

# Temporal dynamics of motion integration for the initiation of tracking eye movements at ultra-short latencies

GUILLAUME S. MASSON, YVES RYBARCZYK, ERIC CASTET, AND DANIEL R. MESTRE

Centre de Recherche en Neurosciences Cognitives, Centre National de la Recherche Scientifique, Marseille 13402, France

(RECEIVED February 24, 2000; ACCEPTED April 18, 2000)

## Abstract

The perceived direction of a grating moving behind an elongated aperture is biased towards the aperture's long axis. This "barber pole" illusion is a consequence of integrating one-dimensional (1D) or grating and two-dimensional (2D) or terminator motion signals. In humans, we recorded the ocular following responses to this stimulus. Tracking was always initiated at ultra-short latencies ( $\approx 85$  ms) in the direction of grating motion. With elongated apertures, a later component was initiated 15–20 ms later in the direction of the terminator motion signals along the aperture's long axis. Amplitude of the later component was dependent upon the aperture's aspect ratio. Mean tracking direction at the end of the trial (135–175 ms after stimulus onset) was between the directions of the vector sum computed by integrating either terminator motion signals only or both grating and terminator motion signals. Introducing an elongated mask at the center of the "barber pole" did not affect the latency difference between early and later components, indicating that this latency shift was not due to foveal versus peripheral locations of 1D and 2D motion signals. Increasing the size of the foveal mask up to 90% of the stimulus area selectively reduced the strength of the grating motion signals and, consequently, the amplitude of the early component. Conversely, reducing the contrast of, or indenting the aperture's edges, selectively reduced the strength of terminator motion signals and, consequently, the amplitude of the later component. Latencies were never affected by these manipulations. These results tease apart an early component of tracking responses, driven by the grating motion signals and a later component, driven by the line-endings moving at the intersection between grating and aperture's borders. These results support the hypothesis of a parallel processing of 1D and 2D motion signals with different temporal dynamics.

**Keywords:** Motion integration, Ocular following, Tracking eye movements, Second-order motion, Aperture problem, Visual cortex

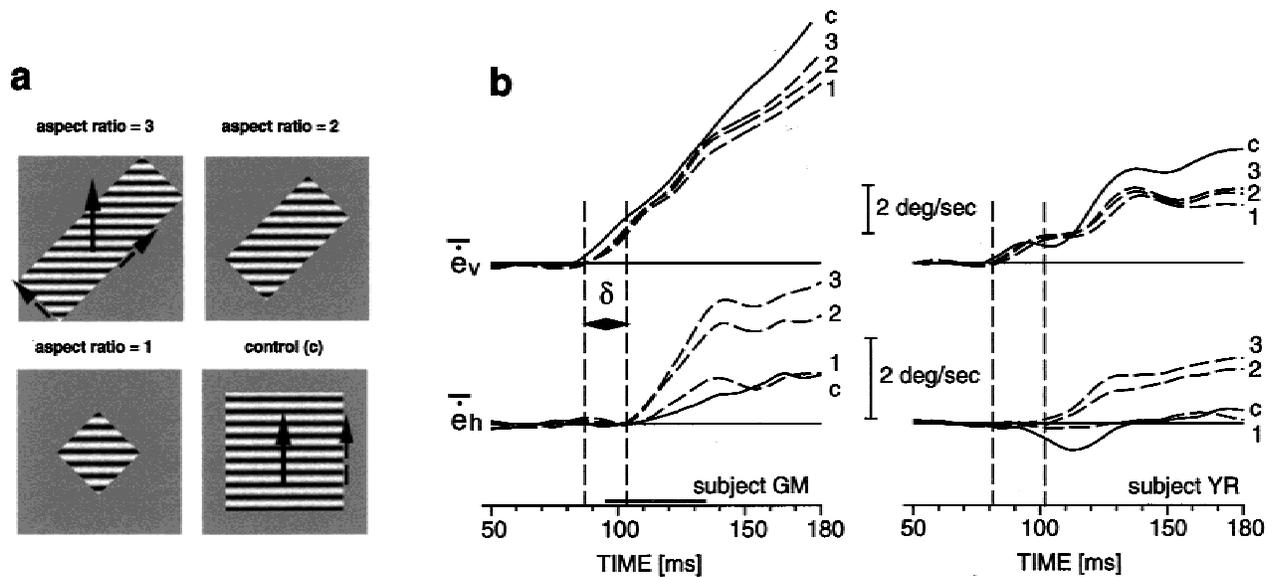
## Introduction

Temporal variations of image intensity provide the only information available from successive retinal images to recover the two-dimensional (2D) vector describing the motion of a visual object in the environment. Single extended contours of the surface are of considerable importance to the visual system. However, because of spatial and temporal limits of any retinal image sampling mechanism, the motion of these one-dimensional (1D) features is inherently ambiguous: there will be always a family of real movements in two dimensions that produce the same motion of this isolated contour (Fennema & Thompson, 1979; Marr & Ullman, 1981). One way to resolve this so-called "aperture problem" is the integration of different 1D motion signals across the visual field (Adelson & Movshon, 1982; Movshon et al., 1985). On the other hand, since the direction of motion of 2D patterns, such as texture elements or moving features (line-endings, corners, dots, . . .) is un-

ambiguous, many psychophysical and computational studies have suggested that these features play a major role in object motion perception (Wallach, 1935; Hildreth, 1984; Nakayama & Silverman, 1988; Lorenceau & Shiffrar, 1992; Castet & Wuerger, 1997).

The "barber pole" illusion is a powerful paradigm for the psychophysical investigation of how the integration of 1D and 2D motion signals give rise to surface motion perception (e.g. Wallach, 1935; Power & Moulden, 1992; Kooi, 1993). Fig. 1a illustrates one instance of the classical "barber pole" stimulus. A horizontal grating is viewed behind a tilted aperture of various aspect ratios. When set into motion in the upward direction, early motion detectors provide two kinds of local motion signals. The component perpendicular to the grating is extracted in regions in which no unique direction can be assigned to the displacement of the 1D luminance profile (continuous arrow). We will call these 1D, or grating motion signals. There are also motion signals generated at the parts of the stimulus where the grating intersects with the aperture border. These motion signals are called 2D, or terminator motion signals because a unique 2D direction can be assigned at these moving bar-endings. Two different motion direction signals are elicited along the two axes of the aperture edges (broken arrows). For aspect ratios higher than 1, subjects report a perceived

Address correspondence and reprint requests to: Guillaume S. Masson, Centre de Recherche en Neurosciences Cognitives, CNRS UPR 9012, 31 Chemin Joseph Aiguier, 13402 Marseille cedex 20, France. E-mail: masson@lnf.cnrs-mrs.fr



**Fig. 1.** Ocular following responses to the “barber pole” stimuli. (a) Snapshot of each type of stimulus: a horizontal grating is viewed behind a tilted aperture of aspect ratio ranging from 1 to 3 (barber-pole stimuli) or behind a large, upright square aperture (control). Motion signals in the direction normal to the grating orientation are always present (continuous arrows). In the three “barber poles” additional terminator motion signals are generated in two different directions (broken arrows). (b) Velocity profiles of horizontal ( $\dot{e}_h$ ) and vertical ( $\dot{e}_v$ ) eye movements in responses to the control (c) or the “barber poles” (1–3). Vertical dotted lines indicate, for each subject, the estimated latencies for each component elicited by a “barber pole” of aspect ratio 3.

direction of surface motion biased towards the longest axis of the aperture (e.g. 45 deg right-upward direction), presumably because more terminator motion signals are generated along the longest aperture edges. This bias is never observed with an upright square aperture (Fig. 1a, bottom-right): all motion signals are in the same direction and then the perceived motion direction is aligned with the axis perpendicular to the grating orientation. It is still unclear to what extent motion signals elicited by moving features can disambiguate grating motion signals (Castet & Wuerger, 1997). For long stimulus duration, “barber pole” stimuli actually yield multistable motion direction perception (Castet et al., 1999) indicating that grating and terminator local motion signals compete to dominate the motion field. Moreover, psychophysical studies have demonstrated that the different inputs to this competitive network may be weighted by several segmentation cues (e.g. Shimojo et al., 1989; Shiffrar et al., 1995; Castet et al., 1999). In understanding how grating and terminator motion signals are integrated, one important step is to tease apart the respective roles of low-level mechanisms, such as spatial interactions between different local motion signals (Power & Moulden, 1992) and higher-level mechanisms, such as occlusion rules or depth perception (Nakayama & Silverman, 1988; Shimojo et al., 1989; Anderson & Sinha, 1997; Castet et al., 1999).

The problem of integrating local motion signals to recover the object motion is instantiated by biological visual systems. In the primate cerebral cortex, the first stage of motion processing occurs in the primary visual cortex. Response selectivity of these early neurons illustrates the aperture problem: motion-sensitive neurons in V1 have a small receptive field tuned to stimuli of a specific spatial frequency, orientation, and/or direction of motion (see Lennie, 1998). As a consequence, V1 neurons respond to the local motion of a 1D pattern in the moving image and cannot individually signal the velocity of the global 2D pattern (Movshon et al., 1985). Integrating responses from multiple V1 neurons is one method

of recovering the actual 2D pattern velocity and indeed there is some evidence that, at a second stage of motion processing, MT neurons integrate multiple 1D motion signals to compute the 2D direction of pattern motion (Movshon et al., 1985; Rodman & Albright, 1989). Most of psychophysical, physiological, and computational studies on the problem of motion integration have dealt with plaid patterns, that is with stimuli from which ambiguous signals of two gratings can be locally combined to determine the direction of rigid motion of the pattern (e.g. Adelson & Movshon, 1982; Movshon et al., 1985; Rodman & Albright, 1989; Simoncelli & Heeger, 1998). Less attention has been paid to the problem of how nonambiguous 2D motion signals can be extracted and combined over space with ambiguous, 1D motion signals to compute a single, accurate estimate of the 2D rigid surface motion. It has been proposed that these features might be detected by specific discontinuity detectors such as end-stopped cells (Hubel & Wiesel, 1968). Alternatively, recent computational studies have pointed out that the same nonlinear process as used as for the so-called “second-order” motion signals might be involved in the detection of moving features and that the output of this non-Fourier mechanism would then be spatially pooled with Fourier mechanisms at the level of MT cells (Löffler & Orbach, 1999). This mechanism is reminiscent of the two pathways combination models (Wilson et al., 1992; Wilson, 1999) for plaid motion perception and is supported by physiological evidence for the existence of a second motion pathway, that can detect higher-order motion signals locally and also lead to MT neurons (Albright, 1992; O’Keefe & Movshon, 1998). A cardinal feature of these models is that the non-Fourier pathway is indirect and slower (Mareschal & Baker, 1998), this latter properties explaining the dynamics of motion integration (Wilson, 1999; Baker, 1999).

In this framework, the temporal dynamics of motion integration need to be clarified. Eye movement studies can tackle this critical problem (Masson & Mestre, 1998). Primates have visual tracking

systems that help vision by stabilizing the eyes on the surroundings, by responding to retinal image motion at ultra-short latencies (see Miles, 1998). These ocular following responses are of special interest here because (1) they exhibit many of the properties generally attributed to low-level motion detectors (Miles et al., 1986; Gellman et al., 1990), (2) they are driven by a velocity error signal built-up by integrating local motion signals over a large portion of the visual field (Miles et al., 1986), and (3) they are mediated by visual stages as early as visual areas MT and MST in monkeys (Kawano et al., 1994). Therefore, by studying the initiation of short-latency tracking responses, we can probe both the properties of early motion processing and the integration of local motion signals. Since tracking eye movement requires continuous visual inflow, we can also probe the temporal dynamics of motion integration by demonstrating early and late changes in the ocular response properties.

