



Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth



Modulus and direction of the neural current vector identify distinct functional connectivity modes between human MT+ areas

Masaki Maruyama^{a,*}, Andreas A. Ioannides^{a,b}

^a Laboratory for Human Brain Dynamics, RIKEN Brain Science Institute, 2-1 Hirosawa, Wakoshi, Saitama 351-0198, Japan

^b Laboratory for Human Brain Dynamics, AAI Scientific Cultural Services Ltd., 33, Arch. Makarios III Avenue, Nicosia 1065, Cyprus

ARTICLE INFO

Article history:

Received 17 January 2010

Received in revised form 7 July 2010

Accepted 10 July 2010

Keywords:

Magnetoencephalography

Functional connectivity

Mutual information

hMT+

Single trials

ABSTRACT

Reconstruction of neural current sources from magnetoencephalography (MEG) data provides two independent estimates of the instantaneous current modulus and its direction. Here, we explore how different information on the modulus and direction affects the inter-hemisphere connectivity of the human medial temporal complex (hMT+). Connectivity was quantified by mutual information values of paired time series of current moduli or directions, with the joint probability distribution estimated with an optimized Gaussian kernel. These time series were obtained from tomographic analysis of single-trial MEG responses to a visual motion stimulus. With a high-contrast stimulus, connectivity measures based on the modulus were relatively strong in the prestimulus period, continuing until 100 ms after stimulus onset. The strongest modulus connectivity was produced with a long lag (19 ms) of the right hMT+ after the left hMT+. On the other hand, connectivity measures based on direction were relatively strong after 100 ms, with a short delay of less than 6 ms. These results suggest that nonspecific and probably indirect communication between the homologous areas is turned, by the stimulus arrival, into more precise and direct communication through the corpus callosum. The orientation of the estimated current vector for the strong connectivity can be explained by the curvature of the active cortical sheet. The temporal patterns of modulus and directional connectivity were different at low contrast, but similar to those at high contrast. We conclude that the modulus and direction indicate distinct functional connectivity modes.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Sensory inputs are processed by functionally segregated brain regions. The resulting segregated sensory information must be integrated to produce unified perceptions, emotions, and behaviors (Damasio, 1989; Singer, 1993). Magnetoencephalographic (MEG) and electroencephalographic (EEG) recordings have the necessary time resolution to characterize the dynamics of integration in humans. Together with theoretical considerations, evidence from EEG and MEG recordings has suggested that linked activity can be modulated on time scales from a few milliseconds (Ioannides et al., 2005) to tens of milliseconds (reviewed by Friston, 2000; Varela et al., 2001) or many hundreds of milliseconds (Ioannides et al., 2004). Rodriguez et al. (1999) identified synchronized gamma oscillations over distant brain regions during periods corresponding to face perception (230 ms after stimulus presentation) and motor response (650 ms), separated by a period of desynchronization (500 ms). This

finding of Rodriguez et al. (1999) implies that temporal characteristics of linked activity can identify distinct integration modes.

Using the high temporal resolution of MEG recordings, the delay in linking between paired regional activities has also been examined, since it is indicative of an activation shift from one area to the other. It is thus possible to identify activation shift processing from areas higher in the visual hierarchy to the primary visual area (V1). Such an activation shift has been identified from the human medial temporal complex (hMT+) to the primary/secondary visual area (V1/V2) with low-contrast stimuli (Maruyama et al., 2009) and, with face stimuli in the upper visual field (Liu and Ioannides, 2006), from the fusiform face area to V1/V2.

Early measures of linked activity were based on linked signals between EEG sensors (Gevins et al., 1981) or MEG sensors (Friston, 1997). Recently, as methods for estimating brain activity from non-invasive electrographic signals, namely EEG (Astolfi et al., 2007) and MEG (Ioannides et al., 2000) signals, have become more established, linked activity between specific brain areas has been estimated. The electromagnetic inverse problem solutions are not unique, (for a review, see Hamalainen et al., 1993) and so raise theoretical questions about source modeling techniques and their ability to localize neural currents from EEG and MEG data. These concerns can be addressed by approaching the biomagnetic inverse problem

* Corresponding author. Present address: CEA/DSV/I²BM/NeuroSpin, INSERM U.992 - Cognitive Neuroimaging Unit, Bât 145 - Point Courrier 156, Gif sur Yvette, F-91191, France. Tel.: +33 01 69 08 95 00; fax: +33 01 69 08 79 73.

E-mail address: masaki1974@gmail.com (M. Maruyama).

as a probabilistic estimation problem (Ioannides et al., 1990) and implementing it by using magnetic field tomography (MFT). This, together with post-MFT statistical analysis, can identify active brain regions with excellent localization accuracy (Moradi et al., 2003). Dynamic imaging of coherent sources (DICS) is a spatial filtering method that can also localize coherently activating areas from MEG and EEG signals (Gross et al., 2001).

It is generally believed that much of the MEG signal is generated by currents along the interior of apical dendrites of postsynaptic pyramidal neurons (Murakami et al., 2002), which are oriented roughly perpendicular to the cortical sheet of gray matter. The polarity of the current vector depends on whether the postsynaptic potentials of the neurons are excitatory or inhibitory (Linás and Nicholson, 1974) and whether the postsynaptic neurons receive the synaptic input near the apical dendrites or the cell body (Kandel et al., 2000). Furthermore, when the banks of opposite sulci are parallel and close to each other, the simultaneous activation of each bank (by the same mechanism) will generate currents with opposite direction. Inevitably, a region of interest (ROI) centered nearby will receive some contributions of one polarity and some of the opposite polarity.

Since the cortex is convoluted with numerous sulci and gyri, the current can flow with any orientation, depending on the location of the current flow in the cortical sheet. Therefore, the orientation of the current vector estimated from MEG signals can be explained by the orientation of the cortical sheet at the location of the current vector. Following this general assumption, Lin et al. (2006) obtained more focal source localization results by applying a loose orientation constraint to the estimation of current distribution. Perfetti et al. (2007) found that the current vector orientations for low (1–15 Hz)- and high (15–30 Hz)-frequency MEG signals were different, suggesting that distinct neural populations underlie the different frequency responses. Ioannides et al. (2005) observed different distributions of current vector direction, depending on saccadic eye movements. Taken together, the findings of these studies converge to the conclusion that current vector direction is potentially very useful for gaining specific knowledge about the neural generators of EEG and MEG signals.

MEG and EEG studies have characterized neural states mostly by measures of activity intensity. Some of these measures were based on the amplitude of the sensor signal, and others on the modulus of the estimated current density vector or its projection onto a certain direction. Our previous study on functional connectivity also indicated a neural state by a projected component of the current vector obtained from a subset of trials onto the averaged direction over all trials (Maruyama et al., 2009). These measures did not reflect purely directional changes of the current vector. As a consequence, possible distinct modes of connectivity could not be disentangled by using the independent estimates of current vector modulus and direction.

Motivated by the considerations above, we independently computed statistically linked activities based on the current vector modulus or its direction. The unique new contribution of this study is its independent estimation of connectivity using either modulus- or direction-based measures. We used MFT estimates of single-trial-activity MEG data for the left and right hMT+, with the corresponding well-separated ROIs already identified by their motion sensitivity in an earlier analysis of the average MEG signal obtained during the same experiment (Maruyama et al., 2009). It is known that the hMT+ consists of two sub-regions: the middle temporal (MT) area which is believed to project to the medial superior temporal (MST) area. The receptive fields of MT areas are restricted to the contralateral visual field while the MST areas receive input from the ipsilateral field as well (Huk et al., 2002). The MST areas are more sensitive than the MT areas to global flow structure, such as the expansion of motion stimuli (Smith et al., 2006). While the dif-

ferent activations of adjacent sub-regions have been distinguished spatially in functional magnetic resonance imaging (fMRI) studies, it has not been easy to do this in MEG studies that focused on spatial distribution of neural current amplitudes. However, when the cortical sheet curved sharply between the adjacent regions so that the current flow in the MT area was restricted toward a different direction from the flow in the MST area, the direction of estimated current vector might be sensitive to different activations of the two regions. We therefore used the current vector direction as a possible indicator for differentiating contributions from distinct sub-regions in the hMT+ areas. The sensitivity differential between measures that make use of the direction and modulus of the current vector becomes more pronounced when the activity of single trials is measured, because variability across trials can be utilized. In contrast, when a signal averaged over multiple trials is used, the estimates of regional activation are necessarily confounded by the mixing of contributions from nearby sources that are not precisely time-locked to the stimulus.

