© 2015 Wiley Periodicals, Inc.

different velocities are superimposed, observers perceive transparent surfaces sliding over each other on planes of different depth. There is ambiguity with respect to the depth ordering of two transparent surfaces.

The strategies generally used to investigate bistable percept are either prolonged observation or brief presentation. The dynamics of bistable alternation was studied by asking subjects to continually report whether they perceived coherency or transparency [Hupe and Rubin, 2003]. However, brief presentation is suitable to reveal its underlying process where neural response happened to be close to stimulus presentation. In this study, we focused on the initial determination of depth order rather than the process involved in rivalry. This is because we would like to explore the cortical activation associated with determination of depth order during transparent motion perception.

Many electrophysiological studies have investigated how the visual system extracts transparent surfaces from 2D motion. One electrophysiological study of nonhuman primates suggested the involvement of middle temporal (MT) neurons during transparent motion perception [Qian and Andersen, 1994]. In addition, a study of response to coherent and transparent stimuli in the fundus of the superior temporal sulcus (FST) area in the macaque suggested that the FST plays a role in motion segmentation based on transparent signals [Rosenberg et al., 2008]. An another study in humans using electroencephalography and magnetoencephalography (MEG) suggested the involvement of several cortices along the dorsal visual pathway and lateral occipital cortex (LO) in processing of dynamic 3D shapes from motion cues [Jiang et al., 2008]. Furthermore, a human functional magnetic resonance imaging (fMRI) study reported the main role of the human MT complex in differentiating transparent motion from single-surface coherent motion [Muckli et al., 2002]. However, the cortical activity associated with determination of depth order during transparent motion perception remains to be elucidated.

Many factors are known to affect the determination of depth order. For example, difference in speed of two visual patterns affects the perception of the depth order (the slowest pattern is perceived to be the nearest) [Mestre and Masson, 1997]. Meanwhile, when a visual pattern drifts across the field of view, involuntary pursuit (optokinetic) eye movements follow the direction of motion of the pattern. In an eye movement study involving a transparent motion task, Maruyama et al. [2003] reported that optokinetic eye movements at a short latency after stimulus onset might determine the perceptual depth order [Maruyama et al., 2003].

To determine the cortical activity associated with determination of depth order during transparent motion perception, hemodynamic techniques, such as positron emission tomography and fMRI, are inadequate because of their low temporal resolution. Electrophysiological methods, such as event-related potential, event-related magnetic

field (ERF), and event-related synchronization and desynchronization (ERS/ERD) offer high temporal resolution and are ideally suited to investigate these processes. For this reason, brain activities elicited by transparent motion perception task were measured using both fMRI and MEG in this study. The results were analyzed using our recently developed integrative fMRI-MEG neuroimaging method [Natsukawa and Kobayashi, 2012], which can estimate dynamic neural activity with high spatiotemporal resolution and extract data regarding cortical activities commonly observed in subjects using the clustering technique and the Montreal Neurological Institute (MNI) coordinates. Hereafter, we refer to this method as the normalized integrative fMRI-MEG method.

The goal of this study was to clarify multiple cortical activations associated with determination of the depth order in a task that induced transparent motion perception. We performed fMRI and MEG experiments during a transparent motion perception task under depth attention. To clarify the time involved mainly in the determination of depth order, we additionally performed the eye movement experiment and estimated latencies before the start of both pursuit eye movements and depth attention.

METHOD

Subjects

Eleven subjects (10 males and one female, age range: 21–27 years) participated in the fMRI measurement portion of the study, and seven subjects (six males and one female, age range: 21–27 years) of the 11 participated in the both MEG and eye movement measurement portion of the study. All subjects were healthy and had normal visual, oculomotor, and vestibular functions. The reason of different number among experience is that fMRI data is used only to localize the cortical areas related to transparent motion perception in MNI coordinate for integrative analysis, therefore, we do not necessarily need to have the same number of subject. Experiments were conducted in accordance with the principles of the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Kyoto University Hospital for Physiological Science.

Visual Stimuli and Tasks

All visual stimuli were presented on a black background. A random dot pattern appeared within a circular aperture 15 degrees in diameter. Individual dots of 0.077×0.077 degrees were displayed in white. The dot density was 5 dots/degree². Optokinetic eye movement during active tracking of specific dots has been reported to differ from the movement during passive viewing of a visual motion stimulus [Honrubia et al., 1968; Niemann et al., 1994]. Thus, dot lifetime was manipulated to preclude voluntary pursuit of any specific dot. Dots had a lifetime of

Figure 1.

Visual stimuli and experimental paradigm. (a) Configuration of the visual stimuli in the transparent motion MEG experiment and (b) the coherent motion MEG experiment. (c) Visual stimulation of the fMRI block design. TM, transparent motion; RM, random motion; CM, coherent motion; ACS, attention control stimulus; AT, answering time.

150 ms and were randomly repositioned at the end of their lifetime. The motion speed was 10 degrees/s.

Transparent, coherent, and random motion visual stimuli were used in the experiments. In the transparent motion stimulus, half of the dots were horizontally moved to the right and the other half to the left. During coherent motion stimulus, all the dots moved in one of the two horizontal directions. During random motion stimulus, each dot moved in a random (but each frame the same) direction. Other visual stimuli included the attention control and answering time stimuli (Fig. 1). For the attention control stimulus, the letter “N” or “F” was displayed denoting near and far depths respectively. For the answering time stimulus, the letter “A” was displayed.

