Spatiotemporal separability in the human cortical response to visual motion speed: a magnetoencephalography study

Lihong Wang, Yoshiki Kaneoke *, Ryusuke Kakigi

Department of Integrative Physiology, National Institute for Physiological Sciences, Myodaiji-cho, Okazaki 444-8585, Japan

Received 27 February 2003; accepted 22 May 2003

Abstract

Humans can estimate the speed of an object’s motion independently of other visual information. Although speed-related neural activity is known to exist in the primate brain, there has been no physiological study that investigated where and how the speed of motion is represented in the human brain. Nine different combinations of spatial and temporal frequencies were used to make drifting sinusoidal grating of five different speeds (from 1.5 to 24 deg/s). Using the stimuli, we evaluated whether the magnetoencephalographic response property changes were due to a speed-tuned mechanism or to separable spatial and temporal frequency detection mechanisms. The latency change was caused mainly by an inseparable speed-tuned mechanism. In contrast, the amplitude was inversely related to the spatial frequency and was also affected by the temporal frequency differently depending on the frequency. Our results support the view that the human visual system has three sets of mechanisms tuned to spatial frequency, temporal frequency, and speed.

© 2003 Elsevier Ireland Ltd and the Japan Neuroscience Society. All rights reserved.

Keywords: Visual motion; Magnetoencephalography; Human; Sinusoidal grating; Spatial and temporal frequency; Speed

1. Introduction

The survival of most animals depends on how well they can process the visual information of motion. Studies have been done to elucidate where and how visual motion is processed in the brain to achieve perceptual and behavioral experiences (Smith and Snowden, 1994). It has long been controversial whether humans can estimate the speed of an object’s motion itself without interference from other attributes that inevitably covary with speed (McKee et al., 1986; Smith and Edgar, 1991; Schrater and Simoncelli, 1998; Seiffert and Cavanagh, 1998; Reisbeck and Gegenfurtner, 1999).

Although a number of studies have provided evidence that neurons in the monkey MT/V5 are tuned to the motion speed (Maunsell and Van Essen, 1983; Lagae et al., 1993; Rodman and Albright, 1987) and that humans perceive the speed of an object’s motion itself (Mandriota et al., 1962; McKee et al., 1986), these were not conclusive because broadband stimuli were used in some studies and the other study did not fully rule out the possibility that the subjects might have measured the spatial and temporal frequencies independently to calculate the speed. The conclusive evidence has been provided by the recent two studies (Reisbeck and Gegenfurtner, 1999; Perrone and Thiele, 2001), both of which used drifting sinusoidal gratings for their visual motion stimuli for the following reason: motion speed is expressed as the slope of a line that passes through the origin (0, 0) in the spatiotemporal frequency domain (see Fig. 1 in this study). If a visual scene with a certain spatial frequency (s cycle/deg) moves at a certain speed (v deg/s), it produces a temporal luminance change. The frequency (t Hz) of such a temporal luminance change can be calculated by the simple relationship: \( t = \frac{s}{v} \), that is \( v = ts \). Thus, the motion of a sinusoidal grating at a spatial frequency of s and a temporal frequency of t is represented by a single point (s, t). Using various spatial and temporal frequency combinations, the authors (Reisbeck and Gegenfurtner, 1999; Perrone and Thiele, 2001) could successfully show that the responses of...
monkey MT neurons and human motion discrimination are related to the speed and are not merely due to independent spatial or temporal frequency changes.

The next question is whether there is a region in the human brain whose activation is related to the speed, and if it exists, how the speed of motion is represented in the neural activity. In our previous studies, we found that the peak latency and amplitude of the MEG response to the motion of a light spot and the coherent and incoherent motions of random dots varied with the motion speed (Kawakami et al., 2002; Maruyama et al., 2002). These observations, however, were not direct evidence because the stimuli used are so-called broadband. Because a broadband stimulus has a wide range of spatial and temporal frequencies, when using such a stimulus it may not be possible to distinguish between a spatiotemporally separable mechanism and a speed-tuned mechanism (Reisbeck and Gegenfurtner, 1999). Here, we used sinusoidal gratings at various temporal and spatial frequencies to evoke MEG responses and assessed to what extent the response changes are related to the speed.