Mutual information (MI) (Shannon, 1948) estimates give a statistical relationship between the activity in two areas without directly implying an anatomical connection between them; they therefore express what has been termed “functional connectivity” (Friston, 1994). Whereas a correlation coefficient analysis detects only linear dependencies, an MI analysis accounts for both linear and nonlinear relationships. It is advantageous to eliminate the need to assume a linear relationship, because fMRI studies have found that nonlinear connectivity is present (Friston et al., 1995; Buchel and Friston, 1997; Friston, 2002). However, MI analysis may be less sensitive to weak linear connectivity than other methods of functional connectivity analysis, such as correlation coefficient or phase synchronization analyses (David et al., 2004). This low sensitivity is a consequence of the difficulty of estimating the joint probability density (JPD) distribution of two areas’ activities from a limited sample size. Although our previous study and other MEG studies have estimated JPD by a histogram method, Moon et al. (1995) have already suggested that a Gaussian kernel estimator can be used for precise MI analysis. Therefore, the current study used a Gaussian kernel estimator. To compute the optimal smoothing length of the Gaussian kernel, Moon et al. (1995) approximated the sample distribution by a normal distribution. However, it is already known that the normal approximation leads to over-smoothing in the case of a multimodal sample distribution (Silverman, 1986), which can result in an underestimation of MI. Moreover, when a statistical linkage exists, the sample distribution does not follow a normal distribution. Here, we optimized the choice of smoothing length using the distribution-free method of likelihood cross-validation (LH; Silverman, 1986).

We hypothesized that distinct modes of communication between two segregated areas can be revealed by an MI analysis based on the modulus and direction of the current vector, possibly highlighting different stages of processing at different latency ranges. We tested this hypothesis for a well-defined pair of neural systems: the left and right hMT+. These areas were expected to communicate strongly, since they are anatomically well-connected via the corpus callosum, as demonstrated in monkeys (Van Essen et al., 1982). A motion stimulus in one visual hemifield is represented in both contralateral and ipsilateral hMT+ areas in humans (Huk et al., 2002), suggesting that the two areas share common visual information. It has been claimed that motion signals are transferred from one hMT+ to the other not only via the corpus callosum but also via pathways through other areas (ffytche et al., 2000). A delay of the direct communication to the ipsilateral from contralateral hMT+ area was estimated as 3 ms (ffytche et al., 2000). We therefore examined a delay in functional connectivity between the responses of these two areas. In general, delays shorter than several milliseconds would be indicative of direct communication, longer

delays allow for the possibility of indirect connectivity. The specific hypothesis we tested here was that direct and indirect communications are disentangled by independent estimations of current modulus and direction.

2. Materials and methods

2.1. MEG experiment and analysis

We used single-trial MEG data from a recent study that focused on the analysis of average signals (Maruyama et al., 2009). Therefore, in this paper we provide only a brief overview of the MEG experiment and signal processing for estimation of the current density vector in the hMT+.

Eight healthy, right-handed men with a mean age of 30 participated in the experiment. All experimental procedures were undertaken with the understanding and written consent of each subject, conformed to the Code of Ethics of the World Medical Association (Declaration of Helsinki), and were approved by the RIKEN Ethics Committee.

A random dot pattern was presented on a screen. One circular region was presented at a distance of 8° along an axis oriented 45° to the horizontal and vertical, and one stimulus was presented centrally. The region had a radius of 4° , and within this circle 250 dots, each 0.15° in size, were placed. For the stationary stimulus, the dot pattern was fixed. For the motion stimulus, each dot moved along a radius toward the outside of the circle at a speed of $15^\circ/s$ and disappeared on reaching the perimeter, to be replaced by another randomly situated dot. High-contrast (0.8, by the Michelson definition) or low-contrast (0.2) stimuli were created by setting the dot luminance to 63 or 10.5 cd/m^2 , respectively, and the background luminance to 7 cd/m^2 . The subjects were instructed to fixate on a black square in the center of the screen. To maintain alertness, the subjects were asked to quickly lift the index finger of their right hand from a fiber optical button when the color of the fixation square changed to red. Both the random dot patterns and the red fixation square were presented for 300 ms with an inter-stimulus interval of $700 \pm 100 \text{ ms}$. During each run, each stimulus condition was presented either in one of the quadrants or in the central visual field 15 times in random order. The color of the fixation square changed 1, 2, or 3 times with random timing during the run. The timing of the color change was not related to any particular random dot pattern. Each subject performed 10 such runs.

The experiment presented high- and low-contrast motion stimuli to determine the feedforward and feedback processing between V1/V2 and hMT+, as published in our recent paper (Maruyama et al., 2009). In addition, each stimulus was presented repeatedly (150 times in total) in order to obtain a rich response variance among the trials. The present study used the trial variance to indicate the functional connectivity. We analyzed those MEG signals obtained when the high- and low-contrast motion stimuli were presented at the bottom-right location, where we had previously obtained responses with greater intensity in the left hMT+ under the high-contrast condition.

The magnetic field was measured with a 151-channel whole-head MEG system (Omega, CTF Systems Inc., Vancouver, Canada). Eye movement artifacts and the subject's heart function were simultaneously measured by electrooculography (EOG) and electrocardiography (ECG) for off-line noise reduction. The signals were sampled at 625 Hz. After DC-offset removal, the signals were band-pass-filtered below 200 Hz and above 3 Hz, and notch-filtered to eliminate power-line noise (50 Hz). We extracted trials between 100 ms before and 300 ms after the onset of the stimulus. The EOG signal was visually inspected and any trials that contained blinks or saccadic eye movements were discarded, leaving 138–148 trials from each subject. Independent component analysis was then

applied and the components correlated with the EOG and ECG signals were eliminated to remove heart-beat and eye-blink artifacts (Jahn et al., 1999). Magnetic resonance imaging (MRI) was performed with either a 1.5-T Magnetom Symphony (Siemens Inc., Erlangen, Germany) or ExcelArt (Toshiba Inc., Tochigi, Japan) MRI machine. The MEG coordinate system was transformed into the MRI coordinate system by the standard method used by our group (e.g., Liu and Ioannides, 2006; Maruyama et al., 2009).

We extracted tomographic estimates of source current from the MEG signals by using MFT. We used the standard MFT algorithm in which the order of the weighted minimum norm is set to zero, i.e., $p=0$ in the generalized mathematical formalization of MFT, as concisely described in Supplementary Materials. [Generalized description is provided by Poghosyan and Ioannides (2008, supplementary materials) and details by Ioannides et al. (1990) and Taylor et al. (1999).] The standard MFT has two adjustable parameters. One is the *a priori* probability weight, which compensates for the lead-field bias for superficial sources. The *a priori* probability weight is uniquely optimized, using computer-generated data, for the data of each run by taking into account head position relative to the sensors (Ioannides, 1994). The second is the regularization parameter, which resolves the conflicting requirements of high spatial accuracy and insensitivity to noise. The regularization parameter in MFT code is determined via a dimensionless quantity (ζ), which is chosen according to the noise level of the recording environment. For the details of mathematical formulation of the regularization parameter, see Ioannides et al. (1990). In the relatively quiet and stable environment at RIKEN, we were able to use the same value ($\zeta = 0.1$) for all our experiments, including the present one.