## Methods

### Subjects

Experiments were performed on four subjects including one naive subject (IB). All subjects were free of neurological or ophthalmological diseases and had eye examination before participating in the experiments. All subjects had normal or corrected-to-normal acuity. All procedures followed CNRS guide for the use of human subjects.

### Visual stimuli generation and presentation

Visual stimuli were 24 frames movies, computer-generated using the HIPS software (Landy et al., 1984) and the OpenGL libraries on a Silicon Graphics Octane workstation. At the beginning of each session, movies were loaded in memory and were backprojected with appropriate timing onto a translucent large screen (viewing distance, 1 m, subtend,  $80 \times 80$  deg) using a trichromatic videoprojector (Marquee 1800) with a refresh rate of 76 Hz. Single spot targets for triggering saccadic eye movements were also backprojected using two laser LEDs. Drifting gratings were displayed within a rectangular window of varying size, aspect ratio (AR), and orientation. Spatial and temporal frequencies of the grating were kept constant along the direction of motion across conditions (0.3 cpd; 10 Hz) so that speed orthogonal to the grating orientation was constant (33 deg/s). These stimulus parameters are optimal for triggering ocular following responses in humans (Gellman et al., 1990). To manipulate the speed of the grating motion, we kept the spatial frequency constant at 0.3 cpd and varied the temporal frequency. Mean grating luminance was of  $22.25 \text{ cd/m}^2$  and Michelson's contrast was of 92%. Stimulus surround was a gray-level background with same mean luminance as the grating. For the "barber pole" stimuli, the width of the elongated aperture was kept constant at 10 deg, while length was varied from 10 ( $AR = 1$ ) to 30 deg ( $AR = 3$ ). Therefore, size of the grating area ranged from 100 to  $300 \text{ deg}^2$ . Except for the first experiment, where a larger control stimulus was used ( $1600 \text{ deg}^2$ ), the size of the control stimulus (upright square aperture) was similar to that of the "barber pole" stimulus.

### Eye movements recording and stimulus control

The vertical and horizontal positions of the right eye were recorded using the electromagnetic search coil technique (Skalar, Delft, Neth-

erlands), using coils embedded in a Silastin scleral ring (Fuchs & Robinson, 1966; Collewijn et al., 1975). Coils were placed in one eye following application of 1–2 drops of anesthetic (Novesine®) and daily wearing time was limited to about 50 min. The subject was seated in a fiberglass chair with his/her head stabilized by means of chin and forehead rests. Presentation of stimuli and collection, storage, and on-line display of data were controlled by a PC 486/66Mhz running the REX software package (Hays et al., 1982). Voltage signals separately encoding horizontal and vertical positions were low-pass filtered (Bessel, 6 poles, 180 Hz) and sampled at 1 kHz with a resolution of 16 bits. After recording sessions, all data were transferred to a Silicon Graphics workstation for off-line signal processing. After linearization with a fifth-order polynomial function derived from the calibration procedure ran before each session, eye position data were fitted with a  $10e^6$  cubic spline function to reduce the noise and eye velocity signals were computed with a two-point differentiation (Busettini et al., 1991).

The PC and the SGI workstation communicated *via* a serial RS232 interface. Synchronization between the two computers was done using the following protocol. On the UNIX machine, the process reading the serial port and displaying movies was launched at the highest nondegrading priority at the beginning of the experiment. This process was executed by one of the two CPUs while the other handled all IRIX system processes. By doing so, we reduced the maximal response latency to the external trigger down to 3 ms. Once the external trigger signal was received, the motion stimulus was displayed, starting with the next vertical retrace and therefore replacing the stationary random-dots pattern (13 ms interframe). Due to the 76-Hz refresh rate, the maximal stimulus onset latency was of  $13+3$  ms, relative to the trigger signal output from the PC. The same vertical sync signal was used to trigger the motion stimulus and to set the time zero of the PC recording file for that particular trial, *via* the serial port, again with a maximum delay of 3 ms. In summary, there was a 16-ms jitter around the selected postsaccadic delay and a 3-ms jitter around the time zero of stimulus onset. Stability of these delays was carefully checked with electronic devices (oscilloscope, photoelectric transistor, TTL signals from PC AD/DA card) and the latencies of typical ocular following responses obtained on one subject (GM) under these conditions were in close agreement with those obtained previously with opto-electronic stimulus devices (mirror galvanometers, M3, General Scanning, Watertown, MA, latency < 1 ms) (Masson et al., 1995).

### Behavioral paradigm

The behavioral paradigm has been described previously (Miles et al., 1986; Gellman et al., 1990). A trial started with a low spatial-frequency random-dot pattern subtending  $50 \times 50$  deg. The subject was required to fixate a small target spot projected onto the screen 10 deg right of the center. After a randomized interval, this spot was extinguished and a second appeared at the center of the screen. Subjects were required to make a saccadic eye movement to acquire this new target, at which time the target was switched off. On-line control of the eye position checked that the final gaze position was within a  $\pm 1$  deg window around the central target position. Otherwise, the trial was canceled and the 10-deg target was turned on again. With gaze now directed at the center of the screen, moving stimuli were displayed (postsaccadic delay:  $50 \pm 16$  ms) for a brief period of time ( $220 \pm 3$  ms) before the screen was blanked, ending the trial. This procedure served to apply the

motion stimuli in the wake of centering saccades to take advantage of the postsaccadic enhancement (Kawano & Miles, 1986). Varying the postsaccadic delay affects the amplitude of the ocular following responses but not their latency (Gellman et al., 1990). Therefore, the jitter due to the synchronization between the PC and the UNIX Workstation introduced some variability in the amplitude of the responses but had no consequence on the temporal parameters. In the different experiments, all conditions were fully randomized and interleaved with catch-trials where the same static random-dot pattern was left after the saccade. Therefore, subjects were able to predict neither the grating motion direction, nor the aperture's aspect ratio or orientation before completing the centering saccade.

#### Data analysis

In a given experiment, it was usual to collect data until each condition had been repeated more than 150 times, permitting good resolution of the responses to be achieved through averaging. Data were then displayed with an interactive visual software to remove remaining small saccadic eye movements and extract both average velocity profiles, latencies, and amplitude measurements. To illustrate the dynamics of the responses, mean horizontal and vertical right eye velocities were calculated for each condition. We used the convention that rightward and upward visual motion directions were positive. To eliminate any effects due to postsaccadic drift, all data shown have the responses to the static pattern (saccade-only condition) subtracted. All the velocity traces in the figures have been so adjusted and upward deflections of these traces represent either rightward or upward tracking velocities. Subtracting the saccade-only trial might disturb the later parts of the response. This was the case with subject GM in whom a large and long-lasting downward postsaccadic drift was observed. As a consequence, the subtraction introduced an upward component in all adjusted velocity traces. Moreover, despite the precise calibration procedure used, cross-coupling artifacts can contaminate two-dimensional eye movement recordings, due to coil misalignments or scleral coil slippage. To tackle this problem, a complete set of control conditions, where each grating motion direction was seen through an upright square aperture, was always interleaved with other experimental conditions. Corresponding control velocity traces are plotted together with the velocity traces of interest for a particular experiment to allow direct comparison between the control and the "barber pole" conditions.

Quantitative estimates of the amplitude of initial tracking responses were obtained by measuring the change in horizontal and vertical position, over either a 40-ms or a 20-ms time interval, starting 95 ms after stimulus onset. These intervals were timed to coincide with either the early or the later parts of the tracking responses, but before closing of the visual feedback loop. These time windows are illustrated by a gray bar on the related velocity profiles, for each experiment. The mean change in horizontal and vertical position were then calculated together with the standard deviation (SD) and standard errors (SE), for each stimulus condition. Since we are focusing on the earliest, open-loop part of the 2D tracking behavior, the instantaneous 2D tracking direction is continuously changing over time. To compute the average 2D tracking direction, we measured the changes in both horizontal and vertical position over a later time window, from 135 to 175 ms that is when both early and later tracking components have been initiated. Amplitude and direction (relative to the horizontal, rightward direction) of the 2D tracking were computed for each trial.

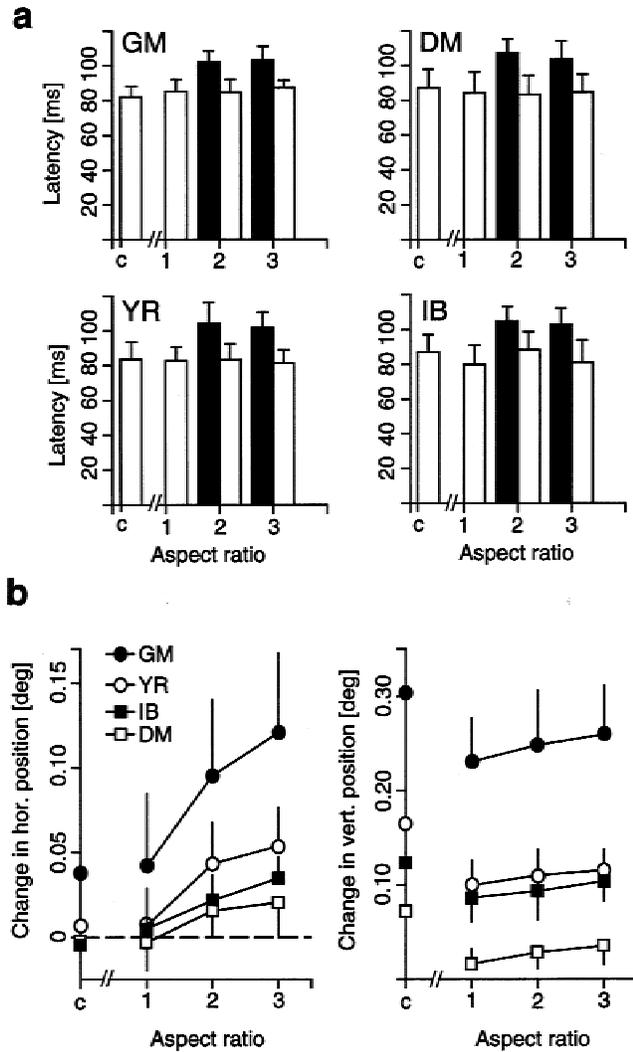
Response latencies were computed using an objective method extensively described in previous publications (Carl & Gellman, 1987). With the aid of the analysis software, the investigator viewed eye velocity signals for each trial. Two intervals were identified. The first interval ("baseline") had a duration of 40 ms and started 20 ms after the stimulus onset. A first regression line was fitted to the eye velocity data, as a function of time, within this interval. The second interval (response) had a duration of 40 ms and began when eye velocity first exceeded 4 SD of the mean measured from the baseline interval. The software computed a second regression line over this response interval and then determined when the two linear functions intersect. This time was defined as the response latency. Vertical and horizontal response latencies were measured independently but analysis was performed on both the mean vertical and horizontal latencies and the latency difference on a trial-by-trial basis.