2. Methods

2.1. Subjects

Seven healthy colleagues (one female and six males, age ranged 33–43 years old) participated in this study. All of them were right-handed and had normal or corrected to normal visual acuity. Informed consent to participate in this study, which was approved by the Ethics Committee of the National Institute for Physiological Sciences, Okazaki, Japan, was obtained from all the participants.

2.2. Visual stimuli

We used vertically oriented drifting sine wave gratings of nine different spatial and temporal frequencies. The gratings were defined by a 16% luminance contrast of Michelson. This relatively low contrast was chosen in order to activate neurons in the magnocellular pathway more than those in the parvocellular pathway (Hadjikhani and Tootell, 2000) and to avoid the activation of neurons non-optimally tuned to the gratings (Perrone and Thiele, 2001). Three temporal frequencies (3, 6, or 12 Hz) and three spatial frequencies (0.5, 1, or 2 cycle/deg) were used to make nine different spatiotemporal frequency combinations and five different speeds (1.5, 3, 6, 12, and 24 deg/s) of drifting (Fig. 1).

The stimulus was presented on a screen in a magnetically shielded room from outside through a small window using a liquid crystal display projector (LP-9200, SANYO). This projector was chosen because all pixels could be refreshed in each frame at the rate of 60 Hz even though they had the same luminance in the next frame. The stimuli subtended a 9 × 9° visual angle in the left visual field, 2° offset from the fixation point. To avoid evoking responses to the luminance and contrast changes, a stationary grating at a determined spatial frequency was presented for 2–3 s before its drifting at a determined temporal frequency. The grating drifted toward the fixation point for 200 ms and remained still for 2–3 s. This short stimulus duration was chosen because most of the first response component to be investigated occurred less than 200 ms after the motion onset. The short stimulus duration could shorten the MEG acquisition time and could therefore avoid fatigue the subject. This presentation of the gratings was repeated 120 times for each experimental session at one of nine different spatiotemporal frequency combinations.

2.3. MEG measurement

We used a 37-channel neuromagnetometer (Magnes, BTI) to record the magnetic fields from the right occipitoparietotemporal region of each subject's brain. Each subject lay on his or her right side on a bed in the dimly lit magnetically shielded room and was instructed to gaze at the fixation point located on the right side of the visual stimuli at a viewing distance of 175 cm. The subjects practiced the timing of blinking beforehand. We did not monitor the eye movements using an electro-oculogram because it cannot reliably measure small eye movements of less than 0.2°. Rather, we could determine the subjects' eye fixations by the high S/N ratio of the averaged responses. Blink artifacts could be discarded in the off-line analysis.

The magnetic field response data of 50 ms before and 500 ms after the onset of motion were amplified, filtered...
(0.1–800 Hz) and digitized at a sampling rate of 2083.3 Hz until 120 epochs of data were collected. The data for each epoch were averaged and the baseline for each channel was corrected at the mean level of 50 ms before the onset of motion. The data were then filtered in the range of 1–50 Hz for further analysis. In the off-line averaging, MEG responses with a drift of more than ±1500 fT were discarded. The peak latency and amplitude of the first component of the magnetic response were measured as root mean square (RMS) values across the averaged response data of the 37 channels (Kawakami et al., 2002). The single equivalent current dipole (ECD) model was used to estimate the location of the cortical activities that produced the magnetic fields (Sarvas, 1987). We had two criteria for the application of the ECD model (Okada et al., 1997). First, the estimated ECD must be stable for at least 10 ms within 20 ms around the peak. Second, the correlation coefficient between the recorded magnetic field and values expected from the ECD model must be more than 0.95. The mean location of the estimated dipole during that period was calculated for each magnetic response. The 95% confidence distance of the dipole location for each subject was calculated as the mean distance +1.96 S.D. between the mean and the estimated dipole location for each response. For the anatomical investigation of the estimated dipole location, T1 weighted sagittal magnetic resonance images (MRI) 1.5-mm thick were obtained with a 1.5 T MR imager (Shimadzu Medical Systems) as in our previous study (Bundo et al., 2000; Kawakami et al., 2002).