Four separate MFT computations were performed, in each case using partially overlapping hemispheric source spaces ($17 \times 17 \times 11$ grid points each) which completely covered the left, right, back and top (superior parts of the brain) (Ioannides, 2002). Each MFT computation uses a spherical conductor model for the conductivity of head (Grynszpan and Geselowitz, 1973), with the center of the conducting sphere in each case chosen to fit the inner surface of the skull in the appropriate hemisphere. Source currents were allowed only within the appropriate source space, i.e., the brain area of the corresponding MFT hemisphere (left, right, back and top). MFT was performed separately for each source space, after choosing 90 MEG sensors from the corresponding side. The solutions from all four source spaces were combined into a single, large source space which covered the whole brain, using the sensitivity-profile-modified current density values of the sensors from nearby points in the individual source spaces. The result was stored at a resolution of 9–12 mm, depending on the size of subject's head.

This study applied MFT to each trial: that is, the source current was estimated for single trials. The current density vector was estimated throughout the brain at each time slice.

We used the hMT+ ROIs defined in our previous MEG study (Maruyama et al., 2009). The ROIs were defined at foci showing a strong current density consistent in amplitude and direction across runs, as quantified by an estimator of the signal-to-noise ratio. We employed a contrast between moving and stationary stimuli known to favor selective activation in hMT+. For further details of the ROI definition, please refer to our previous study. The current density vector was regionally averaged within a 5-mm radius with an exponential weight (decay, 5 mm). This is the minimum ROI size that we could choose under the spatial resolution of current density estimation.

For a spherical conductor, MEG measurements are sensitive only to the tangential direction of the current vector (Grynszpan and Geselowitz, 1973). Therefore, the current density estimates that can be extracted from the MEG data are confined mostly to the

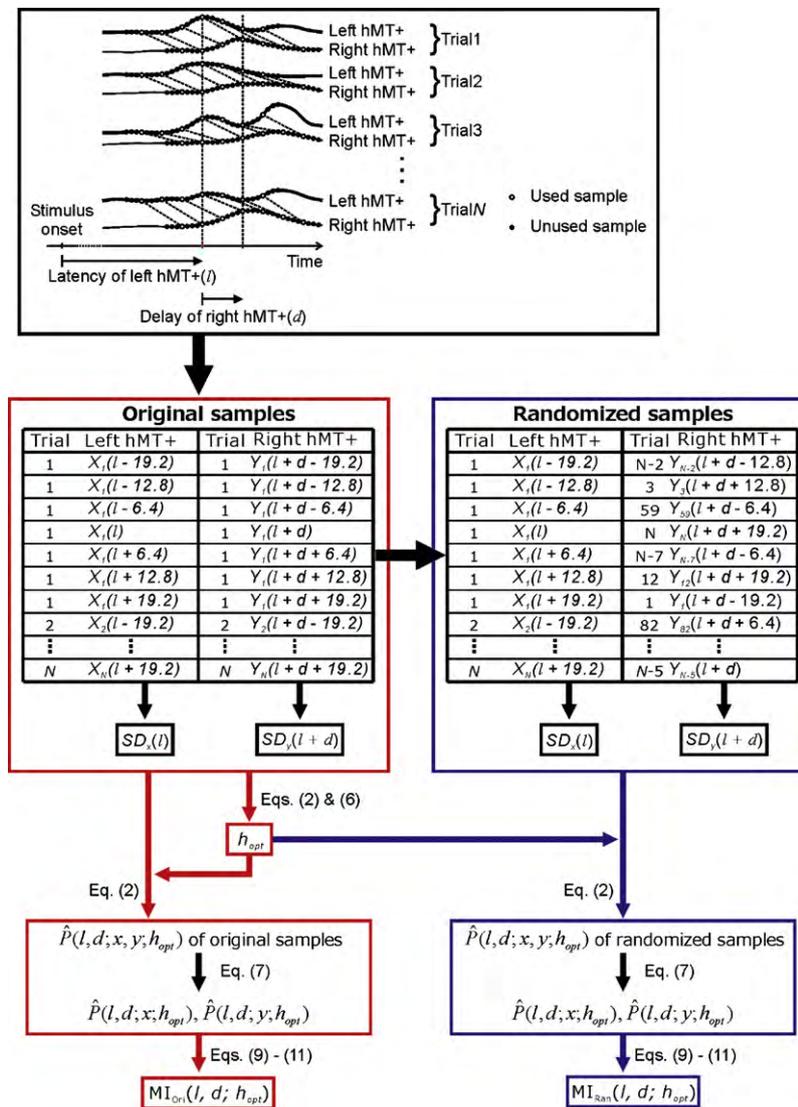


Fig. 1. Schematic diagram illustrating the procedure used for obtaining mutual information between the left hMT+ moduli at latency l and the right hMT+ at $l+d$, using the Gaussian kernel method. The diagram outlined in black shows the extraction of samples from time courses of current vector moduli over all trials in the left hMT+ (thick black curves) and the right hMT+ (thin black curves). First, moduli recorded at a sampling frequency of 625 Hz were resampled at intervals of three-sample (i.e., every fourth value was sampled). Open and filled symbols represent used and unused samples as a consequence of the re-sampling. Next, in each trial, seven samples were extracted within 19.2 ms before and after the left hMT+ latency $[X_i(l+w)]$, and within 19.2 ms before and after the right hMT+ delay $[Y_i(l+d+w)]$. Then the right hMT+ samples were paired with the left hMT+ samples in each trial with delay d , as shown in the diagram outlined in red. After normalization of the left and right hMT+ samples in relation to their standard deviation $[SD_x(l), SD_y(l+d)]$, the optimal smoothing length of the Gaussian kernel (h_{opt}) was computed by using the normalized samples. The optimized Gaussian kernel estimator was used to estimate the joint probability distribution of the left and right hMT+ $[\hat{P}(l, d; x, y; h_{opt})]$ and the probability distribution of each region $[\hat{P}(l, d; x; h_{opt}), \hat{P}(l, d; y; h_{opt})]$. Using the probability distributions, a mutual information value was estimated for the originally paired samples (MI_{ori}). To obtain the degree of overestimation of mutual information for the cases where no relationship was present, we also computed mutual information for samples randomized over time slices and trials (MI_{ran}) using the optimal smoothing length used for the original samples (diagram outlined in blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tangential direction, even though the actual neural current in the brain can flow in any direction. In this study, we computed the MI by using the current density vector confined to the tangential plane of a large sphere whose radius and center was chosen to fit the inner surface of the skull. We made separate computations for the modulus and the direction of this vector and compared them with each other. To gain insight into the anatomical explanation for the current vector direction, orientations of the closest sulci were compared with highly frequent directions for which a relatively high MI value was obtained.

2.2. Paired samples between left and right hMT+

The time-series samples of current vector moduli or directions were paired between the left and right hMT+. Here we describe the

pairing method over the trials. For the schematic illustrating of the pairing method, see Fig. 1.

