifically, V1 neurons have evolved to efficiently encode dynamic stimuli of similar and largely dissimilar orientation structure as these stimuli have the highest probability of occurrence during natural viewing. Indeed, the strongest improvement in stimulus discriminability occurs when individual neurons are adapted to a stimulus of similar or largely dissimilar orientation—the same anisotropy that we found in our eye movement experiments. Thus, visual cortical processing may be well suited for encoding stimuli that vary smoothly between eye movements (such as during local scanning) or vary abruptly after a saccade to a more distant image region.

One issue that could be raised is whether the orientation asymmetry of successive fixations reported here originates from changes in the orientation specificity of the image patches fixated during viewing or is due to changes in other stimulus features, such as local contrast. Indeed, previous studies in humans (Reinagel &

Zador, 1999) have shown that image locations are likely to be foveated if they have a contrast higher than that of the background. We controlled for this possibility and found that, although there is a significant correlation between the contrast and the OSI of fixation patches ( $p < .01$ , Pearson correlation), the local contrast and orientation preference of the foveated patch are unrelated ( $p > .1$ , Wilcoxon sign test). Thus, although locations with higher contrast tend to have higher OSI values, they do not exhibit orientation anisotropy. Furthermore, if contrast would indeed be involved in the orientation anisotropy of consecutive fixations during free viewing, then shuffling fixations, which maintains the same contrast level for the foveated patches, should not change the orientation asymmetry of successive fixations. However, this is contradicted by Figure 1D, which shows that shuffling fixations abolishes the orientation asymmetry of fixations. Altogether, these data indicate that stimulus contrast is unlikely to underlie the orientation anisotropy of successive fixations shown here.

The link between our eye movement and physiology data rests on the assumption that V1 neurons discriminate orientations based on responses near the peak of the tuning curve. However, it has been suggested that V1 neurons could signal small differences in orientation by using responses on the steepest part of the tuning curve (flank responses). Indeed, changes in the slope of the tuning curve have been correlated with long-term improvement in orientation discrimination performance after perceptual learning (Schoups, Vogels, Qian, & Orban, 2001). We argue here that short-term changes in stimulus discriminability, such as changes induced by adaptation, are more likely to be correlated with changes in responses near the peak, not the flank, of the neuronal tuning curve. Specifically, the repulsive shift in orientation preference after adaptation near cells' preferred orientation is equivalent to an increase in the slope of tuning curves near the peak and to a decrease in slope on the flank of tuning curves. If orientation discrimination relies entirely on changes in responses on the flank of the tuning curve (as proposed by Schoups et al., 2001), iso-orientation adaptation will most likely cause a decrease in neuronal discrimination performance (because the slope of the tuning curve flank decreases), which is inconsistent with perceptual improvement in orientation discrimination performance after adaptation. Similarly, orthogonal adaptation decreases the slope of the tuning curve to impair flank orientation discrimination performance, which is also inconsistent with psychophysical evidence (Clifford et al., 2001; Regan & Beverley, 1985). Thus, although we cannot exclude that V1 neurons may use their flank responses to discriminate stimulus orientation after perceptual learning, the fact that adaptation decreases the gain of flank responses to impair neuronal discrimination performance contradicts the available psychophysical data.

It can be argued that, in establishing a relationship between the orientation anisotropy of successive fixations during image scanning and neuronal adaptation during visual fixation, we have largely ignored how the information that is extracted during image scanning is used by the visual system after encoding. That is, because vision is a purposeful process, saccades may primarily target image regions that are behaviorally relevant to the animal. Although these concerns are certainly valid, the free-viewing approach that we have employed offers the advantage that we do not make any assumptions about the specific visual tasks or goals, or the type of information that is being extracted (but see Supplementary Figure 2). However, although focusing on elementary features (such as orientation) is appropriate for constraining models of neuronal function at the early stages of sensory processing (such as retina or V1), more complete hypotheses, such as Bayesian estimation, will need to be used to link the specifics of visual task to neuronal responses in higher cortical areas, such as V4 or the inferotemporal cortex; e.g., see Sheinberg & Logothetis, 2001).