#### Results

We ran several experiments to investigate, in humans, the properties of the short-latency ocular following responses to a drifting grating viewed through an elongated tilted aperture, the so-called "barber pole" stimulus. We first show that we can distinguish between an early response component in the direction orthogonal to the grating orientation and a later component that deviates the tracking responses towards the direction of the elongated aperture. We demonstrate also the respective contribution of 1D, or grating and 2D, or terminator motion signals to these two components.

#### Ocular following responses to the "barber pole" stimulus

In a first experiment, we interleaved drifting horizontal or vertical gratings viewed through a large, square, upright aperture or through a tilted rectangular aperture of different aspect ratio (*AR*) (Fig. 1a). Grating orientation and aperture size/orientation were fully randomized to avoid anticipatory ocular responses (Kowler & Steinman, 1981). To assess the very early, open-loop part of the tracking responses, without confounding effects of fixation, attention shifts, and so on, subjects were asked to make a 10-deg centering saccade. Stimuli were presented 50 ms after the end of the saccade and set into motion for 200 ms, before blanking. Fig. 1b illustrates the velocity profiles of horizontal ( $\dot{e}_h$ ) and vertical ( $\dot{e}_v$ ) eye movements elicited in two subjects by either the control stimulus (continuous lines) or the "barber pole" stimuli with three different aspect ratios (broken lines). Vertical tracking responses were initiated at the usual ultra-short latencies while horizontal tracking components were initiated only 15–20 ms later. With an aspect ratio of 1, horizontal responses were not significantly different from the genuine residual horizontal drift sometimes observed with a pure upward grating motion. Significant responses in the horizontal direction were, however, observed with aspect ratios of 2 and 3 and we measured the latencies of these stimulus-driven responses. Fig. 2a illustrates the latencies of horizontal and vertical eye movements for four subjects and for each stimulus displayed in Fig. 1a. Ultra-short latency of the vertical responses was nearly constant across conditions and ranged from  $79 \pm 11$  to  $88 \pm 10$  ms (mean across subjects:  $83.8 \pm 2.5$  ms). When significant responses were observed in the horizontal direction ( $AR > 2$ ), latency of horizontal eye movements ranged from  $102 \pm 6$  to  $107 \pm 8$  ms (mean across subjects:  $103.5 \pm 1.6$  ms). Mean latency difference between vertical and horizontal components on individual trials ranged from  $15 \pm 10$  and  $23 \pm 13$  ms. Average latency difference



**Fig. 2.** Latency and amplitude response plots for four subjects. (a) Mean ( $\pm$ SD) latency of vertical (white bars) and horizontal (black bars) eye movements elicited by stimuli displayed in Fig. 1a, plotted as a function of the aperture’s aspect ratio. Bars labeled c are the latencies of vertical responses to the control stimulus. (b) Mean ( $\pm$ SD) change in horizontal and vertical position over the 135–175 ms time window, plotted as a function of the aperture aspect ratio, for each subject. Data labeled c correspond to the control stimulus.

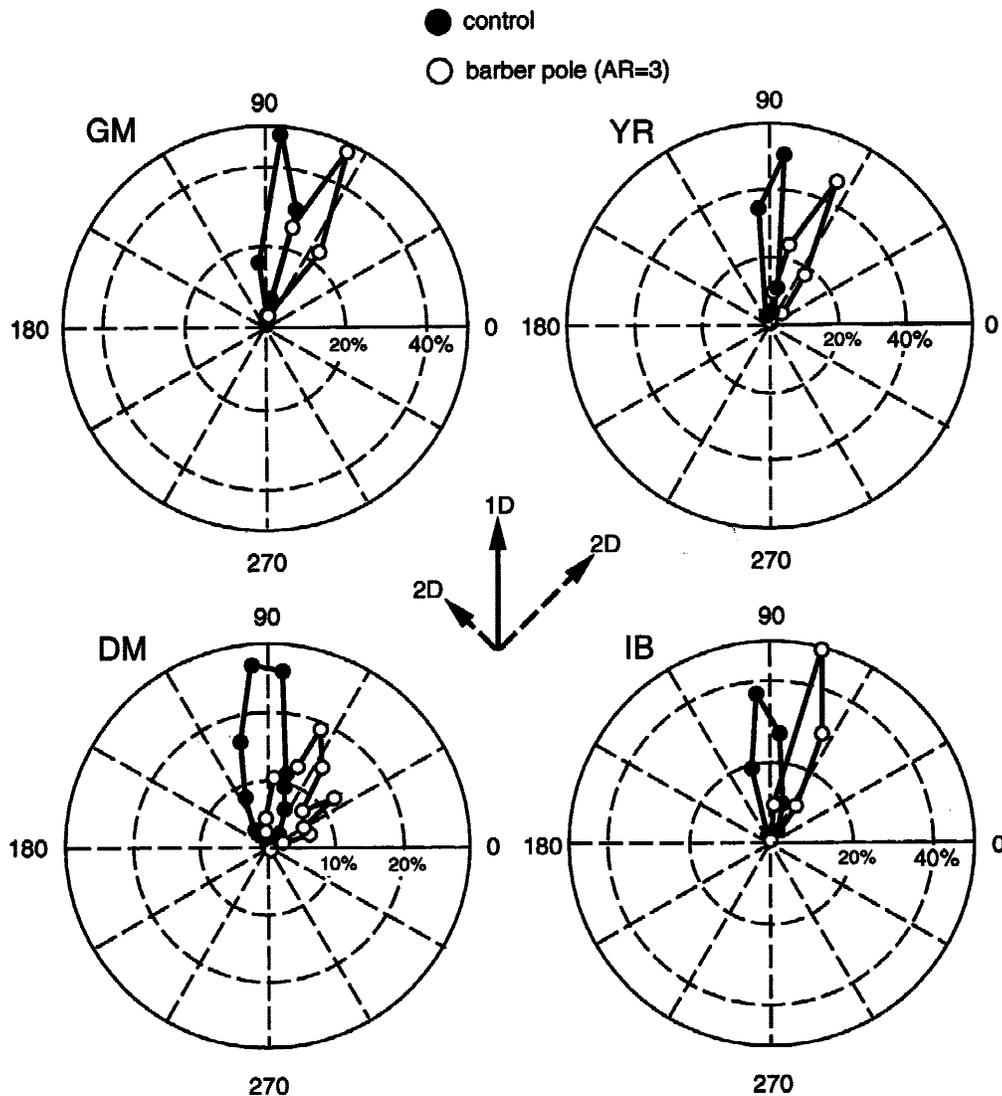
across subjects were of  $18 \pm 5$  ( $AR = 2$ ) and of  $19 \pm 3$  ms ( $AR = 3$ ). Statistical analysis showed that the difference in latency between early and later component was highly significant [ $t(14) = 18.07, P < 0.00001$ ] but that it did not vary between aspect ratios of 2 and 3. On the contrary, Fig. 2b illustrates that amplitude of the horizontal response component was dependent upon the aperture’s aspect ratio: when the aspect ratio increased, the change in horizontal position over a fixed time window increased significantly (ANOVA,  $F(3,9) = 17.92, P < 0.0001$ ). Change in vertical position over the same time window was only marginally modulated. Symbols plotted on the ordinate axis indicate the change in the horizontal and vertical position, respectively, for the control stimulus, labeled c. There was no significant difference in horizontal response amplitude between this latter condition and a “barber pole” of aspect ratio 1, indicating that a tilted square aperture did

not significantly modulate the tracking responses [ $t(6) = 0.12, P > 0.18$ ].

For each individual trial, we computed the 2D direction of the tracking eye movement from the changes in both horizontal and vertical position over a 40-ms time window starting at 135 ms after the stimulus onset. We selected such a later time window to get a robust estimate of the initial tracking direction, which was therefore not dependent on the variability in the latency measurements. We then computed the frequency distribution of the 2D tracking direction across bins of 10 deg width. Fig. 3 plots, for each subject, the frequency distribution of the tracking direction, for both the control (i.e. upright square) and “barber pole” stimulus with an aspect ratio of 3. It is evident that for all three subjects, an elongated aperture resulted in a shifted direction of tracking towards the long axis of the aperture. However, at the end of the trial, the 2D direction of the tracking eye movements was not yet co-linear with the aperture long axis which corresponds to the 45-deg direction in the polar plots.

To further quantify the shift in the initial tracking direction, we fitted the frequency distribution of the tracking direction with a Gaussian function with three parameters ( $k, \mu,$  and  $\sigma$ —being amplitude, mean, and standard deviation of the distribution). We estimated the mean tracking direction from the best-fit  $\mu$  parameter. First, distributions of 2D tracking directions were unimodal indicating that, over the measured time window, tracking behavior was not multistable. Moreover, the estimated mean value of 2D tracking varied with the aperture aspect ratio. Fig. 4 plots both the estimated mean value for each subject and the mean ( $\pm$ SD) across subjects, as a function of the aperture’s aspect ratio. Data labeled c correspond to the control condition. On the same graph, the direction of the different motion signals or vector summations of motion signals are also plotted. Clearly, an upright square and a tilted aperture of  $AR = 1$  resulted in a similar tracking direction (means across subjects:  $89.93 \pm 4.04$  and  $88.21 \pm 6.83$  deg, respectively), that was very close to the direction of the upward grating motion (90 deg). Increasing the aspect ratio deviated, for all subjects, the estimated mean initial tracking direction towards the long axis direction (45 deg). However, with a mean direction of  $65.8 \pm 6.8$  and  $65.6 \pm 5.4$  for aspect ratios of 2 and 3, respectively, at the end of the trial the 2D tracking was not co-linear with the aperture long axis. Mean estimated directions fall between the two dashed lines indicating the direction of the two different vector sums, computed either from the terminator motion vectors only ( $\Sigma_{2D}$ ) or from both terminator and grating motion vectors ( $\Sigma_{2D,1D}$ ). Identical patterns of results were observed with other directions of grating motion, grating, and aperture orientations, as illustrated by Figs. 5–9.

We tested the speed dependency of this new phenomenon by keeping the spatial frequency of the grating constant but varying its temporal frequency from 3 to 15 Hz. Therefore, the speed of the grating motion signals ranged from about 10 to 50 deg/s while the speed of the terminator motion signals along the aperture edges ranged from about 14 to 71 deg/s, because of the 45-deg orientation difference between the two axis. In two subjects, we found that changing the speed of the grating motion had no significant effects of the difference in latency between early and late components. Fig. 5 illustrates horizontal and vertical eye velocity profiles in response to a horizontal grating, drifting upward and viewed through an elongated aperture ( $AR = 3$ ) tilted counterclockwise, so that later ocular following responses deviated towards the left-upward direction. The temporal dynamics of the ocular following responses were not dif-

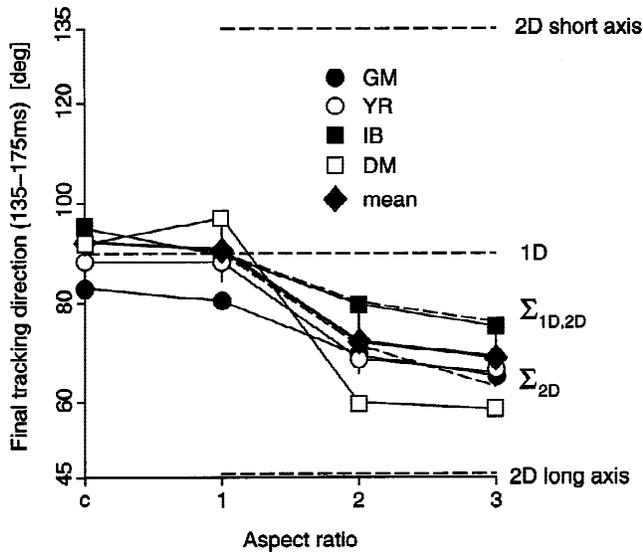


**Fig. 3.** Frequency distribution of tracking directions. For each subject, polar plots represent the frequency distribution (expressed in %) of the two-dimensional tracking direction, computed from each individual trial over a time window from 135 to 175 ms after stimulus onset. Closed symbols illustrate responses to the control grating motion. Open symbols illustrate responses to a “barber pole” of aspect ratio 3.

ferent across speeds. However, changing the grating temporal frequency (speed) affected the amplitude of both components. This effect was larger with the early component than with the later component, probably because the speeds along the aperture edges spanned a higher speed range, exceeding the optimal speed range of ocular following responses. Right-end plots display the initial changes in horizontal and vertical positions, as a function of grating and line-endings speeds, respectively. As a comparison, broken lines indicate the same changes in position elicited by a horizontal grating moving upward through an upright square aperture. Change in vertical position increased monotonically with grating speeds up to 50 deg/s, in both control and “barber pole” conditions. The change in horizontal position illustrates the later component. It also increased with the terminators’ speed, up to 45–50 deg/s. Comparatively, responses to a horizontal grating moving upward showed no horizontal component.

