How Motion Signals Are Integrated Across Frequencies: Study on Motion Perception and Ocular Following Responses Using Multiple-Slit Stimuli

Ryusuke Hayashi,1,2,3 Yuko Sugita,1 Shin’ya Nishida,4 and Kenji Kawano1

1Department of Integrative Brain Science, Graduate School of Medicine, Kyoto University, Kyoto; 2Laboratory for Integrative Neural Systems, RIKEN, Brain Science Institute, Saitama; 3PRESTO, Japan Science and Technology Agency, Saitama; and 4NTT Communication Science Laboratories, NTT Corporation, Kanagawa, Japan

Submitted 21 January 2009; accepted in final form 30 October 2009

INTRODUCTION

Visual motion is initially detected by a bank of spatiotemporal band-pass filters, such as those mathematically described by the motion energy model, which execute a wavelet-like decomposition of the input image stream (Adelson and Bergen 1985; Watson and Ahumada 1985). Motion signals extracted at different spatiotemporal frequencies, however, can be incongruent with each other in many viewing situations. For example, when a moving object is occluded by other surfaces and visible only intermittently through the gaps of the occluders, the partial lack of visual information causes conflicting motion signals relative to the movement of the overall object, known as aliasing noise. Despite such potential incongruence, the motion system seems to incorporate the motion signals from multiple filters, enabling us to judge the direction of the moving objects and to control our eyes. Although the frequency property of the motion detection process has been investigated mainly by measuring contrast threshold (Burr and Ross 1982; Gomi et al. 2006; Hutchinson and Ledgeway 2006), how the visual system integrates motion signals across different frequencies after detection (i.e., supra-threshold motion signals) remains poorly understood. In the present paper, we describe behavioral experiments that address this problem.

We investigated the human visual motion responses in an idealized situation of occlusion, called multiple-slit viewing (MSV; see Fig. 1A) (Burr 1979; Nishida 2004; Ross 1977). In MSV, the moving pattern is occluded by rectangular masks and visible only through an array of slits. As shown in Fig. 2A, the MSV stimuli have a different distribution of motion energy as a function of the interval between the slits (inter-slit interval). Therefore using such MSV stimuli, we can systematically test how conflicting motion signals affect human visual motion processing.

In addition, to explore human visual motion processing in contexts of more complicated patterns of conflicting motion energy distributions, we developed a novel viewing condition, called contrast-alternating MSV (CA-MSV). The MSV stimuli always contain coherent motion signals distributed along a line through the origin in frequency space, the slope of which corresponds to the velocity of the pattern, regardless of the inter-slit interval. On the other hand, in CA-MSV (Fig. 1B), when a moving pattern is observed through an array of multiple slits, the contrast polarity of the pattern flips as it passes from one slit to the next. The manipulation of the contrast alternation has the effect of canceling the motion energy distributed along the line through the origin (a line of constant object speed; Fig. 2B). Thus the CA-MSV stimuli provide a further variation of motion energy distributions that is potentially useful for investigating the mechanism of cross-frequency motion integration.

We measured direction discrimination performance of human observers with the MSV and CA-MSV stimuli while changing the inter-slit interval and the frame rate of the image stream. In addition to the participants’ reports on the perceived motion direction, we also measured their tracking eye movements, or ocular following response (OFR). OFRs are reflex-like eye movement responses, elicited by abrupt presentation of moving visual patterns with short latencies (<70 ms in monkeys (Miles et al. 1986); <100 ms in humans (Gellman et al. 1990)]. The mechanism for eliciting OFRs is exclusively driven by the luminance-based motion energy of the stimuli.
Hayashi et al. 2008; Masson et al. 2002; Miura et al. 2006; Sheliga et al. 2005). On the other hand, motion perception is also mediated by complex processes known as a second-order or non-Fourier motion system, the processing of which cannot be understood directly in terms of the motion energy of visual inputs (Cavanagh and Mather 1989; Chubb and Sperling 1988; Fleet and Langley 1994; Lu and Sperling 2001; Nishida and Sato 1992; Zanker 1993). The measurement of the OFRs therefore allows us to specifically investigate the inter-frequency integration of the motion energy signals.

Our results show that both the OFR and the perceived motion direction were altered in rather complicated ways. Our question was whether the observed behavior pattern could be explained by changes in the motion energy distribution introduced by the stimulus parameter change.

Here we simply assumed that motion signals were linearly integrated across frequencies with a given weighting function (tuning map) and that behaviors including both OFR and perception were determined by the motion contrast, i.e., the difference in the total motion energy between the opposing motion directions normalized by the sum of opposing motion energy, which is known to be a good predictor of the motion direction perceived for narrow-band stimuli (Georgeson and Scott-Samuel 1999; Rainville et al. 2005). We then estimated the shape of the tuning map that could best account for the behaviors using a two-dimensional (2D) log-scaled Gaussian function.

Our results indicate that this simple model could account for the behaviors elicited by both the MSV and CA-MSV stimuli reasonably well, but only when different tuning maps were separately estimated for the two types of slit stimuli. This supports the validity of our assumption of linear cross-frequency integration, at least as a first approximation, while it also suggests a change in the mechanism of cross-frequency integration depending on stimulus conditions. We next measured the frequency tuning maps using more neutral sinusoidal gratings and compared them with the tuning map estimated with the slit stimuli. The frequency tuning map for the OFR was measured from the response to a suprathreshold drifting grating. A ceiling effect makes it difficult to estimate the suprathreshold frequency tuning map for motion perception using the proportion of correctly perceived directions for a suprathreshold sinusoidal grating. Therefore we generated stimuli that consist of multiple moving gratings of randomly selected frequencies and applied the reverse correlation technique (Ahumada 1996; Jones and Palmer 1987; Neri and Levi 2006; Ringach and Shapley 2004) to investigate how much each frequency contributed to the observed motion perception. The results showed that the measured tuning maps had peaks at low spatial frequency in both the OFR and motion perception. They corresponded very well with the maps estimated with the CA-MSV. On the other hand, different tuning maps estimated with the MSV stimuli suggest that the cross-frequency integration mechanism is modulated by some stimulus structure factors, such as inclusion of constant speed components across frequencies. In light of the obtained results, we will examine how motion energy signals are integrated across frequencies for both suprathreshold motion perception and ocular motor control.

