Systems/Circuits

A Motion-from-Form Mechanism Contributes to Extracting Pattern Motion from Plaids

Christian Quaia, ⬤Lance M. Optican, and Bruce G. Cumming

Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland 20892

Since the discovery of neurons selective for pattern motion direction in primate middle temporal area MT (Albright, 1984; Movshon et al., 1985), the neural computation of this signal has been the subject of intense study. The bulk of this work has explored responses to plaids obtained by summing two drifting sinusoidal gratings. Unfortunately, with these stimuli, many different mechanisms are similarly effective at extracting pattern motion. We devised a new set of stimuli, obtained by summing two random line stimuli with different orientations. This allowed several novel manipulations, including generating plaids that do not contain rigid 2D motion. Importantly, these stimuli do not engage most of the previously proposed mechanisms. We then recorded the ocular following responses that such stimuli induce in human subjects. We found that pattern motion is computed even with stimuli that do not cohere perceptually, including those without rigid motion, and even when the two gratings are presented separately to the two eyes. Moderate temporal and/or spatial separation of the gratings impairs the computation. We show that, of the models proposed so far, only those based on the intersection-of-constraints rule, embedding a motion-from-form mechanism (in which orientation signals are used in the computation of motion direction signals), can account for our results. At least for the eye movements reported here, a motion-from-form mechanism is thus involved in one of the most basic functions of the visual motion system: extracting motion direction from complex scenes.

Key words: flicker; motion; MT; unikinetic

Significance Statement

Anatomical considerations led to the proposal that visual function is organized in separate processing streams: one (ventral) devoted to form and one (dorsal) devoted to motion. Several experimental results have challenged this view, arguing in favor of a more integrated view of visual processing. Here we add to this body of work, supporting a role for form information even in a function— extracting pattern motion direction from complex scenes—for which decisive evidence for the involvement of form signals has been lacking.

Introduction

At the earliest stage of cortical visual processing [primary visual cortex (V1) in primates] visual information is filtered by orientation-selective units with small receptive fields. As a result, V1 units generally cannot respond to motion parallel to their preferred orientation and only generate one-dimensional (1D)

motion signals (signed speed in the direction orthogonal to their preferred orientation). Recovering the motion direction of rigid patterns containing more than one orientation then requires additional processing (Movshon et al., 1985) and involves later areas [e.g., middle temporal area (MT)]. This type of serial processing plays an important role in many models of brain function. How this integration of 1D signals is accomplished for patterns containing two orientations (plaids) has been the subject of intense study.

With some plaids (type I; Fig. 1A), summing the velocity vectors of the two gratings is sufficient to recover the direction (although not necessarily the speed) of pattern motion (Fig. 1B). With other plaids (type II; Fig. 1C) this simple solution does not work (Fig. 1D), and a more complex rule, such as the intersection-of-constraints (IOC) rule (Fennema and Thompson, 1979; Adelson and Movshon, 1982), described in the legend for Figure 1B, must be invoked. Models based on a neural implementation of the IOC rule have been proposed (Simoncelli and
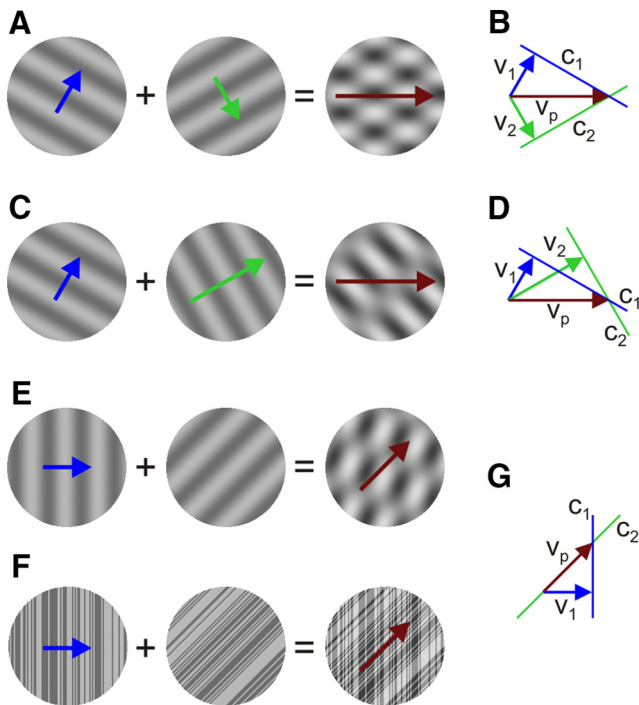
Figure 1. Pattern motion in plaids. The pattern motion of a plaid is a function of both the orientation and speed of the two gratings. *A*, The two gratings have the same drifting speed, and their motion directions are 120° apart (blue and green arrows). In this, like in all sinusoidal plaids, the pattern velocity (dark red arrow) equals the velocity of the high and low luminance blobs formed where the minima and maxima of the two gratings coincide. *B*, The pattern motion direction can be computed using a construct called IOC. From the tip of each grating motion vector (V_1 and V_2), one draws the orthogonal constraints lines (C_1 and C_2). These lines (which are parallel to the orientation of the respective grating) define the locus on which all velocities compatible with the motion of the 1D grating lie. The intersection of these two lines represents the only velocity that is compatible with the motion of both gratings: the pattern motion velocity (V_p). For this plaid, the pattern motion direction equals the average of the gratings motion directions. This is then a so-called type I plaid. *C*, In this plaid, the two sinusoidal gratings drift at different speeds. *D*, Notably, with this configuration, the pattern motion velocity vector (V_p) does not lie between the two grating velocity vectors (V_1 and V_2), as shown by the IOC construct. This is then a so-called type II plaid. *E*, In a sinusoidal unikinetic plaid, one grating drifts, whereas the other is static. *F*, In a binary noise unikinetic plaid, one line grating drifts, whereas the other is static; however, there are no clear (i.e., pre-attentively salient) blobs to track. *G*, In a unikinetic plaid, made up of sinusoidal or noise gratings, the constraint line for the static grating (C_2) is still parallel to the orientation of the grating, but because the speed of the grating is zero, it passes through the origin. The pattern motion direction (V_p) must thus lie on this constraint line and is thus parallel to the orientation of the static grating. Unikinetic plaids are a limit case of type II plaids.

Heeger, 1998; Perrone, 2004; Rust et al., 2006), but other mechanisms, which we will describe in Results, would also work. What these other mechanisms have in common is that they extract, in each frame, features that arise from the combination of the two gratings and then track how those features move across frames. In contrast, an IOC-based mechanism only relies on the 1D velocity vectors of the two gratings in the plaid. A corollary of this distinction is that the IOC explicitly uses, when present, information about the orientation of static gratings (i.e., form information). Accordingly, an IOC operator embeds a so-called motion-from-form mechanism, i.e., a mechanism in which form information, typically orientation, affects motion processing. The importance of such mechanisms in human motion perception is well established (Geisler, 1999; Ross et al., 2000; Badcock et al., 2003), but the role that motion-from-form might play in pattern motion extraction has been neither recognized nor experimentally evaluated.

Plaids composed of two drifting gratings, used in the majority of studies performed thus far, are obviously ill-suited to test theories of pattern motion extraction that use static form information. Unikinetic sinusoidal plaids (Fig. 1E), formed by adding a drifting and a static sinusoidal grating having different orientations (Gorea and Lorenceau, 1991), are more useful but fall short of isolating the role of form signals. We thus designed a novel type of plaid, obtained by summing two random line stimuli (RLS; Fig. 1F), which allows manipulations that cleanly discriminate the predictions of different models. For example, when one RLS drifts and the other is dynamically refreshed on each video frame (i.e., flickering), the stimulus, which we call a flicker plaid, does not contain 2D pattern motion (a different pattern is present in each frame), seriously challenging existing models.

To infer how the brain processes our novel plaids, we recorded the ocular following responses (OFRs; Miles et al., 1986; Gellman et al., 1990) that they elicit in human subjects. We found that flicker plaids induce OFRs directed parallel to the orientation of the flickering RLS. By means of simulations, we then show that a model based on the IOC (Simoncelli and Heeger, 1998), with parameters derived from the literature (Nishimoto and Gallant, 2011) and not fit to our data, can reproduce these results. We argue that this, and a range of additional results reported here, are only compatible with models in which a motion-from-form mechanism plays an important role in pattern motion analysis.

Materials and Methods

Subjects. Three male subjects (aged 30–55 years) participated in all the experiments; one was an author (C.Q.), and the others were experienced as OFR subjects but unaware of the experimental questions being investigated. All had normal or corrected-to-normal visual acuity and normal stereo-acuity. Experimental protocols were approved by the Institutional Review Board concerned with the use of human subjects.

Visual apparatus. The subjects sat in a dark room and were positioned so that their eyes were located approximately in the center of a cubic box (70 cm side) containing orthogonal magnetic-field generating coils. Their chin and forehead rested on padded supports, and their head was stabilized using a head band. The coil box was surrounded by three monitors, one in front of the subjects and one on each side. The latter two were part of a Wheatstone mirror stereoscope, with removable mirrors.

For most of the experiments, visual stimuli, generated through a EVGA GEForce GTX 680 graphic card, were presented on the CRT monitor (NEC MultiSync FP2141SB) facing the subject. The monitor screen was located 525 mm from the corneal vertex, covered 41° (horizontal) × 31° (vertical) of visual angle, was set at a resolution of 1024 columns × 768 rows, and had a refresh rate of 150 Hz. Only the red channel was used, because red phosphors had the shortest persistence (800 μs rise time, 3.2 ms fall time), guaranteeing the absence of motion streaks on the monitor.

For the dichoptic recordings, each eye saw one of the two side CRT monitors (Sony GDM-FW900) through a 45° mirror, creating a binocular image straight ahead at a distance of 521 mm from the corneal vertex (which was also the optical distance to the two monitor screens). In this case, each monitor screen covered 50° (horizontal) × 32° (vertical) of visual angle, was set at a resolution of 1280 columns × 800 rows, and had a refresh rate of 140 Hz. Again, only the red channel was used to minimize persistence (1 ms rise time, 4 ms fall time). A single video card (EVGA GEForce GTX 580 Classified) was used to provide the inputs to both monitors. Using the NVIDIA Control Center configuration tool, the two monitors were set up so as to appear to the operating system (Microsoft Windows XP) as a single monitor with a resolution of 2560 columns × 800 rows. Using two photocells connected to a digital oscilloscope, we verified that the refresh timing of the two monitors was tightly synchronized, with the left-eye image consistently preceding the right-eye image by <50 μs.

In both settings, luminance linearization was performed by interpolation after dense luminance sampling (using a Konica Minolta LS100 luminance meter), independently for each monitor.

Eye movement recording. A scleral search coil embedded in a silastin ring (Skalar; Collewijn et al., 1975) was placed in one of the subject's eyes, or both for the dichoptic experiments, after application of topical anesthetic (proparacaine HCl). The horizontal and vertical orientations of the eyes were recorded using an electromagnetic induction technique (Robinson, 1963). These outputs were calibrated at the beginning of each recording session by having the subject look at targets of known eccentricity. Peak-to-peak noise levels resulted in an uncertainty in eye position recording of $<0.03°$. Coil signals were sampled at 1000 Hz.

Experiment control. The experiment was controlled by two computers, communicating over an Ethernet with TCP/IP. The Real-time EXperimentation software package (Hays et al., 1982), running on the master computer under the QNX operating system, was responsible for providing the overall experimental control, as well as acquiring, displaying, and storing the eye movement data. The other machine, directly connected to the CRT displays, ran under the Windows XP operating system and generated the required visual stimuli in response to REX commands. This was accomplished using the Psychophysics Toolbox 3.0.8, a set of MATLAB (MathWorks) scripts and functions (Brainard, 1997).

Behavioral paradigm. Trials were presented in blocks; each block contained one trial for each stimulus condition. All conditions within a block were interleaved randomly. Each trial began with the appearance of a central fixation cross on a blank, midluminance (6.0 cd/m^2) background. The subject was instructed to look at the center of the cross and avoid making saccadic eye movements. After the subject maintained fixation within a small (1° on the side) invisible window around the fixation point for 800–1100 ms, the fixation cross disappeared, and the visual stimulus sequence (24 frames) was presented. Subsequently, the screen was blanked (again at 6.0 cd/m^2), signaling the end of the trial. After a short intertrial interval, a new trial was started. If the subject blinked or if saccades were detected during the stimulus presentation epoch, the trial was discarded and repeated within the block. With few exceptions, a single experiment required multiple daily recording sessions to collect enough trials from each subject (we collected between 150 and 450 trials for each condition, depending on the signal-to-noise ratio, and between 1500 and 3000 trials in a session; the number of conditions varied across experiments).

Visual stimuli. In all experiments described here, we presented plaids obtained by summing the contrasts (i.e., deviations from the mean luminance) of a vertical and an oblique (tilted +45° or −45° from horizontal) grating. All patterns had a mean luminance of 6.0 cd/m^2, and patterns were presented within a circular aperture, centered on the screen, with a 28° diameter. Outside the aperture, the screen was blank, with a luminance of 6.0 cd/m^2.