#### *Later component versus response anisotropy*

With our experimental setup, grating motion signals were always along the cardinal axis and terminator motion signals were along the diagonal axis. In three subjects, we verified that the delay between the two components was not generated by some anisotropy in either the motion direction processing or the oculomotor system. We interleaved drifting gratings of four orientations, 45 deg apart, viewed through a circular Gaussian window which removed any aperture effects. Both horizontal and vertical components of 2D tracking responses driven by gratings moving in the diagonal direction had ultra-short latencies of about 80–85 ms. Therefore, all 2D tracking responses elicited by grating motion were of similar ultra-short latencies, indicating that later response components to “barber pole” stimuli were not due to response anisotropy between cardinal and diagonal directions.



**Fig. 4.** Estimated mean two-dimensional tracking direction. For each subject, the frequency distributions of tracking directions were fitted by a Gaussian function and mean initial tracking direction was estimated from the best-fit  $\mu$  parameter, for each condition. Diamond symbols are the mean value ( $\pm$ SD) across subjects. Dashed lines indicate the direction of grating motion signals, terminator motion signals along either short or long axis of the “barber pole,” and the vector sum of either the terminator motion signals alone ( $\Sigma_{2D}$ ) or both grating and terminator motion signals ( $\Sigma_{2D,1D}$ ).

In a second control experiment, we recorded ocular following responses to a high density (50%) random-dots pattern moving behind either an upright or an elongated oblique aperture were also recorded. As expected, responses were always driven in the direction of the random-dots pattern motion and no later components were observed in both conditions, suggesting that the aperture itself had no effects on the initiation of the tracking responses. These two control experiments indicate that the latency difference observed between early and later components cannot be explained by anisotropy in either the motion detection or the oculomotor subsystems.

#### *Integrating motion over central and peripheral visual fields*

At the time of the stimulus motion onset, subjects’ gaze was located at the center of the “barber pole” stimulus. Thus, one might argue that while grating motion signals covered the foveal part of the images, aperture edges and therefore terminator motion signals were located more in the periphery. Anisotropy in latency between the fovea and the peripheral parts of the retina might then explain the observed latency shift. We tested this hypothesis by inserting an elongated mask of the same luminance as the background at the center of the stimulus (Fig. 6a). Now, there were terminator motion signals both inside and outside the fovea. Moreover, the total stimulus area was kept constant between control and test conditions. Fig. 6b illustrates, for one subject, the velocity profiles of horizontal and vertical tracking responses elicited by a vertical grating moving rightward and viewed behind different apertures. Continuous lines indicate responses to the full-field stimulus while broken lines indicate responses to the fovea-masked stimulus. It is evident that adding a mask of the same geometry as the aperture

inside the fovea did not change the responses. The delay between horizontal (i.e. grating-driven) and vertical (i.e. terminator-driven responses) eye movements was again of 15–20 ms and was not changed by introducing line-endings near the fovea. The same results were obtained with two other subjects.

#### *Contribution of 1D motion signals*

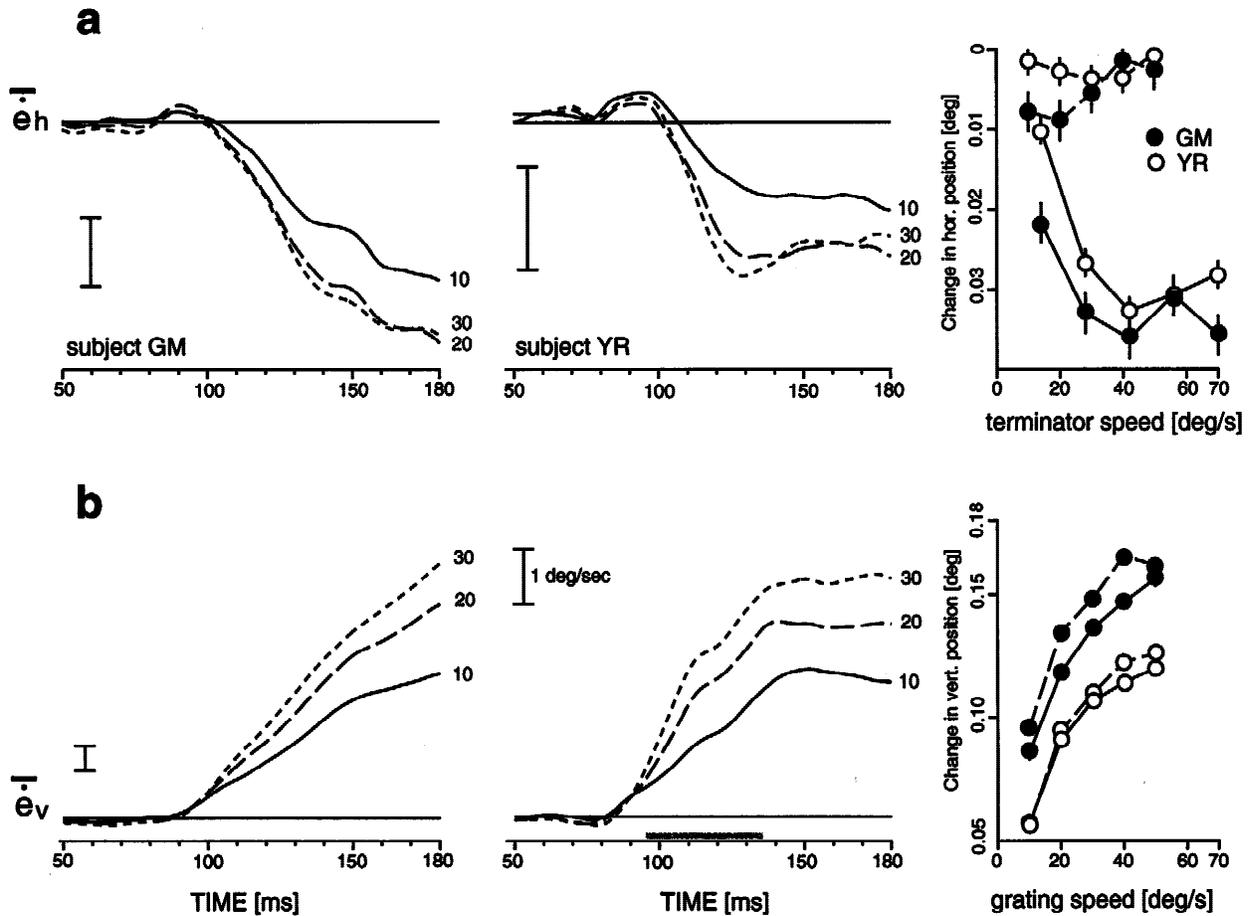
The results presented above suggest that a second motion signal, presumably the terminator motion signals generated at the edges of the aperture, started to cause a change in tracking direction only 20 ms after the initiation of ocular following responses in the direction orthogonal to the grating. Moreover, the late tracking direction seemed to be dependent upon some vector combination of the potential different motion signals present in the “barber pole” stimulus. We further investigated the respective contribution of these different motion signals. First, we concentrated on the role of grating motion vectors located at the center of the stimulus, by removing more and more of these local, ambiguous motion signals.

We varied the size of the foveal mask, now covering from 9 to 90% of the “barber pole” for a reduced set of grating/aperture orientations (Fig. 7a). Hence, we further reduced the weight of the grating motion signals and therefore favored the influence of localized 2D motion signals over the tracking initiation. Figs. 6b and 6d illustrate that latencies of the early and late component remained remarkably constant across all foveal mask sizes. The very small shift in the later component that can be observed in the vertical eye velocity profile were not found with other directions of grating motion and orientations of the aperture. Furthermore, subject YR was the only subject that showed this fairly small trend. On the contrary, Figs. 7b and 7d clearly illustrate that increasing the size of the mask reduced the amplitude of the early component whereas the amplitude of the late component remained largely unchanged. Figs. 7c and 7e plot the change in both horizontal and vertical position over a shorter time window (20 ms), defined to further illustrate the earliest dynamics. It is clear that for the three subjects, increasing the size of the foveal mask from 0 to 90% decreased the change in horizontal position down to about 0 deg, but had no significant effects upon the change in the vertical position. The data point plotted on the right end of the plot indicate the changes in both horizontal and vertical position obtained when the rightward moving grating was seen behind a square, upright aperture of same area as the tilted elongated aperture. It can be seen that the earliest parts of the horizontal change in position are not different between conditions where a grating is viewed behind either an upright square or a tilted aperture without foveal mask. These results indicate that the amplitude of the initial acceleration in the direction of the grating motion signals are controlled by, and only by, the area covered by the moving grating, suggesting a spatial summation mechanism of local 1D motion signals.

#### *Contribution of 2D motion signals*

We have shown that there is a delay between the responses driven by grating and, presumably, terminator motion signals, respectively. We have also demonstrated that magnitude of terminator-driven components was dependent upon the aperture aspect ratio, suggesting that this late component was driven by motion signals generated at the aperture edges. We further investigated the role of terminators motion signals with two experiments.

First, we specifically decreased the contrast of the terminators by windowing a “barber pole” stimulus ( $AR = 3$ ) with an elon-



**Fig. 5.** Speed dependency. Ocular following responses elicited by a counterclockwise “barber pole” ( $AR = 3$ ), where the horizontal grating is moving upward. For two subjects, horizontal (a) and vertical (b) eye velocities are illustrated as a function of time, for three grating speeds (10, 20, and 30 deg/s). Right-end panels plot the mean ( $\pm$ SE) change in horizontal or vertical position, as a function of terminator speed or grating speed, respectively, for both “barber pole” (continuous lines) and control (broken lines) conditions. Error bars are smaller than the symbol size for changes in vertical position.