2.4. Data analysis

To investigate how the MEG peak response latency/amplitude changes according to the spatial and temporal frequency combination, we investigated the effects of both frequencies on the data by two-way ANOVA (in Systat 8.0, spss) with repeated measure. Discriminant analysis was used to investigate the difference in the distribution of the estimated MEG response sources. P < 0.05 was considered to be significant.

3. Results

Fig. 2 (top) shows the waveforms of the MEG responses to all the stimuli for one subject. The first component was always prominent and its peak latency and amplitude changed with the stimulus conditions. This was the case for all the subjects’ responses. The time courses of the RMS values for each response are shown in Fig. 2 (bottom). As can be seen, the peak amplitude tended to increase and the peak latency tended to decrease as the motion speed increased. Fig. 3 shows the mean (±S.E.M) latency and the amplitude data across all the subjects. The latency was clearly inversely related to the speed and the data was well described by the equation ($r^2 = 0.933$): $RT = a + b/V^\gamma$, where $a$, $b$, and $\gamma$ are positive constants. The value of $\gamma$ was 0.31. One-way ANOVA revealed a significant effect of the speed on the latency data ($P < 0.001$, df = 4, $F = 25.1$). Although the amplitude was generally related to the speed, there was no significant effect of the speed revealed by one-way ANOVA ($P > 0.05$, df = 4, $F = 2.2$).

The separate effects of both the spatial and temporal frequencies on the response latency and amplitude are shown in Fig. 4. The latency increased with the spatial frequency and decreased as the temporal frequency increased. The effects of both spatial and temporal frequencies were significant ($P < 0.001$, df = 2, $F = 38.9$ and 15.6, respectively) by two-way ANOVA. There was no significant interaction ($P > 0.05$, df = 4, $F = 0.45$) indicating that the effects of both frequencies on the...
The effect of the temporal frequency on the amplitude values was different depending on the spatial frequency. When the spatial frequency was low (0.5 and 1.0 cycle/deg), the amplitude increased with the temporal frequency but the value tended to decrease at 2.0 cycle/deg. The overall effects of both frequencies on the amplitude were evaluated by two-way ANOVA and the effect of the spatial frequency was found to be significant ($P < 0.05$, df = 2, $F = 3.8$). The temporal frequency had no significant effect ($P > 0.05$).

The cortical sources of the responses from six of the seven subjects could be estimated using the single ECD model. For all the subjects, the estimated locations were in the extrastriate area as our previous studies (the subjects participating in this study were the same as those in our previous study) (Maruyama et al., 2002).

Fig. 5 shows the dipole locations estimated for all the responses from one subject. The region was always around the occipitotemporal area and did not vary with the stimulus conditions. Discriminant analysis revealed no significant differences in the distribution of the dipole locations among the different stimulus conditions ($P = 0.883$) for the data of all six subjects (Table 1).

4. Discussion

The present study investigated the human cortical responses to visual motion using drifting sinusoidal gratings of various temporal and spatial frequencies. The stimuli enabled us to assess to what extent the response was related to the motion speed and whether the response was also affected separately by the temporal and spatial frequencies.

The peak latency of the first component in the MEG response was inversely related to the speed (Fig. 3) as shown by the previous MEG studies (Kawakami et al., 2002; Maruyama et al., 2002). The separate effects of both the temporal and spatial frequencies were assessed by two-way ANOVA with repeated measure. Although the effects of both frequencies on the latency data were significant ($P < 0.001$) (see Fig. 4), there was no significant interaction ($P > 0.05$). This fact indicates that neither frequencies significantly affect the latency data independently because there was a significant effect of the speed (temporal frequency divided by spatial frequency) as shown by one-way ANOVA ($P < 0.001$).