In some experiments, the two gratings were sinusoidal modulations of luminance, with a spatial frequency (SF) that varied across experiments, and was often different for the two gratings in a plaid. The vertical grating drifted either to the right or to the left, with a temporal frequency (TF) of 20 Hz (48° phase shift across frames), whereas the oblique grating was static. Each grating had a 32% Michelson contrast, unless noted otherwise.

In other experiments, the gratings were 1D RLS. For vertical RLS, each line in a pattern (two-pixel, 0.08°, wide) was assigned randomly either a high or a low luminance value (32% Michelson contrast). An oblique RLS was obtained by generating a vertical RLS and then rotating it using an anti-aliasing algorithm. Motion of the vertical RLS was simulated by shifting, with a speed of 40°/s either to the right or to the left, a stimulus larger than the screen behind a fixed aperture (i.e., the stimulus did not wrap around). The oblique RLS was either static or different on each frame. In some experiments, the drifting and static RLS were only visible in a subset of the frame sequence.

Data analysis. All the measures reported here are based on either the velocity of the instrumented eye or, when the stimuli were presented dichoptically, on the version velocity (average of the velocity of the two eyes). The calibrated eye position traces (see above, Eye movement recording) were differentiated using a 21-point finite impulse response acausal filter (47 Hz cutoff frequency). Trials with saccadic intrusions and unstable fixation that went undetected at run time were removed by the following automatic procedure aimed at detecting outliers. For each

velocity signal (for monocular recordings: horizontal, vertical; for binocular recordings: left eye horizontal, right eye horizontal, left eye vertical, right eye vertical, horizontal vergence, vertical vergence) at each time point (0–169 ms from stimulus onset, in 1 ms increments), trials for which the velocity deviated more than ±4.5 SDs from the mean (across all the valid trials for a given condition) were excluded. This was repeated iteratively until no trials were excluded. Average temporal profiles, time-locked to stimulus onset, were then computed over the remaining trials, separately for each stimulus condition.

Opposite stimulus motion directions are known to elicit OFRs that are not simply opposites (Boström and Warzecha, 2010; Quaia et al., 2012). This asymmetry is attributable to the presence of a so called default response (Quaia et al., 2012) and is probably partly related to the disengagement of fixation. To remove this component and to increase the signal-to-noise ratio, most ocular following studies report not the raw OFR but rather the difference between the OFRs to opposite motion directions. We did so here as well. The traces and measurements reported here are thus based on the difference between the average response to stimuli containing motion energy in opposite directions.

We used bootstrap-based methods (Efron, 1982) for all our statistical analyses. For the component of the OFR orthogonal to the drifting grating, the latency was determined as the time at which the response (mean of response to +45° and −45° static grating orientation) became significantly ($p < 0.05$) different from zero. For the components of the OFR parallel to the drifting grating, the latency was determined as the time at which the response to the +45° static grating first differed significantly ($p < 0.05$) from that to the −45° static grating. Bootstrap techniques are used to compute these measures and to compute confidence intervals around the measures themselves. A description of the procedures used has been published in detail previously (Quaia et al., 2012).

The direction of eye movements. Most perceptual and neurophysiological experiments that investigated the mechanism(s) that extracts pattern motion direction from plaids naturally described their observations in terms of motion direction (either perceived or neurally coded). We have not described the output of the ocular following system in terms of the direction of the eye movement. We have instead used stimuli in which component motion signals are all horizontal, so that any vertical eye movement directly indicates a response to pattern motion. The reason for doing so is that, in the ocular following system, the visual motion signals present in area MT and the medial superior temporal (MST) area are processed by a premotor circuitry, which then innervates the eye muscles (Takemura et al., 2007; Miles and Sheliga, 2010; Masson and Perrinet, 2012). This operates in a Cartesian coordinate frame, because the eye muscles are organized in approximately horizontal and vertical planes. Unfortunately, the gain of this sensory-to-motor transformation for the OFRs is typically not the same for vertical and horizontal movements (and these differences are idiosyncratic). As a consequence, the magnitude of the OFRs induced by the same stimulus drifting horizontally or vertically can be very different. In our subjects, J.H. had stronger horizontal than vertical movements, C.Q. had the opposite, and B.M.S. was fairly balanced. Because of this unbalance, the direction of oblique movements can be misleading (Blum and Price, 2014).

To verify that this is indeed what makes eye movement direction measures unreliable, we took the subject that had the largest unbalance (C.Q.) and recorded his OFRs in response to two unikinetic plaids. In both cases, the static RLS was oblique (45°), but the drifting RLS could be either vertical or horizontal. Theoretically, the motion direction for these two plaids should be the same (45°). A directional bias in the visual system would cause a different population of pattern motion neurons to be activated, resulting in different eye movement directions. With a premotor horizontal/vertical bias, the same population of pattern motion neurons would be activated, resulting in the same eye movement direction (not necessarily 45°). We found that the movement directions were very similar with the two plaids (data not shown), supporting the view that the same oblique motion signal was fed to a premotor stage with different gains for vertical and horizontal movement.

Perception with noise plaids. The plaid stimuli that we have used in our experiments are quite different from those used normally in perceptual studies of pattern motion perception. They are much larger (28° in di-

ameter) and are presented for much shorter intervals (160 ms). In the case of sine plaids, they also have much lower SF (0.25 cpd) and higher TF (20 Hz). As far as we know, neither noise plaids nor flicker plaids have been used previously. Accordingly, there are no perceptual studies on which to rely, and we have not systematically measured perceived motion direction with any of our stimuli. Qualitatively, unikinetic sine plaids appear coherent (i.e., moving as a single plaid in a single direction), but all others do not (at least at the short exposures we used). Asking a subject to report perceived motion direction under conditions of poor coherence is problematic, and we have not attempted it. However, after informal questioning, all subjects reported that, with flicker plaids, they clearly perceived vertical and oblique gratings (i.e., transparency), but the vertical grating did not move horizontally. Rather, they perceived motion in an oblique direction. When the oblique grating was presented only for a very short time (<20 ms), some subjects perceived it only in a limited region of the visual field, and one did not perceive its presence at all. Nevertheless, all perceived oblique motion of the vertical grating. With dichoptic plaids, each subject perceived a (quite idiosyncratic) patchwork, with some areas of the visual fields dominated by the stimulus seen by one eye, others by the stimulus seen by the other eye, and still others in which the stimuli coexisted.

Models and simulations. As noted in Introduction, several models have been proposed to extract pattern motion from plaid stimuli. Some of them rely on extracting features from the images and tracking their motion across frames. As explained in Results, because our flicker plaids have been designed specifically to remove all features from the images, those models will not, by definition, respond to our stimuli. Determining how models inspired by the IOC would respond to our stimuli is instead less straightforward and, as we will show, depends to some extent on model parameters. To better illustrate these aspects, we thus simulated how two models respond to our stimuli.

The first model we used is an implementation of the IOC rule. The IOC is a graphical construct (Fig. 1). It is based on the orientation and signed speed of the 1D gratings that compose the stimulus (their contrast and SF play no role). Constraint lines for each grating are determined from the respective speed vectors, and their intersection determines pattern speed. We are not aware of any proposed implementation of the IOC mechanism that can be applied directly to image sequences, so we devised an algorithm that represents a straightforward implementation of the graphical construct. Assuming continuous motion, all the information needed is contained in the first two frames. First, each orientation present in the stimulus is extracted from each frame. Because we built the images ourselves, we directly took the two individual RLS that were then summed to make the plaid. If these are not available, one would have to Fourier transform each frame, convert the horizontal and vertical SF coordinates into polar coordinates (or, alternatively, compute the Fourier transform directly in polar coordinates), apply, for each orientation of interest, a narrow-band filter in the orientation domain to the amplitude spectrum, and finally reconstruct the components through an inverse Fourier transform. With sinusoidal plaids, one can also directly filter in the image domain by convolving the images with oriented Gabor filters. Next, for each orientation, the two filtered frames are crosscorrelated. This will typically reveal a ridge of high cross-correlation, corresponding to the displacement of the filtered stimulus across the two frames (a proxy for speed). This ridge, parallel to the orientation of the filter, represents the constraint line for that orientation. Finally, the cross-correlation images for all orientations present in the stimulus are multiplied. The location of the peak represent the intersection of the constraint lines, i.e., the IOC solution. Because any sequence of noise patterns will contain some spurious correlations, we average the crosscorrelations over 10 randomly generated sequences and only plot the central part of the cross-correlation matrix.

The second model we used is an implementation of the model by Simoncelli and Heeger (1998) (S&H), based on the experimental results and simulations of Nishimoto and Gallant (2011) (N&G). We considered using directly the reference implementation of the S&H model (made openly available by Dr. Simoncelli on his institutional website) but decided against it for two reasons. First, the weights between V1 and MT neurons in the original S&H model were proposed on a theoretical basis; Nishimoto and Gallant (N&G) have instead inferred those weights by fitting them to the response of real MT neurons to naturalistic movies. The weights so found are thus experimentally validated. Second, the reference implementation of the S&H model only samples stimuli at a single scale and thus is not suited for processing broadband noise stimuli. We also considered the model proposed by Rust et al. (2006), but it is not available in a version that can accept arbitrary image streams. Note that, just as in the original S&H model, we only modeled so-called pattern cells, which are a fraction of all MT cells (Movshon et al., 1985; Rodman and Albright, 1989; Kumano and Uka, 2013).

The V1 stage in the S&H model is represented by a population of ideal V1 complex cells, each with a spherical (in spatiotemporal Fourier space) receptive field; the receptive fields of the population tile the entire Fourier space. Because such ideal neurons approximately compute the Fourier energy in the region of spatiotemporal space covered by their receptive field (Adelson and Bergen, 1985; Watson and Ahumada, 1985), this stage can be bypassed, and the model can be implemented directly in Fourier space (at least for spatially and temporally homogeneous stimuli). For the MT stage, we then defined a population of MT neurons, each characterized by a preferred velocity vector; we tiled the velocity space evenly, covering the range between +12 and −12 pixels/frame. For each neuron, we then defined the regions of the Fourier space that excite and inhibit the cell, based on the results of N&G. S&H suggested that the excitatory region should lie on a plane in Fourier space, the so-called velocity plane. This is the plane on which all the oriented gratings that are compatible with motion in the preferred direction and at the preferred speed of the MT cell lie. Note that this plane will contain the grating orthogonal to the preferred velocity vector of the MT cell and moving at its preferred speed, but also gratings that have different orientations and move, in the direction orthogonal to their orientation, at a lower speed (Fig. 1B) than the preferred speed of the MT neuron (Simoncelli and Heeger, 1998; Okamoto et al., 1999). If we define the velocity of interest (i.e., the preferred velocity of the model pattern cell under examination) as a 2D vector, $\mathbf{v} = [v_x, v_y]$, then the velocity plane, in Fourier space, is described by the following equation (Watson and Ahumada, 1985; Grzywacz and Yuille, 1990; Simoncelli and Heeger, 1998):

$$TF = -v_x SF_x - v_y SF_y,$$

where SF_x is the SF in the x direction, and SF_y is the SF in the y direction (i.e., for a vertical grating, $SF_y = 0$). N&G confirmed that the excitatory inputs to each MT cell indeed sample the Fourier space around the velocity plane. Accordingly, we set the excitatory weights for each MT neuron as decreasing monotonically based on the distance from the velocity plane for the cell. The distance of a generic point, $P_0 = (sf_x, sf_y, tf)$, from a generic plane, $ASF_x + BSF_y + CTF + D = 0$, is equal to the following:

$$d(sf_x, sf_y, tf) = \frac{|Asf_x + Bsf_y + Ctf + D|}{\sqrt{A^2 + B^2 + C^2}}$$

In our specific case, $A = v_x$, $B = v_y$, $C = 1$, and $D = 0$. Having so determined d for each point in Fourier space, we then set the excitatory weight from that point in Fourier space to our MT neuron according to the following equation:

$$E_w(sf_x, sf_y, tf) = e^{-\frac{d(sf_x, sf_y, tf)^2}{2\sigma^2}},$$

with $\sigma = 0.025$ (note that all three coordinates in Fourier space vary between −0.5 and 0.5, the Nyquist limit).