Since the MEG signal recorded at the sampling frequency of 625 Hz was low-pass-filtered below 200 Hz during preprocessing, neighboring samples were dependent on each other. Given the potential risk that this analysis-induced sample dependency would distort the results, we first resampled at 6.4-ms intervals. The aliasing effect that this re-sampling may introduce will be confined to the frequency range between the effective Nyquist frequency due to re-sampling (78 Hz) and the original upper limit of 200 Hz. This effect is negligible because the strong biological signal is likely to be at lower frequencies, while the noise from the environment is effectively attenuated for frequencies higher than a few Hz by the passive reduction of environmental noise provided by the shielded room housing the MEG hardware. Next, the resampled time series

was segmented in order to examine the dynamics of functional connectivity. Longer segments include more samples, allowing more detailed estimation of the probability distribution with shorter smoothing lengths. In addition, long segments would be beneficial if a change of slow activation was too small to detect within a short period. However, longer segments also decrease the time resolution of our functional connectivity results. In order to choose a segment length, we considered the suggestion of Friston (1997), who showed in a MEG study that the significance of MI was increased with a segment length of around 40 ms. Liu and Ioannides (2006) have also shown the effectiveness of 48-ms MI segments. In view of these findings, we chose a segment length of 44.8 ms. We defined l as the latency of the left hMT+ from the stimulus onset, and d as the delay of the right hMT+ with respect to the left hMT+. We extracted seven resamples within 19.2 ms before and after the left hMT+ latency $[X_t(l+w); w = -19.2, -12.8, \dots, 0, \dots, 19.2 \text{ ms}; t = 1, 2, \dots, N_t; N_t$ is the number of trials] and within 19.2 ms before and after the right hMT+ delay $[Y_t(l+d+w)]$. Finally, we paired the left and right hMT+ samples of each trial, as $\{X_t(l+w), Y_t(l+d+w)\}$.

We used seven sample pairs from each trial. The total number of paired samples for each calculation of MI ranged between 966 (7 samples \times 138 trials) and 1036 (7 samples \times 148 trials). A set of paired samples was made in 6.4-ms steps over a latency range from -100 to 300 ms and a delay range from -64 to 64 ms. MI was computed for each latency and delay. The paired samples were used to estimate the JPD distribution, as described in the following subsection.

2.3. Estimation of the joint probability density distribution

The estimation of the JPD distribution from a limited sample size is a key issue in MI analysis. The current study used a Gaussian kernel estimator. We optimized the smoothing length by the LH method. Here, we first defined the kernel for the modulus samples. Next, we defined the kernel for the tangential direction, which is not identical to the kernel for the modulus because direction is a circular variable (i.e., the angles φ and $\varphi + 2\pi$ are equivalent in radians). We next describe the method of kernel optimization.

The JPD of the current vector moduli in the left hMT+, x , and in the right hMT+, y , was calculated by using the Gaussian kernel estimator, defined as

$$\hat{P}(l, d; x, y; h_x, h_y) = \frac{1}{N_w N_t h_x h_y} \sum_{t=1}^{N_t} \sum_w \frac{1}{2\pi} \exp \left[-\frac{\{x - X_t(l+w)\}^2}{2 \cdot h_x^2} - \frac{\{y - Y_t(l+d+w)\}^2}{2 \cdot h_y^2} \right], \quad (1)$$

where $X_t(l+w)$ and $Y_t(l+d+w)$ denote the samples of the current vector moduli in the left and right hMT+, w indicates one of seven time slices in the segmented window, N_w is the number of time slices in each trial (i.e., seven time slices), and N_t is the number of trials. h_x and h_y represent the smoothing lengths of the Gaussian kernel. The optimal smoothing lengths depend on the dispersion of modulus samples as well as their scales of amplitudes. We know empirically that the dispersion of current density distributions varies between regions in MEG studies; the optimal smoothing length should therefore be different among regions. However, we wanted to reduce the number of free parameters to be optimized in order to reduce the computational load. To do this, we normalized the samples in relation to their standard deviation. As a result, the Gaussian kernel estimator becomes as follows:

$$\hat{P}(l, d; x, y; h) = \frac{1}{N_w N_t h^2} \sum_{t=1}^{N_t} \sum_w \frac{1}{2\pi} \exp \left[-\frac{\{(x - X_t(l+w))/SD_x(l)\}^2 + \{(y - Y_t(l+d+w))/SD_y(l+d)\}^2}{2 \cdot h^2} \right], \quad (2)$$

where $SD_x(l)$ and $SD_y(l+d)$ are the standard deviations of samples $X_t(l+w)$ and $Y_t(l+d+w)$, respectively, and h represents the common smoothing length between the two regions. Note that the h is optimized for the normalized samples, and thus its scale does not need to be comparable to the original samples, unlike the h_x and h_y .

The Gaussian kernel estimator of the JPD between the current vector direction in the left hMT+, φ_x , and that in the right hMT+, φ_y , is given as

$$\hat{P}(l, d; \varphi_x, \varphi_y; h) = \frac{1}{N_w N_t h^2} \sum_{t=1}^{N_t} \sum_w \frac{1}{2\pi} \exp \left[-\frac{\{Dx_t(l+w)/SD_x(l)\}^2 + \{Dy_t(l+d+w)/SD_y(l+d)\}^2}{2 \cdot h^2} \right],$$

$$Dx_t(l+w) = |\pi - |\pi - |\varphi_x - \Phi_{x_t}(l+w)|||,$$

$$Dy_t(l+d+w) = |\pi - |\pi - |\varphi_y - \Phi_{y_t}(l+d+w)|||, \quad (3)$$

where $\Phi_{x_t}(l+w)$ and $\Phi_{y_t}(l+d+w)$ denote the samples of current vector direction in the left and right hMT+. $Dx_t(l+w)$ represents the circular difference between φ_x and $\Phi_{x_t}(l+w)$, and similarly for $Dy_t(l+d+w)$. $SD_x(l)$ is the standard deviation of the circular samples $\Phi_{x_t}(l+w)$, given by Mardia and Jupp (2000) and Batschelet (1981):

$$SD_x(l) = \sqrt{-2 \log \bar{R}(l)},$$

$$\bar{C}(l) = \frac{1}{N_w N_t} \sum_{t=1}^{N_t} \sum_w \cos \Phi_{x_t}(l+w),$$

$$\bar{S}(l) = \frac{1}{N_w N_t} \sum_{t=1}^{N_t} \sum_w \sin \Phi_{x_t}(l+w),$$

$$\bar{R}(l) = \sqrt{\bar{C}^2(l) + \bar{S}^2(l)}. \quad (4)$$

and similarly for $SD_y(l+d)$. These standard deviations were used for normalization of the circular differences in Eq. (3) for the same reason as in the modulus analysis.

The LH optimization method is identical for the modulus and circular samples. We describe the method by referring to the modulus variables (x and y), and the method for the circular samples can be found by replacing those variables with directional ones (φ_x and φ_y).

The difference between the JPD estimated from the samples $[\hat{P}(l, d; x, y; h)]$ and the true JPD $[P(l, d; x, y)]$ is quantified as the information distance (Silverman, 1986), defined as

$$ID(l, d; h) = \iint P(l, d; x, y) \ln \left\{ \frac{P(l, d; x, y)}{\hat{P}(l, d; x, y; h)} \right\} dx dy. \quad (5)$$

Although $ID(l, d; h)$ contains the function $P(l, d; x, y)$, which we cannot compute because the number of samples was limited, we can minimize ID without knowing it because the optimal smoothing length maximizes the score function:

$$CV(l, d; h) = \sum_{t=1}^{N_t} \sum_w \ln \hat{P}_{-t, -w}[X_t(l+w), Y_t(l+d+w); h], \quad (6)$$

where $\hat{P}_{-t,-w}[X_t(l+w), Y_t(l+d+w)]$ is the probability density for $x = X_t(l+w)$ and $y = Y_t(l+d+w)$, estimated by using all samples except the paired samples $\{X_t(l+w), Y_t(l+d+w)\}$. We sought the optimal value of $h(h_{opt})$ that maximized the score function.