**METHODS**

**Participants**

Apart from one of the authors (R. Hayashi), all participants (total \( n = 6 \)) were naive to the hypotheses and goals of the experiments. Each participant had normal or corrected-to-normal vision. All exper-

---

**FIG. 1.** Stimulus configuration of the multiple-slit viewing (MSV) stimuli (A) and the contrast-alternating MSV (CA-MSV) stimuli (B). A: MSV is a viewing condition in which a moving picture is occluded by masks and visible only through an array of thin slits. Although only a fraction of the pattern is actually present at any instant in time, we can tell which direction the pattern moves behind the slits. The inter-slit interval is defined as the interval between the slits or the width of the mask. B: CA-MSV is a similar viewing condition to the standard MSV except that the contrast of the moving pattern flips from the original to the negative when it passes the slits. If the base pattern is a black ring with a white background, then it appears in its original form on even-numbered slits. When the pattern moves to odd-numbered slits, it changes to a white ring with a black background. The actual MSV and CA-MSV stimuli used in our experiments, however, were random-dot patterns whose contrast ranged from 0 to 100%. See Movie 1 for the demo of the MSV and CA-MSV stimuli consisting of random-dot patterns.
ments were conducted in accordance with the principles embodied in the Declaration of Helsinki (code of ethic of the World Medical Association) under the approval of the ethics committee of Graduate School of Medicine, Kyoto University.

**Apparatus**

Visual stimuli were presented using the Psychophysics Toolbox (Brainard 1997) in the Matlab programming environment on Windows XP. Participants’ head movements were restricted using a bite bar fixed on a head holder located 65 cm away from a 22-in CRT monitor (Iiyama HM204D A). The monitor resolution was set to 640 × 480 pixels, and maximum monitor luminance was 74.66 cd/m². EyeLink 1000 (SR Research) with a sampling frequency of 1,000 Hz was used for the eye movement recording. The communication between the eye tracker and the display PC was established via Ethernet using the Eyelink Toolbox.

**Visual stimulus sequence and task**

At the beginning of each trial, a fixation point (bulls-eye marker, 0.4° diam) was shown at the center of the blank screen (neutral gray, 36.77 cd/m²). After a 500-ms fixation period, the fixation point was removed. Then, a random-dot pattern (240 × 240 pixels, 25 × 25 deg, 10 pixel/°), moving at a speed of 1 pixel shift per frame, was presented for 20 frames under either the MSV or CA-MSV condition (see Movie 1 for the demo of the MSV and CA-MSV stimuli used in our experiments). The contrast of each dot ranged from 0 to 100%, and a new pattern was generated for each trial. The direction of the pattern shift was randomly set to be either upward or downward for each trial. The width of the inter-slit interval was also randomly chosen to be 0–8 pixels for each trial. An inter-slit interval of zero pixels means no mask occluded the moving pattern. The width of each slit was fixed at 1 pixel, and masking bands were uniform neutral gray. The frame rate of the stimuli was set at one of five values (100, 75, 60, 50, 25 Hz) and fixed throughout one session. To investigate the property of cross-frequency integration of luminance-based motion signals, we attempted to minimize the contribution of additional motion mechanisms (e.g., the 2nd-order motion system), the processing of which cannot be understood directly in terms of the luminance motion energy of visual inputs. Because presenting the motion stimuli in the peripheral visual field is known to reduce the strength of second-order motion relative to first-order motion signals (Edwards and Nishida 2004), we presented the stimuli in the periphery with occlusion of the central visual field by a disk of neutral gray (5° radius). (See Supplementary Material1 for results of the condition in which stimuli were presented on the whole screen without the central mask, at a frame rate of 100, 75, and 60 Hz). One session consisted of 360 trials.

# FIG. 2

The Fourier power spectra for the MSV stimuli (A) and the CA-MSV stimuli (B). Each figure shows how the power spectra change as inter-slit interval and the frame rate of the image stream change. Spectra are the average power spectra of 1,000 random-dot textures moving at a constant velocity. A: because a random-dot texture contains broadband frequency components, its power spectrum aligns on a line passing through the origin in frequency space. Viewing a moving pattern through the array of slits corresponds to the sparse sampling of the image stream with its inter-slit interval, thereby generating aliasing of the original power spectrum. In our experiment, the dot pattern was shifted by one pixel every frame. Therefore the change in the frame rate results in the change of the speed, which is represented as the change of the slope of the lines. Power spectra distributed in the 1st and 3rd quadrants represent forward motion, whereas those distributed in the 2nd and 4th quadrants represent opposite-direction motion. The interval of aliasing components, corresponding to the spatial frequency of the inter-slit interval, becomes narrower as the inter-slit interval becomes wider.  

B: the contrast-reversal of the moving pattern at every slit has the effect of canceling the original power spectrum aligned on a line passing through the origin. Instead the contrast-reversing manipulation creates aliasing components flanking the original power spectrum, resulting in the bias of the power spectrum toward the 2nd and 4th quadrants, and generating reversed motion perception. The interval of the aliasing components, then, becomes narrower as the inter-slit interval becomes wider.

1 The online version of this article contains supplemental data.

J Neurophysiol • VOL 103 • JANUARY 2010 • www.jn.org
starting from the response onset of each condition, because it was difficult to reliably specify the response onsets for all the conditions. The use of the fixed time window could underestimate the amplitude under the conditions with delays, although we observed that response onset delays were generally followed by reductions in overall response magnitude.