Although excitatory connections are obviously crucial, inhibition plays a key role in how MT pattern cells behave (Movshon et al., 1985; Simoncelli and Heeger, 1998; Rust et al., 2006; Perrone and Krauzlis, 2008a; Tsui et al., 2010; Nishimoto and Gallant, 2011; Solomon et al., 2011; Perrone and Krauzlis, 2014). S&H proposed that any region of the Fourier space that does not excite an MT cell, inhibits it, and that the algebraic sum of all weights is zero. However, N&G found that the inhibitory region is in fact much more concentrated and mostly lies around the opposite-velocity plane. This implies that a stimulus moving in the di-

rection opposite to that preferred by the neuron inhibits the cell much more than a stimulus moving in an orthogonal direction: motion opponency (Adelson and Bergen, 1985). We used this principle to determine the inhibitory weights for our MT cells and thus computed the inhibitory weights I_w using the same Gaussian equation (with the same dispersion parameter) used for the excitatory weights but based on the distance from the opposite-velocity plane, which is naturally described by the following equation:

$$TF = v_x \text{SF}_x + v_y \text{SF}_y.$$

Finally, N&G found that the lowest TFs do not contribute very strongly to the response of MT neurons. In the S&H model, a static stimulus parallel to the preferred direction of motion of the MT neuron activates it as strongly as a stimulus that is orthogonal to it and moves at its preferred speed. This is not observed in real MT cells: static parallel stimuli activate MT pattern cells, but they do so not as robustly as moving stimuli (Albright, 1984; Rodman and Albright, 1989; Wallisch and Movshon, 2011; Khawaja et al., 2013). To account for these findings, we reduced the weights for low TFs according to the following equations:

$$E'_w(\text{sf}_x, \text{sf}_y, \text{tf}) = E_w(\text{sf}_x, \text{sf}_y, \text{tf}) \left(1 - K e^{-\frac{\text{tf}^2}{2\sigma^2}} \right)$$

$$I'_w(\text{sf}_x, \text{sf}_y, \text{tf}) = I_w(\text{sf}_x, \text{sf}_y, \text{tf}) \left(1 - e^{-\frac{\text{tf}^2}{2\sigma^2}} \right), \qquad (1)$$

with $\sigma = 0.025$ and $K = 0.5$. The resulting pattern of excitatory and inhibitory connections matches quite well the behavior of MT pattern cells, as recorded by N&G (their Figs. 4A, 5A). Importantly, it provides some overall excitatory input from low TF components, necessary for MT neurons to exhibit a (weak but nonzero) response to stationary stimuli. In one model simulation, in which the aim was to test the behavior of the model without excitation or inhibition from this region, we set $K = 1$ in Equation 1.

The weight matrix was then computed by subtracting the inhibitory weights from the excitatory ones:

$$W(\text{sf}_x, \text{sf}_y, \text{tf}) = E'_w(\text{sf}_x, \text{sf}_y, \text{tf}) - I'_w(\text{sf}_x, \text{sf}_y, \text{tf}).$$

To introduce the nonlinear operators at the level of V1 neurons that have been shown to be crucial to reproduce the activity of MT neurons (Simoncelli and Heeger, 1998; Rust et al., 2006; Nishimoto and Gallant, 2011), we added a static nonlinearity (expansive or compressive) to modulate the Fourier energy at each spatiotemporal frequency. Within the context of these models, this is approximately equivalent to adding the nonlinearity to the output of V1 complex cells projecting to area MT, which in turn is functionally similar to the tuned normalization mechanism invoked by others (Simoncelli and Heeger, 1998; Rust et al., 2006). Note that we did not include the untuned normalization (i.e., a contrast gain control mechanism) present in the N&G model, because it only improved slightly their fits and was not critical for the simulations presented here.

The response of each MT pattern cell in our implementation of the S&H model was then as follows:

$$R = \left\lfloor \sum_{\text{SF}_x} \sum_{\text{SF}_y} \sum_{\text{TF}} W(\text{sf}_x, \text{sf}_y, \text{tf}) \left| F(\text{sf}_x, \text{sf}_y, \text{tf}) \right|^{2\alpha} \right\rfloor$$

where $\lfloor \ \rfloor$ indicates half-rectification (i.e., negative values are set to zero), $\mathbf{F}$ is the 3D Fourier transform, and α is the exponent of the static nonlinearity mentioned above.

We made no attempt to readout from this population an overall velocity signal, which could then be used to drive perception or eye movements. The readout problem was left open by the S&H model and its derivations, and our intent here was to test how these models respond to our novel stimuli, not to further extend the models themselves.

Numerical simulations of the models were performed using custom Python scripts, executed under the Canopy Python Distribution (Enthought).

Results

Using eye movements to study pattern motion extraction

In unikinetic plaids (Fig. 1E), the pattern moves parallel to the orientation of the static grating, with a speed and sign that are determined by the speed and direction of the drifting grating (Fig. 1G). For example, with the plaid shown in Figure 1E, if the vertical grating drifts to the right at 80°/s, the pattern moves up and to the right at 113.1°/s.

The ultrashort latency reflexive ocular following system can be used to great effect to study how the visual system extracts a pattern motion signal from such plaids. As shown previously by Masson and colleagues (Masson and Castet, 2002; Masson, 2004; Barthélemy et al., 2008; Barthélemy et al., 2010), when presented with such a stimulus, the eyes first move in the direction orthogonal to the drifting grating (i.e., horizontally with the stimulus shown in Fig. 1E), but after a delay of ∼15 ms, the eyes start tracking the pattern.

In Figure 2A, we show the OFRs induced by two unikinetic sinusoidal plaids in one of our subjects. In both stimuli, the vertical grating drifted to the right (at 20 Hz), but the oblique grating was tilted either +45° (thick lines) or −45° (thin lines) off the horizontal. The horizontal eye traces (gray) were very similar in the two cases; the vertical eye traces (black) were instead approximately opposite in the two cases, following the orientation of the oblique grating (and thus of the pattern motion direction). This is an invariable finding with such symmetric stimulus configurations. In all the experiments described below, we used both of these stimulus configurations (i.e., a drifting vertical grating paired with an oblique grating tilted either +45° or −45°). However, for clarity of presentation, from now on, we plot (Fig. 2B) the so-called folded responses (Rambold and Miles, 2008):

$$\left(\dot{H} = \frac{\dot{H}_{+45} + \dot{H}_{-45}}{2} \quad \dot{V} = \frac{\dot{V}_{+45} - \dot{V}_{-45}}{2} \right)$$

where $\dot{H}$ and $\dot{V}$ represent the horizontal and vertical components of the eye velocity, respectively. The vertical component is the signature of the pattern motion computation and is the focus of our study.

Figure 2 highlights a critical advantage of measuring eye movements over collecting perceptual reports: the evolution of responses can be observed with high temporal resolution. As noted by Masson and colleagues, this allows to easily identify the presence of two signals: an early one related to the component motion and a significantly delayed one (in this case by 21 ms, $p = 0.001$, unpaired t test) related to pattern motion. However, it is also important to stress that ocular following eye movements and perception often behave differently in response to motion stimuli, with perception being more sensitive to some stimuli and eye movements to others (Boström and Warzecha, 2010; Simoncini et al., 2012; Blum and Price, 2014; Glasser and Tadin, 2014; Price and Blum, 2014). It is generally agreed that eye movements are less sensitive to cognitive expectations and thus reflect low-level computations more faithfully (Spering and Carrasco, 2015), but much remains to be understood. Accordingly, it must be recognized that the neural computations that we uncover here are not necessarily accessible to perception, and their effect on perceptual processes would need to be assessed separately. A brief, qualitative description of how our subjects perceived the stimuli used in this study is provided in Materials and Methods (see Perception with noise plaids).

Pattern motion extraction with RLS plaids

As mentioned in Introduction, besides the IOC, there are several mechanisms that have been proposed to extract the pattern motion signal from plaids. For example, the visual system might do the following: (1) track the 2D features of maximum/ minimum luminance ("blobs") in the patterns (Derrington and Badcock, 1992; Alais et al., 1994; Masson and Castet, 2002; Barthélemy et al., 2008; Dimova and Denham, 2010); (2) track the intersections between the zero-crossing (or mean luminance) lines of each grating (Bowns, 1996, 2011, 2013); (3) extract the motion of contrast modulations in the pattern (Wilson et al., 1992); and (4) respond to the local orientation of 2D features (Dimova and Denham, 2010). With unikinetic sinusoidal plaids, the IOC as well as mechanisms that track either the high and low luminance "blobs" or the intersections of the zero-crossings/mean-luminance lines (provided that the interframe phase shift of the moving grating is <90°, because such lines have twice the SF of the gratings) would all work. A mechanism based on contrast modulations (often referred to as second-order) could also work: alternate bands of high and low contrast, tilted 22.5° clockwise from horizontal, are clearly visible in Figure 1E (right). However, this mechanism would only provide an approximate solution, because the orientation of the contrast modulations (and hence their direction of motion) varies as a function of the relative SF of the two gratings (if the gratings have the same SF, as in Fig. 1E, the direction of motion of the contrast modulation differs from that of the pattern by 22.5°). Finally, the 2D orientation of the local features (blobs) is shifted, relative to the direction orthogonal to the moving grating, in the wrong direction (in Fig. 1E clockwise from vertical) and can be excluded. Thus, there are at least four mechanisms (IOC, tracking of blobs, zero crossings, or contrast modulations) that in principle can, at least approximately, extract pattern motion direction in unikinetic sine plaids and account for the eye movements shown in Figure 2.

A unikinetic sinusoidal plaid thus has limited value in discriminating between an IOC mechanism and the alternatives, as recognized previously (Gorea and Lorenceau, 1991). However, some simple stimulus manipulations can lead to a much more powerful tool. Replacing sinusoidal gratings with RLS eliminates from the plaid all preattentively selectable (i.e., salient) features, produces a large number of zero-crossings/mean-luminance lines, making them effectively untrackable, and eliminates any contrast modulations (Fig. 1F). An IOC operator would still succeed. To demonstrate this, in Figure 3 (top row), we show simulations of an implementation of the IOC that is image computable (see Materials and Methods, Models and simulations) based on cross-correlating orientation-filtered images. From left to right, we show a frame in the stimulus, the next frame, the constraint line for the vertical component (extracted by cross-correlation of orientation-filtered images), the constraint line for the oblique component, and finally the product of the last two images: the intersection of the constraint lines. The pattern motion direction (up and to the right) is correctly extracted. In Figure 4, we show the OFRs induced in three human subjects by such a stimulus (gray lines). Just like with unikinetic
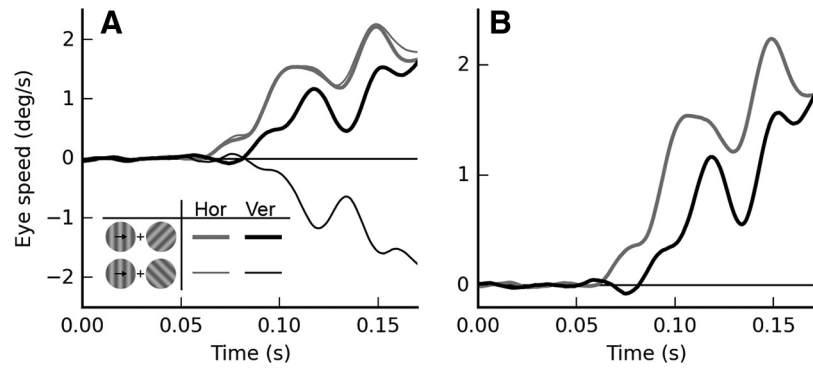

Figure 2. OFRs to unikinetic sine plaids. **A**, When a sinusoidal unikinetic plaid formed by summing a drifting vertical grating and a static oblique grating is presented, a short-latency OFR ensues. Initially, the OFR is driven by the drifting grating only, and the eye velocity is horizontal (Hor; gray lines). After 15–20 ms, a vertical component of the eye movements emerges (Ver; black lines). Notably, this component has a different sign depending on the orientation of the static grating (thick vs thin lines). **B**, Because the horizontal component of the eye movement is the same regardless of the orientation of the static grating and the vertical component is very nearly opposite, the data can be summarized by plotting the semi-sum of the former and the semi-difference of the latter: the folded OFRs. Subject B.M.S.
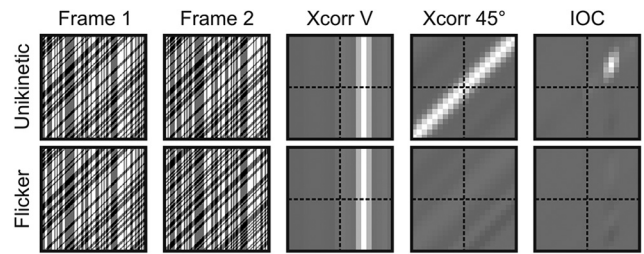

Figure 3. Computing the IOC with noise plaids. Expected response of an IOC operator to unikinetic (top row) and flicker (bottom row) noise plaids. The IOC operator we implemented infers speed by extracting the displacement over two frames; two frames of each pattern are shown in the first two columns. The third and fourth columns show the cross-correlation between the two frames after filtering the images with vertical or oblique-preferring filters, respectively. This operation extracts the constraint lines for the two orientations. The fifth column shows the product of the previous two columns: the IOC. The IOC extracts a pattern velocity signal for unikinetic, but not flicker, plaids. Therefore, the response to flicker plaids should be driven only by component-selective processes.

sinusoidal plaids (Fig. 2), the eyes initially moved horizontally, tracking the motion direction orthogonal to the drifting RLS. After a delay of 10–20 ms (Table 1), a vertical component of the eye velocity emerged, reflecting the extraction of a pattern motion signal. The aforementioned trackable features, absent from these stimuli, are thus not necessary for extracting pattern motion.