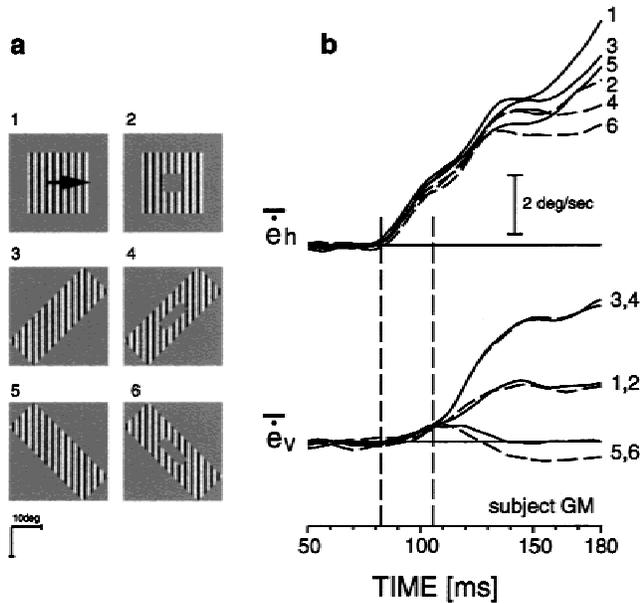
gated 2D Gaussian function of the same aspect ratio. By doing so, we strongly reduced the contrast at the aperture edges. Fig. 8 illustrates for three subjects the vertical and horizontal eye velocity profiles of the tracking responses elicited by each type of stimulus. In all three subjects, horizontal tracking of the rightward grating motion was initiated at the usual ultra-short latencies of about 80 ms. Lower panels show the vertical eye velocity profiles. For all subjects, no significant difference were noticed between symmetrical and elongated Gaussian windows conditions indicating that the latter component was no longer initiated in the direction of the aperture long axis.

In the next experiment, we indented the four aperture edges by adding small, square masks of background luminance (Fig. 9a). These indentations caused the local motion direction at the edges to be in the same direction as the grating motion signals. Fig. 9b displays velocity profiles of horizontal and vertical eye movements elicited by a horizontal grating, drifting upward and viewed behind a counterclockwise tilted aperture ( $AR = 3$ ), for various sizes of indentation. With a nonindented stimulus, end-line signals along the longest axis of aperture drove the horizontal responses in the leftward direction (continuous line). Increasing the size of the indentation from 0 to 1.42 grating period dramatically decreased

the magnitude of leftward responses. Latency of the horizontal component was not affected by the indentation. Amplitude of the vertical responses driven by the grating motion signals perpendicular to the grating orientation was only marginally modulated by indenting the aperture edges. Results are summarized in Fig. 9c for three subjects, including naive subject IB. Response amplitudes over a 40-ms time window starting at 95 ms after the stimulus onset were normalized relative to the no-indentation condition. Normalized changes in horizontal and vertical position are plotted against indentation size, expressed as a fraction of the grating period. Amplitude of terminator-driven horizontal responses decreased with increasing indentation size, down to an asymptote found at about one-half of the grating period. The change in vertical position only slightly increased with indentation size. Identical results were observed with other directions of grating motion and aperture orientation.

## Discussion

In this series of experiments, we measured the time course of tracking eye movements to probe the temporal dynamics of motion integration, a key question that is difficult to investigate directly in



**Fig. 6.** Effects of an elongated mask in the fovea. (a) One frame of “barber poles” with (left column) or without (right column) a central mask of same geometry as the aperture. Masks covered  $27 \text{ deg}^2$  that is 11% of the grating area. Vertical grating are drifted rightward (arrow). (b) Horizontal ( $\dot{e}_h$ ) and vertical ( $\dot{e}_v$ ) velocity profiles of tracking responses. Numbers refer to the type of stimulus showed in the left panel. Vertical broken lines indicate the latency of the horizontal responses elicited by grating motion signals and vertical responses due to terminator motion signals, respectively. Note that this latter is the point at which responses to either “barber pole” or control stimuli diverge.

humans using either psychophysical or physiological methods. Present data highlight two important characteristics of motion integration for tracking eye movements. First, elongated apertures can bias the early phase of tracking eye movements in humans. Change in tracking direction exhibits the same type of dependency upon line-endings as that previously shown for perceived direction in psychophysical studies (Power & Moulden, 1992; Kooi, 1993). This suggests that the “barber pole” illusion is a low-level phenomenon, reflecting early and fast motion integration in the human visual system. Second, there is a 15–20 ms delay between the responses driven by 1D, or grating motion signals and those driven by 2D, or terminator motion signals. We suggest that it can be attributed to the different dynamics of grating and terminator motion signals processing.

#### Early and late components of tracking initiation

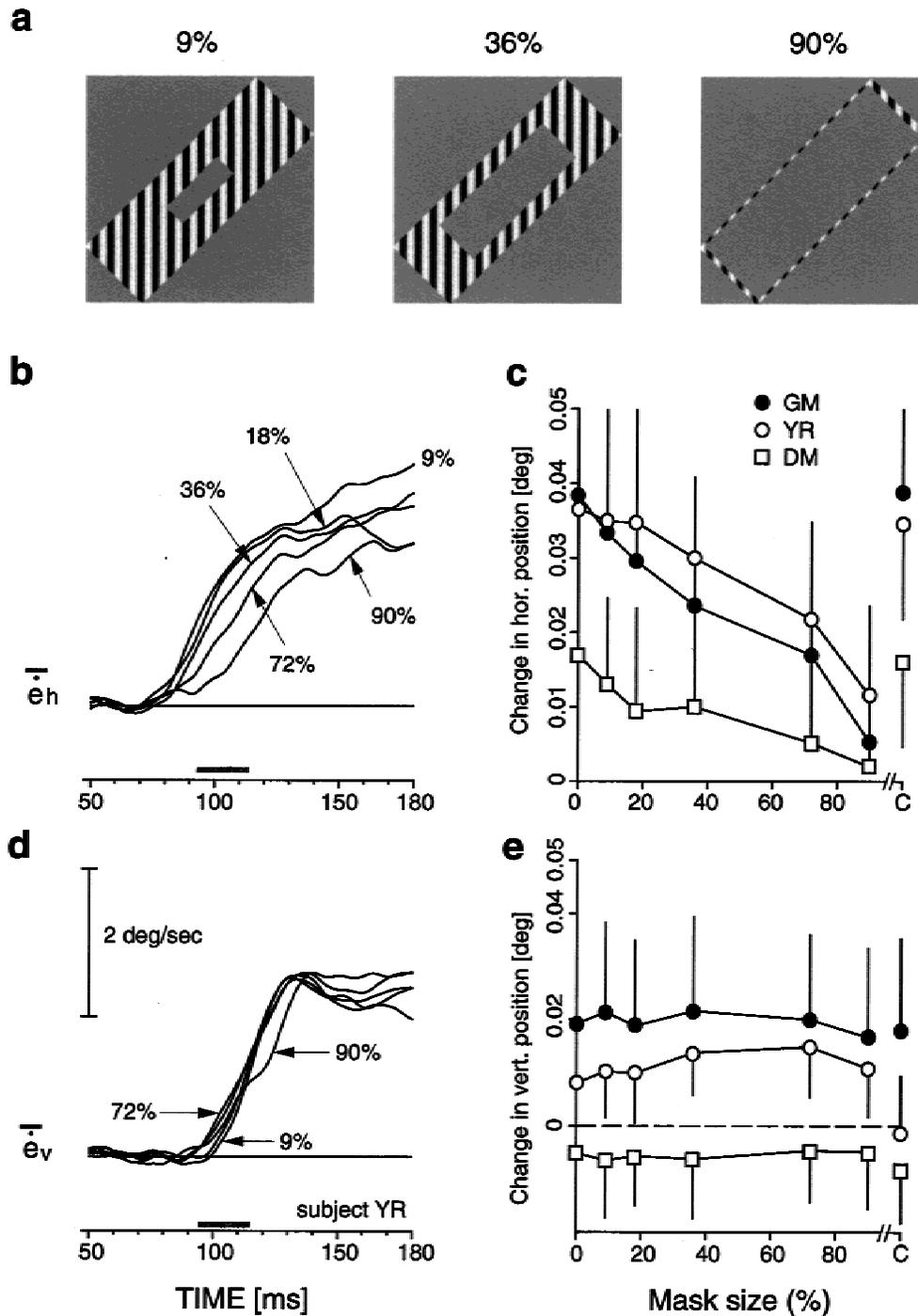
To recover and represent trajectories of objects moving in the real-world, an adequate description of the object retinal image motion is an unambiguous 2D vector that can be used as an error signal to be canceled out by the tracking oculomotor system (Lisberger et al., 1987). A crucial question is how such 2D vector is elaborated by the visual motion system. Psychophysical and physiological studies have suggested that this computation is done by integrating piecewise, 1D motion signals sensed by V1 motion detectors (Movshon et al., 1985). Such visual computation can also integrate 2D motion signals indicated by localized features (e.g. Lorenceau & Shiffrar, 1992; Shiffrar et al., 1995). Psycho-

physical studies of the “barber pole” illusion indeed suggest that local 1D and 2D motion signals compete to drive the motion direction perception (Power & Moulden, 1992; Kooi, 1993; Castet et al., 1999). In that sense, the “barber pole” illusion is a powerful tool to investigate surface motion processing.

Herein, we demonstrated that tracking eye movements are driven at short latency in the direction of the elongated aperture of the “barber pole.” This phenomenon occurs in the early, open-loop phase of tracking eye movements and not only in steady-state, closed-loop pursuit behavior (Beutter & Stone, 1997). Moreover, we uncovered two components in the early phase of tracking initiation. The very first response is always driven in the direction perpendicular to the grating with the usual ultra-short latency ( $\approx 85 \text{ ms}$ ) of ocular following responses in humans (Gellman et al., 1990). A change in the tracking direction towards the direction of the longest aperture edges occurred only 15–20 ms later. This difference was consistent across trials and was found independent upon the relative orientations of gratings and apertures. This delay cannot be attributed to some direction-selective anisotropy in the processes underlying either motion detection or oculomotor control, since 2D tracking can be elicited at ultra-short latency ( $\approx 85 \text{ ms}$ ) by drifting tilted gratings. Moreover, the delay was independent upon the location of grating and terminator motion signals relative to the fovea (Fig. 6). Increasing the size of the foveal mask reduced the amplitude of the earlier component but did not affect the latency difference between the early and later component (Fig. 7). Furthermore, the temporal dynamics of these two components was not changed when indenting the aperture edges (Fig. 8). Finally, the latency difference was found independent upon the speed of the drifting grating (Fig. 5). Therefore, we suggest that this 15–20 ms latency difference between early and later components reveals a pure difference in temporal dynamics of 1D, or grating and 2D, or terminator motion signals processing within the motion stream.