In the transparent motion task, the attention control stimulus, random motion stimulus, transparent motion stimulus, and answering time stimulus were presented in sequence. The duration of these stimuli were 1.0 s, 1.25–1.75 s, 1.5 s, and 1.0 s, respectively. The subjects were instructed to stare at the center of the stimulus and focus

their attention on one particular depth plane according to the attention control stimulus (near or far) and to report the direction of motion of the surface on which their attention was focused by pressing the appropriate button while the answering time stimulus was displayed. In the control experiment, the random and coherent motion stimuli were displayed for 1.25–1.75 s and 1.5 s, respectively. Figure 1b shows the MEG experimental paradigm. In the eye movement experiment, the same experimental paradigm as MEG experiment was used.

In the fMRI experiment, a blocked design was used to maximize the signal-to-noise ratio and to identify the regions related to perception of transparent motion (Fig.1c). To detect brain activity, fMRI data were acquired and compared among the transparent motion, coherent motion, and resting conditions. Under resting conditions, a fixation point was shown at the center of the display. This experiment required subjects to fixate on a target in blocks of 21 s duration. The activated areas detected by fMRI were designated as the regions of interest.

Eye Movement Experiment

Optokinetic eye movement, which is a combination of involuntary pursuit and quick return eye movements, were recorded with limbus tracking method (Takei Scientific Instruments Co., T.K.K.2901). Before each session, the eye movement recording system was calibrated by asking the subject to fixate several times at each of eight positions on a square and at the center position. This procedure was repeated if the calibration was not adequately performed. The head of each subject was stabilized by supporting the chin and forehead. Eye movement data of right eye was recorded at a frequency of 1,000 Hz. Eye movement data were sorted into four different conditions based on depth order (a combination of attention control stimulus and response data during answering time: far right, far left, near right, and near left) and moving average filter (10 sampling points) was applied to each eye position data, and they were averaged over the condition. Data for quick returns were detected and removed. Subsequently, to determine the properties in the slow phase of nystagmus, eye movement data were replaced with a line connecting the point preceding the quick return to the following point.

The goal of eye movement analysis is to estimate behavioral indicators to specify the period related to determination of depth order during transparent motion perception for a further analysis. It is known that the slow phase of nystagmus tended to be in the same direction as the perceived motion of the surface at the attended depth [Watanabe, 1999]. Therefore, the difference in slow phase properties between two attention control conditions is observed when the subject starts attending the surface. In the eye movement analysis, we estimate this latency as the behavioral indicator (BI1) of attention onset. Conversely, it is reported that the optokinetic eye movements during early latency may determine the perceptual depth order [Maruyama et al., 2003]. Thus, if eye movements during early latency are important in the determination of depth order, the latency of the onset of eye movement becomes a useful indicator (BI2) to elucidate the dynamics underlying the determination of depth order. Using these indicators, cortical activities associated with the determination of depth order in transparent motion perception are considered to occur before BI1 because subjects cannot attend a surface before they perceive transparent motion surfaces, whereas they are considered to occur after BI2 if eye movements during early latency affect the depth ordering.

MEG Experiment

MEG data were recorded using a 306-channel whole-head system (VectorView; Elekta Neuromag, Finland) located inside a magnetically shielded room. All data were obtained at a frequency of 600.6 Hz with an online 0.1–200 Hz band-pass filter. The MEG system was equipped with 102 magnetometers and 102 pairs of planar gradiometers. Eye movements were recorded using vertical and horizon-

tal electrooculography. Epochs containing eye blinks, muscle artifacts, or signal jumps were removed to eliminate artifacts. Individual head shapes and the sensor frame coordinate system were coregistered by digitization (Isotrak; Polhemus Navigation Sciences, Colchester, VT) of individual landmarks (nasion, left, and right preauricular points), whose locations relative to the sensor positions were derived based on signals provided by four head coils with a fixed spatial position relative to the landmarks. These landmarks enabled coregistration of MEG data to individual anatomical MR scans for each subject.

The raw MEG data were acquired using a time window of –300 to 700 ms and were determined by trigger onset. MEG data were sorted into four different conditions based on depth order: far right, far left, near right, and near left. MEG signals were averaged across trials for the near left and far right conditions to focus on the response at stimulus onset underlying transparent motion perception. Root sum square values were calculated from the gradiometer responses [Hämäläinen, 1993a; Hämäläinen et al., 1993b].

fMRI Experiment

The fMRI measurements were performed on a 1.5-T MR imager (Vantage; Toshiba Medical Systems, Tokyo, Japan) using standard echo-planar imaging with a standard radio frequency head coil for signal transmission and reception. The acquisition parameters of the imaging sequences were as follows: repetition time 3 s, echo time 40 ms, flip angle 90°, and voxel size $3.44 \times 3.44 \times 5.00 \text{ mm}^3$.

The fMRI data were analyzed by statistical parametric mapping with the SPM software using MATLAB (<http://www.fil.ion.ucl.ac.uk/spm/>). Preprocessing included motion corrections and the removal of linear trends from the time series. Functional 2D images were coregistered with 3D high-resolution structural images. Statistical analysis was performed using a general linear model [Friston et al., 1995] by convolving the blood oxygen level dependent (BOLD) time series with a standard hemodynamic function corrected for the duration of the task block. A statistical parametric map was created for each subject. After individual analysis, group analysis was performed. To estimate the activated areas correlated to experimental tasks, the transparent motion condition was compared with the coherent motion condition (TM–CM) and the transparent motion condition was compared with the resting condition (TM–rest). The areas in the TM–CM beyond a voxel level threshold of $P < 0.001$ uncorrected for multiple comparisons and within clusters of >30 voxels were considered to be significantly activated. The areas in the TM–rest beyond a threshold of $P < 0.0005$ uncorrected and within clusters of >30 voxels were considered to be significantly activated. These parameters were established to ensure appropriate cluster size. The activation maps were then superimposed on the high-resolution normalized 3D anatomical images. Note that the fMRI result is used only

Figure 2.