A limit of theories of visual processing based on feature detection is that what constitutes a "feature" is often not strictly defined. This makes it difficult to falsify all possible feature-based accounts. For example, although our unikinetic noise plaids do not contain contrast modulations or zero crossings that might be tracked, one might argue that they contain features, because any intersection of the vertical and oblique grating could be considered a feature and thus tracked. There would be countless such features in our stimuli, and they would not be generally salient. Nonetheless, a sufficiently broad definition of feature could in principle impute the responses to these noise patterns to a feature-tracking mechanism. To deal with this possibility, we devised a new stimulus, obtained by summing a vertical drifting RLS and an oblique RLS that is created anew on each frame (dynamic, or temporally uncorrelated, noise pattern). With such a stimulus, which we call a flicker plaid, each frame might contain
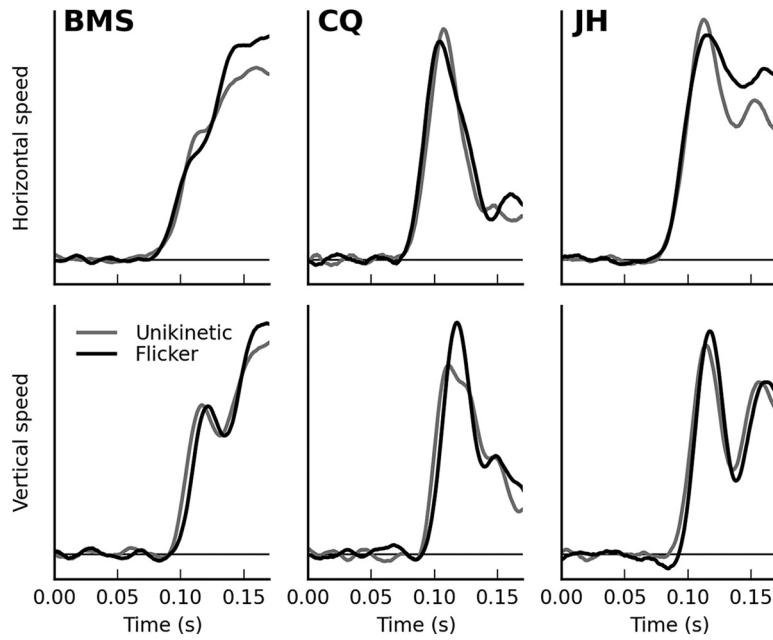
Figure 4. OFRs to unikinetic and flicker noise plaids. RLS unikinetic plaids (gray) and flicker plaids (black) induced very similar OFRs in our three subjects. The horizontal component of eye velocity is shown in the top panels and the vertical in the bottom ones. Folded responses are shown. Latency measures for the two components are reported in Table 1.

Table 1. Latency of OFRs to unikinetic and flicker noise plaids (Fig. 4)

Stimulus	Horizontal	Vertical	Delay	p (delay = 0)
B.M.S.				
Unikinetic	75 (6.02)	94 (1.33)	19 (6.12)	0.001
Flicker	81 (2.53)	97 (1.86)	16 (3.22)	0.001
C.Q.				
Unikinetic	79 (3.60)	91 (0.99)	12 (3.73)	0.001
Flicker	78 (2.13)	94 (1.85)	16 (2.80)	0.002
J.H.				
Unikinetic	80 (2.04)	90 (2.43)	10 (3.18)	0.006
Flicker	79 (3.39)	95 (0.98)	16 (3.49)	0.001

The latencies of the horizontal and vertical components of the OFRs, and the delay of the latter relative to the former, are listed separately for each experimental condition and subject. The values are expressed in milliseconds. SEM estimates for each value are also indicated (in parentheses). The p values for the test that the delay is significantly different from zero based on bootstrapping techniques (see Materials and Methods) is indicated in the last column. Note that, because we used 1000 samples for our bootstrap analysis, we cannot estimate p values <0.001. Hence, we conservatively indicate a value of 0.001 whenever our algorithm returns a p value of 0.

mechanism is responsible for both. The human visual system is thus able to extract a putative pattern motion signal even from stimuli that do not have one and that do not engage an IOC mechanism. Importantly, this response can only arise by combining the signed speed (i.e., motion information) of the moving RLS with the orientation (i.e., form information) of the flickering RLS: a motion-from-form mechanism.

No feature-tracking mechanism can be responsible for the OFRs induced by these stimuli, and neither can an IOC mechanism (Fig. 3). However, as we will now show, a model that has been inspired by the IOC mechanism (Simoncelli and Heeger, 1998) responds similarly to unikinetic and flicker plaids. In Figure 5, we show simulations of our implementation of that model (see Materials and Methods, Models and simulations), based on a recently proposed refinement of it (Nishimoto and Gallant, 2011). Each panel shows how a population of model MT pattern cells, each tuned to a different preferred horizontal and vertical speed, responds to unikinetic noise plaids (left column) and flicker plaids (right column). In the top row, we show simulations obtained using the model parameters that have been found to be necessary to best reproduce the response of MT cells to natural movies (Nishimoto and Gallant, 2011). Clearly, the model responds in a similar manner to unikinetic and flicker plaids, and a pattern motion velocity signal could be easily read out from this population of neurons for both stimuli.

To gain insights into why this model behaves differently from the IOC mechanism on which it is based (Fig. 3), we altered the model in two ways. First, we eliminated the excitatory contributions to the MT pattern cells from the region of Fourier space around zero TF (by setting $K = 1$ in Eq. 1). The results of this manipulation are shown in Figure 5 (middle row): the model now is incapable of extracting a pattern motion signal from either unikinetic or flicker plaids. In both cases, it only responds to the motion of the drifting RLS. Thus, the responses of the S&H model to our stimuli depend on excitatory inputs carrying information about the orientation of static stimuli (in a more neuromimetic implementation of our model, this information would be carried by V1 neurons tuned to zero TF; below, we will use the shorthand "V1 form cells" to refer to such neurons and "V1 motion cells" to describe directionally selective neurons tuned to nonzero TF). Next, with all the excitatory connections back in place, we removed the compressive nonlinearity. The result of this manipulation (Fig. 5, bottom row) is that the model now loses its ability to extract a pattern motion signal from the flicker plaids but not from the unikinetic plaids. The explanation is quite simple. In a unikinetic plaid, the Fourier energy of the oblique static RLS, which is broadband in SF and narrowband in TF, is concentrated around a line on the zero-TF plane, and the excitatory connections from that region activate model MT pattern cells according to the IOC rule. With a flicker plaid, energy from the oblique grating, which is broadband in both SF and TF, is likewise present in this zero-TF region, but it is also scattered in the plane orthogonal to the zero-TF plane that includes it. The Fourier energy
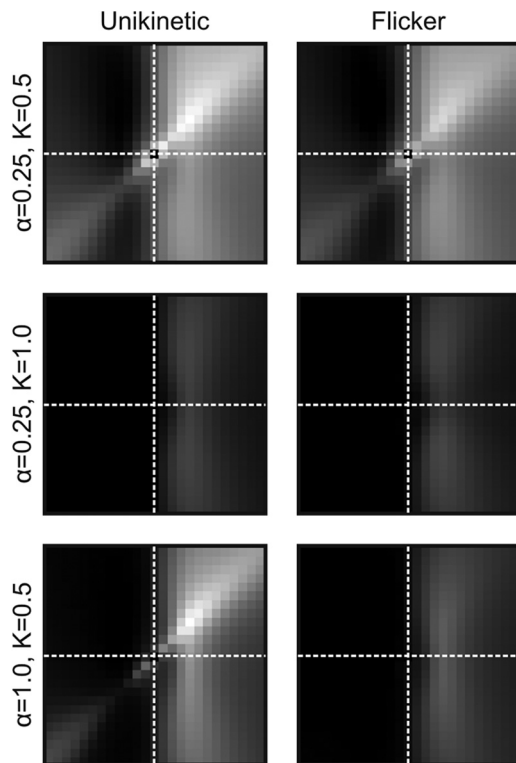
a set of features, but it would be a different set in each frame: no feature, however defined, could thus be tracked.

In a flicker plaid, the dynamic oblique RLS, while maintaining a constant and well defined orientation, has motion energy spread across all speeds, in both directions orthogonal to its orientation. Even an IOC mechanism would thus not be expected to extract a pattern motion signal from such stimulus, because the constraint line for the oblique RLS would be undefined. To demonstrate this, in Figure 3 (bottom row), we show simulations of our IOC implementation applied to a flicker plaid: only the motion of the drifting RLS is extracted by the mechanism. Notice that this is not a failure of the IOC, because in such a stimulus, there is no conserved 2D pattern across frames and thus no "true" pattern motion to be extracted.

In Figure 4, we show the OFRs induced in three human subjects presented with flicker plaids (black lines), in which the oblique grating changed every frame, i.e., every 6.67 ms. In all subjects, both the horizontal (top row) and the vertical (bottom row) components were very similar to those obtained with unikinetic noise plaids (gray lines), and it is thus likely that a single

Figure 5. Model responses to noise plaids. RLS unikinetic plaids (left column) and flicker plaids (right column) were fed to our implementation of a well established model of MT pattern cells. In each plot, we show the response of a population of model MT pattern cells, each tuned to a different preferred speed (horizontal speed on the x-axis, vertical speed on the y-axis). The activity of the cells is described using a grayscale (brighter for stronger activity). The dashed white lines indicate zero horizontal and vertical preferred speed. In all simulations, one noise grating is vertical and drifts to the right, and the other is oblique and is either static or different on each frame. Each row represents a different set of model parameters. In the top row, we used a standard set of parameters, determined on the basis of MT responses to natural movies. Cells that prefer motion up and to the right were activated in both cases. In the middle row, connections to the MT pattern cells from the region of Fourier space around zero TF were removed. The model responded to neither stimulus, only weakly indicating the direction of motion of the drifting grating. In the bottom row, the contribution of the zero-TF region was reintroduced, but the output compressive nonlinearity was removed. The model responded to unikinetic plaids but not to flicker plaids.

around zero TF is thus much smaller with a flicker plaid than with a unikinetic plaid. A strong output compressive nonlinearity reduces the difference between large and small energies, enabling this region of Fourier space to contribute, through its excitatory effect on model pattern cells, even with a flicker plaid.

To summarize, the model responds to unikinetic and flicker plaids because it receives input from the zero TF region, which carries form signals. The strong compressive nonlinearity following the extraction of Fourier energy is important for the response to flicker plaids, because it allows strong responses to weak stimulus energy at zero TF. Form information and tuned normalization are thus necessary and sufficient to extract pattern motion signals from our stimuli. Note that, although we have implemented only a version of the S&H model, other models based on the IOC (embedding a motion-from-form mechanism), such as the one developed by Perrone and colleagues (Perrone, 2004; Perrone and Krauzlis, 2008a,b, 2014), might work just as well.

In an effort to further characterize these mechanisms, which will aid in identifying the underlying neuronal circuits, we now describe OFRs to several additional manipulations of the stimuli described above.

Temporal properties

The OFRs with flicker plaids and our simulations of the S&H model reveal that the mechanism that extracts the orientation of the oblique grating (i.e., form) does not strictly require a static signal. Therefore, it may be possible to activate this orientation-selective mechanism even with briefly presented stimuli. To test this hypothesis, we presented stimuli in which the drifting vertical RLS was present for the entire sequence (24 frames), whereas the dynamic oblique RLS was presented only at the beginning of each sequence, for a variable number of frames. Strictly speaking, we thus briefly presented a flicker plaid, followed by a drifting RLS (see graphic legend in Fig. 6, bottom right corner). We found that (Fig. 6A, bottom row), in all three subjects, presenting the oblique grating on the first frame only (i.e., for 3 ms given the persistence of our monitor) was sufficient to elicit an almost maximal response (dark red). Three frames (green) were as effective as a full presentation (magenta). The latencies of the horizontal and vertical components of the responses were not significantly affected by this manipulation.

To quantify this phenomenon, in Figure 6B, we plot, for all subjects and conditions, the mean vertical velocity in a 30 ms window starting at the onset of the vertical eye deviations (slightly different latency in the three subjects; see figure legend), normalized relative to the same measure for the 24-frame flicker plaid. Note that, for clarity of presentation, we have slightly offset horizontally the data points. For all subjects, in this early time window, the response is not significantly different ($p > 0.05$, unpaired t test) when either 3 or 24 frames of the oblique grating are presented. Unsurprisingly, later, the sustained presence of the flickering oblique grating results in a stronger response (Fig. 6A).

Having established that even a very short presentation of the form signal is sufficient to strongly engage the mechanism, we next attempted to quantify how large a temporal asynchrony between the two signals is tolerated. We presented a series of stimuli in which the oblique grating was presented only for the initial frame, and the vertical moving grating appeared after a variable delay (see bottom right corner in Fig. 7). Because the two components are separated in time, this is technically not a plaid but a pseudo-plaid (Kumbhani et al., 2015). In Figure 7A, we show the OFRs induced by these stimuli, aligned relative to the onset of the vertical (moving) grating. As the onset delay of the vertical grating increased, the magnitude of the horizontal component of the OFRs barely changed (top row). However, the magnitude of the vertical component (bottom row), the signature of the pattern motion computation, dropped monotonically, although by quite different amounts across subjects.