Using the optimal smoothing length, we estimated the JPD distribution according to Eq. (2) for the modulus samples [and according to Eq. (3) for the directional samples]. The probability of each modulus variable, $\hat{P}(l, d; x; h_{opt})$ and $\hat{P}(l, d; y; h_{opt})$, was obtained by the integration of the JPD over y or x :

$$\hat{P}(l, d; x; h_{opt}) = \int \hat{P}(l, d; x, y; h_{opt}) dy,$$

$$\hat{P}(l, d; y; h_{opt}) = \int \hat{P}(l, d; x, y; h_{opt}) dx. \quad (7)$$

The estimated probabilities $[\hat{P}(l, d; x, y; h_{opt}), \hat{P}(l, d; x; h_{opt}),$ and $\hat{P}(l, d; y; h_{opt})]$ were then used to calculate MI values.

2.4. Mutual information analysis

MI analysis measures statistical relatedness in an entropy form. The entropy of the current vector modulus in the left hMT+ is defined as

$$H_L(l, d; h_{opt}) = - \int \hat{P}(l, d; x; h_{opt}) \ln \hat{P}(l, d; x; h_{opt}) dx, \quad (8)$$

and similarly for the right hMT+ [$H_R(l, d)$]. The entropy of the joint current vector moduli in the two regions is

$$H_{LR}(l, d; h_{opt}) = - \iint \hat{P}(l, d; x, y; h_{opt}) \ln \hat{P}(l, d; x, y; h_{opt}) dx dy. \quad (9)$$

MI quantifies the relatedness between the two regions as

$$MI(l, d; h_{opt}) = H_L(l, d; h_{opt}) + H_R(l, d; h_{opt}) - H_{LR}(l, d; h_{opt}). \quad (10)$$

MI for the circular samples can be found by replacing the modulus variables by the directional ones (φ_x and φ_y). The MI value increases as the probability distribution of the modulus or direction of the current vector in the left hMT+ (right hMT+) becomes more dependent on that in the right hMT+ (left hMT+). The MI value reaches a maximum when the states of the two areas are identical to each other. The upper bound of MI is $H_L(l, d; h_{opt}) + H_L(l, d; h_{opt}) - H_{LL}(l, d; h_{opt})$, which depends on the probability distribution $\hat{P}(l, d; x; h_{opt})$. The MI value approaches zero as the states of the two areas become independent.

It is known that measurements of low MI are typically overestimated because of limited sample size (Samengo, 2002). The overestimation tends to be larger for smaller sample sizes relative to the bin number in histogram methods. Our simulations based on computer-generated samples also indicated overestimation of low MI values, but only by small amounts (< 0.06). It nevertheless implied that the overestimation could be reduced by using the Gaussian kernel. To facilitate the removal of overestimation, we used a randomization method that estimated the bias for cases where no relationship was present. First, we paired the samples of the left hMT+ and right hMT+ after randomizing across the trial order ($t' = 1, 2, \dots, N_t$ in random order) and the time series of the seven samples within the segment period ($w' = -19.2, -12.8, \dots, 0, \dots, 19.2$ ms in random order), i.e., $\{x_t(l+w), y_{t'}(l+d+w')\}$. Then an MI value of the randomized samples [$MI_{Ran}(l, d; h_{opt})$] was computed, using the smoothing length optimized based on the original samples. This procedure helps shorten the computational time of JPD estimation, and we confirmed that the MI analysis was not degraded. We repeated the randomization 100 times, producing 100 values of $MI_{Ran}(l, d; h_{opt})$ for each MI value of the original samples, $MI_{Ori}(l, d; h_{opt})$. The mean of the

MI_{Ran} values ($\overline{MI}_{Ran}(l, d; h_{opt})$) reflects the expectation of overestimation given no statistical relationship. We defined the corrected MI value, MI_{Cor} , as the difference between the MI_{Ori} and \overline{MI}_{Ran} values: $MI_{Cor}(l, d; h_{opt}) = MI_{Ori}(l, d; h_{opt}) - \overline{MI}_{Ran}(l, d; h_{opt})$. As a result of the subtraction operation, both the upper and lower limits of our MI estimation (MI_{Cor}) are lower than those of MI defined by Eq. (10).

The set of MI_{Ran} values was also used to determine the presence of a statistical relationship. We deemed a MI_{Ori} value above the 0.95 quantile of MI_{Ran} values to be significant (one-tailed test, $p < 0.05$). We did not deem a low MI_{Ori} value to be significant, since an MI value close to zero indicates no relationship.

2.5. Simulation for comparisons of analysis methods

The optimization of Gaussian kernel by the LH method has already been discussed from a theoretical perspective, but its effectiveness in MI analysis has been poorly understood. Moreover, the method has never been applied to circular samples. Therefore, we first tested the method by using computer-generated samples. Its effectiveness for real samples was examined by comparing with a histogram method which has often been used in previous MEG studies. Next, we compared with a Gaussian kernel method optimized using a least-squares cross-validation (LS) method in which the sample distribution is approximated by the normal distribution (Moon et al., 1995). Finally we compared with correlation coefficient analyses, which might be more sensitive than MI analysis to weak linear connectivity (David et al., 2004). Summarizing the results of our simulations, we conclude that MI analysis using LH kernel optimization reduced the MI underestimation that often appears when a multimodal JPD distribution is approximated by the normal distribution. The simulation also showed that our method can reduce the standard deviation (SD) of MI estimation relative to the previous histogram and kernel methods, when connectivity is weak ($MI < 0.1$). For the details of our simulation, see [Supplementary Materials](#).

3. Results

3.1. Joint probability density distribution of current vector moduli

JPD distributions obtained from the modulus samples are shown in Fig. 2 (left panels) for three representative subjects presented with the high-contrast stimulus. We chose the JPD distribution in which a peak MI value was found in relation to the latencies and delays (arrows in Fig. 3A, left panels). The distributions we obtained were unimodal and skewed. They were elongated toward the upper right (panels in Fig. 2), producing relatively high MI_{Ori} values (Sub. 1, 0.098; Sub. 2, 0.044; Sub. 3, 0.073). This elongation was slight when the probability distribution of the original samples is compared with the probability distribution of the randomized samples (Fig. 2, right panels). However, a significant MI can appear even for such slight difference of JPD distributions. Indeed, the MI_{Ori} values were significantly high relative to the set of 100 MI_{Ran} values ($p < 0.01$). In all eight subjects, the JPD distributions were almost always unimodal and skewed at any latencies and delays.

To examine the influence of stimulus presentations repeated 150 times during the experiment, we analyzed statistical relationships between the current vector modulus and the order of trials. The influence was seldom found to be significant over the subjects. For the details, see [Supplementary Materials](#).

3.2. Functional connectivity at high contrast based on current vector moduli

The time courses of MI_{Cor} values at high contrast for subjects 1–3 are shown in Fig. 3A (left panels). The horizontal axes show