Our results argue that the processing of visual features in V1 takes advantage of orientation anisotropy in the selection of visual targets during eye movements. Brief adaptation at the time scale of fixation affects the signaling capabilities of V1 neurons by improving orientation discrimination performance when the adapting stimulus is either parallel or orthogonal to neurons' preferred orientation. The quantitative link between the orientation statistics of successive fixations and the performance of the visual system is further strengthened by the perceptual consequences of orientation adaptation. Thus, short-term adaptation (a fraction of a second to minutes) to a grating of fixed orientation improves the discrimination of similar and orthogonal orientations and impairs the discrimination of intermediate orientations in both human and nonhuman primates. Our V1 analysis (Figure 5) shows that brief exposure to similar or largely dissimilar orientations increases the dynamic range of single neurons near the peak of their tuning curve to improve orientation discrimination performance, whereas discrimination performance is only mildly improved after adaptation at intermediate orientations.

Why should the visual system be optimized for orientation discrimination? It is conceivable that a more natural and important task would be object and shape processing, hence, the visual system might have evolved to optimize the processing of high-level complex stimuli instead of simple oriented bars. However, this view ignores the fact that natural scene categorization is a remarkably fast process that develops within the time frame of only one fixation (VanRullen & Thorpe, 2001; Thorpe, Fize, & Marlot, 1996). This very high speed relative to time delays of information transmission and processing in different cortical areas indicates that simple feature detectors in early visual areas, such as

orientation-selective neurons, must operate at an astonishing speed to allow accurate image recognition after just one saccade. Our results also raise the possibility that changes in orientation discrimination performance in V1 could lead to changes in object perception in higher cortical areas. This proposal is consistent with recent findings demonstrating that short-term exposure to an individual face generates a significant and precise bias in the subsequent perception of face identity (Leopold, O'Toole, Vetter, & Blanz, 2001). Although these effects have not been investigated at the fixation time scale, they nonetheless raise the issue of whether there is a link between changes in orientation and object discrimination performance between successive fixations.

We conclude that sampling an image is an active process that exploits the image structure by allocating more time to similar and largely dissimilar image regions, possibly optimizing feature and shape extraction. The link between the orientation anisotropy of successive fixations and the adaptive capacity of V1 neurons to encode orientation is functional, not mechanistic. The exact mechanism involved in selecting particular spatial locations during image scanning is unknown, although it is possible that a combination of rapid saliency-based bottom-up mechanisms (Itti & Koch, 2000; Riesenhuber & Poggio, 1999) and slower top-down attentional mechanisms (Nakayama & Mackeben, 1989; Treisman, 1988) could bring selected image regions under cognitive control. Nonetheless, whatever the mechanism of target selection, we suggest that V1 neurons employ an adaptive coding strategy to dynamically improve orientation discrimination performance between successive fixations to similar and largely dissimilar oriented targets.

## Acknowledgments

This work was supported by the McDonnell-Pew and Pew Charitable Trust (V. D.) and the National Institutes of Health (M. S.).

Reprint requests should be sent to Valentin Dragoi, Department of Neurobiology and Anatomy, University of Texas–Houston Medical School, Suite 7.166, 6431 Fannin Street, Houston, TX 77030, or via e-mail: v.dragoi@uth.tmc.edu.

## REFERENCES

- Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, *103*, 62–70.
- Atick, J. J., & Redlich, A. N. (1990). Towards a theory of early visual processing. *Neural Computation*, *2*, 308–320.
- Attneave, F. Informational aspects of visual processing. (1954). *Psychological Review*, *61*, 183–193.
- Barlow, H. B. (1990). A theory about the functional role and synaptic mechanism of after-effects. In C. Blakemore (Ed.), *Vision: Coding and Efficiency* (pp. 363–375). Cambridge: Cambridge University Press.
- Barlow, H. B. (2001). Redundancy reduction revisited. *Network*, *12*, 241–253.
- Betsch, B. Y., Einhauser, W., Kording, K. P., & Konig, P. (2004). The world from a cat's perspective—Statistics of natural videos. *Biol Cybernetics*, *90*, 41–50.