**Estimation of the spatiotemporal frequency tuning map**

To estimate the weighting function of the inter-frequency integration that explained the behavioral response to the MSV and CA-MSV stimuli, we carried out the following analysis. First, we hypothesized that motion signals were linearly integrated across frequencies with a given weighting function [tuning map, \( \text{Map}(f_x, f_t) \)] where \( f_x \) and \( f_t \) are spatial and temporal frequency, respectively). Second, we assumed that both OFR and the perceptual decision in the motion direction discrimination task is based on a variable known as *motion contrast* because previous psychophysical studies suggested that the motion contrast can account very well for the performance of the direction judgment when the two narrow-band sinusoidal gratings move in opposite directions (Georgeson and Scott-Samuel 1999; Rainville et al. 2005). Motion contrast of the ith condition [\( M_{\text{netcont}}(i) \)] is given by normalizing the opposing energy with the sum of the directional motion energy, as described in Eq. 1.

\[
M_{\text{netcont}}(i) = \frac{\int \int \text{Map}(\pm f_x, \pm f_t) \times E(f_x, f_t, i)df_x df_t - \int \int \text{Map}(\pm f_x, \mu f_t) \times E(f_x, f_t, i)df_x df_t}{\int \int \text{Map}(\pm f_x, \pm f_t) \times E(f_x, f_t, i)df_x df_t + \int \int \text{Map}(\pm f_x, \mu f_t) \times E(f_x, f_t, i)df_x df_t}
\]

Here, \( E(f_x, f_t, i) \) is the distribution of the stimulus’ power spectrum of the ith condition. Motion energy for forward motion is calculated by integrating \( E(f_x, f_t, i) \) in the first and third quadrants of the frequency space with a weighting function of \( \text{Map}(f_x, f_t) \), while motion energy in the opposite direction is calculated by those in the second and fourth quadrants. The power spectrum, \( E(f_x, f_t, i) \), was computed as the average power spectrum of 1,000 random-dot slit stimuli generated using the parameters of the ith stimulus condition (as shown in Fig. 2). Because we computed the correlation of the stimulus spectrum with the observers’ performance averaged over many trials, we did not take into account any variation in the power spectrum among stimuli made with the same parameters.

We assumed that the tuning map [\( \text{Map}(f_x, f_t) \)] can be fitted with the following 2D log-scaled Gaussian function

\[
G(f_x, f_t, c_{fx}, c_{ft}, \sigma_{fx}, \sigma_{ft}) = \exp \left( -\frac{[\log(f_x) - \log(c_{fx})]^2}{2\sigma_{fx}^2} \right) \\
\times \exp \left( -\frac{[\log(f_t) - \log(c_{ft})]^2}{2\sigma_{ft}^2} \right)
\]

\( c_{fx}, c_{ft}, \sigma_{fx}, \sigma_{ft} \) are the center spatial frequency, center temporal frequency, width of the Gaussian window in the spatial frequency direction, and width of the Gaussian window in the temporal frequency, respectively. The four parameters of the Gaussian function \( (c_{fx}, c_{ft}, \sigma_{fx}, \sigma_{ft} \text{ of Eq. 2} \) were estimated to best fit the net motion contrast with the behavioral responses of the experiments in terms of the least RMS errors for the MSV and CA-MSV conditions, respectively. The motion contrast, which ranges from –1 to 1, was linearly scaled to range from 0 to 100 (percentage scale) to compare it with the performance of motion direction discrimination (the correct response rate).

**Measurement of the spatiotemporal frequency tuning of OFR**

The spatiotemporal frequency tuning map of the OFR was directly measured using a single grating stimulus for the purpose of validating the tuning map estimated using a Gaussian function of free parameters. A horizontal sinusoidal grating (50% Michelson contrast) with a spatial frequency chosen randomly from [2.09, 1.04, 0.52, 0.26, 0.13] c/deg and temporal frequency chosen randomly from [25, 16.67, 12.5, 8.33, 6.25, 3.13, 1.56] Hz was presented for 20 frames in the same manner as previous experiments. Forty trials were completed for each condition. To interpolate the sparsely sampled spatiotemporal frequency map, the spatiotemporal frequency tuning maps were fit using the 2D log-scaled Gaussian function described in Eq. 2. Monitor refresh rate was fixed at 100 Hz. The frequency tuning maps were measured with the central visual field of the gratings occluded.

**Estimating the spatiotemporal frequency tuning map using randomly sampled stimuli**

Because it is not easy to directly measure the perceived motion strength, we developed the following reverse correlation method to validate the estimated tuning map of motion perception using a different set of stimuli. First, we sampled the spatiotemporal frequency space using a log scale, sampling the spatial frequency at eight points ([1.48, 1.04, 0.74, 0.52, 0.37, 0.26, 0.18, 0.13] c/deg) and the temporal frequency at nine points ([25, 17.68, 12.5, 8.84, 6.25, 4.42, 3.13, 2.21, 1.56] Hz). We then randomly chose 10 spatiotemporal frequencies from 72 possible points to generate 10 sinusoidal gratings moving upward, and randomly chose other 10 frequencies to generate 10 gratings moving downward, all of which were at 5% contrast. Figure 3 depicts how we generated such a random grating stimulus. The stimulus was 480 × 480 pixels, 25 × 25°, presented with the...
The focus of this study was to explore how motion energy signals are integrated across frequencies when the direction information at different frequencies conflict with each other. The OFR is considered to be a good behavioral index to this end as it is not heavily affected by second-order motion signals, and it is genuinely open-loop controlled by the first-order motion energy (Hayashi et al. 2008; Masson et al. 2002; Miura et al. 2006; Sheliga et al. 2005). Moreover, the frequency tuning of the OFR can be measured directly. The neural mechanism that drives the OFR has been well investigated (Kawano et al. 1994; Takemura et al. 2007). Therefore we first examined the OFR when the participants were viewing the standard MSV and CA-MSV stimuli with variable inter-slit intervals and frame rates of the image stream.