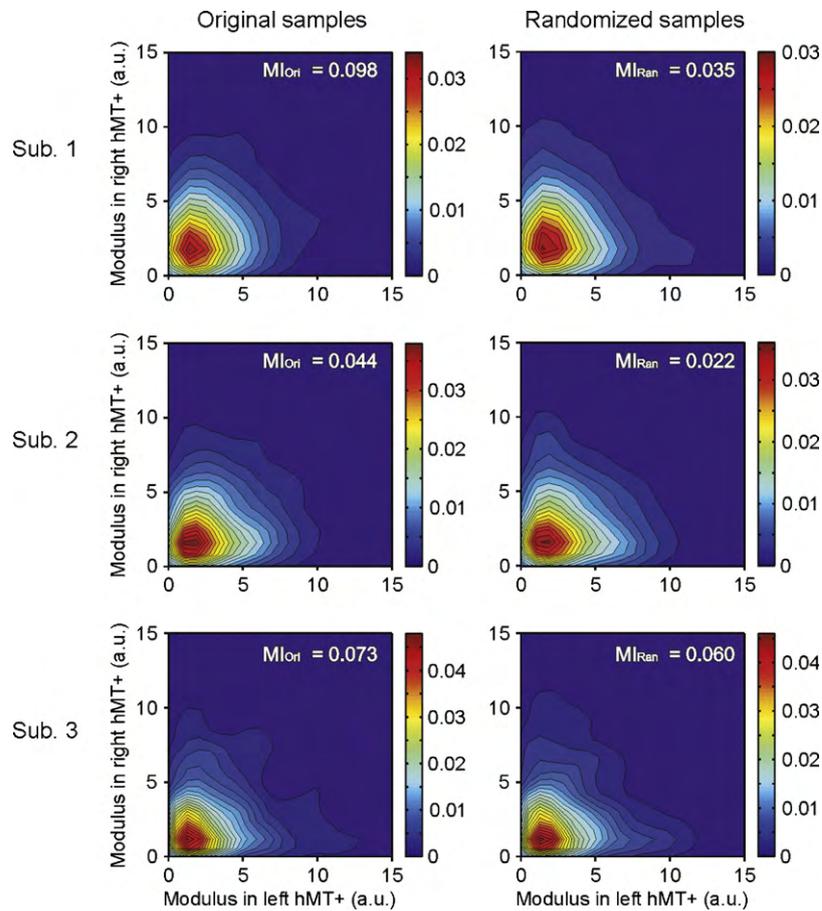


Fig. 2. Joint probability density (JPD) distributions of current vector moduli in three subjects at representative latencies and delays for which a peak corrected MI value (MI_{Cor}) was found (see Fig. 4A, open arrows). A high-contrast motion stimulus was presented at the bottom-right location. Horizontal and vertical axes represent the moduli of the current density vector in the left and right hMT+, respectively, in a.u. (arbitrary units defined by the MFT analysis). The JPDs are indicated by the colors as specified in the color bars. *Left panel:* Results from the original paired samples. *Right panel:* Results from the pairs of samples after randomization. The latency of the left hMT+ from stimulus onset and the delay of the right hMT+ with respect to the left hMT+ were, respectively, 104.8 and -19.2 ms in Sub. 1, 85.6 and -12.8 ms in Sub. 2, and 104.8 and -12.8 ms in Sub. 3. The MI value of each JPD distribution is noted at the top right of each panel. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the latency of the left hMT+ from the onset of the stimulus, and the vertical axes denote the delay of the right hMT+ relative to the left hMT+. Relatively high MI_{Cor} values (compared with other values obtained in this study) were found in all three subjects during the prestimulus period, and the MI_{Cor} values continued to be relatively high until a latency of around 100 ms. Afterward, the functional connectivity tended to vanish. We obtained MI_{Cor} values of less than 0.1.

The MI_{Ori} value at each latency and delay was statistically tested by comparing it with the set of 100 MI_{Ran} values. The latencies and delays of significant connectivity are shown in the right panels of Fig. 3A as yellow ($p < 0.01$) and orange ($p < 0.05$) areas. The dark brown areas are those where the connectivity was not significant ($p \geq 0.05$). The MI_{Ori} values obtained in the late period (after about 100 ms) tended not to be significant.

The average across all eight subjects also shows relatively strong connectivity during the prestimulus period, which continued until a latency of around 100 ms (Fig. 3B, left panel). Interestingly, the maximum MI_{Cor} was found at a negative delay (latency, 104.8 ms; delay, -19.2 ms; $MI_{Cor} = 0.012$). In other words, the strongest functional connectivity corresponds to the ipsilateral hMT+ (i.e., right hMT+ relative to the bottom-right visual field) leading the contralateral hMT+. The right panel of Fig. 3B shows the percentage of subjects exhibiting a significant MI_{Ori} ($p < 0.05$). Six of eight subjects exhibited a significant MI_{Ori} at the identical latency and delay to the maximum MI_{Cor} , shown by the yellow area. Significant

connectivity was found in more subjects during the prestimulus and early stimulus periods (orange and red areas) than during the late stimulus period (dark brown areas). These results are entirely consistent with the connectivity patterns identified in individual subjects.

3.3. Joint probability distribution of current vector directions

The JPD distributions of the tangential current vectors in the left and right hMT+ were often multimodal. Fig. 4 (left panels) shows representative JPD distributions of the original samples for which a peak of MI value was found in relation to the latencies and delays (open arrows in Fig. 6A, left panels). The high-contrast motion stimulus was presented. Each JPD distribution possessed two peaks (Subs. 1 and 2) or more than two (Sub. 3). It should be noted that the smoothing length of Gaussian kernel was optimized to minimize a difference between estimated and true JPD, hence the multimodal distributions were not likely to be an artifact of our estimation method. Indeed, our simulation results of MI estimation demonstrated a successful optimization for multimodal distributions (see Supplementary Materials). Rather, our results demonstrated that an approximation of sample distribution by the normal distribution could be erroneous for directional samples in MEG studies, and thus a distribution-free method such as the LH method is very important for the optimization of Gaussian kernel. The most frequent direction (open arrow) was nearly opposite the second-most fre-

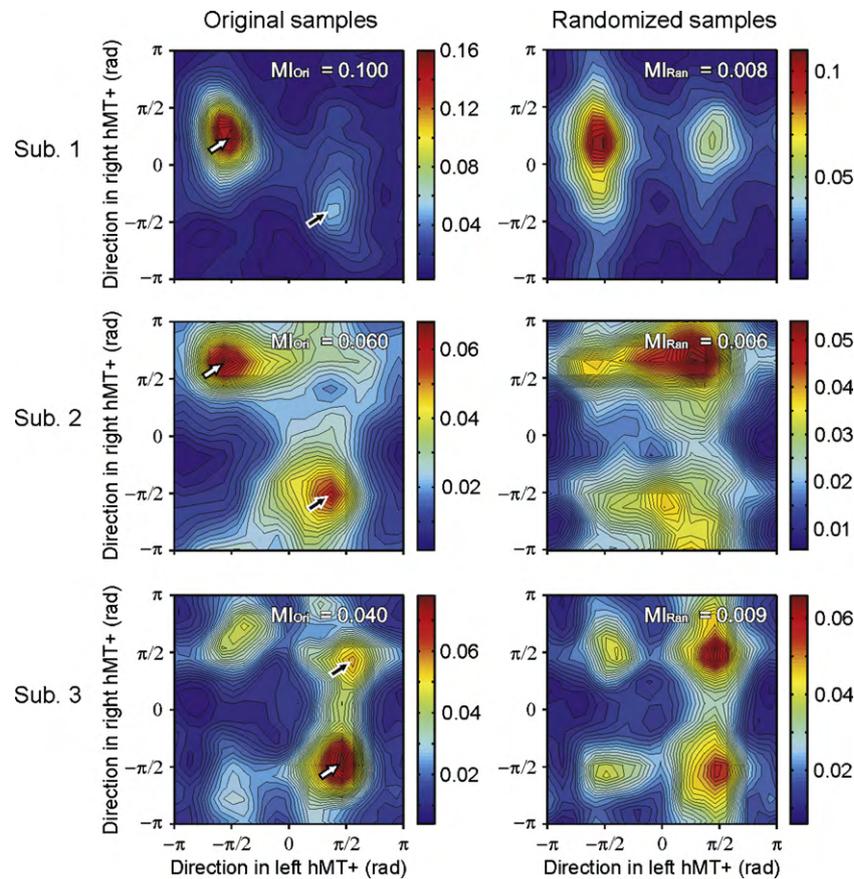


Fig. 4. JPD distributions of the tangential directions at representative latencies and delays for which a peak of MI_{Cor} value was found (see Fig. 6A, open arrows). A high-contrast motion stimulus was presented. Horizontal and vertical axes represent the direction of the current vector in the left and right hMT+, respectively. The JPDs are indicated by the colors in units of $(rad)^{-2}$, as specified in the color bars. *Left panel:* Results from the original paired samples. *Right panel:* Results from the pairs of samples after randomization. The latency of the left hMT+ from stimulus onset and the delay of the right hMT+ with respect to the left hMT+ were, respectively, 181.6 and -6.4 ms in Sub. 1, 143.2 and 0 ms in Sub. 2, and 111.2 and 0 ms in Sub. 3. Open and filled arrows in the left panels denote the most and second-most frequent directions of the JPD distributions; the highly frequent directions are superimposed on the MRI images in Fig. 5. The MI value of each JPD distribution is indicated at the top right of each panel. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