Flowchart of the normalized integrative fMRI-MEG method. This method consists of three separate analyses: MEG, fMRI, and integrative analyses. In this figure, (a) represents an example of topographical 3D maps at the peak latency of the root sum square

MEG waveform; (b) represents an example of an activation pattern of the fMRI group analysis; (c) illustrates the estimation of dipole direction in a spherical conductor model; and (d) is an example of reconstructed MEG source activation of the activated cluster.

to localize the cortical areas related to transparent motion perception in MNI coordinate for the following integrative analysis. Although it can be the first choice to determine the regions of interest (ROIs) from fMRI individual analysis, we consider that it is reasonable and effective to use the fMRI results of group analysis, which show significant differences in the interindividual variabilities, because it is difficult to determine the ROIs from individual analysis with due consideration of interindividual variabilities. When we use the results of group analysis, there is a problem of position gap between individual results and group results. In our fMRI group analysis, however, we set the liberal threshold to ensure the nonfocal activities and renormalized the relatively broad ROIs to maintain consistency with the individual results.

It is known that the hMT complex has an important role in visual motion processing and is activated by this stimulus [Muckli et al., 2002]. Furthermore, the results of previous studies suggested that activity in LO region was induced by this type of visual stimulus. As we described later in the Results and Discussion section, hMT and LO, which observed in the contrast of TM - rest, were included in the activated clusters. In this clustering, the location of LO was determined with reference to the literature [Malach et al., 1995; Kourtzi et al., 2003; Welchman et al., 2005].

Normalized Integrative fMRI-MEG Method

To analyze dynamic cortical activation, we used our newly developed normalized integrative fMRI-MEG method [Natsukawa and Kobayashi, 2012] illustrated in

Figure 2. In this method, activated regions (clusters) determined by fMRI group analysis were treated as common spatial constraints for the reconstruction of temporal activation from individual MEG data. This method assumes that the current sources of measured MEG responses exist within the activated clusters. To solve the inverse problem, the geometries of the sensor array and the activated clusters in each subject's brain were required. Accordingly, activated clusters in the MNI coordinates were transformed to individual head coordinates (the inverse operation of normalization [Ashburner and Friston, 1999]). This transformation, referred to as renormalization, enables common clustering among all subjects.

The time-varying (dynamic) dipole moments in all voxels included in the clusters determined by fMRI were estimated by applying a spatial filtering technique to event-related fields in the transparent motion and coherent motion conditions. We used the linearly constrained generalized least square (GLS) method to acquire the best linear unbiased estimation [Natsukawa and Kobayashi, 2012]. This method achieves accurate reconstruction of MEG source activity even when fMRI-invisible sources, which have a high temporal correlation to sources detected by fMRI, are present. An outline of the method of estimation of time courses of dipole moments is described in the Appendix.

To determine the common neural dynamics associated with determination of depth order during transparent motion perception, we computed the average values of reconstructed source activity for all subjects in each cluster. The results and discussion of individual analysis using the data before averaging for subjects were shown in

Supporting Information (Table S1, Fig. S1). Hereafter, if not otherwise qualified, the phrase “reconstructed activity” implies the averaged reconstructed activity for all subjects in each cluster. To investigate the dynamic brain processes involved in determination of depth order during transparent motion perception, we compared the reconstructed activity induced by transparent motion to that induced by coherent motion. Data were analyzed using paired *t*-tests, with a time bin of every 10 ms during the period between the two behavioral indicators described below. When it comes to the integration analysis of different imaging modalities, we have to consider the interindividual variability of both modalities. In this study, therefore, to estimate the location and activities which show the significant differences in the interindividual variabilities, we used two-stage statistical test for fMRI analysis and MEG analysis. The ROIs were determined by fMRI results of group analysis (second-level fMRI analysis) and statistical tests were performed for the reconstructed waveforms (MEG analysis).

RESULTS

All subjects reported that they perceived two transparent motion surfaces at different depth planes during the experiments. In all subjects, no reversal of perceptual depth order occurred in most cases during the tasks because of the short duration and iterative nature of the experimental design. The optokinetic eye movements were observed in response to both transparent and coherent motion stimuli, whereas they were not observed in response to random motion stimuli.

Figure 3 shows the root sum square values of the ERFs for the transparent and coherent motion stimuli in a representative subject. In all subjects, prominent ERFs were commonly observed in the transparent motion condition after approximately 200 ms with the sensors in the vicinity of the occipital, middle occipital, and parietal cortices. In the coherent motion condition, prominent ERFs were also observed at approximately 200 ms. However, the amplitudes of the ERFs in the coherent motion condition were larger than those in the transparent motion condition.