#### Motion integration for 2D tracking

The fact that magnitude, but not latency, of the later component was dependent upon the aperture aspect ratio (Fig. 2b), the contrast of line-endings (Fig. 8) or the direction of local motion signals at the aperture edges (Fig. 9) indicate that it was driven by 2D motion signals arising at the intersects between the grating and the aperture edges. Psychophysical studies have already demonstrated the role of line-endings (or terminators) in generating 2D motion perception (Kooi, 1993; Castet et al., 1999). By increasing the aspect ratio, we increased the number of line-endings along one direction. Since grating motion signals compete with the two existing terminator motion signals for driving the eyes, changing the aspect ratio changes the weight between terminator motion signals at the long and short aperture’s edges and therefore deviates the later components towards the aperture long axis. The role of the terminators is further supported by two results. First, the later component was abolished when the luminance profile of the aperture edges was smoothed out by filtering them with a 2D Gaussian window of the same aspect ratio. Second, it was also reduced by cutting aperture edges to give them a staircase profile. Cutoff was found around 0.25 cycle of grating period. This is in close agreement with psychophysical studies (Power & Moulden, 1992; Kooi, 1993), which found that bias in the perceived direction was minimal with an indentation of about 0.25 cycle of the grating period (Kooi, 1993). This result suggests that similar mechanisms are involved for driving both the perception and the initiation of tracking responses (Beutter & Stone, 1997). By contrast, we found that decreasing the amount of grating

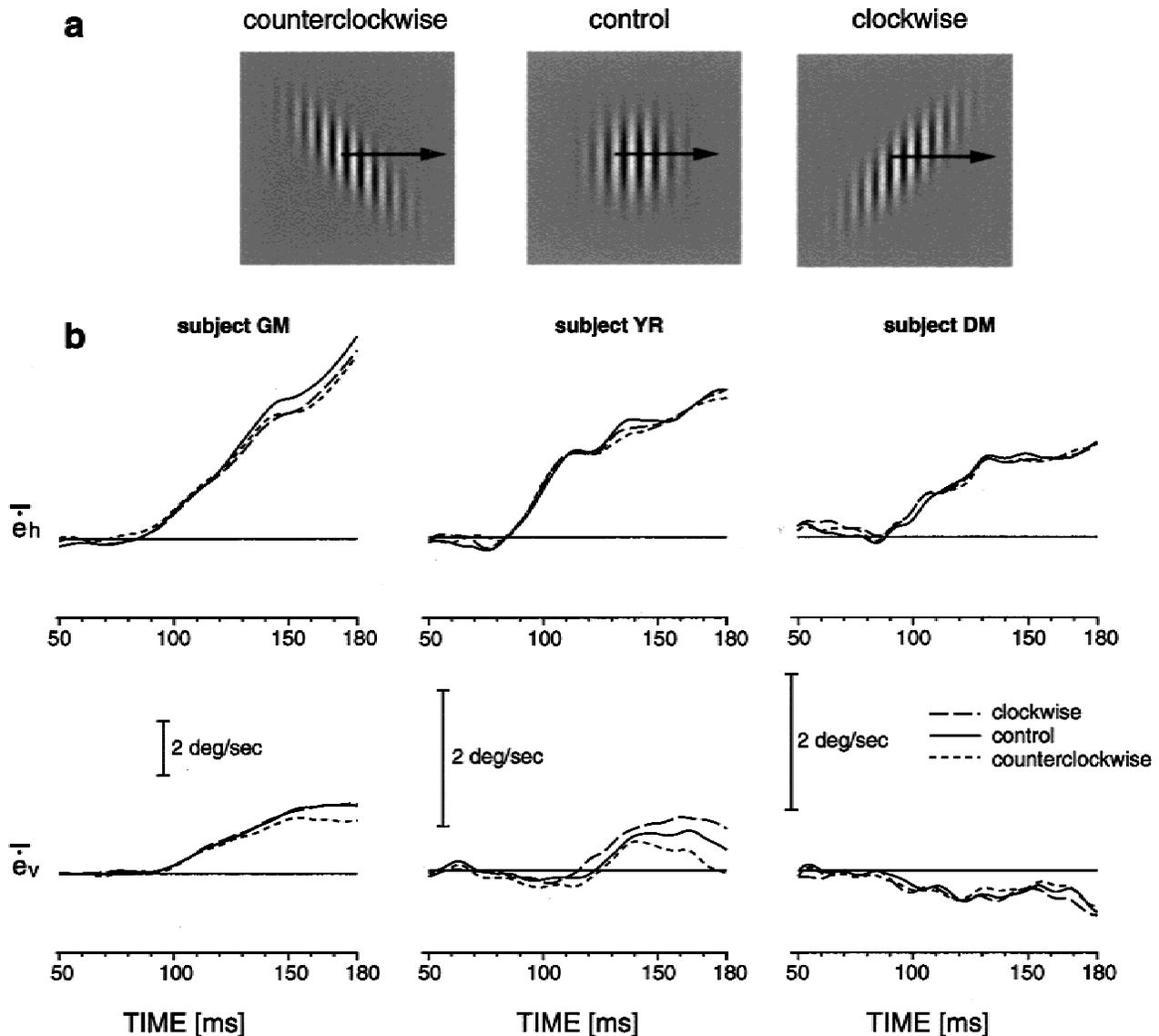


**Fig. 7.** Effects of foveal mask size. (a) One frame of the stimuli with different mask sizes. Notice that aspect ratios of both the mask and the aperture are identical. (b) Horizontal eye velocities of tracking responses elicited by a rightward drifting grating, for different mask sizes. The gray bar indicates the 20-ms time window (95–115 ms) selected to measure the effect of the mask size on the responses amplitude. (c) Mean ( $\pm$ SD) change in horizontal position as a function of mask size, for three subjects. Right-end symbols indicate the change in horizontal position induced by a rightward drifting grating viewed behind an upright square aperture. (d) Vertical eye velocity profile of the responses induced by the “barber pole” stimuli. Numbers indicate mask sizes. (e) Mean ( $\pm$ SD) change in vertical position, as a function of mask size. Right-end symbols illustrate the genuine cross-talk observed with a pure rightward drifting grating presented within a square, upright aperture.

motion signals by increasing the size of the mask at the center of the stimulus decreased the amplitude of the early component, but did not affect the magnitude of the later component. Altogether, these findings indicate that 2D tracking eye movements are controlled by

a 2D velocity vector built up by integrating different local motion signals across the visual field.

As plotted in Fig. 4, mean tracking direction at the end of the trial (time window: 135–175 ms, that is when both components

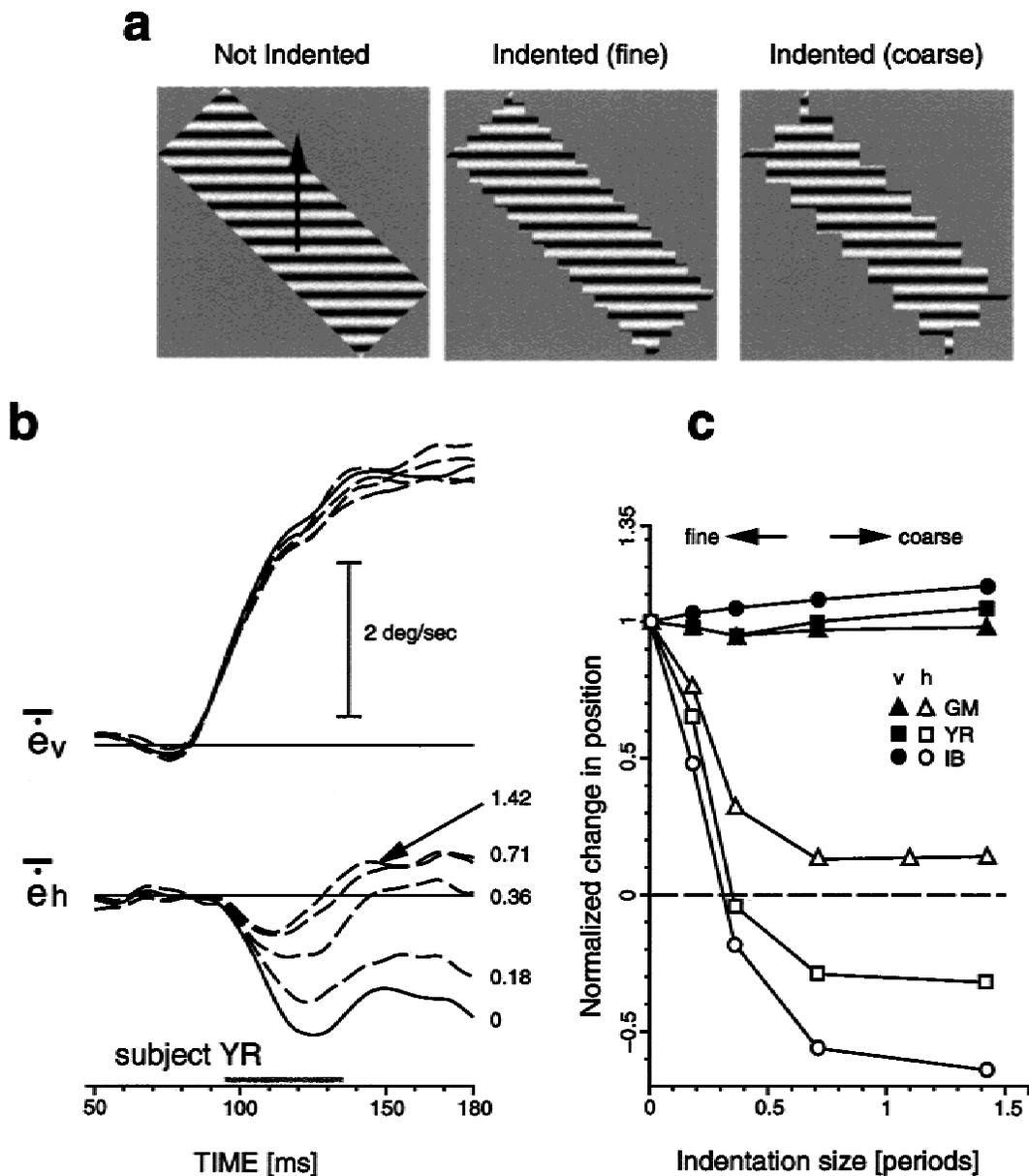


**Fig. 8.** Effect of reducing the contrast of line-endings. (a) A vertical grating, drifting rightward, is filtered by a symmetrical or an elongated Gaussian window, producing circular or elongated tilted Gabor patches, respectively. (b) Horizontal and vertical velocity of tracking responses for each type of stimulus. No significant difference in vertical eye movements were elicited by elongating the Gabor patch along either of the two diagonal axis, as compared to the circular symmetric patch.

have been fully initiated) was found to be dependent upon the aspect ratio of the “barber pole.” Clearly, with  $AR > 1$ , the mean tracking direction was no longer colinear with the direction of the grating motion, that is with the 1D motion signal. It was also not colinear with the aperture’s long axis (i.e. +45 deg in Fig. 4). We found that the mean tracking direction was in between the directions of the two vector sums computed either from the two different terminator motion vectors only (i.e. terminator motion signals along the short and long axis) or both terminator and grating motion vectors. Notice that vector summation and vector averaging give identical direction of the resulting vector and that therefore our analysis was not designed to disentangle these two sorts of motion vectors combination. In brief, our findings suggest that short-latency ocular following responses are driven by a mechanism that integrates the different types of local motion signals

across the visual field but that this integration is constrained by the temporal dynamics of the processing of each type of motion signal.

Previous studies in monkeys have shown that motion averaging is the most likely mechanism for converting a distributed representation of image motion into commands for tracking eye movements (e.g. Groh et al., 1997; Lisberger & Ferrera, 1997). Similarly, initial phases of both optokinetic and smooth pursuit eye movements in humans exhibit a similar motion averaging computation in either the speed (Mestre & Masson, 1997) or the direction (Watamaniuk & Heinen, 1999) domains. Most of these studies used dots (either single spot or random-dots flowfields) moving in different directions across different parts of the visual field. Lisberger and Ferrera (1997) assumed that “the pursuit system uses the same computation to combine information from the two spatial locations [we used] as it does for a single location” (page 7500).