Figure 4 depicts the average temporal profiles of eye velocity. The ocular movements consistent with the actual pattern shift are represented as positive velocity, whereas the eye movements in the opposite direction are represented as negative velocity. As shown in Fig. 4A, when the inter-slit interval increased, the amplitude of the OFR monotonically decreased for the MSV stimuli. On the other hand, the change in OFR amplitude for the CA-MSV stimuli was more complicated. Figure 4B demonstrates that the eye moved toward the direction opposite to the actual pattern shift when the inter-slit interval was 0, 1, or 2 pixels. However, when the inter-slit interval was 3 pixels, the participants’ eyes moved in the forward direction.

The amplitude of the OFRs as a function of the inter-slit interval is shown in Fig. 5A. The amplitude is normalized by the maximum OFR amplitude across all conditions for each participant. For the MSV stimuli (open triangles and thick gray lines), the OFR amplitude was initially high and positive, then monotonically decreasing as the inter-slit interval increased. In contrast, the OFR amplitude for the CA-MSV stimuli (asterisks and thin gray lines) was initially negative, i.e., the participants’ eyes moved in the direction opposite to the direction of pattern shift. The amplitude of the OFR gradually increased and significantly exceeded zero as the inter-slit interval increased, then returned to zero. While the data shown here were acquired with the central field of the stimulus occluded by a uniform gray disk (5° radius), similar results were obtained with full-screen presentation. (See the 1st section of the supplementary material for the details of the difference between the full-screen and the center-occluded conditions.)

**Results**

**OFR elicited by the MSV and CA-MSV stimuli**

The focus of this study was to explore how motion energy signals are integrated across frequencies when the direction information at different frequencies conflict with each other. The OFR is considered to be a good behavioral index to this end as it is not heavily affected by second-order motion signals, and it is genuinely open-loop controlled by the first-order motion energy (Hayashi et al. 2008; Masson et al. 2002; Miura et al. 2006; Sheliga et al. 2005). Moreover, the frequency tuning of the OFR can be measured directly. The neural
Finally the correct response rate fell to chance level when the stimuli was almost at chance level at the same inter-slit interval. Quite accurately, although the direction discrimination of MSV to be for the CA-MSV stimuli (x marks and thin black lines) tended the image stream. On the other hand, the correct response rate fall, however, differed slightly depending on the frame rate of inter-slit interval at which the correct response rate started to chance level, regardless of the frame rate. The value of the row. As the inter-slit interval increased, the response rate fell to chance level, whereas negative velocity means the eye moved in the opposite direction.

shows motion discrimination performance as a function of the inter-slit interval. The performance in the motion discrimination was evaluated using the percentage of trials in which the participants reported the perceived direction to be consistent with the overall pattern shift. Chance level was 50%; a correct response rate lower than 50% indicates that the participant perceived motion moving in the direction opposite to the actual shift of the image. For the MSV stimuli, the correct response rate (indicated as open circles and thick black lines) was almost 100% when the inter-slit interval was narrow. As the inter-slit interval increased, the response rate fell to chance level, regardless of the frame rate. The value of the inter-slit interval at which the correct response rate started to fall, however, differed slightly depending on the frame rate of the image stream. On the other hand, the correct response rate for the CA-MSV stimuli (x marks and thin black lines) tended to be <50% when the inter-slit interval was relatively narrow, indicating that the participants’ reports were biased toward opposite-direction motion. This is obvious in the condition where the inter-slit interval was 1 pixel, and the monitor refresh rate was 100 Hz. As the inter-slit interval increased, the response rate became >75%, i.e., participants reported the forward motion quite accurately, although the direction discrimination of MSV stimuli was almost at chance level at the same inter-slit interval. Finally the correct response rate fell to chance level when the inter-slit interval became very large. Therefore the perceived motion direction changes as a function of the stimulus parameters in a way somewhat resembling the response pattern observed in the OFR. A notable difference, however, is that the reported motion perception at zero or small slit intervals was more in the forward direction than OFR was. Interestingly, this trend was more evident when the central mask was removed (see Supplementary Fig. S1 of the supplementary material). We therefore speculate that a major factor of the discrepancies of motion perception from OFR is the involvement of a higher-order motion mechanism that detects the motion of the pattern features despite changes in luminance-contrast polarity over time (Chubb and Sperling 1988; Lu and Sperling 2001).

Estimating the tuning map of the OFRs elicited by the slit stimuli

The results indicated that the OFR and the reports of the perceived motion direction vary depending on the inter-slit interval and the frame rate of the image stream. We wondered whether the changing patterns of the OFR and the motion perception could be attributed to the change of the motion energy distribution. We hypothesized that the motion energy was integrated with a certain weighting function and that the net motion contrast (see methods) calculated from the weighted motion energy determine the behavior of both the OFR and motion perception. Then we estimated the shape of the tuning map using a 2D log-scaled Gaussian function (with 4 free parameters, defined in the Eq. 2) that best fit the net motion contrast with the OFR profiles.

The response pattern obtained with the MSV and CA-MSV stimuli could not be well fit with a single common Gaussian function (see the section 2 of the supplemental material). However, the fit was fairly good when the Gaussian function was estimated separately for each slit-viewing condition. The peak of the Gaussian function estimated for the MSV stimuli was 0.13 cpd and 6.22 Hz (Fig. 6A). On the other hand, the Gaussian function estimated for the CA-MSV stimuli has a peak at 0.07 cpd and 10.23 Hz (Fig. 6B). Figure 7, A and B, indicates how well the net motion contrast can account for the experimental data when the tuning map was fitted using the Gaussian function shown in Fig. 6. Figure 7A shows the fitted net motion contrast as a function of inter-slit interval. It should be noted that the exact scaling of the net motion contrast to the OFR amplitude is unknown. In Fig. 7A, we simply normalized the maximum absolute (unsigned) OFR amplitude to 1. We evaluated the consistency in the sign and order by plotting in Fig. 7B the net motion contrast as a function of the OFR responses. In Fig. 7B, most of the plots appear in the first and third quadrants, which show that the directions indicated by the net motion contrast were consistent with the actual responses. The correlation coefficient between the data and the fitted net motion contrast was 0.96 for the MSV stimuli and 0.86 for the CA-MSV stimuli.