In the eye movement analysis, to estimate BI1, we used a paired *t*-test to determine the significance of the difference in slow phase properties between two attention control conditions. BI1 was defined as the latent period before which a significant difference in properties due to attention was first observed. In other words, BI1 was the time at which a subject first paid attention to a particular surface. Similarly, BI2 was defined as the latent period before which a significant difference in the motion direction of stimuli in the control experiment was first observed. Thus, BI2 was the time at which eye movement was induced by motion stimuli and started to drift. The latency of BI1 was 405 ± 144 ms (mean $\pm$ SD), whereas the latency of BI2 was 216 ± 35 ms (mean $\pm$ SD). Cortical activity associated with

the determination of depth order in transparent motion perception is considered to occur between BI2 and BI1 as described before. Among the results of this analysis using our developed integrative method, we focused on the period between the two behavioral indicators.

By fMRI, the contrast in activation between the transparent and coherent motion conditions yielded multiple indications of activity in the following functional areas: the right V2, anterior cingulate cortex (ACC), supplementary eye field, bilateral BA19, and bilateral intraparietal sulcus (IPS) (Fig. 4). These areas were selected as activated clusters. In addition, the bilateral hMT complex and the right LO, which were observed in the contrast between transparent motion and resting conditions, were included in the activated clusters. Thus, the coordinates of 10 activated clusters were identified: BA19, IPS, and hMT in the right and left hemispheres, and BA32, supplementary eye field, right V2, and LO. These ten clusters are determined by second-level statistical analysis (group analysis) with due consideration of interindividual variability. Table I shows the range of activated clusters and the number of voxels included in the clusters. In addition, the supporting information about the interindividual variability of fMRI results was shown in Supporting Information (Fig. S2). In this fMRI group-analysis, no sources derived from the coherent motion condition compared with the transparent motion condition using same statistical threshold were observed.

Figure 5a shows reconstructed MEG source activations of 10 activated clusters in transparent and coherent motion conditions. Differences in the levels of neural activity among these 10 clusters were detected successfully. Significant differences in activation were particularly prominent in the bilateral IPS, left hMT, right LO, right BA19, and ACC. Differences in neural activity in the right IPS were observed at approximately 246, 316, and 381 ms. In the left IPS, these differences peaked at 331 ms, in the left hMT at 371 ms, and in the right LO at 326 and 401 ms, in the right BA19 at 256 ms, and in the ACC at 261 and 291 ms (Fig. 5b). Differences in neural activation in both conditions were mainly observed during the period between the two behavioral indicators.

DISCUSSION

Transparent motion perception requires different neural processes related to detection of dot movement, eye movement, and perception of transparency. The integrative fMRI-MEG method, which elucidates dynamic processes in multiple areas, was used to evaluate these neural activities. In a previous study, the reliability and usefulness of the normalized integrative fMRI-MEG method were demonstrated in a saccade experiment with high spatial and temporal resolution for reconstruction of cortical activity commonly occurring in subjects [Natsukawa and Kobayashi, 2012].

Figure 3.

Representative results of MEG root sum square values for the magnetic fields measured from two orthogonal gradiometers in transparent motion and coherent motion conditions. MEG sensor-level topographical maps at representative peak latencies of the upper root sum square waveform are shown. Upper, right and left topographical maps are drawn as back view, right and left views, respectively.

Figure 4.

fMRI results of group analysis. Activation patterns (activated clusters) estimated from the results of fMRI group analysis: activated clusters of transparent motion vs. coherent motion, and transparent motion vs. rest condition, where BOLD signals are significantly larger. V2, visual area 2; BA19, brodmann area 19; LO, lateral occipital cortex; ACC, anterior cingulate cortex; SEF, supplementary eye field; IPS, intraparietal sulcus; hMT, human middle temporal area.

In this study, event-related magnetic fields were observed after approximately 200 ms in both experiments. Since MEG data were recorded during optokinetic eye movement, the influence of eye movement artifacts localized in the region of the extraocular muscles needs to be considered. One study reported no effect of the saccadic spike artifact in MEG on posterior sensors [Carl et al., 2012]. Consequently, because a saccadic spike has a much larger amplitude than optokinetic eye movement in this setting, it is considered that oculomotor neuromagnetic artifacts are not prone to contaminate the MEG signal reflecting visual processes in the parietal and occipital areas.

In the fMRI experiment, 10 activated clusters were identified. These results were consistent with those of previous fMRI studies on transparent motion perception [Muckli et al., 2002; Peuskens et al., 2004]. In contrast with the difference between transparent and coherent motion conditions, no significant increase in BOLD signal was observed in the hMT complex. This result indicates that the hMT was activated in a manner similar to that in the coherent motion task. Although no difference between tasks was observed in the hMT by fMRI analysis, the hMT may be an fMRI-invisible MEG source that reflects the differences between the two visual perception tasks. In practice, the hMT is a key structure for motion processing and

TABLE I. Range of activated clusters and voxel numbers included in the activated clusters

	Region	Contrast	Range (mm) MNI coordinate						Voxel number
			min x	max x	min y	max y	min z	max z	
1	right V2	$TM > CM$	9	26	-91	-77	-23	0	83
2	right BA 19	$TM > CM$	12	33	-96	-67	2	30	145
3	left BA 19	$TM > CM$	-41	-28	-86	-70	4	16	34
4	right IPS	$TM > CM$	18	36	-64	-42	57	71	108
5	left IPS	$TM > CM$	-36	-16	-75	-41	45	67	162
6	right hMT	$TM > Rest$	40	56	-71	-55	-13	4	213
7	left hMT	$TM > Rest$	-53	-37	-80	-57	-8	6	191
8	right LO	$TM > Rest$	35	48	-75	-59	-24	-13	122
9	ACC	$TM > CM$	-12	1	38	49	15	30	44
10	SEF	$TM > CM$	-18	0	17	44	40	61	117