In Figure 7B, we plot the mean vertical velocity in a 45 ms window starting at the onset of the vertical eye deviations, normalized relative to the simultaneous onset condition, together with quadratic fits (see figure legend; again, the data points are slightly offset horizontally for clarity). This measure decreases as the gap increases; a 33 ms onset delay reduced the response by as much as 80% and by as little as 30% across subjects. We do not have a good explanation for such a large intersubject variability, because OFRs are usually highly reproducible across subjects. However, we have observed that subjects differ in how sustained their OFRs are, even in response to simple grating stimuli, and that, in our three subjects, there is a strong inverse correlation between how sustained their OFRs are and how quickly the pattern response decreases with the extent of the onset delay in this experiment.
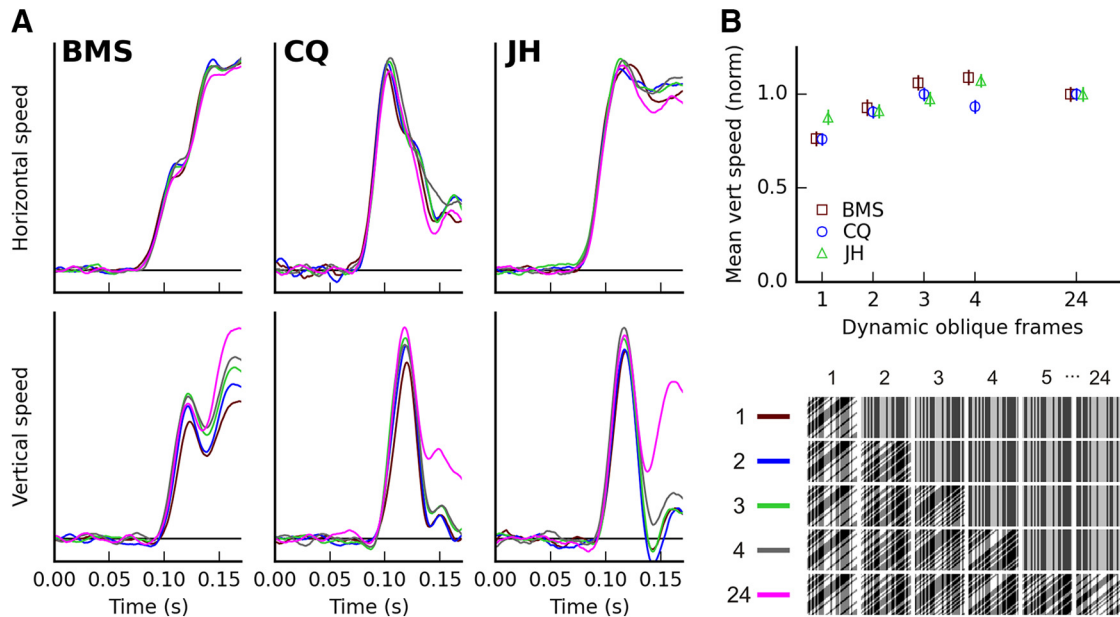
Figure 6. OFRs to flicker plaids with limited presentation of the oblique grating. The stimuli used are represented graphically on the bottom right (actual stimuli were presented in a circular aperture). Each row represents a different stimulus, and each column represents a frame in the stimulus sequence (frames 6 –23 omitted). **A**, The strength of the horizontal component of eye velocity (top row) was very similar regardless of the number of initial frames during which the oblique grating was present. The vertical component (bottom row) increased but only moderately; a single frame of the oblique grating was sufficient to induce a strong response. Folded responses are shown. The latency of the responses (both horizontal and vertical components) was not significantly different in the various conditions. For the three subjects, average latency was 81 ms (B.M.S.), 77 ms (C.Q.), and 78 ms (J.H.) for the horizontal component and 97 ms (B.M.S.), 94 ms (C.Q.), and 96 ms (J.H.) for the vertical component. These latter values are used as starting points for the time windows in which vertical speed is averaged in **B**. **B**, Mean vertical speed from the data plotted in **A** in a 30 ms time window starting at the onset of the vertical component. This short time window is used because late in the movement, after the visual loop has closed, the absence of the oblique RLS affects the response. The measure is normalized, separately for each subject, relative to the response obtained with a flicker plaid in which the oblique RLS is present for all 24 frames. Note that the x-axis is categorical, not linear.
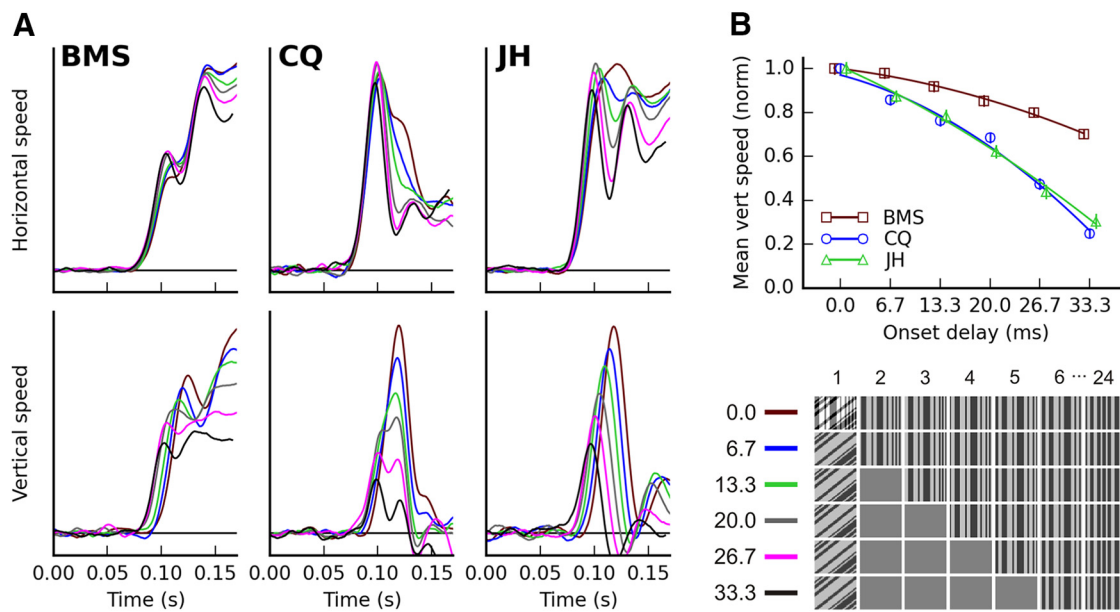

Figure 7. OFRs to flicker plaids with temporal asynchrony between oblique and vertical gratings. The stimuli used are represented graphically on the bottom right. Same format as in Figure 6. **A**, The OFR responses are aligned relative to the time at which the vertical grating appeared (time 0). The horizontal component of the eye velocity (top row) is initially only marginally affected by the delay between the presentation of the oblique grating (present on the first frame only) and the onset of the vertical moving grating. The latency is not affected at all and is thus time-locked to the onset of motion. However, the vertical component (bottom row) is strongly affected, decreasing in both magnitude and latency (again, relative to motion onset) as the onset delay increased. The progression is very orderly. Folded responses are shown. **B**, Mean vertical speed from the data plotted in **A** in a 45 ms time window starting at the onset of the vertical component. The measure is normalized, separately for each subject, relative to the response obtained when there is no onset delay. Quadratic fits to the data are also indicated.

Pattern motion signal delay

A striking aspect of the previous experiment is that, although the horizontal eye velocity was time-locked to the onset of motion (i.e., constant latency as plotted in Fig. 7A), the vertical compo-

nent was not. Notably, the longer the onset delay, the shorter the latency of the vertical components (the only exception to this rule is represented by the condition with the longest onset delay in subject C.Q.; this is probably attributable to the small magnitude

of the vertical response in this condition). To better visualize the extent of this effect, in Figure 8, we plot the horizontal and vertical velocity components in the same panel, for the simultaneous onset and the 33.3 ms onset delay conditions (or 26.7 ms in subject C.Q., for the reason indicated above). In all subjects, in the simultaneous onset condition (top row), the vertical component is delayed relative to the horizontal one, as noted above. However, this delay is much smaller (and never significantly different from zero, $p > 0.1$ for all subjects; Table 2) in the delayed onset condition (bottom row). This suggests that the delay between component and pattern responses does not simply reflect the time required to compute pattern motion from the components but might instead reflect, or at least be significantly affected by, a difference in latency between the extraction of form and motion signals.

To confirm this novel finding, we performed two additional experiments. First, we evaluated the effect of the presence of blank frames. We did so by replicating (in two subjects, C.Q. and J.H.) the above experiment, but with the oblique flickering grating visible up until the moving vertical grating appeared. The delay between the horizontal and vertical traces was affected just as in the original configuration (data not shown). In a second experiment with one subject (C.Q.), we attempted to alter the delay by manipulating contrast rather than temporal onset. The stimulus was composed of a vertical drifting sine grating (20 Hz TF, 0.25 cpd SF) and a flickering oblique RLS. We then varied the Michelson contrast of the sine grating between 4 and 32%, while keeping the contrast of the oblique grating constant (32%). These contrast variations were aimed at altering the response latency of the neurons that respond to the motion signal without changing the latency of the neurons that detect the oblique grating. Once again, the large delay (28 ± 5.62 ms, $p = 0.001$) observed when both gratings had 32% contrast (Fig. 9A) all but disappeared (8 ± 8.64 ms, $p = 0.164$) when the drifting grating had 4% contrast (Fig. 9B). Both the horizontal and vertical components are delayed at low compared with high contrast, but the horizontal more so. Note that lowering the contrast of both components in a sine (or noise) plaid would not have yielded this result, and the delay would have been essentially unchanged, in agreement with previous studies (Barthélemy et al., 2008, 2010). Two manipulations were critical for our results. First, to delay only the moving component, only the contrast of the moving grating was lowered. Second, to prevent the high contrast component from suppressing the low contrast component ahead of the motion integration stage (Sheliga et al., 2008; Masson and Perrinet, 2012), a moving sine grating had to be paired with a flickering RLS, yielding a "hybrid plaid."

The model described above uses as input the Fourier energy of the stimulus, a global spatiotemporal measure, and thus cannot be used to simulate the temporal evolution of the pattern motion signal. In a more realistic implementation, whose development is beyond the scope of this study, the presence of the delay with the standard plaids, and its dependence on the contrast and temporal
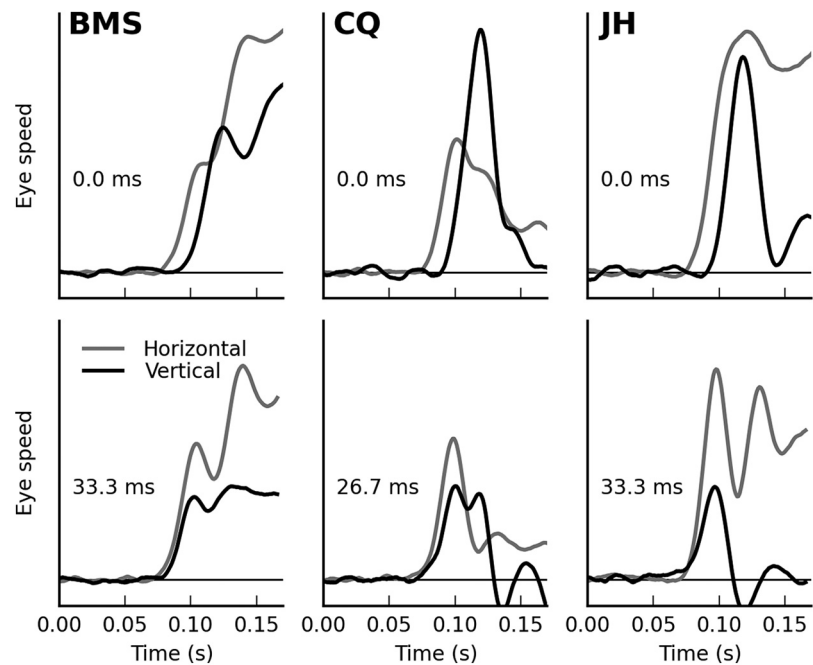

Figure 8. Pattern motion extraction delay for different onset delay conditions. Subset of the traces plotted in Figure 7, displayed in a different format to highlight the changing delay between the horizontal and vertical OFR components (associated with the extraction of component and pattern motion, respectively). In the top row, OFRs under the simultaneous onset condition are shown. In all cases, the vertical component (black) is significantly (Table 2) delayed relative to the horizontal one (gray). In the bottom row, OFRs under the 33.3 ms (B.M.S. and J.H.) or 26.7 ms (C.Q.) onset delay conditions are shown. The delay between the onset of the two components is smaller and not statistically significant.