The simulated net motion contrast, however, tends to over-estimate the OFR amplitude when the inter-slit interval is >4 pixels (Fig. 7A). This might be due to the nonlinear scaling between the net motion contrast and the OFR amplitude when the net motion contrast is small. It is also possible that the OFR amplitude estimated using the fixed time window might be underestimated for larger inter-slit intervals because weak
motion stimuli are known to delay the onset of OFR. Nonetheless, in general, the signs of the data and net motion contrast agree with each other very well with the existing differences being only a matter of scaling, which may vary depending on the individual. This supports the validity of our assumption that the OFR elicited by the slit stimuli can be explained by the net motion contrast weighted with a tuning function. The results also suggest that the tuning of the inter-frequency integration varies depending on the stimulus configuration. The tuning maps estimated here are also consistent with the frequency tuning derived from a totally different analysis. (Please see Supplementary Fig. S3 or section 3 of the supplementary material).

**Estimating the tuning map of the slit motion perception**

In a similar manner to estimating the tuning map of the OFR, we also estimated the tuning map for slit motion perception: we estimated the shape of the tuning map using a Gaussian function (with 4 free parameters, defined in Eq. 2) that best fit the net motion contrast with the responses to the slit stimuli. We could not, however, fit the results of both MSV and CA-MSV stimuli with a single Gaussian function. Instead tuning maps were estimated separately for the MSV and CA-MSV stimuli. The peak of the Gaussian function estimated from the responses to the MSV stimuli was 0.62 cpd and 3.45 Hz (Fig. 6D). On the other hand, the tuning map estimated to explain the responses to the CA-MSV stimuli had a peak at 0.21 cpd and 8.38 Hz, showing a bias for
very low spatial frequencies and moderate temporal frequencies (Fig. 6).

Figure 8, A and B, depicts how well the net motion contrast weighted with a Gaussian function, which was separately estimated for the MSV and CA-MSV stimuli (Fig. 6), can account for the experimental data. The correlation coefficient between the data and the net motion contrast was 0.96 for the MSV stimuli and 0.68 for the CA-MSV stimuli. Figure 8B shows a significant number of plots in the second and fourth quadrants, which implies that the directions predicted by the net motion contrast are often opposite to the actual responses. The weighted motion contrast model therefore showed larger deviation from the participants' reports on the perceived motion direction than from the OFR, especially when the CA-MSV was presented at a frame rate of 25 Hz. This seems to reflect a complex aspect of motion processing related to the perception of motion, which cannot be completely implemented in our simple model. Nonetheless the fitting was good over the wide range of tested inter-slit intervals and frame rates. The results also indicate that the tuning was modulated according to the viewing situation, as we found with the OFR. The tuning maps estimated here are also consistent with the frequency tuning derived from a totally different analysis. (Please see Supplementary Fig. S3 or section 4 of the supplementary material.)

Spatiotemporal frequency tuning of OFR amplitude

We had to estimate different tuning maps to account for the changing pattern of OFR for the MSV and the CA-MSV stimuli. We interpret this finding to indicate that some inter-frequency interactions alter the frequency tuning map depending on the stimulus structure. For instance, MSV, but not CA-MSV contains motion signals coherently moving at a particular speed across frequencies. We suspect this is an inter-frequency stimulus structure that might drive a specific interaction. To gain further insight into this problem, we next directly measured how the OFR amplitude changes as a function of the spatiotemporal frequency without introducing specific inter-frequency structures. That is, a single suprathreshold sinusoidal moving grating was presented for this measurement. Figure 6C depicts a color-scale plot of such tuning. The peak of the tuning map directly measured using sinusoidal gratings was located at a relatively low spatial frequency (0.22 cpd) and moderate temporal frequency (11.3 Hz). The result corre-
sponds very well to the tuning map estimated from the responses to the CA-MSV stimuli using a 2D log-scaled Gaussian function (Fig. 6B) in that both are sensitive to temporal frequencies around 11 Hz and to very low spatial frequencies. Our finding that the tuning maps estimated from the responses to the MSV stimuli differ from those estimated from the CA-MSV stimuli and those measured using single gratings indicated that the cross-frequency integration varies depending on the viewing condition or the structure of the stimulus’s motion energy.

The results are more or less consistent with a previous study that measured the OFRs in humans (Gellman et al. 1990). However, the previous study showed a steeper decay of OFR amplitude as a function of spatial frequency in some participants. A study done by the same group on monkey OFR (Miles et al. 1986) reported a further low-pass spatial frequency characteristic of the response. The discrepancy between the present study and the previous studies might be partially due to differences in the sizes of the visual fields (Hayashi et al. 2008). It has been reported that spatial and temporal frequency tuning of OFRs varies depending on the stimulus size (Barthelemy and Masson 2006). We used stimuli subtending 25 × 25° on a CRT monitor, whereas the previous studies used stimuli projected onto a large screen (subtending 90 × 90° in the monkey study and 85 × 85° in the human study). The smaller display in our study may have stimulated less of the peripheral visual field, which is sensitive to lower spatial frequencies, resulting in slightly higher spatial frequency tuning for the OFRs in the present study. Another notable difference in the paradigm between the present and previous studies is the timing of saccadic eye movements. The original OFR studies measured the eye movements after saccades to the center of the screen with variable postsaccadic delays and reported that the OFRs were best when the stimulus motion began soon after saccades and attenuated as the postsaccadic delay interval increased: postsaccadic enhancement of OFR (Kawano and Miles 1986; Takemura and Kawano 2006). On the other hand, in the present study, eye movements were elicited by the motion stimuli that started after a period of fixation, and thus the measured OFRs were presumably little affected by the postsaccadic enhancement. The present results suggest an intriguing possibility that the postsaccadic enhancement might weight the spatiotemporal frequencies differently.
Frequency tuning map estimated from the randomly sampled stimuli