- Clifford, C. W., Wyatt, A. M., Arnold, D. H., Smith, S. T., & Wenderoth, P. (2001). Orthogonal adaptation improves orientation discrimination. *Vision Research*, *41*, 151–159.
- Coppola, D. M., White, L. E., Fitzpatrick, D., & Purves, D. (1998). Unequal representation of cardinal and oblique contours in ferret visual cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *95*, 2621–2623.
- Delorme, A., Rousset, G. A., Mace, M. J., & Fabre-Thorpe, M. (2004). Interaction of top-down and bottom-up processing in the fast visual analysis of natural scenes. *Brain Research: Cognitive Brain Research*, *19*, 103–113.
- Dragoi, V., Sharma, J., Miller, E. K., & Sur, M. (2002). Dynamics of neuronal sensitivity in visual cortex and local feature discrimination. *Nature Neuroscience*, *5*, 883–891.
- Dragoi, V., Sharma, J., & Sur, M. (2000). Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron*, *28*, 287–298.
- Dragoi, V., & Sur, M. (2003). Orientation plasticity in visual cortex and its significance for vision. In L. Chalupa & J. S. Werner (Eds.), *The visual neurosciences*. Cambridge: MIT Press.
- Dragoi, V., Turcu, C. M., & Sur, M. (2001). Stability of cortical responses and the statistics of natural scenes. *Neuron*, *32*, 1181–1192.
- Felsen, G., Shen, Y. S., Yao, H., Spor, G., Li, C., & Dan, Y. (2002). Dynamic modification of cortical orientation tuning mediated by recurrent connections. *Neuron*, *36*, 945–954.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America, A*, *4*, 2379–2394.
- Gallant, J. L., Connor, C. E., & Van Essen, D. C. (1998). Neural activity in areas V1, V2 and V4 during free viewing of natural scenes compared to controlled viewing. *NeuroReport*, *9*, 2153–2158.
- Green, D. M., & Swets, J. A. (1996). *Signal detection theory and psychophysics*. New York: Wiley.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*, 1489–1506.
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, *4*, 219–227.
- Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, *4*, 89–94.
- Locher, P. J., & Nodine, C. F. (1973). Influence of stimulus symmetry on visual scanning patterns. *Perception & Psychophysics*, *13*, 408–412.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman.
- Mazer, J. A., & Gallant, J. L. (2003). Goal-related activity in V4 during free viewing visual search. Evidence for a ventral stream visual salience map. *Neuron*, *40*, 1241–1250.
- Muller, J. R., Metha, A. B., Krauskopf, J., & Lennie, P. (1999). Rapid adaptation in visual cortex to the structure of images. *Science*, *285*, 1405–1408.
- Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research*, *29*, 1631–1647.
- Peli, E. (1996). Test of a model of foveal vision by using simulations. *Journal of the Optical Society of America, A*, *13*, 1131–1138.
- Regan, D., & Beverley, K. I. (1985). Postadaptation orientation discrimination. *Journal of the Optical Society of the American Academy*, *2*, 147–155.
- Reinagel, P., & Zador, A. M. (1999). Natural scene statistics at the centre of gaze. *Network*, *10*, 341–350.
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*, 1019–1025.
- Schoups, A., Vogels, R., Qian, N., & Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature*, *412*, 549–553.
- Sheinberg, D. L., & Logothetis, N. K. (2001). Noticing familiar objects in real world scenes: The role of temporal cortical neurons in natural vision. *Journal of Neuroscience*, *15*, 1340–1350.
- Simoncelli, E. P. (2003). Vision and the statistics of the visual environment. *Current Opinion in Neurobiology*, *13*, 144–1449.
- Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*, 1193–1216.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.
- Tollhurst, D. J., Tadmor, Y., & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmic & Physiological Optics*, *12*, 229–232.
- Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology A*, *40*, 201–237.
- van Hateren, J. H., & Ruderman, D. L. (1998). Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, *265*, 2315–2320.
- VanRullen, R., & Thorpe S. J. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience*, *13*, 454–461.
- Yarbus, A. L. (1967). *Eye movement and vision*. New York: Plenum.