Figure 5.

fMRI-MEG integrative results. (a) Reconstructed MEG source activation of 10 activated clusters in transparent and coherent motion conditions. We focused on the results in the gray areas, defined as the latency between behavioral indicators, which may indicate the period before the occurrence of transparent motion perception. **(b)** Statistical results of integrative analysis. Asterisks marked on the figure indicate significant differences between conditions at a corresponding time bin.

perception. In another fMRI study, Muckli et al. varied the spatial constellation of dots moving in opposite directions and observed a correlation between hMT and IPS activity and the perception of coherent motion. The hMT played a key role in differentiating transparent motion from single-surface coherent motion [Muckli et al., 2002]. Peuskens et al. [2004] suggested in their fMRI study that only a small subpart of the LO was involved in 3D motion processing. In this study, activity in both the hMT and LO was detected in the contrast between transparent motion and resting conditions. Therefore, these two regions were included among the activated clusters in the integrative fMRI-MEG analyses conducted here. From our fMRI group-analysis, no activated areas derived from the coherent motion condition compared with the transparent motion condition (CM-TM) using same statistical threshold were observed. Although it seemed to be a contradictory result, we consider that this contradiction is caused by the

differences of signals observed between fMRI and MEG. fMRI signals arise from the BOLD effect having relatively slow response, and integrated value of the BOLD signals during blocks are obtained as a result. As indicated in the Result section, the amplitudes of the ERFs in CM condition were larger than those in TM condition. We consider that response in TM was larger than that in CM in terms of integrated value, and smaller than that in terms of peak value, because TM task requires endogenous processes including perceptual rivalry, which is not time-locked, more than CM task over whole time of blocks.

Using our normalized integrative fMRI-MEG method, differences in neuronal activity were observed after stimulus onset in the IPS, hMT, LO, part of the frontal cortex, and other occipital visual areas. These results were consistent with the results of a study using electroencephalography and MEG [Jiang et al., 2008]. The IPS may play a role in extracting higher level motion information including

the perception of structure from motion and eye movement planning [Andersen, 1989]. Recent neuroimaging investigations on the neural correlates of bistable perception have indicated that activity in the parietal and prefrontal cortices may be associated with visual awareness [Lumer et al., 1998; Lumer and Rees, 1999; Kleinschmidt et al., 1998]; thus, activations observed in IPS clusters may reflect transparent motion perception. Another neuroimaging study investigating cortical activity related to the perception of 3D objects defined by motion suggested that the LO may be the key region for the assembly of object features [Jiang et al., 2008]. In addition, a previous study on human 3D shape perception combining two depth cues reported that extrastriate areas such as the hMT and LO reflected differences in perceived 3D shape, suggesting the existence of “combined-cue” representations in higher visual areas [Welchman et al., 2005]. Therefore, the LO difference observed here may indicate that an assembly of structural components, including shape and depth, into a unitary whole occurs in the LO area.

In this study, differences in activation in the IPS, hMT, and LO regions were prominent during the period between the two behavioral indicators. As mentioned earlier, cortical activity associated with the determination of the depth order during transparent motion perception is considered to occur from BI2 to BI1 for the following two reasons: first, subjects cannot perform a depth-based attention task before they perceive transparent motion surfaces at approximately BI1 (405 ms) and second, optokinetic eye movement during early latency, which starts at approximately BI2 (216 ms), may determine the perceptual depth order [Maruyama et al., 2003]. Therefore, differences in activation in the IPS, hMT, and LO areas between BI2 and BI1 indicate that these areas are closely related to the determination of depth order during transparent motion perception.

Differences in neural activity were observed several times and had differences from 246 to 401 ms (Fig. 5b). Iterative nature of these neural activities in several cortical hierarchy can indicate that hierarchical cortico-cortical network including bottom-up and top-down effects are associated with the determination of depth order during transparent motion perception. However, to make further reference to underlying network related to determination of depth order, we need further investigation such as connectivity analysis. In addition, although the behavioral indicators are useful to interpret the neural activities, their relevance may depend on where in the cortical hierarchy a signal is derived. For example, the dynamic processing in the SEF might be more tightly time-locked to the eye movement latencies than other cortical responses are. That can be why the SEF has no significance of its response.

An fMRI study indicated that accurate prediction of perceptual states is possible on the basis of the activation taken from area hMT and dorsal, parietal areas during prolonged presentation of stimulus [Brouwer and van Ee,

2007]. Although strategies to investigate bistable percept are different, cortical activation area corresponds to our results. Our results moreover indicated the importance of cortical activities after 200 ms for formation of a particular percept. The other psychophysiological study reported that perceptual memory influenced the initial and prolonged perception of bistable percept [De Jong et al., 2012]. In this paper, perceptual memory can affect the processing of determination of depth ordering. However, previous experiences cause no harm to our results due to perceptual stabilization by iterative brief presentation.

Some previous studies have examined factors influencing depth order in transparent motion [Moreno-Bote et al., 2008; Schütz, 2011]. The study of Schütz indicated that perceived depth order was not determined solely by eye movement and that global motion signal strength determined from dot number and motion adaptation can induce a perceptual bias [Schütz, 2011]. In our study, no cue regarding depth order or motion adaptation was provided; therefore, the effect of global motion signal strength could not be determined. However, during transparent motion perception, differences in activation were observed after the initial eye movement (representing a more reflexive response) and before the later eye movement (representing a volitional selection of a transparent surface). Thus, our results suggest that the initial eye movement and accompanying cortical activation in the IPS, hMT, LO, and ACC are important in determining the perceptual depth order of transparent motion.