Table 2. Latency of OFRs to flicker noise plaids with asynchronous components (Fig. 8)

Stimulus	Horizontal	Vertical	Delay	p (delay = 0)
B.M.S.				
0.0 ms	79 (1.26)	93 (3.29)	14 (3.56)	0.010
33.3 ms	75.67 (2.20)	81.67 (3.75)	6 (4.37)	0.146
C.Q.				
0.0 ms	77 (1.71)	93 (1.56)	16 (2.31)	0.002
26.7 ms	69.33 (5.26)	76.33 (3.69)	7 (6.48)	0.110
J.H.				
0.0 ms	77 (1.43)	94 (1.47)	17 (2.02)	0.001
33.3 ms	75.67 (1.91)	73.67 (8.91)	−2 (9.05)	0.664

The latencies of the horizontal and vertical components of the OFRs and the delay of the latter relative to the former are listed for each subject and for two gap conditions. The values are expressed in milliseconds. SEM estimates for each value are also indicated (in parentheses). The p values for the test that the delay is significantly different from zero based on bootstrapping techniques (see Materials and Methods) is indicated in the last column.

offset of the two gratings, could be reproduced if the output of V1 form cells was delayed (or had a longer latency) relative to that of V1 motion cells. Manipulations that selectively delay the output of motion cells relative to that of form cells would then be expected to reduce, and eventually abolish, this delay, as we observed. It is worth noting that similar delays have been reported in most studies of pattern motion computations (Yo and Wilson, 1992; Bowns, 1996; Pack and Born, 2001; Masson and Castet, 2002; Barthélemy et al., 2008; Smith et al., 2010; Tlapale et al., 2010; Beck and Neumann, 2011; Perrone and Krauzlis, 2014). Obviously, our variable delays are specific to our stimuli and to the modality (reflexive eye movements) we investigated. Nevertheless, our findings highlight the need to more carefully investigate the nature of such delays and, in particular, to verify how robust they are to various stimulus manipulations.
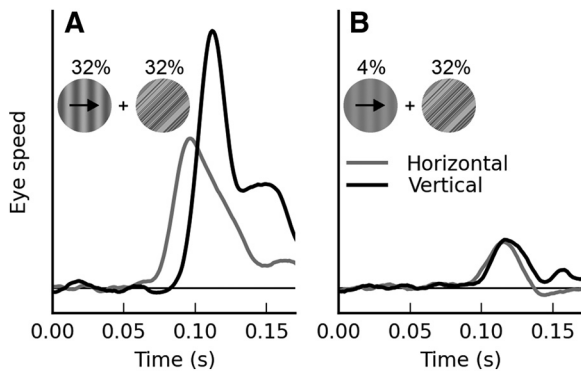
Figure 9. Pattern motion extraction delay is affected by the contrast of the moving grating. Here, one subject (C.Q.) is presented with the sum of a vertical drifting sinusoidal grating and an oblique flickering RLS. ***A***, Both gratings have a Michelson contrast of 32%, and the vertical component is significantly delayed relative to the horizontal one. ***B***, The contrast of the sine grating has been lowered to 4%, whereas that of the oblique grating is still 32%. As a result, both components are weaker and delayed. However, the horizontal component is delayed much more, so that the two components have now essentially the same latency. Folded responses are shown.

Spatial properties

When Masson and Castet (2002), in their seminal work on OFRs to unikinetic sine plaids, paired a drifting grating (6 Hz TF, 0.25 cpd SF) with static oblique gratings of different SFs, they found that the strongest pattern motion response was observed when the static oblique grating had an SF that was two to three times higher than that of the moving grating. Given the low TF used in their experiments, pattern motion extraction might have been based on feature-tracking mechanisms, an explanation that they favored. We were interested in verifying whether a similar tuning also applies to the IOC-based pattern motion mechanism that appears to underlie the OFRs to our noise plaids and whether it varies with the SF of the drifting grating. In an effort to minimize the potential effect of mechanisms that track features, zero crossings/mean luminance and/or contrast modulations, we used unikinetic sine plaids in which the drifting grating (and thus all the potential features) has a high TF (20 Hz). Because no tracking mechanism has been ever reported to operate above 10 Hz (Lu and Sperling, 2001), this stimulus should not engage any such mechanism.

In Figure 10, we plot, for each pair of SFs tested, the mean vertical component of the OFR in a 50 ms time window starting 80 ms after stimulus onset, together with log-Gaussian fits for each SF of the moving grating tested (parameters for the fits are listed in Table 3). We replicated Masson and Castet's original finding with the 0.25 cpd drifting grating (blue line), but we additionally discovered that the strength of the pattern motion signal is barely affected by the SF of the moving grating and is instead mostly determined by the SF of the oblique grating. The most effective SF of the static grating is thus approximately constant (mean and SEM for the location of the fit peak are plotted above each graph, color-coded for each SF of the moving grating), not a constant multiple of the SF of the moving grating. It is also obviously not equal to the SF of the moving grating (which varies much more; see bars at the bottom of each graph), as would have been expected on the basis of perceptual coherence experiments (Adelson and Movshon, 1982; Movshon et al., 1985; Clifford and Vaina, 1999).

To account for this type of behavior, the model described above would have to be modified by restricting the near-zero TF

region projecting to MT cells to a limited band of SFs, excluding the lowest SFs. This would require the MT cells to sample energy in the 3D Fourier space not across the entire velocity plane but rather in a toroiodal (i.e., doughnut-like) region coplanar with the velocity plane. At least some MT cells do in fact appear to sample inputs in exactly this way (Nishimoto and Gallant, 2011).

Another question regarding the spatial properties of this mechanism is whether the two gratings in the plaid must occupy the same region of the visual field. Recent results in anesthetized monkeys revealed that, with type I plaids, pattern neurons in area MT integrate the motion signals from the two gratings only if they are physically superimposed or separated by a small distance (comparable with the size of V1 receptive fields): falling within the receptive field of the MT pattern cell being recorded is not sufficient (Majaj et al., 2007; Kumbhani et al., 2015). Based on these results, Movshon and colleagues proposed that a pattern motion signal is computed only for gratings that are overlapping, or nearly so.

To verify whether this hypothesis holds for unikinetic plaids, we spatially separated the drifting and static gratings. Once again, a question such as this is better addressed using stimuli localized in SF, because it is well known that the size of early cortical filters is directly related to their preferred SF (De Valois et al., 1982; De Valois and De Valois, 1988). Thus, we presented a center-surround configuration, in which a vertical sine grating (0.5 cpd) drifted (at 20 Hz) within a central square aperture (7.5° side), whereas a static oblique sine grating (1.0 cpd) was presented in a surrounding square frame. The frame was 8.4° thick, and a blank (at mean luminance) gap of variable width (between 0.08° and 2.7°) was added between the edges of the central drifting grating and the frame's inner border. Because the two gratings are physically separated in space, these are pseudoplaids.

We found that the magnitude of the horizontal component of the OFRs was only marginally affected by the width of the gap; the magnitude of the vertical component was instead strongly affected. In Figure 11, we plot the mean vertical eye velocity in a 40 ms window starting 90 ms after stimulus onset (the approximate latency in all three subjects in this experiment) as a function of the gap between the central moving grating and the surrounding static grating. Small horizontal offsets have been introduced to avoid overlapping of data points. The measures have been normalized to the value of the same measure obtained when both gratings are superimposed centrally. In all three subjects, the response decreased as the gap increased. In two subjects, the response was very small once the gap was wider than 2° (one wavelength of the moving grating), but in one it was sizeable, albeit small, even then. We fit a line through the data points, with the gap expressed on a logarithmic scale. For each doubling of the gap, the response decreased by 4.8% (B.M.S.), 6.1% (C.Q.), and 8.9% (J.H.). Note that this drop cannot be imputed to a lower effectiveness of the periphery: when we presented both gratings superimposed in the periphery (with the stimulus covering the same area as the static grating did in the largest gap condition), a robust pattern response was induced in all three subjects (data not shown).

Although a physical overlap between the moving and static stimuli in a unikinetic plaid is thus not necessary to compute a pattern motion signal, the separation tolerated is relatively small and compatible with the idea that the detectors of the two gratings must have overlapping receptive fields in early visual cortex (as opposed to area MT).

Binocular properties

Previous evidence with type I sinusoidal plaids (Tailby et al., 2010) indicates that pattern motion computations are partially disrupted when the two gratings in the plaid are presented separately to each eye (dichoptic viewing). More precisely, in anesthetized monkeys, the vast majority of cells in area MT that are pattern selective with monocular plaids lose their selectivity with dichoptic plaids. Similarly, human subjects perceive pattern motion only 30% of the time with dichoptic plaids.

Because OFRs to pattern motion arise even when stimuli do not cohere, they might be a sensitive tool to study pattern motion extraction under dichoptic viewing. Therefore, we presented flicker plaids dichoptically to our subjects and measured their version eye velocity (i.e., mean of left and right eye velocity). We compared two viewing conditions: monocular (one eye sees a flicker plaid and the other eye sees a blank screen, at mean luminance) and dichoptic (one eye sees the drifting vertical RLS and the other eye sees the flickering oblique RLS). Because we found no significant differences in OFRs depending on which eye was seeing what, we averaged the response obtained under the two monocular conditions and separately under the two dichoptic conditions. We found (Fig. 12) a clear horizontal and vertical eye deviation under both conditions, in all three subjects. Initially the OFR was identical in the two conditions; starting ~110 ms (listed as split time in Table 4) after stimulus onset, the OFRs observed under dichoptic and monocular viewing diverged, with the former becoming weaker. This was true for both components but more extensively for the vertical one. In all subjects, the split time was not significantly different for the horizontal and vertical components ($p > 0.1$), although the absolute latencies of the horizontal and vertical components were significantly different ($p < 0.01$) under both viewing conditions. We repeated these experiments with unikinetic noise plaids (in two subjects, C.Q. and B.M.S.) and with unikinetic sine plaids (in all three subjects). The results were identical: initially, the OFR was the same in the monocular and dichoptic conditions, but, in the dichoptic condition, an attenuation quickly ensued (data not shown).

We do not have a simple explanation for the delayed weakening under dichoptic viewing. The most obvious explanation would be that it is the result of binocular rivalry. The delay could be explained if rivalry takes time to ensue. Indeed, this seems to be the case, but previous estimates of rivalry onset delay (150–450 ms) are much longer than the delay we observe (Wolfe, 1983; Blake, 1989). We believe that a more nuanced explanation, requiring additional experimentation, will be needed.

Discussion

The rules by which the brain extracts pattern motion from plaids have been studied previously using sinusoidal plaids. Here we introduced plaids obtained by summing two RLS, which offer distinct advantages. In particular, they do not engage several higher-order mechanisms, thus highlighting the role of more narrowly defined mechanisms.
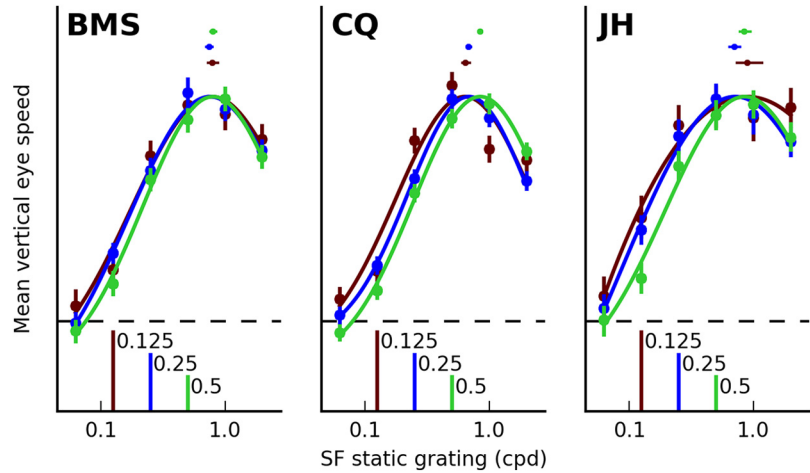

Figure 10. SF tuning of the pattern motion response. Mean vertical speed in a 50 ms time window starting 80 ms after stimulus onset. The stimuli are unikinetic sinusoidal plaids, in which the SF of each grating is varied. The SF of the static oblique grating is indicated on the abscissa, and a separate log-Gaussian fit is plotted for different values of the SF of the drifting grating (indicated by the colored bars at the bottom of the graph). Fits and data are normalized relative to the fit peak to highlight the minimal changes in peak location and bandwidth of the three curves. Values of the (un-normalized) fit parameters are listed in Table 3. The SF (mean and SEM bars) of the static grating associated with peak vertical OFR, for each SF of the moving grating, is indicated using color-coded dots at the top of the graph.