Unlike in the case of the OFR, it is difficult to directly estimate the suprathreshold magnitude of perceived motion strength using single grating stimuli. Instead we used multiple grating stimuli that had no specific frequency structure. Specifically, we applied a reverse correlation analysis to the psychophysical responses of observers in judging the dominantly perceived direction for compound gratings made of 20 different spatiotemporal frequencies, half moving upward and half downward, that were randomly and uniformly sampled from the spatiotemporal frequency space (see also Fig. 3). We then compared the tuning map obtained from such a different set of stimuli, the spectral components of which are less biased and structured, to the tuning map estimated from the slit stimulus responses. Figure 6F depicts the frequency tuning map evaluated by this experiment (the map for the center-occluded condition is shown). As indicated in Fig. 6F, the participants had a tendency to respond by relying on the motion cues defined by very low spatial frequencies (0.23 cpd) and moderate temporal frequencies (8.1 Hz); this corresponds very well to the peak of the tuning map estimated from the CA-MSV stimuli (Fig. 6E). Therefore the results show that the cross-frequency integration mechanism observed in CA-MSV has a peak sensitivity at a similar spatiotemporal frequency to those measured in a more direct manner using sinusoidal grating(s). Like the results of OFR, the different tuning map estimated with the MSV stimuli indicated that the tuning varies depending on the viewing conditions.

We also found a negatively tuned peak around 1.2 cpd and 6.9 Hz. This indicates that ~2.2-octave-higher spatial frequency components are negatively correlated with the perceptual decision based on the low spatial frequency cues in this stimulus configuration.

Note that the frequency tuning map depicted here was acquired when the center of the visual field was occluded so as to reduce the effects of second-order motion signals. The overall tuning did not change, however, when the stimuli were presented across the full screen. The positive tuning peak in this condition is around 0.22 cpd and 8.8 Hz, whereas the
negative peak is around 1.0 cpd and 6.6 Hz (data not shown), indicating that the participants made their decision mainly by relying on the motion cues at lower spatial frequencies.

**OFR versus perception**

The comparison of the tuning maps between OFRs and motion perception showed that the temporal frequencies preferred by OFR were higher than those preferred by motion perception, indicating a discrepancy between the two types of responses. Figure 9 is a plot of the normalized OFR amplitude as a function of correct response rate (reports on the perceived motion direction). The plots located in the second and fourth quadrants indicate that OFR direction and perceived motion direction were incongruent. The discrepancy between eye movements and motion perception was observed for some conditions, such as when the CA-MSV stimuli with the inter-slit interval of 3 pixels were presented at the monitor refresh rate of 75 Hz with its center occluded. Although the participants reported reverse-direction motion perception (Fig. 5B), their eyes moved in the forward direction (Fig. 5A). This discrepancy is depicted in the Fig. 9 by the plots located in the second quadrant and may reflect the difference in frequency tuning between the perception and visuo-oculomotor response because both perception and OFRs were successfully predicted from the weighted motion signals. In addition, although the reported motion perception and OFRs agree well with regard to the motion direction (Fig. 9), the quantitative relation of the two responses is highly nonlinear. This is mainly due to a rapid saturation of the direction discrimination performance relative to the OFR magnitude, a tendency observed for the MSV stimuli but not for the CA-MSV stimuli. This stimulus difference might suggest that coherent motion signals across spatial frequencies, included in the MSV spectrum, might control the direction of motion perception more effectively than it does the magnitude of OFRs.

**FIG. 9.** Plots of the normalized OFR amplitudes as a function of the correct response rate (the response on the perceived motion direction). Plots located in the second and forth quadrants indicate that the direction of perceived motion was opposite to the direction of the elicited OFR, indicating the discrepancy between the frequency tuning of the motion perception and the ocular motor control or the different effects of the 2nd-order motion.

**DISCUSSION**

We have developed a novel motion display, called CA-MSV, in addition to MSV stimuli, to investigate motion processing when directional information is incongruent across frequencies. We hypothesized that the OFR and the motion perception elicited by the slit stimuli were determined by the motion contrast of the weighted motion energy, and we estimated the shape of the tuning function of cross-frequency integration that can best account for the observed results. The model based on our hypothesis can account for the responses very well when different frequency tuning maps are separately estimated for the MSV and CA-MSV stimuli but not when a single common frequency tuning map was estimated for the two slit-viewing conditions. We also investigated the frequency tuning map more directly using sinusoidal gratings and obtained similar frequency maps as those estimated from the CA-MSV rather than those estimated from the MSV stimuli. OFRs and motion perception had similar response profiles and thus similar frequency tuning maps, except for the higher temporal frequencies preferred by OFRs. Therefore linear inter-frequency integration mechanisms can account for the results in many respects, although the tuning property of such mechanisms varies depending on whether the stimulus is MSV or CA-MSV. In the following section, we will discuss why the stimulus dependency observed here is likely to be mediated by a robust speed detection mechanism that integrates coherent motion signals and/or the suppressive interaction between frequencies.

**Speed tuning mechanism**

Our results suggest that the underlying frequency map was different for the MSV stimuli as compared with the CA-MSV and sinusoidal stimuli. The question is what kind of mechanism could cause such a stimulus-dependent modulation in the tuning map? The spatiotemporal spectrum of the MSV stimuli always contains motion signals coherently moving at a particular speed (as indicated by the oblique line passing through the origin in Fig. 2A). One possible explanation for the discrepancy between the viewing conditions is that a mechanism integrating the coherent motion signals was involved. In support of this view, the in-line frequency integration mechanism was suggested from psychophysical experiments and a theoretical framework (Schrater et al. 2000; Simoncelli and Heeger 1998). In addition, electrophysiological studies indicate that many MT neurons possess speed tuning, that is, frequency selectivity along an oriented line in frequency space (Perrone and Thiele 2001; Priebe et al. 2006). However, we do not exclude the possibility that other types of frequency interactions, such as higher-order motion processing or suppressive connections (see following text), which are inherently nonlinear, are involved in this stimulus dependency.