CONCLUSION

By applying our developed normalized integrative fMRI-MEG method, dynamic recurrent neural activities in multiple cortical areas associated with determination of depth order during transparent motion perception could be detected successfully. Significant differences in neural activity in the bilateral IPS, left hMT, right LO, right BA19, and ACC were observed. Activity in all these areas was observed during the period between the two behavioral indicators determined by eye movement experiment. These results indicate that initial eye movement and accompanying cortical activations in the IPS, hMT, LO, and ACC regions during focused duration can play an important role in the determining the perceptual depth order.

APPENDIX : INTEGRATIVE FMRI-MEG METHOD BASED ON LC-GLS FILTERING

This appendix describes the outline of the integrative fMRI-MEG method based on a linearly constrained generalized least squares method (LC-GLS) spatial filter. In this method, the locations and orientations of equivalent current dipoles (ECDs) placed at the center of gravity in individual activated voxels were first estimated within the

fMRI-activated clusters by maximizing the following evaluation function:

$$f_r(\varphi) = \frac{1}{n} \sum_{t=p_e-\frac{n-1}{2}}^{p_e+\frac{n-1}{2}} \left| \frac{\mathbf{b}_S(t)}{\|\mathbf{b}_S(t)\|} \cdot \frac{\hat{\mathbf{I}}_{Sr}(\varphi)}{\|\hat{\mathbf{I}}_{Sr}(\varphi)\|} \right| \quad (\text{A1})$$

where φ is the plane angle perpendicular to the direction of the radial line. $\mathbf{b}_S(t)$ is the root mean sum (RSS) of magnetic field signals measured by the m th sensor at time t , $\hat{\mathbf{I}}_{Sr}(\varphi)$ is the lead field of the sensor array S for a source located at $\mathbf{r}$ and directed in direction φ . p_e is the peak latency of the evoked magnetic field, and n is the number of sample points used for evaluation.

$f_r(\varphi)$ is the average of the inner product of the lead field and measured field vectors at n sample points centering on the p_e . We used five sample points to represent the peak signal from the perspective of a sampling frequency ($n = 5$). The orientation that maximizes this $f_r(\varphi)$ is selected as the source orientation. In the estimation of dipole orientation, $f_r(\varphi)$ is degraded when the number of sensor arrays used is too large or too small. Thus, we optimized the number of sensor arrays to maximize $f_r(\varphi)$. The lead field matrix was calculated using the Sarvas' equation [Sarvas, 1987], which is a simple analytic expression for magnetic fields using a spherical conductor model.

After determining orientation, the location of a representative ECD was estimated in each activated cluster by maximizing the following evaluation function P_j .

$$P_j = \frac{1}{\mathbf{I}_{Sj}^T \mathbf{R}_{Sb}^{-1} \mathbf{I}_{Sj}}, j = 1, \dots, m_i \quad (\text{A2})$$

where $\mathbf{R}_{Sb}$ is the covariance matrix of the measurements on sensor array S and m_i is the number of voxels included in the cluster i . In the calculation of multiple source activities, a diverging solution and computational instability occurs when multiple ECDs exist in proximity. Therefore, in our integrative method, a linear constraint based on fMRI clusters was applied. This constraint approximates brain activity of the cluster by a representative ECD to simplify the calculation. When we calculate the amplitude of current dipoles included in a cluster, the linear constraint eliminates the contribution of representative ECDs of other clusters, which indicates that it eliminates all activities other than that of the cluster of interest. This calculation can reconstruct all ECDs in each cluster as described below. Therefore, although it could be the case that multiple ECDs were activated serially within each cluster, it is possible to reconstruct activities of the multiple ECDs within a cluster by this integrative method assuming one representative ECD in each nonfocused cluster.

P_j indicates reconstructed signal power obtained by minimum variance spatial filtering. The locations of the representative ECDs that maximize this power reconstruction P_j are determined. In this method, we can determine the location of representative ECDs based on the contribution of the ECDs in the cluster to the measurement data.

We let the vector $\mathbf{b}$ represent measured MEG data. The lead field of a representative ECD in the i^{th} cluster was

defined as $\mathbf{l}_i^r$, and its ECD amplitude as $s_i^r (i = 1, \dots, q)$. In addition, the lead field of the ECD included in the j^{th} voxel of the i^{th} cluster was defined as $\mathbf{l}_{ij}$, and its ECD amplitude as $s_{ij} (j = 1, \dots, m_i)$. Then, we assumed that the fundamental equation for measurements performed with a sensor array can be expressed as:

$$\begin{aligned} \mathbf{b} &= \mathbf{L}_{ij} \boldsymbol{\zeta}_{ij} + \mathbf{n}, \\ \text{where } \mathbf{L}_{ij} &= [\mathbf{l}_1^r \cdots \mathbf{l}_{i-1}^r \mathbf{l}_{ij} \mathbf{l}_{i+1}^r \cdots \mathbf{l}_q^r], \\ \boldsymbol{\zeta}_{ij} &= [s_1^r \cdots s_{i-1}^r s_{ij} s_{i+1}^r \cdots s_q^r]^T. \end{aligned} \quad (\text{A3})$$

where $\mathbf{n}$ is an additive noise vector. In this method, the number of discrete fMRI clusters, q , is smaller than the number of sensors, M . Consequently, it is not possible to obtain a unique solution in general. Therefore, approximate solutions need to be found in some way. The most popular way is the least-squares method minimizing the squared error term. If the noise covariance is the unit matrix multiplied by a constant, this estimation acts as the best linear unbiased estimator (BLUE) [Hyvärinen et al., 2001].