Table 3. Spatial frequency tuning of pattern motion response (Fig. 10)

SF moving	Peak response	SF at Peak	Bandwidth
B.M.S.			
0.125	0.73 (0.03)	0.79 (0.73–0.88)	4.28 (0.38)
0.25	0.97 (0.04)	0.75 (0.70–0.79)	4.09 (0.25)
0.5	0.97 (0.04)	0.80 (0.77–0.85)	3.82 (0.21)
C.Q.			
0.125	0.65 (0.03)	0.65 (0.61–0.70)	3.94 (0.26)
0.25	0.96 (0.03)	0.68 (0.66–0.71)	3.60 (0.15)
0.5	0.94 (0.03)	0.85 (0.82–0.88)	3.74 (0.16)
J.H.			
0.125	0.59 (0.03)	0.89 (0.73–1.17)	5.75 (0.79)
0.25	0.83 (0.04)	0.70 (0.65–0.78)	4.70 (0.36)
0.5	0.87 (0.04)	0.85 (0.78–0.95)	4.21 (0.36)

Parameters for the log-Gaussian fits shown in Figure 10 are listed here. For the peak value of the log-Gaussian function and the bandwidth (full-width at half-height, expressed in octaves) SEM estimates are also listed (in parentheses). For the SF corresponding to the peak, an SEM cannot be indicated (the distribution is symmetric in log space and hence skewed in linear space), and the 68% confidence interval is listed instead.
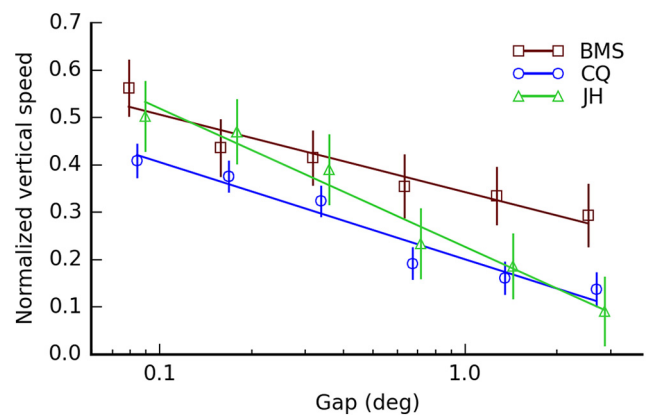

Figure 11. Pattern motion response varies with the spatial separation between oblique and vertical gratings. Mean vertical speed in a 40 ms time window starting 90 ms after stimulus onset (the approximate latency of the response). The measure is normalized, separately for each subject, relative to the response obtained when both gratings are centrally located and fully overlapping. Note that the abscissa is logarithmic. Log-linear fits to the data are also indicated.
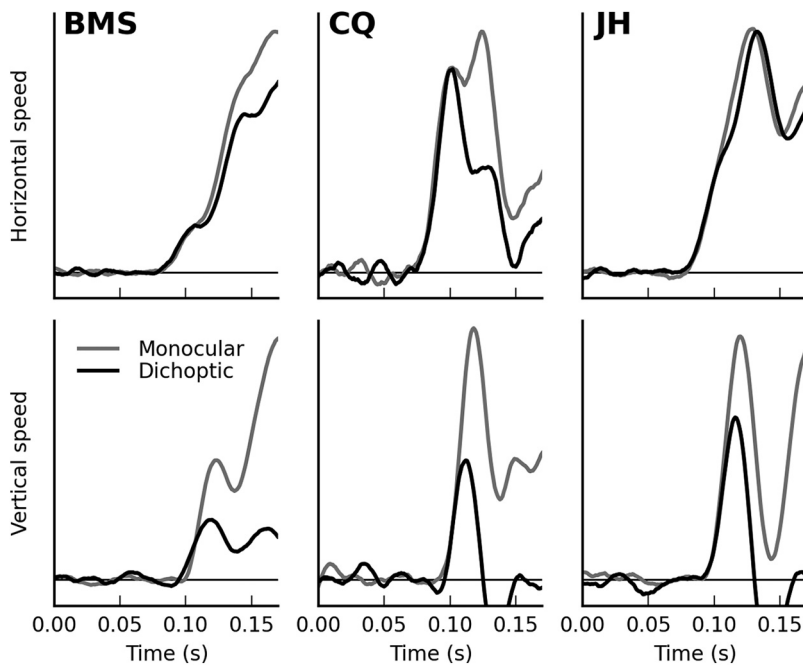
Figure 12. OFRs to monocular and dichoptic flicker noise plaids. Even dichoptically presented flicker plaids induce sizeable pattern motion responses, indicating that the combination of the signals representing the two gratings occurs at a cortical level. For the first 110 ms, the OFRs were identical in the monocular and dichoptic conditions, in both their horizontal (top row) and vertical (bottom row) components. Afterward, the dichoptic viewing condition resulted in weaker responses. Folded responses are shown. Timing statistics are reported in Table 4.

Table 4. Latencies under different viewing conditions (Fig. 12)

Viewing	Latency		Split time	
	Horizontal	Vertical	Horizontal	Vertical
B.M.S.				
Monocular	87 (3.65)	103 (1.71)		
Dichoptic	85 (3.16)	100 (1.78)	117 (4.84)	110 (1.56)
C.Q.				
Monocular	77 (3.36)	98 (2.16)		
Dichoptic	79 (2.57)	99 (3.01)	107 (3.24)	108 (2.52)
J.H.				
Monocular	83 (1.86)	97 (2.50)		
Dichoptic	82 (3.02)	97 (3.05)	108 (10.94)	109 (4.12)

The latencies of the horizontal and vertical components of the OFRs, separately for each subject and for two viewing conditions, are listed. The time at which the horizontal and vertical conditions first differ significantly across the two viewing conditions are also shown. All values are expressed in milliseconds. SEM estimates for each value are indicated (in parentheses).

We found that unikinetic RLS plaids induce robust reflexive OFRs in a direction that reveals the operation of a mechanism for computing pattern motion (Fig. 4). Similar responses were also observed with flicker plaids, implying the operation of a motion-from-form mechanism. Although none of the many feature-tracking mechanisms proposed can account for this type of behavior, we showed that a model based on the IOC mechanism can. Our simulations (Fig. 5) reveal that its ability to do so depends on nonlinear processing of Fourier energy, mediated by a compressive nonlinearity (or by tuned normalization), and on the explicit use of information about the orientation of components at zero TF (motion-from-form). What makes this mechanism compelling is that it correctly predicts responses even for stimuli, such as flicker plaids, that do not contain true pattern motion.

We performed several additional experiments to further characterize the neural mechanism responsible for this computation.

In agreement with recordings from MT pattern cells in monkeys tested with type I plaids (Majaj et al., 2007; Kumbhani et al., 2015), we found that moderate temporal (Fig. 7) and/or spatial (Fig. 11) separation of the gratings are the manipulations most effective at impairing the computation of pattern motion. In contrast, even short presentations of the oblique grating, barely perceived, are sufficient to support the computation (Fig. 6). We also found that the earliest 30 ms of the OFRs elicited by our plaids are just as strong under monocular and dichoptic presentation of the gratings (Fig. 12). The mechanism thus operates binocularly, and motion and orientation signals are combined regardless of eye of origin. Later, responses are attenuated in the dichoptic case, in agreement with previous physiological and perceptual findings that pattern motion signals are disrupted under these conditions (Tailby et al., 2010). This finding highlights the sensitivity of eye movement recordings to neural processes that produce short-lived responses, often hard to capture with other techniques.

Next, we found that the latency difference between component and pattern responses is sensitive to the relative timing (Fig. 8) or contrast (Fig. 9) of the two components. This observation is very important, because it demonstrates that the delay is not simply attributable to the time required to compute the pattern motion signal. Furthermore, it implies that form and motion information are treated differently. This idea is reinforced by our results with sinusoidal plaids in which the static and moving components had different SFs. We found that it is the absolute SF of each component, not the relative SF, that determines the strength of the pattern response (Fig. 10). The motion and form pathways that cooperate to compute pattern motion, at least for the eye movements measured here, have thus distinct SF tuning properties. For the reflexive eye movements we studied, motion information is most effective at SFs ~0.25 cpd (Sheliga et al., 2005; Miura et al., 2006), whereas the sensitivity to form information peaks at ~0.8 cpd (Table 3).

One way to account for these results is to hypothesize that form information is first encoded by V1 sustained (V1s) neurons and motion information by V1 transient (V1t) neurons. The different SF tuning we reported could then be explained if V1s neurons are, on average, tuned to higher SFs than V1t neurons, as reported in cats (Ikeda and Wright, 1975b,a). The temporal dependencies we observed would be reproduced by a model in which the form signal normally arrives in area MT after (i.e., with a longer latency than) the motion signal. There are two ways in which such a delay might arise. First, the V1s neurons might have longer latencies than the V1t neurons, as shown in cats (Ikeda and Wright, 1975a). Second, the form and motion signals might reach MT through different routes, with the one for the form signal being slower. The currently available evidence (Movshon and Newsome, 1996) suggests that V1 neurons projecting directly to area MT tend to be direction selective and thus transient (Hawken et al., 1996). We tentatively suggest that the motion signal might be carried mostly by the V1t–MT pathway. The form

signal might instead be routed through area V2, in which many nondirection-selective V1 neurons project (El-Shamayleh et al., 2013). From there, it might either go to MT directly, as proposed by others to account for the barber-pole illusion (Lidén and Mingolla, 1998), or through area V4, as proposed by Ross et al. (2000). Therefore, the V1s–V2–MT or V1s–V2–V4–MT pathways are both potential candidates for carrying the form signal. This longer routing of the form signal could then contribute to the latency differences discussed above.

Our interpretation of the data presented here has focused mainly on a model of MT pattern cells, but neurophysiological studies describing how they actually respond to unikinetic sinusoidal plaids are also becoming available. Only one study has been published so far, and it has reported that only a small fraction of MT pattern cells are tuned to the pattern motion direction of unikinetic plaids and that this fraction increases significantly in area MST (Khawaja et al., 2013). Thus, it is possible that the pattern motion signal might take shape only in area MST, in which the overall fraction of pattern cells is larger than in area MT (Khawaja et al., 2009). However, another study, which used a more sensitive measure to test the behavior of MT cells to unikinetic plaids, found that most pattern cells in area MT are at least partially affected by the presence of the static grating in a way that is compatible with coding of pattern motion direction (Wallisch and Movshon, 2011). Additional recordings, ideally with patterns such as those used here, will be needed to understand the extent to which real MT cells behave like ideal model MT cells.

The many facets of motion-from-form
The term motion-from-form is used to indicate mechanisms in which form information, typically orientation, affects motion processing. The first concrete proposal that this type of mechanism might play a role in human vision is usually considered the motion streak (or speed lines) mechanism proposed by Geisler and colleagues (Geisler, 1999; Geisler et al., 2001). The basic idea behind this mechanism is that, thanks to their finite temporal integration window, sustained V1 neurons extract the path (i.e., unsigned direction) along which 2D features move. Such a mechanism would work better at high rather than low speed, a prediction experimentally confirmed by Geisler. Since then, other motion-from-form mechanisms have been proposed, and experimentally supported, to account for a number of motion perception phenomena, including the barber-pole illusion (Badcock et al., 2003), and the motion percept that arises with Glass patterns (Ross et al., 2000; Burr and Ross, 2002; Krekelberg et al., 2003; Ross, 2004; Badcock and Dickinson, 2009). Therefore, the presence of motion-from-form mechanisms in the computational repertoire of the visual system is well established and generally accepted.

Although not widely recognized, the S&H model (Simoncelli and Heeger, 1998) also includes a motion-from-form component. In fact, in the reference implementation of the S&H model, each MT pattern cell receives excitatory inputs from four V1 complex cells, and one of these is tuned to the orientation parallel to the preferred direction of the MT cell and to zero TF. This is the key element of the motion streak mechanism proposed by Geisler. Much like the motion streak model, the S&H model also extracts form information even when a static orientation signal is not physically present (e.g., with flicker plaids). This type of behavior is also observed in human OFRs: flicker plaids elicit strong OFRs, and even a single 3 ms flash of the oriented stimulus, perceptually invisible to some of our subjects, is sufficient to produce a strong pattern motion response. However, the role played by form information in the S&H model had never been validated experimentally. By ruling out the potential contribution of mechanisms that track stimulus features, to which previous results with unikinetic plaids could have (and often have) been attributed, our study is the first to do so. This demonstrates that a motion-from-form mechanism plays a role in extracting pattern motion from complex patterns.