**Suppressive interaction between frequencies?**

Another type of stimulus-dependent change in the tuning map we observed was the suppressive interaction between frequencies when the tuning map was estimated from the randomly sampled gratings: the participants judged the motion direction based on motion signals from low spatial frequencies, whereas signals from high spatial frequency were negatively
correlated with motion judgments. The tuning map involving the frequencies showing negative correlation, however, failed to duplicate the slit motion perception if the map was directly used to calculate the net motion contrast in our model (see section 6 of the supplementary material for detail). The negative effect was estimated by a reverse correlation method with localized spatiotemporal frequency components randomly chosen. The power spectra of the slit stimuli had a more linearly aligned structure. Therefore the nature of this negative interaction is also considered to vary depending on the distribution or structure or organization of the power spectra of the input motion stimuli.

There are some psychophysical findings that may be related to this suppressive interaction. Serrano-Pedraza et al. (2007) reported that direction perception based on a low spatial frequency grating is facilitated by a 3-octave-higher spatial frequency grating moving in the opposite direction and inhibited by the same grating moving in the same direction (see also (Derrington and Henning 1989; Nishida et al. 1995; Yanagi et al. 1995). These findings are consistent with the results of the present study in that low spatial frequencies dominate the ongoing motion perception, whereas high spatial frequencies are negatively related to the dominant low frequency.

Winner-take-all frequency interaction in OFR

Previous studies on OFRs have shown that two sinusoidal gratings moving in opposite directions interact with each other in a winner-take-all fashion, rather than in a simple summation manner, to elicit the OFR (Matsuura et al. 2008; Sheliga et al. 2006). Although the configuration of our slit stimulus differs from the previous studies in many aspects, both present and previous studies investigated the OFR when the direction of the multiple frequency components conflicted with each other. In the present study, we found that the tuning map of the OFR estimated with the slit motion stimuli (Fig. 6, A and B) showed a narrower bandwidth than the tuning map directly measured using a single sinusoidal grating where there are no directional conflicts across frequencies (Fig. 6C). In addition, the tuning map in Fig. 6C could predict the OFR to the CA-MSV stimuli only when the frequencies that elicited very large OFRs (normalized amplitude >0.85) were used (see section 5 of the supplementary material). Therefore in accord with previous studies, our results support the idea of a winner-take-all interaction between frequencies where only certain frequency components contribute to elicit the OFR.

Low spatial frequency preference for supra-threshold motion perception?

Although the tuning map is essentially stimulus-dependent and cross-frequency inhibition exists, our results also indicate that the frequency tuning has a peak at low spatial frequency for both motion perception and OFRs. The frequency dependence of motion processing has been conventionally measured by contrast threshold, and it has been reported that it has a peak at a moderate spatial frequency (0.5–1 cpd) and moderate temporal frequency (5–10 Hz) (Burr and Ross 1982; Gomi et al. 2006; Hutchinson and Ledgeway 2006). As an alternative way to measure contrast threshold, we proposed using a reverse correlation method to estimate the frequency tuning of motion processing above threshold. Our results revealed that there is a general tendency for spatiotemporal frequency components around 0.23 cpd and 8.1 Hz, which are lower than those measured by contrast-threshold sensitivity, to be preferentially integrated for direction judgment. In line with the idea that suprathreshold motion perception preferentially uses low spatial frequency components, when two moving gratings of different frequencies move in opposite directions, the movement of the lower-frequency grating captures the motion perception of the higher-frequency grating, which is known as low-frequency dominance in motion capture (Ramachandran and Cavanagh 1987).

Imaging studies have reported that motion-evoked activities localized around hMT+ (the human homologous area of monkey MT) have a peak at a spatial frequency lower than those expected from contrast threshold sensitivity [MEG response (Amano et al. 2007; Anderson et al. 1996); EEG response (Korth et al. 2000); fMRI response (Singh et al. 2000)]. Moreover, the optimal speed for direction discrimination estimated here is around 30–40°/s and is reasonably consistent with the range of preferred speeds of MT neurons (around 16–32°/s) (Cheng et al. 1994; Lagae et al. 1993; Liu and Newsome 2003; Maunsell and Van Essen 1983; Perrone and Thiele 2001; Pribe et al. 2003). It is also noteworthy that the distribution of frequency selectivity of MT neurons, shown in Fig. 5 of Perrone and Thiele (2001), is very similar to our tuning map (Fig. 6F), indicating that majority of MT neurons are sensitive for 0.2–0.4 cpd and 8 Hz. In addition, the response of each MT neuron to multiple motion energy components can be described well using a competitive interaction model (Krekkelberg and Albright 2005) similar to the motion contrast used here. We therefore speculate that suprathreshold motion sensitivity revealed by the reverse correlation method might primarily reflect the property of postdetection motion processing in such areas as MT, whereas contrast threshold sensitivity might primarily reflect early motion detection in areas such as V1 (Foster et al. 1985; Hawken et al. 1996; Singh et al. 2000).

Our results suggest that spatial frequency tuning might be significantly different between threshold and suprathreshold motion perception and that reverse correlation might be a promising way to measure suprathreshold sensitivity. A weakness of this argument, however, is that we have not successfully revealed the standard suprathreshold sensitivity function free from such complexities as stimulus dependency and inhibitory effects. Further investigation is necessary to fully understand the relationship between the contrast threshold sensitivity and inter-frequency integration above threshold.