However, this assumption cannot be made using actual MEG data. Therefore, we applied a prewhitening operation to sensor noise $\mathbf{n}$. A zero-mean random vector is said to be white if its elements are uncorrelated and have unit variance. In terms of the covariance matrix, this implies that its covariance matrix becomes the unit matrix. For this operation, the least-squares estimation described above meets the conditions necessary for BLUE. Thus, the $\mathbf{C}$ noise covariance matrix, the least squares solution with prewhitening, is obtained using the following constrained minimizations:

$$\arg \min_{\boldsymbol{\zeta}_{ij}} \|\mathbf{C}^{-1/2} (\mathbf{b} - \mathbf{L}_{ij} \boldsymbol{\zeta}_{ij})\|^2. \quad (\text{A4})$$

Then, the solution of Eq. (A4) is obtained as

$$\boldsymbol{\zeta}_{ij} = (\mathbf{L}_{ij}^T \mathbf{C}^{-1} \mathbf{L}_{ij})^{-1} \mathbf{L}_{ij}^T \mathbf{C}^{-1} \mathbf{b}. \quad (\text{A5})$$

This least squares method with prewhitening is equivalent to the generalized least squares method (GLS) in the field of statistics [Waldorp et al., 2001].

Therefore, the estimated dipole moment of ECD included in the j^{th} voxel in the i^{th} fMRI cluster is expressed as:

$$\begin{aligned} s_{ij} &= \mathbf{w}_{ij}^T \mathbf{b}, \\ \text{where } \mathbf{w}_{ij} &= \mathbf{C}^{-1} \mathbf{L}_{ij}^T (\mathbf{L}_{ij}^T \mathbf{C}^{-1} \mathbf{L}_{ij})^{-1} \mathbf{e}_i \end{aligned} \quad (\text{A6})$$

where $\mathbf{e}_i$ is $q \times 1$ matrix in which the i^{th} element is one and the others are zero. The dipole moments S_{ij} of all ECDs located at fMRI clusters are obtained by the j^{th} iterative operation on the i^{th} cluster.

Note that the plus and minus signs of reconstructed activities S_{ij} (time-varying dipole moments) indicate the direction of the equivalent current dipoles. The sign depends on the orientation of each dipole with each subject's voxel. This in turn depends on local gyral anatomy,

which may differ across subjects. Therefore, we calculated the absolute values of reconstructed dipole moments to discount individual differences in local orientation.

When the GLS filter is applied, it is necessary to calculate the noise covariance matrix C . In the low-rank signal model, the measurement covariance matrix R can be decomposed into signal and noise subspace components, such that

$$R = E_S \Lambda_S^{-1} E_S^T + E_N \Lambda_N^{-1} E_N^T, \quad (A7)$$

where Λ_S and Λ_N are diagonal matrices defined as

$$\begin{aligned} \Lambda_S &= \text{diag}[\lambda_1, \dots, \lambda_Q], \\ \Lambda_N &= \text{diag}[\lambda_{Q+1}, \dots, \lambda_M], \end{aligned} \quad (A8)$$

where diag denotes a diagonal matrix whose diagonal elements are equal to the entries in the square brackets [Sekihara and Nagarajan, 2008]. λ_j is the j^{th} eigenvalue of the matrix R . Furthermore, we define matrix E_S as $E_S = [e_1, \dots, e_Q]$ and E_N as $E_N = [e_{Q+1}, \dots, e_M]$, where e_j is the j^{th} eigenvector of the matrix R . Therefore, the second term of the right side in Eq. (A7) corresponds to the noise covariance matrix C .

In our method, MEG data were decomposed into signal and noise components by applying principal component analysis. It was assumed that the cumulative proportion of the signal components is α , while the other components which are not included in the signal component are considered to be noise components. Then, the number of signal components can be determined by estimating minimum Q to satisfy the following criterion.

$$\left(\sum_{j=1}^Q \lambda_j / \sum_{i=1}^M \lambda_i \right) \times 100 \geq \alpha. \quad (A9)$$

The noise subspace component is then reconstructed using the basis vectors of the noise subspace E_N . By projection of MEG signals into the noise subspace, the reconstructed waveform $\hat{n}(t)$ is expressed as

$$\hat{n}(t) = E_N E_N^T b(t). \quad (A10)$$

When $\sigma_p(p = 1, \dots, 204)$ is the standard deviation of MEG averaged data of prestimulus duration and $\hat{n}_{p,\max}$ is the maximum value of the absolute $\hat{n}(t)$ value of poststimulus duration, with the number of signal component increments until the minimum Q which does not satisfy the following equation is determined,

$$\hat{n}_{p,\max} \geq \alpha \sigma_p. \quad (A11)$$

Thus, α is determined as the cumulative proportion of Q components.

ACKNOWLEDGMENTS

This work was partially supported by the Innovative Techno-Hub for Integrated Medical Bio-imaging of the Project for Developing Innovation Systems, from MEXT,

Japan. The authors thank Dr. Masao Matsushashi and Dr. Hidenao Fukuyama for their assistance in MEG measurements and Dr. Takenori Oida for his assistance in fMRI measurements. The authors declare no conflict of interest associated with this manuscript.