References

Adelson EH, Movshon JA (1982) Phenomenal coherence of moving visual patterns. Nature 300:523–525. CrossRef Medline

Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. J Opt Soc Am A 2:284–299. CrossRef Medline

Alais D, Wenderoth P, Burke D (1994) The contribution of one-dimensional motion mechanisms to the perceived direction of drifting plaids and their after effects. Vision Res 34:1823–1834. CrossRef Medline

Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. J Neurophysiol 52:1106–1130. Medline

Badcock DR, Dickinson JE (2009) Second-order orientation cues to the axis of motion. Vision Res 49:407–415. CrossRef Medline

Badcock DR, McKendrick AM, Ma-Wyatt A (2003) Pattern cues disambiguate perceived direction in simple moving stimuli. Vision Res 43:2291–2301. CrossRef Medline

Barthélemy FV, Perrinet LU, Castet E, Masson GS (2008) Dynamics of distributed 1D and 2D motion representations for short-latency ocular following. Vision Res 48:501–522. CrossRef Medline

Barthélemy FV, Fleuriet J, Masson GS (2010) Temporal dynamics of 2D motion integration for ocular following in macaque monkeys. J Neurophysiol 103:1275–1282. CrossRef Medline

Beck C, Neumann H (2011) Combining feature selection and integration–a neural model for MT motion selectivity. PLoS One 6:e21254. CrossRef Medline

Blake R (1989) A neural theory of binocular rivalry. Psychol Rev 96:145–167. CrossRef Medline

Blum J, Price NS (2014) Reflexive tracking eye movements and motion perception: one or two neural populations? J Vis 14:23. CrossRef Medline

Boström KJ, Warzecha AK (2010) Open-loop speed discrimination performance of ocular following response and perception. Vision Res 50:870–882. CrossRef Medline

Bowns L (1996) Evidence for a feature tracking explanation of why type II plaids move in the vector sum direction at short durations. Vision Res 36:3685–3694. CrossRef Medline

Bowns L (2011) Taking the energy out of spatio-temporal energy models of human motion processing: the component level feature model. Vision Res 51:2425–2430. CrossRef Medline

Bowns L (2013) An explanation of why component contrast affects perceived pattern motion. Vision Res 86:1–5. CrossRef Medline

Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10:433–436. CrossRef Medline

Burr DC, Ross J (2002) Direct evidence that "speedlines" influence motion mechanisms. J Neurosci 22:8661–8664. Medline

Clifford CW, Vaina LM (1999) Anomalous perception of coherence and transparency in moving plaid patterns. Brain Res Cogn Brain Res 8:345–353. CrossRef Medline

Collewijn H, van der Mark F, Jansen TC (1975) Precise recording of human eye movements. Vision Res 15:447–450. CrossRef Medline

De Valois RJ, De Valois KK (1988) Spatial vision. Oxford, UK: Oxford UP.

Derrington AM, Badcock DR (1992) Two-stage analysis of the motion of 2-dimensional patterns, what is the first stage? Vision Res 32:691–698. CrossRef Medline

De Valois RL, Albrecht DG, Thorell LG (1982) Spatial frequency selectivity of cells in macaque visual cortex. Vision Res 22:545–559. CrossRef Medline

Dimova KD, Denham MJ (2010) A model of plaid motion perception based on recursive bayesian integration of the 1-D and 2-D motions of plaid features. Vision Res 50:585–597. CrossRef Medline

Efron B (1982) The jackknife, the bootstrap, and other resampling plans. Philadelphia: SIAM.

El-Shamayleh Y, Kumbhani RD, Dhruv NT, Movshon JA (2013) Visual response properties of V1 neurons projecting to V2 in macaque. J Neurosci 33:16594–16605. CrossRef Medline

Fennema CL, Thompson WB (1979) Velocity determinantion in scenes containing several moving images. Comp Graph Image Proc 9:301–315. CrossRef

Geisler WS (1999) Motion streaks provide a spatial code for motion direction. Nature 400:65–69. CrossRef Medline

Geisler WS, Albrecht DG, Crane AM, Stern L (2001) Motion direction signals in the primary visual cortex of cat and monkey. Vis Neurosci 18:501–516. CrossRef Medline

Gellman RS, Carl JR, Miles FA (1990) Short latency ocular-following responses in man. Vis Neurosci 5:107–122. CrossRef Medline

Glasser DM, Tadin D (2014) Modularity in the motion system: independent oculomotor and perceptual processing of brief moving stimuli. J Vis 14:28.

Gorea A, Lorenceau J (1991) Directional performances with moving plaids: component-related and plaid-related processing modes coexist. Spat Vis 5:231–252. CrossRef Medline

Grzywacz NM, Yuille AL (1990) A model for the estimate of local image velocity by cells in the visual cortex. Proc R Soc Lond B Biol Sci 239:129–161. CrossRef Medline

Hawken MJ, Shapley RM, Grosof DH (1996) Temporal-frequency selectivity in monkey visual cortex. Vis Neurosci 13:477–492. CrossRef Medline

Hays AV, Richmond BJ, Optican LM (1982) A UNIX-based multiple process system for real-time data acquisition and control. Presented at the WESCON Conference, El Segundo, CA, September 14–16.

Ikeda H, Wright MJ (1975a) Retinotopic distribution, visual latency and orientation tuning of "sustained" and "transient" cortical neurones in area 17 of the cat. Exp Brain Res 22:385–398.

Ikeda H, Wright MJ (1975b) Spatial and temporal properties of "sustained" and "transient" neurones in area 17 of the cat's visual cortex. Exp Brain Res 22:363–383.

Khawaja FA, Tsui JM, Pack CC (2009) Pattern motion selectivity of spiking outputs and local field potentials in macaque visual cortex. J Neurosci 29:13702–13709. CrossRef Medline

Khawaja FA, Liu LD, Pack CC (2013) Responses of MST neurons to plaid stimuli. J Neurophysiol 110:63–74. CrossRef Medline

Krekelberg B, Dannenberg S, Hoffmann KP, Bremmer F, Ross J (2003) Neural correlates of implied motion. Nature 424:674–677. CrossRef Medline

Kumano H, Uka T (2013) Responses to random dot motion reveal prevalence of pattern-motion selectivity in area MT. J Neurosci 33:15161–15170. CrossRef Medline

Kumbhani RD, El-Shamayleh Y, Movshon JA (2015) Temporal and spatial limits of pattern motion sensitivity in macaque MT neurons. J Neurophysiol 113:1977–1988. Medline

Lidén L, Mingolla E (1998) Monocular occlusion cues alter the influence of terminator motion in the barber pole phenomenon. Vision Res 38:3883–3898. CrossRef Medline

Lu ZL, Sperling G (2001) Three-systems theory of human visual motion perception: review and update. J Opt Soc Am A Opt Image Sci Vis 18:2331–2370. CrossRef Medline

Majaj NJ, Carandini M, Movshon JA (2007) Motion integration by neurons in macaque MT is local, not global. J Neurosci 27:366–370. CrossRef Medline

Masson GS (2004) From 1D to 2D via 3D: dynamics of surface motion segmentation for ocular tracking in primates. J Physiol Paris 98:35–52. CrossRef Medline

Masson GS, Castet E (2002) Parallel motion processing for the initiation of short-latency ocular following in humans. J Neurosci 22:5149–5163. Medline

Masson GS, Perrinet LU (2012) The behavioral receptive field underlying motion integration for primate tracking eye movements. Neurosci Biobehav Rev 36:1–25. CrossRef Medline

Miles FA, Sheliga BM (2010) Motion detection for reflexive tracking. In: Dynamics of visual motion processing: neuronal, behavioral, and computational approaches (Ilg UJ, Masson GS, eds), pp 141–160. New York: Springer.

Miles FA, Kawano K, Optican LM (1986) Short-latency ocular following responses of monkey. I. Dependence on temporospatial properties of visual input. J Neurophysiol 56:1321–1354. Medline

Miura K, Matsuura K, Taki M, Tabata H, Inaba N, Kawano K, Miles FA (2006) The visual motion detectors underlying ocular following responses in monkeys. Vision Res 46:869–878. CrossRef Medline

Movshon JA, Newsome WT (1996) Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. J Neurosci 16:7733–7741. Medline

Movshon, JA, Adelson EH, Gizzi MS, Newsome WT (1985) The analysis of moving visual patterns. In: Pattern recognition mechanisms (Chagas C, Gattass R, Gross C, eds), pp 117–151. Rome: Vatican.

Nishimoto S, Gallant JL (2011) A three-dimensional spatiotemporal receptive field model explains responses of area MT neurons to naturalistic movies. J Neurosci 31:14551–14564. CrossRef Medline

Okamoto H, Kawakami S, Saito H, Hida E, Odajima K, Tamanoi D, Ohno H (1999) Mt neurons in the macaque exhibited two types of bimodal direction tuning as predicted by a model for visual motion detection. Vision Res 39:3465–3479. CrossRef Medline

Pack CC, Born RT (2001) Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. Nature 409:1040–1042. CrossRef Medline

Perrone JA (2004) A visual motion sensor based on the properties of V1 and MT neurons. Vision Res 44:1733–1755. CrossRef Medline

Perrone JA, Krauzlis RJ (2008a) Spatial integration by MT pattern neurons: a closer look at pattern-to-component effects and the role of speed tuning. J Vis 8:1.1–14. CrossRef Medline

Perrone JA, Krauzlis RJ (2008b) Vector subtraction using visual and extraretinal motion signals: a new look at efference copy and corollary discharge theories. J Vis 8:24.1–14. CrossRef Medline

Perrone JA, Krauzlis RJ (2014) Simulating component-to-pattern dynamic effects with a computer model of middle temporal pattern neurons. J Vis 14:19.1–17. CrossRef Medline

Price NS, Blum J (2014) Motion perception correlates with volitional but not reflexive eye movements. Neuroscience 277:435–445. CrossRef Medline

Quaia C, Sheliga BM, Fitzgibbon EJ, Optican LM (2012) Ocular following in humans: spatial properties. J Vis 12:13.1–29. CrossRef Medline

Rambold HA, Miles FA (2008) Human vergence eye movements to oblique disparity stimuli: evidence for an anisotropy favoring horizontal disparities. Vision Res 48:2006–2019. CrossRef Medline

Robinson DA (1963) A method of measuring eye movement using a scleral search coil in a magnetic field. IEEE T Bio-Med Eng 10:137–145.

Rodman HR, Albright TD (1989) Single-unit analysis of pattern-motion selective properties in the middle temporal visual area MT. Exp Brain Res 75:53–64. Medline

Ross J (2004) The perceived direction and speed of global motion in Glass pattern sequences. Vision Res 44:441–448. CrossRef Medline

Ross J, Badcock DR, Hayes A (2000) Coherent global motion in the absence of coherent velocity signals. Curr Biol 10:679–682. CrossRef Medline

Rust NC, Mante V, Simoncelli EP, Movshon JA (2006) How MT cells analyze the motion of visual patterns. Nat Neurosci 9:1421–1431. CrossRef Medline

Sheliga BM, Chen KJ, Fitzgibbon EJ, Miles FA (2005) Initial ocular following in humans: a response to first-order motion energy. Vision Res 45:3307–3321. CrossRef Medline

Sheliga BM, Fitzgibbon EJ, Miles FA (2008) Spatial summation properties of the human ocular following response (OFR): evidence for nonlinearities due to local and global inhibitory interactions. Vision Res 48:1758–1776. CrossRef Medline

Simoncelli EP, Heeger DJ (1998) A model of neuronal responses in visual area MT. Vision Res 38:743–761. CrossRef Medline

Simoncini C, Perrinet LU, Montagnini A, Mamassian P, Masson GS (2012) More is not always better: adaptive gain control explains dissociation between perception and action. Nat Neurosci 15:1596–1603. CrossRef Medline

Smith, MA, Majaj N, Movshon JA (2010) Dynamics of pattern motion computation. In: Dynamics of visual motion processing: neuronal, behavioral, and computational approaches (Ilg UJ, Masson GS, eds), pp 55–72. New York: Springer.

Solomon SS, Tailby C, Gharaei S, Camp AJ, Bourne JA, Solomon SG (2011) Visual motion integration by neurons in the middle temporal area of a

new world monkey, the marmoset. J Physiol 589:5741–5758. CrossRef Medline

Spering M, Carrasco M (2015) Acting without seeing: eye movements reveal visual processing without awareness. Trends Neurosci 38:247–258. CrossRef Medline

Tailby C, Majaj NJ, Movshon JA (2010) Binocular integration of pattern motion signals by MT neurons and by human observers. J Neurosci 30: 7344–7349. CrossRef Medline

Takemura A, Murata Y, Kawano K, Miles FA (2007) Deficits in short-latency tracking eye movements after chemical lesions in monkey cortical areas MT and MST. J Neurosci 27:529–541. CrossRef Medline

Tlapale E, Masson GS, Kornprobst P (2010) Modelling the dynamics of motion integration with a new luminance-gated diffusion mechanism. Vision Res 50:1676–1692. CrossRef Medline

Tsui JM, Hunter JN, Born RT, Pack CC (2010) The role of V1 surround suppression in MT motion integration. J Neurophysiol 103:3123–3138. CrossRef Medline

Wallisch P, Movshon JA (2011) Responses of neurons in macaque mt to unikinetic plaids. Soc Neurosci Abstr 37:577.16.

Watson AB, Ahumada AJ Jr (1985) Model of human visual-motion sensing. J Opt Soc Am A 2:322–341. CrossRef Medline

Wilson HR, Ferrera VP, Yo C (1992) A psychophysically motivated model for two-dimensional motion perception. Vis Neurosci 9:79–97. CrossRef Medline

Wolfe JM (1983) Influence of spatial frequency, luminance, and duration on binocular rivalry and abnormal fusion of briefly presented dichoptic stimuli. Perception 12:447–456. CrossRef Medline

Yo C, Wilson HR (1992) Perceived direction of moving two-dimensional patterns depends on duration, contrast and eccentricity. Vision Res 32: 135–147. CrossRef Medline