**Fig. 9.** Role of line-endings motion signals. (a) One frame of the “barber pole” stimuli when the aperture edges are either not, finely, or coarsely indented. Grating is drifted upward (arrow). (b) Horizontal ( $\dot{e}_h$ ) and vertical ( $\dot{e}_v$ ) velocity profiles of tracking responses evoked by indented “barber poles.” Size of indentation increased from 0 (continuous line, no indentation) to 3.83 deg, that is, 1.42 cycles of grating period (broken lines, coarse indentation), as indicated by numbers. (c) Normalized change in horizontal (open symbols) and vertical (closed symbols) position, plotted as a function of indentation size, expressed as a fraction of grating period, for three subjects.

The present study indicates that a vector combination strategy is used to recover the 2D direction of object motion. Furthermore, such strategy develops over time as it needs to integrate ambiguous (1D) and nonambiguous (2D) motion signals that are connected to form a single motion surface.

#### *From following to pursuing*

The hypothesis that the early part of tracking eye movements consists of two phases has been already suggested (Lisberger & Westbrook, 1985; Miles et al., 1986). Herein, we show that the first

20 ms of tracking are triggered by a mechanism which integrates direction-selective local motion signals from local changes in the luminance profile. The later component is driven by a visual integration process which computes an estimate of the global motion direction of an object, by integrating different local motion signals over a large part of the visual field. These results also pertain to the long-lasting controversy on the distinction between optokinetic and pursuit smooth eye movements (Steinman, 1986; Lisberger et al., 1987). It is of interest to note that the latency of our later component was very close to the latency of the voluntary smooth pursuit eye movements in humans (Carl & Gellman, 1987). One

might then argue that while the earlier component were ocular following responses the later were smooth pursuit eye movements.

We think such interpretation is unlikely for several reasons. First of all, subjects were never instructed to track any particular features within the motion stimulus. Trials were of short duration (200 ms) and all conditions were fully randomized so that both attentional selection or anticipatory mechanisms are minimized. Second, the amplitude of this later component was found dependent upon several parameters specifically affecting the terminator motion signals. Moreover, the very first velocity raising phase of both components were found modulated by the speed of the grating and terminator motion signals, respectively (Fig. 5). Such early speed sensitivity is a signature of the machine-like, ocular following responses (Gellman et al., 1990; Masson et al., 1995) while in both monkeys and humans the very first 20 ms of smooth pursuit responses to single moving spot are insensitive to target speed (Lisberger & Westbrook, 1985; Heinen & Watamaniuk, 1998). Moreover, mean tracking direction at the end of the trial was not systematically aligned with neither single 2D feature motion direction nor the vector combination of them as would be expected if the subjects were picking up a local features or a combination of them. These results suggest that the later component is indeed dependent upon a motion integration process and does not simply reflect the fact that the observers picked up a local 2D motion feature and actively tracked it, ignoring the other competing 1D motion signals.

It is in fact rather unclear whether or not ocular following and smooth pursuit tracking responses are separate types of eye movements. A close correlation has been suggested between these two conjugate visual tracking systems at the neurophysiological level (see Kawano, 1999). However, involvement of extra-retinal signals carrying attentional selection mechanism for instance might help in teasing apart pre-attentive (i.e. reflexive) and attentive (i.e. voluntary) smooth eye movements (Keating et al., 1996). We suggest that considering the motion processing hierarchy and its temporal dynamics would help to define the contribution of reflexive and intentional components of the tracking behavior in primates (Mestre & Masson, 1997).

#### Neural mediation

Our results stress the role of a local mechanism detecting the motion direction at the line-endings. This process is slower than the local process detecting grating motion. The nature of the mechanism extracting 2D motion signals at the aperture edges is unclear. The simplest way to recover a motion signal in that direction would be to extract the Fourier-like motion signals generated by the local changes in luminance along the aperture. A linear motion energy detector performs a local Fourier analysis (Watson & Ahumada, 1985). This type of detector is optimally activated by a motion perpendicular to its preferred orientation. With a "barber pole" stimulus, a grating tilted by 45 deg relative to this orientation is a suboptimal motion input for a detector tuned to the direction of motion along the aperture edges. One would suspect such a weak motion signal to elicit a weaker and, presumably, later response of the population of neurons tuned for this particular direction of motion. In fact, Celebrini et al. (1993) showed that, in area V1, many neurons exhibit a marked tendency to respond at somewhat longer latencies to flashed, nonoptimally oriented stimuli. To our knowledge, no such evidence are available for moving gratings and from higher stages of the primate visual motion pathways. Such a dependency might however explain why motion

angle 45 deg away from the preferred direction resulted in slightly longer (<10 ms) and smaller ( $\approx -30\%$ ) responses.

It is unlikely however that only this weak Fourier-like motion signal can explain our results. First, it is rather difficult to compare the strength of the grating and the terminator motion signals within the "barber pole" stimulus. Their local luminance contrast is the same, a factor which is known to affect the latency of both monkey ocular following responses (Miles et al., 1986) and human smooth pursuit responses (O'Mullane & Knox, 1998). Their speeds are different (33 vs. 46 deg/s) but within this range the changes in latency of the human ocular following responses is rather negligible (<5 ms, Gellman et al., 1990). Nevertheless, further studies shall investigate whether or not, and in what proportion, the motion signal strength in the Fourier domain has an effect on the latency of human ocular following responses as well as on MT neurons (Britten et al., 1993).

Second, a more crucial problem related to analyzing the line-endings motion with a pure Fourier-like motion detection mechanism is its lack of reliability. In a very recent computational study, Löffler and Orbach (1999) showed that a pure 2D Fourier analysis of a moving terminator cannot accurately signal the physical motion direction of the line-ending. Fig. 4 of the present study illustrates that the initial tracking direction ends up between the direction predicted by the (unweighted) vector summation of either terminator (2D) signals alone or both grating (1D) and terminator (2D) signals together. Given the large population of neurons activated by the drifting grating within the aperture, one would expect only a very minor contribution of another 1D motion signal at the aperture edges and initial tracking direction would end up very close to the grating motion direction.

Löffler and Orbach (1999) suggested that a second, non-Fourier mechanism is necessary for the computation of veridical features motion direction, within the 5-deg error range reported in both psychophysical (Ben-Av & Shiffrar, 1995) and steady-state smooth pursuit (Beutter & Stone, 1997) studies. Their model is similar in structure to popular models originally proposed for motion perception with plaid patterns (Wilson et al., 1992) and have been extended to nonlinear cortical processes for motion perception (see Wilson, 1999; Baker, 1999). This class of models postulates two parallel motion pathways, a Fourier and a non-Fourier motion pathways followed by a combination of these responses in a network that exhibits many of the properties of area MT, such as pattern-selective neurons (Movshon et al., 1985; Rodman & Albright, 1989). The first pathway involves a linear spatio-temporal filtering of the moving image. It recovers the local 1D motion signals such as the components of plaid patterns and project directly to the second stage of the motion pathway, area MT (Movshon & Newsome, 1996). This direct route from V1 to MT is expected to be as fast as the 1D motion processing, and in fact it has been recently shown in monkeys that ocular following responses elicited by either drifting grating or plaids have similar latencies (Guo & Benson, 1998). Aside from this linear route, a second, nonlinear mechanism is able to extract stimulus elements that are not represented by any Fourier component or sum of components in the stimulus. Consequently, this type of processing is frequently termed "non-Fourier" processing (Chubb & Sperling, 1988), although it has also been referred to as "second-order" processing (Cavanagh & Mather, 1990; Baker, 1999). Psychophysical studies examining the processing of "second-order" stimuli have suggested a slower, nonlinear stream for motion perception (Yo & Wilson, 1992; Derrington et al., 1993). Physiological studies in cat area 18 (Marschall & Baker, 1998) have evidenced longer latencies of neuronal

responses to second- as compared to first-order stimuli. Therefore, it has been suggested that this indirect route includes an additional stage, presumably area V2. This non-Fourier pathway is expected to be slower. These two parallel motion streams converge onto MT area which computes the vector sum direction between direct and indirect routes and therefore signals a nonambiguous 2D velocity vector. In fact, Born and colleagues recently recorded responses from MT neurons to patches of bars whose orientation was deviated from motion direction by 45 deg or more (Pack et al., 2000). They found that the earliest response (<100 ms) of most MT cells primarily encode the component of motion perpendicular to the orientation of the bars while the later response encode the actual direction of bar motion, irrespective of bar orientation. These neurons might implement this convergent stage computing the veridical object motion direction.

Measuring the motion of a visual target is essential to control tracking eye movements and extra-striate areas MT and MST are implicated as a major step for this computation. Electrophysiological studies in monkey suggest that the short-latency tracking responses are mediated, in part, by area MST (Kawano et al., 1994). The main input to area MST is from area MT that is also involved in the visual motion processing for smooth tracking eye movements (Lisberger & Movshon, 1999). Neurons in area MT are activated by a wide range of, loosely speaking, second-order stimuli (Rodman & Albright, 1989; Albright, 1992; O'Keefe & Movshon, 1998). However, latencies of the responses to Fourier or non-Fourier stimuli have not yet been compared in primates extra-striate cortex, as already done in cats visual cortex (Mareschal & Baker, 1998). The hypothesis that non-Fourier motion processing is slower is, however, supported by recent behavioral studies. First, in monkeys, the latency of ocular following responses to a second-order stimuli is 10–15 ms longer than the response to a grating motion stimulus (Benson & Guo, 1999). Second, the initial saccade during voluntary smooth pursuit responses to a pure second-order motion stimulus is delayed when compared to the latency of responses to a Fourier motion (Bützer et al., 1997). Area MT main output is area MST. It remains unknown firstly whether or not motion-selective neurons in area MST do respond to plaids or to other types of second-order motion stimuli; and secondly, if neuronal responses in areas MT and MST are also delayed since that critical piece of information is still lacking in primates.

### Conclusion

Further experimental studies shall investigate whether later parts of tracking responses reflect higher-order processing yielding to multivalued representation of motion direction and multistable tracking direction as previously evidenced with optokinetic responses to multiple speeds flowfield (Mestre & Masson, 1997). First, by recording eye movements we are therefore able to dissociate the various hierarchical stages from motion detection to motion representation. Second, present results call for the need to design models of oculomotor control that are based on the representation of object motion rather than on retinal velocity-error signals. As suggested by Stone and colleagues (Beutter & Stone, 1997; Krauzlis & Stone, 1999), such representation-based models need a more sophisticated front-end which can perform the spatio-temporal integration which is necessary to recover object motion in an highly complex natural visual scene. We have demonstrated here that a timed, hierarchical processing is a major constraint for any visual model of oculomotor control.

### Acknowledgments

We thank Bernard Arnaud and Raymond Fayolle for technical assistance. We thank Dr. Richard Krauzlis for providing us with the Idea software. We thank Isabelle Barbet from carrying out the experiments; Drs. Richard Krauzlis, Leland S. Stone, and Simona Celebrini for their comments on an earlier version of this manuscript; and Cheryl Frenck-Mestre for editing the final draft. This work was supported by the CNRS and by a grant from C.G 13.