were candidate sources of the regionally averaged current. However, of these sulci, we identified an ALITS segment that was nearly orthogonal to the high-frequent directions, suggesting that it was the main current source. Similarly, among the three sulci within the ROI of the left hemisphere in Sub. 2, a vertical segment of the ALITS was almost perpendicular to both the most and second-most frequent directions (Fig. 5C, thick red line on cartoon panel); thus, this segment is most likely the current source.

In the right hMT+ of Sub. 1, the most frequent direction of the regionally averaged current was nearly orthogonal to the second-most frequent one (Fig. 5B). These directions probably reflected current sources along different segments of the curved ALITS, as shown by the thick red lines on the cartoon panels. In the right hMT+ of Sub. 2, the most frequent direction can be explained by the composition of two current vectors from separate segments of the ALITS (Fig. 5D). Possible current sources were not always identified on the ALITS. In Sub. 3, a segment nearly orthogonal to the most and second-most frequent current vectors was identified on the LOS in each hemisphere (Fig. 5E and F).

The results for the other subjects are similar to the results for Subs. 1–3, as can be seen by inspecting Fig. 5G–P. In each ROI we could identify sulcus segmentations orienting nearly orthogonal to the frequent directions, as expected from the general assumption that the neural currents that are responsible for the generation of the MEG signal flow perpendicularly to the cortical surface of a sulcus. The variation across subjects for the orientations of the sulci, associated with the frequent directions of current vectors, is con-

sistent with the study of Dumoulin et al. (2000). The origin of the variability in the frequently occurring current directions is unclear at this point; it is very likely a reflection of anatomical individual-variance in the terms of sulcus orientation.

3.5. Functional connectivity at high contrast based on current vector direction

Fig. 6A (left panels) shows the time courses of the MI_{Cor} value obtained from the analysis of the direction of the current vectors in Subs. 1–3 when they were presented with the high-contrast stimulus. Relatively high MI_{Cor} values were found after 100 ms. This is in sharp contrast to the results based on the moduli of the currents, where significant linked activity was identified before 100 ms (Fig. 3, left panels). The MI_{Cor} values of direction also ranged below 0.1, but they were generally higher than the values obtained by using the moduli. Accordingly, the periods of significant MI_{Ori} were longer with respect to the current vector directions (Fig. 6A, right panels) than with respect to the moduli (Fig. 3, right panels).

Averaging across all subjects showed that the relatively strong connectivity started at a latency of 100 ms and continued until a latency of 200 ms (Fig. 6B, left panel). The maximum MI_{Cor} value was 0.028 at a latency of 117.6 ms, a delay range between -10 and 10 ms centered on zero. Fig. 6B (right panel) displays the percentage of subjects exhibiting a significant MI_{Ori} ($p < 0.05$), showing that a significant MI_{Ori} was identified in all subjects in the latency range between 90 and 200 ms. There were also two other peaks with pos-