In contrast, the effect of speed on the amplitude was not significant ($P > 0.05$ by one-way ANOVA) (see Fig. 3). Two-way ANOVA, however, revealed that the separate effect of the spatial frequency on the amplitude was significant ($P < 0.05$). Although the effect of the temporal frequency was not significant, this may be due to the difference in the effect on the amplitude depending on the spatial frequency as shown in Fig. 4. Thus, we consider that the response amplitude change is caused by separate mechanisms for the detection of the spatial and temporal frequencies.

These results are consistent with previous psychophysical studies which also indicated that humans can judge the speed of an object’s motion by the speed itself, but their judgments are affected by the temporal frequency to a small extent (McKee et al., 1986). Recent psychophysical studies have provided more direct evidence that there exists a speed-tuned mechanism as well as separable space and time mechanisms in the human visual motion detection system (Reisbeck and Gegenfurtner, 1995).
The increase in the thresholds for direction discrimination with the spatial and temporal frequencies (Reisbeck and Gegenfurtner, 1999) well corresponds to the decrease in the magnetic response amplitude with the frequencies (see Fig. 4).

1999; Schrater and Simoncelli, 1998; Seiffert and Cavanagh, 1998). The increase in the thresholds for direction discrimination with the spatial and temporal frequencies (Reisbeck and Gegenfurtner, 1999) well corresponds to the decrease in the magnetic response amplitude with the frequencies (see Fig. 4).

Our previous study compared the human reaction time (RT) to the motion onset of a light spot with the MEG response latency evoked by the motion of the same light spot (Kawakami et al., 2002). The relationship between RT and the speed ($V$) of motion has often been described by the equation: $RT = \alpha + \beta V^{-\gamma}$, where $\alpha$, $\beta$, and $\gamma$ are positive constants. The exponential value of $\gamma$ for that study at speeds between 0.4 and 100 deg/s was estimated to be 0.51 ($r^2 = 0.999$), which corresponded to values reported previously (Hohnsbein and Mateeff, 1992; Dzhafarov et al., 1993), even though the motion stimuli (random dots) were different from ours. The MEG response latency was also described well by this equation ($r^2 = 0.989$) but the value of $\gamma$ was 0.2, which is much smaller than that for RT. The value of $\gamma$ for the mean latency of the MEG response in the present study was 0.3 ($r^2 = 0.933$), which is similar to the value for the response to the light spot motion if the differences in the experimental conditions and the visual stimuli are taken into account. The latency change for the MEG response to the light spot motion also was considered to be due to the speed itself because the...
stimulus duration and the distance (trajectory length) did not affect it (Kawakami et al., 2002). Similar response properties, that is the speed-related latency change and the similar $g$ values, may indicate that the broadband visual motion stimuli of the light spots are detected by the same speed-tuned mechanism as theoretically predicted and actually observed in the monkey MT neurons (Perrone and Thiele, 2001). Although our results support the existence a low-level speed detection mechanism in the human brain, it is not known whether the MEG response we measured is directly related to the neural basis of speed perception. Probably, the MEG response we measured was evoked by neural activity around the occipitotemporal area which includes the human homologue of MT/V5 because the signal source estimation by the single ECD model revealed that the dipole locations were always around that area (Fig. 5 and Table 1) as in our previous MEG and fMRI studies which used various visual motion stimuli (Kaneoke et al., 1997; Kawakami et al., 2002; Maruyama et al., 2002). Although some subjects’ locations were more posterior or anterior than the anatomical landmark of human MT/V5 (Watson et al., 1993; Tootell and Taylor, 1995; Dumoulin et al., 2000), such areas may have functions of MT/V5 as we discussed before (Bundo et al., 2000).