### References

- ADELSON, E.H. & MOVSHON, J.A. (1982). Phenomenal coherence of moving visual pattern. *Nature* **300**, 523–525.
- ALBRIGHT, T.D. (1992). Form-cue invariant motion processing in primate visual cortex. *Science* **255**, 1141–1143.
- ANDERSON, B.L. & SINHA, P. (1997). Reciprocal interactions between occlusion and motion computations. *Proceedings of the National Academy of Sciences of the U.S.A.* **94**, 3477–3480.
- BAKER, C.L., JR. (1999). Central neural mechanisms for detecting second-order motion. *Current Opinion in Neurobiology* **9**, 461–466.
- BEN-AY, M.B. & SHIFFRAN, M. (1995). Disambiguating velocity estimates across image space. *Vision Research* **35**, 2889–2895.
- BENSON, P.J. & GUO, K. (1999). Stages in motion processing revealed by the ocular following response. *NeuroReport* **10**, 3803–3807.
- BEUTTER, B.R. & STONE, L.S. (1997). Human motion perception and smooth eye movements show similar directional biases for elongated apertures. *Vision Research* **38**, 1273–1286.
- BRITTEN, K.H., SHADLEN, M.N., NEWSOME, W.T. & MOVSHON, J.A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience* **10**, 1157–1170.
- BUSETTINI, C., MILES, F.A. & SCHWARZ, U. (1991). Ocular responses to translation and their dependence on viewing distance. II. Motion of the scene. *Journal of Neurophysiology* **66**, 865–878.
- BÜTZER, F., ILG, U.J. & ZANKER, J.M. (1997). Smooth-pursuit eye movements elicited by first-order and second-order motion. *Experimental Brain Research* **115**, 61–70.
- CARL, J.R. & GELLMAN, R.S. (1987). Human smooth pursuit: Stimulus-dependent responses. *Journal of Neurophysiology* **57**, 1446–1463.
- CASTET, E. & WUERGER, S. (1997). Perception of moving lines: Interaction between local perpendicular signals and 2D motion signals. *Vision Research* **37**, 705–720.
- CASTET, E., CHARTON, V. & DUFOUR, A. (1999). The extrinsic/intrinsic classification of 2D motion signals in the barberpole illusion. *Vision Research* **39**, 915–932.
- CAVANAGH, P. & MATHER, G. (1990). Motion: The long and the short of it. *Spatial Vision* **4**, 103–129.
- CELEBRINI, S., THORPE, S., TROTTER, Y. & IMBERT, M. (1993). Dynamics of orientation coding in area V1 of awake primate. *Visual Neuroscience* **10**, 811–826.
- CHUBB, C. & SPERLING, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America* **5**, 1986–2007.
- COLLEWIJN, H., VAN DER MARK, F. & JANSEN, T.C. (1975). Precise recordings of human eye movements. *Vision Research* **15**, 447–450.
- DERRINGTON, A.M., BADCOCK, D.R. & HENNING, B.G. (1993). Discriminating the direction of second-order motion at short stimulus duration. *Vision Research* **33**, 1785–1794.
- FENNEMA, C.L. & THOMPSON, W.B. (1979). Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing* **9**, 301–315.
- FUCHS, A.F. & ROBINSON, D.A. (1966). A method for measuring horizontal and vertical eye movement chronically in the monkey. *Journal of Applied Physiology* **21**, 1068–1070.
- GELLMAN, R.S., CARL, J.R. & MILES, F.A. (1990). Short-latency ocular-following responses in man. *Visual Neuroscience* **5**, 107–122.
- GROH, J.M., BORN, R.T. & NEWSOME, W.T. (1997). How is a sensory map read out? Effects of microstimulation in visual area MT on saccades and smooth pursuit eye movements. *Journal of Neuroscience* **17**, 4312–4330.
- GUO, K. & BENSON, P.J. (1998). Involuntary eye movements in responses to first- and second-order motion. *NeuroReport* **9**, 3543–3548.
- HAYS, A.V., RICHMOND, B.J. & OPTICAN, L.A. (1982). A UNIX-based multiple process system for real-time data acquisition and control. *WESCON Conference Proceedings* **2**, 1–10.

- HEINEN, S.J. & WATAMANIUK, S.N.J. (1998). Spatial integration in human smooth pursuit. *Vision Research* **38**, 3785–3794.
- HILDRETH, E. (1984). *The Measurement of Visual Motion*. Cambridge, Massachusetts: MIT Press.
- HUBEL, D. & WIESEL, T. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology* (London) **195**, 215–243.
- KAWANO, K. (1999). Ocular tracking: Behavior and neurophysiology. *Current Opinion in Neurobiology* **9**, 467–473.
- KAWANO, K., SHIDARA, M., WATANABE, Y. & YAMANE, S. (1994). Neural activity in cortical area MST of alert monkey during ocular following responses. *Journal of Neurophysiology* **71**, 2305–2324.
- KAWANO, K. & MILES, F.A. (1986). Short-latency ocular following responses of monkey. II. Dependence on a prior saccadic eye movement. *Journal of Neurophysiology* **56**, 1355–1380.
- KEATING, E.G., PIERRE, A. & CHOPRA, S. (1996). Ablation of the pursuit area in the frontal cortex of the primates degrades foveal but not optokinetic smooth eye movements. *Journal of Neurophysiology* **76**, 637–641.
- KOOI, F.L. (1993). Local direction of edges motion causes and abolishes the barberpole illusion. *Vision Research* **33**, 2479–2489.
- KOWLER, E. & STEINMAN, R.M. (1981). The effects of expectations on slow oculomotor control. III. Guessing unpredictable target displacements. *Vision Research* **21**, 191–203.
- KRAUZLIS, R.J. & STONE, L.S. (1999). Tracking with the mind's eye. *Trends in Neuroscience* **22**, 544–550.
- LANDY, M.S., COHEN, Y. & SPERLING, G. (1984). HIPS: Image processing under UNIX. Software and applications. *Behavioral Research. Methods, Instruments & Computers* **16**, 199–216.
- LENNIE, P. (1998). Single units and visual cortical organization. *Perception* **27**, 889–935.
- LISBERGER, S.G. & WESTBROOK, L.E. (1985). Properties of visual inputs that initiate horizontal smooth pursuit eye movements in monkeys. *Journal of Neuroscience* **5**, 1662–1672.
- LISBERGER, S.G. & MOVSHON, J.A. (1999). Visual motion analysis for pursuit eye movements in area MT of macaque monkeys. *Journal of Neuroscience* **19**, 2224–2246.
- LISBERGER, S.G. & FERRERA, V.P. (1997). Vector averaging for smooth pursuit eye movements initiated by two moving targets in monkeys. *Journal of Neuroscience* **17**, 7490–7502.
- LISBERGER, S.G., MORRIS, E.J. & TYCHSEN, L. (1987). Visual motion processing and sensorimotor integration for smooth pursuit eye movements. *Annual Review of Neuroscience* **10**, 97–129.
- LÖFFLER, G. & ORBACH, H.S. (1999). Computing feature motion without feature detectors: A model for terminator motion without end-stopped cells. *Vision Research* **39**, 859–871.
- LORENCEAU, J. & SHIFFRAN, M. (1992). The influence of terminators on motion integration across space. *Vision Research* **32**, 263–273.
- MARESCHAL, I. & BAKER, C.L., JR. (1998). Temporal and spatial response to second-order stimuli in cat area 18. *Journal of Neurophysiology* **80**, 2811–2823.
- MARR, D. & ULLMAN, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society B* (London) **211**, 151–180.
- MASSON, G.S. & MESTRE, D.R. (1998). A look into the black box: Eye movements as a probe of visual motion processing. *Cahiers de Psychologie Cognitive* **17**, 807–829.
- MASSON, G.S., BUSETTINI, C. & MILES, F.A. (1995). Initial tracking of motion-in-depth: Contribution of short latency version (ocular following) and vergence. *Society for Neuroscience Abstracts* **21**, 104–108.
- MESTRE, D.R. & MASSON, G.S. (1997). Ocular responses to motion parallax stimuli: The role of perceptual and attentional factors. *Vision Research* **37**, 1627–1641.
- MILES, F.A. (1998). The neural processing of 3-D visual information: Evidence from eye movements. *European Journal of Neuroscience* **10**, 811–822.
- MILES, F.A., KAWANO, K. & OPTICAN, L.M. (1986). Short-latency ocular following responses of monkey. I. Dependence on temporo-spatial properties of the visual input. *Journal of Neurophysiology* **56**, 1321–1354.
- MOVSHON, J.A. & NEWSOME, W.T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *Journal of Neuroscience* **16**, 7733–7741.
- MOVSHON, J.A., ADELSON, E.H., GIZZI, M.S. & NEWSOME, W.T. (1985). The analysis of visual moving patterns. In *Pattern Recognition Mechanism*, eds CHAGAS, C., GATTASS, R. & GROSS, C., pp. 117–151. New York, New York: Springer.
- NAKAYAMA, K. & SILVERMAN, G.H. (1988). The aperture problem. II. Spatial integration of velocity information along contours. *Vision Research* **28**, 747–753.
- O'KEEFE, L.P. & MOVSHON, J.A. (1998). Processing of first- and second-order motion signals by neurons in area MT of the macaque monkey. *Visual Neuroscience* **15**, 305–317.
- O'MULLANE, G.M. & KNOX, P.C. (1998). Contrast modification of smooth-pursuit latency. *Perception* **27**, Suppl. 1, 145a.
- PACK, C., ABRAMS, P.L. & BORN, R.T. (2000). Neural and behavioral correlates of ambiguous local motion measurements in cortical visual area MT. *Perception* **29** (Suppl. 1), 816.
- POWER, R.P. & MOULDEN, B. (1992). Spatial gating effects on judged motion of grating in apertures. *Perception* **21**, 449–463.
- RODMAN, H.R. & ALBRIGHT, T.D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research* **75**, 53–64.
- SHIFFRAN, M., LI, X. & LORENCEAU, J. (1995). Motion integration across differing image features. *Vision Research* **35**, 2137–2146.
- SHIMOJO, S., SILVERMAN, G. & NAKAYAMA, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research* **29**, 619–626.
- SIMONCELLI, E.P. & HEEGER, D.J. (1998). A model of neuronal responses in visual area MT. *Vision Research* **38**, 743–761.
- STEINMAN, R.M. (1986). The need for an eclectic, rather than systems, approach to the study of the primate oculomotor system. *Vision Research* **26**, 101–112.
- WALLACH, H. (1935). Ueber Visuell Wahrgenommene Bewegungsrichtung. *Psychologische Forschung* **20**, 325–380 (Translated by WUERGER, S. & SHAPLEY, B. 1996). *Perception* **11**, 1317–1367.
- WATAMANIUK, S.N.J. & HEINEN, S.J. (1999). Human smooth pursuit direction discrimination. *Vision Research* **39**, 59–70.
- WATSON, A.B. & AHUMADA, A.J. (1985). Model of human visual motion sensing. *Journal of the Optical Society of America* **A2**, 322–342.
- WILSON, H.R. (1999). Non-Fourier cortical processes in texture, form and motion perception. In *Cerebral Cortex Vol. 13, Models of Cortical Circuits*, ed. ULINSKY, P.S., pp. 445–477. New York: Kluwer Academic/Plenum Publishers.
- WILSON, H.R., FERRERA, V.P. & YO, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience* **9**, 79–98.
- YO, C. & WILSON, H.R. (1992). Perceived direction of moving two-dimensional patterns depends on duration, contrast and eccentricity. *Vision Research* **32**, 135–147.