Comparison between the frequency tuning of motion perception and visuomotor responses

It has been suggested that visuomotor control systems have spatiotemporal tunings different from those for motion perception (Amano et al. 2007; Gomi et al. 2006; Miles et al. 1986). Indeed the present study showed directional dissociations between motion perception and OFRs, in particular for the CA-MSV stimuli, which can be ascribed at least partially to preference of the OFR responses to higher speeds. Our group recently showed that eye movements in the open-loop mode and those in the closed-loop mode are driven by different
mechanisms that are tuned to a fixed temporal frequency and to a stimulus speed, respectively (Tabata et al. 2009). Therefore the disparity between the OFR and the reports on the perceived motion (made by integrating the whole presentation time of the stimulus) may reflect the difference between open- and closed-loop motion processing. However, the dissociation in spatiotemporal tuning between suprathreshold motion perception and OFRs shown in the present study seems to be smaller than that found between threshold motion perception (contrast sensitivity function) and the OFR (see Fig. 6 of Gomi et al. 2006). The similarity between motion perception and the OFR is further suggested by our finding that, in either case, the responses to the MSV stimuli, which only contain coherent motion signals across frequencies in our stimuli, cannot be explained by using the same tuning map estimated with the CA-MSV stimuli and those directly measured by the gratings. The OFR amplitudes elicited by random dot textures tend to be larger than those elicited by sinusoidal gratings (Miles et al. 1986), suggesting that signal integration follows the coherent motion found with suprathreshold motion perception (Schrater et al. 2000; Simoncelli and Heeger 1998). Electrophysiological studies have shown that MT/MST is involved in the initiation of OFRs (Kawano et al. 1994; Takemura et al. 2007) as well as in motion perception (Salzman et al. 1990). Therefore OFRs and motion perception may both participate in cross-frequency integration at intermediate stages of motion processing, like MT/MST, the activity of which is considered to be associated with the frequency and speed tuning of motion-related performance (O’Keefe andMovshon 1998).

CA-MSV as the space-time-interchanged version of reversed-phi motion

In an experimental situation where a pattern alternates between positive and negative contrast as it moves (Chubb and Sperling 1988, 1989), it has been found that such contrast-reversing motion stimuli yield a sensation of motion in the direction opposite to the physical displacement of the pattern. This is known as reversed-phi motion (Anstis 1970). Previous studies have applied contrast reversing manipulations only in the time domain, i.e., the contrast of moving dots flips every frame (discrete time point) to generate the reversed-phi motion. From a mathematical viewpoint, however, contrast reversal in the space domain is also supposed to produce a motion sensation similar to reversed-phi motion. Here we developed the CA-MSV display by extending the already established MSV display. Because standard MSV is equivalent to sparse sampling of the image stream in the space domain, alternating the contrast at every slit is equivalent to the time-space-interchanged version of reversed-phi motion display. Our experiments clearly demonstrate that the CA-MSV stimuli evoke the reversed motion perception and OFRs as well as reversed-phi motion (for the study on OFRs using reversed-phi motion, see Masson et al. 2002). We also showed that the perceived direction of the CA-MSV stimuli can be explained by the changes in the distribution of motion energy and propose that the property of the cross-frequency integration of motion energy may reflect properties of MT neurons. In accordance with this view, the physiological studies using reversed-phi motion stimuli showed that the responses of MT neurons are mainly driven not by the tracking of features (mediated by a higher-order motion system) but by the motion energy of the stimuli (Krekelberg and Albright 2005; Livingstone et al. 2001). Considering the asymmetry of space and time in motion perception, however, there might be some quantitative differences between the motion perception of CA-MSV and reversed-phi stimuli.

Acknowledgments

We are grateful to C. Chubb and B. Tse for discussions, to A. Anderson for comments, to S. Miura for technical assistance, and to Y. Sugita and S. Nishida for support. This research was supported by KAKENHI 16GS0312, KAKENHI 19700289 from JSPS and JST PRESTO program.

References

Anderson SJ, Holliday IE, Singh KD, Harding GF. Localization and motion perception, however, there might be some quantitative differences between the motion perception of CA-MSV and reversed-phi stimuli.

CA-MSV as the space-time-interchanged version of reversed-phi motion

In an experimental situation where a pattern alternates between positive and negative contrast as it moves (Chubb and Sperling 1988, 1989), it has been found that such contrast-reversing motion stimuli yield a sensation of motion in the direction opposite to the physical displacement of the pattern. This is known as reversed-phi motion (Anstis 1970). Previous studies have applied contrast reversing manipulations only in the time domain, i.e., the contrast of moving dots flips every frame (discrete time point) to generate the reversed-phi motion. From a mathematical viewpoint, however, contrast reversal in the space domain is also supposed to produce a motion sensation similar to reversed-phi motion. Here we developed the CA-MSV display by extending the already established MSV display. Because standard MSV is equivalent to sparse sampling of the image stream in the space domain, alternating the contrast at every slit is equivalent to the time-space-interchanged version of reversed-phi motion display. Our experiments clearly demonstrate that the CA-MSV stimuli evoke the reversed motion perception and OFRs as well as reversed-phi motion (for the study on OFRs using reversed-phi motion, see Masson et al. 2002). We also showed that the perceived direction of the CA-MSV stimuli can be explained by the changes in the distribution of motion energy and propose that the property of the cross-frequency integration of motion energy may reflect properties of MT neurons. In accordance with this view, the physiological studies using reversed-phi motion stimuli showed that the responses of MT neurons are mainly driven not by the tracking of features (mediated by a higher-order motion system) but by the motion energy of the stimuli (Krekelberg and Albright 2005; Livingstone et al. 2001). Considering the asymmetry of space and time in motion perception, however, there might be some quantitative differences between the motion perception of CA-MSV and reversed-phi stimuli.

Acknowledgments

We are grateful to C. Chubb and B. Tse for discussions, to A. Anderson for comments, to S. Miura for technical assistance, and to Y. Sugita and S. Nishida for support. This research was supported by KAKENHI 16GS0312, KAKENHI 19700289 from JSPS and JST PRESTO program.

References


J Neurophysiol • VOL 103 • JANUARY 2010 • www.jn.org