A recent physiological study demonstrated a speed-tuned response property in monkey MT neurons that was clearer with a low contrast stimulus (10% Michelson contrast) (Perrone and Thiele, 2001). Our study was done using a similar contrast (16%) and clearly revealed the presence of speed-tuned MEG response properties. The high contrast sensitivity of monkey MT and human MT+ has been demonstrated by physiological (Able, 1984; Sclar et al., 1990) and fMRI studies (Hadjikhani and Tootell, 2000). The speed-tuned response properties in the monkey MT (Perrone and Thiele, 2001) and those of human MT+ found in the present study do not correspond to those of a psycho-physical study (Reisbeck and Gegenfurter, 1999) that used similar visual stimuli (the drifting sinusoidal gratings) in that the human speed judgment was based on the speed-tuned mechanism only when the stimulus

### Table 1
Mean coordinates ($X$, $Y$, $Z$) of the estimated MEG response sources

<table>
<thead>
<tr>
<th>Speed</th>
<th>$X$</th>
<th>$Y$</th>
<th>$Z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>-1.77</td>
<td>-4.98</td>
<td>6.12</td>
</tr>
<tr>
<td>3</td>
<td>-1.45</td>
<td>-4.32</td>
<td>6.32</td>
</tr>
<tr>
<td>6</td>
<td>-1.84</td>
<td>-4.52</td>
<td>6.51</td>
</tr>
<tr>
<td>12</td>
<td>-1.83</td>
<td>-4.29</td>
<td>6.69</td>
</tr>
<tr>
<td>24</td>
<td>-1.85</td>
<td>-4.37</td>
<td>6.24</td>
</tr>
</tbody>
</table>

*a The origin is the midpoint between the left and right tragi. The positive X-axis extends from the origin through the nasion. The positive Y-axis extends from the origin through the top of the head, such that it is perpendicular to the plane formed by the nasion and both tragi. The positive Z-axis extends from the origin through the left side of the head, such that it is perpendicular to the X and Z axes.*
contrast was high (50%). The existence of this discrepancy suggests the importance of other cortical areas in a higher hierarchy for the perceptual experience.

The physiological implications of the magnetic response latency and amplitude change with speed and their underlying neural mechanisms are not known. A recent physiological study also indicated that the latency of the population activity of monkey MT neurons was related to the behavioral reaction times (Cook and Maunsell, 2002), which corresponds to our previous MEG study (Kawakami et al., 2002). Although that physiological study clearly showed a difference in the response property between single neurons and the population activity and a difference between MT and VIP, the basic question of why the response latency changes remains to be solved. It is noteworthy, however that the neuronal activities do not necessarily correspond to the magnetic responses measured by a biomagnetometer. The magnetic response is considered to represent the synchronized excitatory synaptic inputs to the millions of cortical pyramidal neurons (Hamalainen et al., 1993; Okada et al., 1997). Because excitatory input does not always cause an action potential, the neuronal activities measured by extracellular electrodes do not necessarily correspond to the local field potential (Logothetis et al., 2001) and the magnetic response. Although an increase in the excitatory synaptic inputs would cause an increase in the neuronal discharge rates, it would not necessarily be reflected in the magnetic response because a temporal summation of the intracellular currents is necessary to evoke a magnetic response. Thus, the excitatory synaptic inputs must be synchronized to cause magnetic responses. This idea leads to the hypothesis concerning the magnetic response latency and amplitude change. Fig. 6 shows a schematic illustration of the possible time course of the synchronization of the neural population. Even though the synchronization starts at the same time, the latency and amplitude of the detectable magnetic response due to the synchronization can vary depending on the time constant of the synchronization speed because the response can be detected only after the magnitude of the synchronization reaches a certain level.

The delay of the magnetic response to slow speed cannot simply be caused by the time to activate motion detectors as discussed in our previous study (Kawakami et al., 2002). Animals might have developed a neural mechanism that responds more quickly to objects moving at a higher speed for ecological reasons. The higher amplitude of the response to faster motions may also be related to a neural process for the selection of a moving object that has to be responded to more quickly than do other objects. Because the attention to the stimuli increases the neuronal response magnitude to the stimuli (Maunsell and Cook, 2002), one can assume that the stimuli which evoke a higher response amplitude will receive greater attention. Further extensive physiological and psychophysical studies including human imaging studies may reveal a conclusive answer to the fundamental question of why we are late in responding to environmental stimuli, in other words how our brains work in such a process.

Acknowledgements

We thank Mr. O. Nagata and Y. Takeshima for their technical assistance.

References