REFERENCES

- Andersen RA (1989): Visual and eye movement functions of the posterior parietal cortex. *Ann Rev Neurosci* 12:377–403.
- Ashburner J, Friston KJ (1999): Nonlinear spatial normalization using basis functions. *Hum Brain Mapp* 7:254–266.
- Brouwer GJ, van Ee R (2007): Visual cortex allows prediction of perceptual states during ambiguous structure-from-motion. *J Neurosci* 27:1015–1023.
- Carl C, Aık A, Konig P, Engel AK, Hipp JF (2012): The saccadic spike artifact in MEG. *NeuroImage* 59:1657–1667.
- De Jong MC, Knapen T, van Ee R (2012): Opposite Influence of perceptual memory on initial and prolonged perception of sensory ambiguity. *PLoS One* 7:art.no.e343.
- Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ (1995): Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp* 2: 189–210.
- Hamalainen M (1993a): Functional localization based on measurements with a whole-head magnetometer system. *Brain Topogr* 7:283–289.
- Hamalainen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV (1993b): Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev Modern Phys* 65:414–497.
- Honrubia V, Downey WL, Mitchell DP, Ward PH (1968): Experimental studies on optokinetic nystagmus. II. Normal humans. *Acta Otolaryngol* 65:441–448.
- Hupe JM, Rubin N (2003): The dynamics of bi-stable alternation in ambiguous motion displays: A fresh look at plaids. *Vis Res* 43:531–548.
- Hyvarinen A, Karhunen J, Oja E (2001): Independent component analysis. New York: John Wiley & Sons, Inc.
- Jiang Y, Boehler CN, Nonnig N, Duzel E, Hopf J, Heinze H, Schoenfeld MA (2008): Binding 3-D object perception in the human visual cortex. *J Cogn Neurosci* 20:553–562.
- Kleinschmidt A, Buchel C, Zeki S, Frackowiak RSJ (1998): Human brain activity during spontaneously reversing perception of ambiguous figures. *Proc R Soc Lond B* 265:2427–2433.
- Kourtzi Z, Erb M, Grodd W, Bulthoff HH (2003): Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cereb Cortex* 13:911–920.
- Lumer ED, Friston KJ, Rees GE (1998): Neural correlates of perceptual rivalry in the human brain. *Science* 280:1930–1934.
- Lumer ED, Rees GE (1999): Covariation of activity in visual and prefrontal cortex associated with subjective visual perception. *Proc Natl Acad Sci USA* 96:1669–1673.
- Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RBH (1995): Object-related activities revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci USA* 92:8135–8139.
- Maruyama M, Kobayashi T, Katsura T, Kuriki S (2003): Early behavior of optokinetic responses elicited by transparent motion stimuli during depth-based attention. *Exp Brain Res* 151:411–419.

-
- Mestre DR, Masson GS (1997): Ocular responses to motion parallax stimuli: The role of perception and attention factors. *Vis Res* 37:1627–1641.
- Moreno-Bote R, Shpiro A, Rinzel J, Rubin N (2008): Bi-stable depth ordering of superimposed moving gratings. *J Vis* 8:1–13.
- Muckli L, Singer W, Zanella FE, Goebel R (2002): Integration of multiple motion vectors over space: An fMRI study of transparent motion perception. *NeuroImage* 16:843–856.
- Natsukawa H, Kobayashi T (2012): Dynamical cortical activations associated with saccade execution: A normalized integrative fMRI-MEG study. *Adv Biomed Eng* 1:27–35.
- Niemann T, Ilg UJ, Hoffmann K-P (1994): Eye movements elicited by transparent stimuli. *Exp Brain Res* 98:314–322.
- Peuskens H, Claeys KG, Todd JT, Norman JF, Hecke PV, Orban GA (2004): Attention to 3-D shape, 3-D motion, and texture in 3-D structure from motion displays. *J Cogn Neurosci* 16: 665–682.
- Qian N, Andersen RA (1994): Transparent motion perception as detection of unbalanced motion signals II. *Physiology. J Neurosci* 14:7367–7380.
- Rosenberg A, Wallisch P, Bradley DC (2008): Responses to direction and transparent motions stimuli in area FST of the macaque. *Vis Neurosci* 25:187–195.
- Sarvas J (1987): Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys Med Biol* 32:11–22.
- Schütz AC (2011): Motion transparency: Depth ordering and smooth pursuit eye movements. *J Vis* 11:1–19.
- Sekihara K, Nagarajan SS (2008): Adaptive spatial filters for electromagnetic brain imaging. Berlin Heidelberg: Springer-Verlag.
- Waldorp LJ, Huizenga HM, Dolan CV, Molenaar PCM (2001): Estimated generalized least squares electromagnetic source analysis based on a parametric noise covariance model. *IEEE Trans Biomed Eng* 48:737–741.
- Wallach H, O'Connell DN (1953): The kinetic depth effect. *J Exp Psych* 45:205–217.
- Watanabe K (1999): Optokinetic nystagmus with spontaneous reversal of transparent motion perception. *Exp Brain Res* 129: 156–160.
- Welchman AE, Deubelius A, Conrad V, Bühlhoff HH, Kourtzi Z (2005): 3D shape perception from combined depth cues in human visual cortex. *Nat Neurosci* 8:820–827.