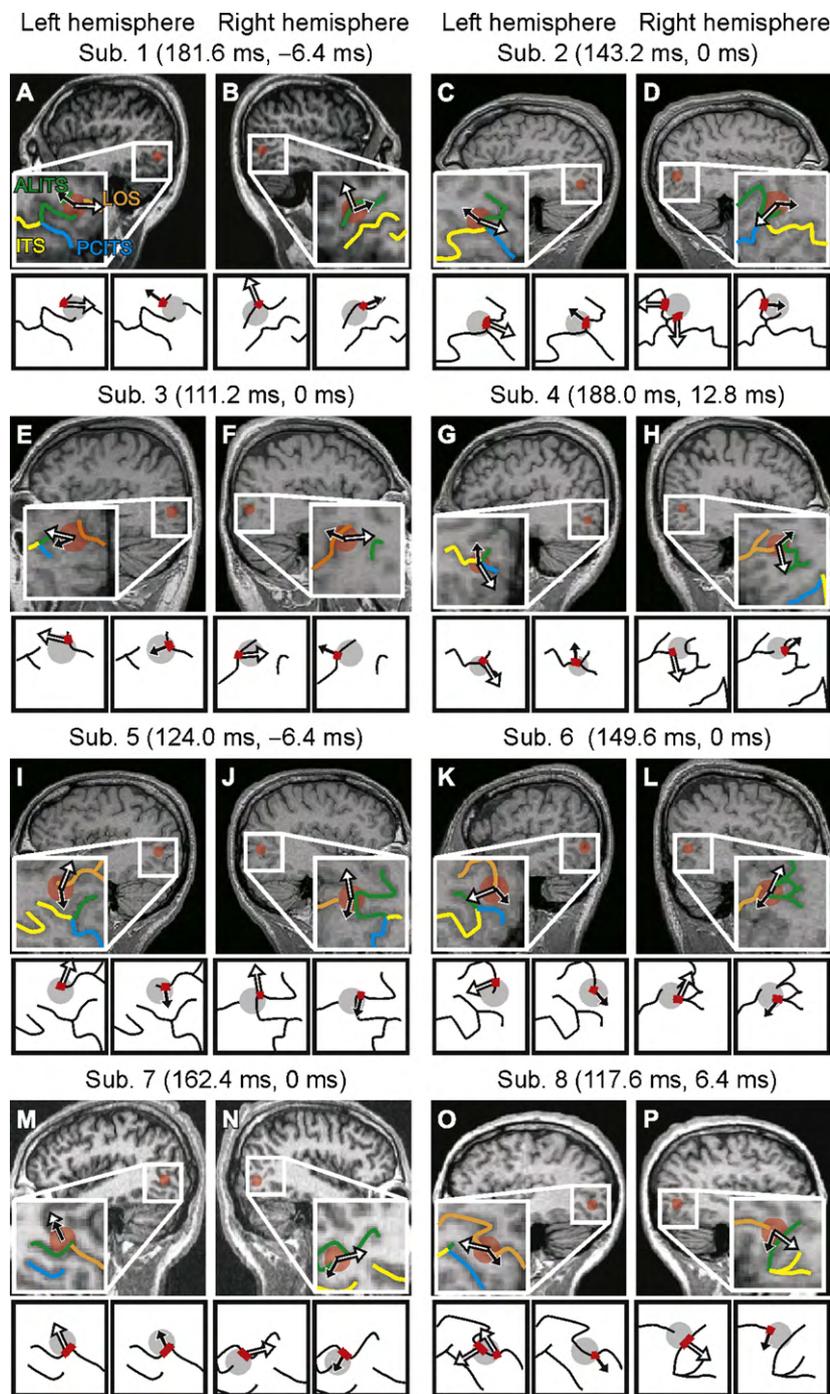


Fig. 5. Frequent directions of the current vector relative to its closest sulci. Each panel shows a sagittal MRI view with an enlargement of hMT+ area. Below is a cartoon to highlight the position of the sulci. ITS, inferior temporal sulcus; ALITS, ascending limb of the inferior temporal sulcus; PCITS, posterior continuation of the inferior temporal sulcus; LOS, lateral occipital sulcus. The open and filled arrows indicate the most and second-most frequent directions of the current vector in the left hMT+ (A, C, E, G, I, K, M and O) and the right hMT+ (B, D, F, H, J, L, N and P). The subject numbers are above each panel. The latency and delay at which the frequent directions were obtained are noted in parentheses for each subject. The directions in Subs. 1-3 were obtained from the JPD distributions shown in Fig. 4, and similarly in the other subjects. The volume of the region of interest (ROI) of hMT+ is indicated by the red areas, which denote the weight decay used in the regional averaging of current density (radius = 5 mm). These ROIs were defined in our previous study (Maruyama et al., 2009) by the contrast of response between motion and stationary stimuli. The sulcus denoted by the green lines in the left hemisphere of Sub. 1 appears as separate segments on this MRI slice, but it was continuous on a more lateral MRI slice. Thus, both segments were deemed the ALITS. On the cartoon panels, the thick red lines denote the sulcus segments whose orientation can explain the frequent directions. A sulcus segment nearly orthogonal to the frequent direction of the current vector was identified within the ROIs, although the most frequent direction of the right hMT+ in Sub. 2 (D) and the left hMT+ in Sub. 8 (O) was explained by the composition of two current vectors on separate segments. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

itive (+50 ms) and negative (-60 ms) delays near the latency of the highest MI_{Cor} . Although MI_{Ori} values were continuously significant over the delays from -60 to 50 ms in most of the subjects (Fig. 6B, right panel), the three peaks of MI_{Cor} were distinguishable from each other based on the MI_{Cor} troughs lying between the peaks

(left panel). The short delay of less than 10 ms for the highest peak of MI_{Cor} suggests direct communication between the left and right hMT+ areas via the corpus callosum, whereas the long delay of several tens of milliseconds for the secondary peaks was possibly a reflection of indirect communication.

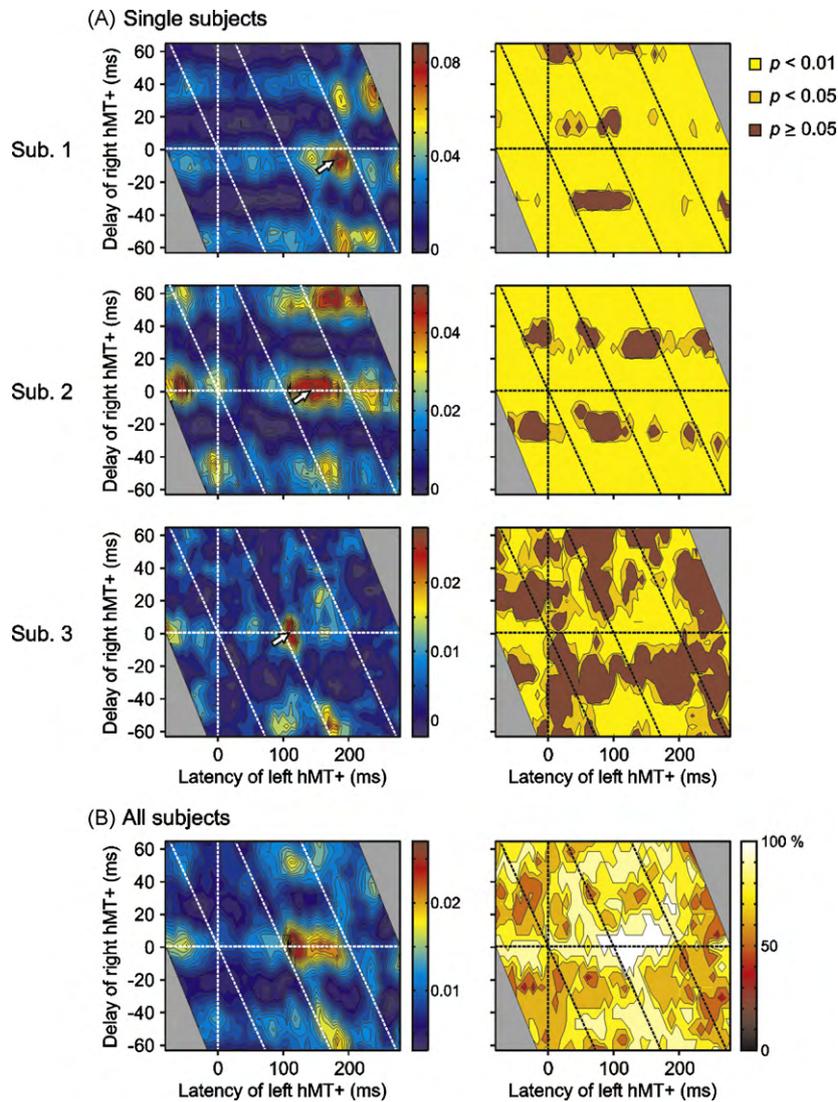


Fig. 6. (A) *Left panel:* Time courses of MI values in single subjects, based on the tangential direction of the current vector. A high-contrast motion stimulus was presented at the bottom-right location. Horizontal axes show the latency of the left hMT+ from the onset of stimulus. Vertical axes show the delay of the right hMT+ with respect to the latency of the left hMT+. Open arrows indicate latencies and delays at which a peak of MI_{Cor} was found. JPD distributions at these latencies and delays are shown in Fig. 5.) *Right panel:* Latencies and delays of significant MI_{Ori} values, tested in relation to the distributions of 100 MI_{Ran} values. Yellow areas denote that MI_{Ori} values were significant at the level of $p < 0.01$, and orange areas indicate that they were significant at the level of $p < 0.05$. Dark brown areas indicate that the MI_{Ori} values were not significant ($p \geq 0.05$). (B) *Left panel:* MI_{Cor} values averaged across all eight subjects. *Right panel:* Percentage of subjects exhibiting significant MI values ($p < 0.05$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.6. Functional connectivity at low contrast

Fig. 7 shows the time courses of functional connectivity under the low-contrast condition. The left panels show the average across all subjects of the MI_{Cor} value computed using the Gaussian kernel estimator. The functional connectivity based on the moduli was relatively strong during the prestimulus period, and continued to be strong until a latency of around 50 ms (Fig. 7A, left panel), whereas it continued until the latency of 100 ms at high contrast (Fig. 3B, left panel). In the early stimulus period, a relatively high MI_{Cor} was often obtained at negative delays, as under the high-contrast condition. The right panel of Fig. 7A shows the percentage of subjects exhibiting a significant MI_{Ori} ($p < 0.05$). Significant connectivity was found in more subjects during the prestimulus and early stimulus periods (yellow and orange areas) than during the late stimulus period (dark brown areas).

For the connectivity based on the directions, averaging across all subjects showed that the relatively strong connectivity started after a latency of 120 ms (Fig. 7B, left panel). Like the high-contrast condi-

tion, two peaks of MI_{Cor} were observed with long delays in positive (+50 ms) and negative (−60 ms). The delay range of the strongest connectivity was negative after a latency of 200 ms at low contrast, whereas the highest peak of MI_{Cor} at high contrast centered on zero until it vanished at the latency of 200 ms. A significant MI_{Ori} was identified in all subjects (white areas in Fig. 7B, right panel) at more latencies and delays after the latency of 120 ms, which is consistent with the temporal pattern of the averaging.

Overall, the temporal patterns of connectivity computed with moduli and directions were different at low contrast, as they were under the high-contrast condition.

3.7. Comparisons of analysis methods using real samples

Using the real MEG data, the MI analysis with the optimized Gaussian kernel was compared with relatively simple methods: MI analysis with a histogram method and correlation coefficient analyses. The simple methods indicated similar temporal patterns of connectivity to those indicated in the MI analyses with the Gaussian

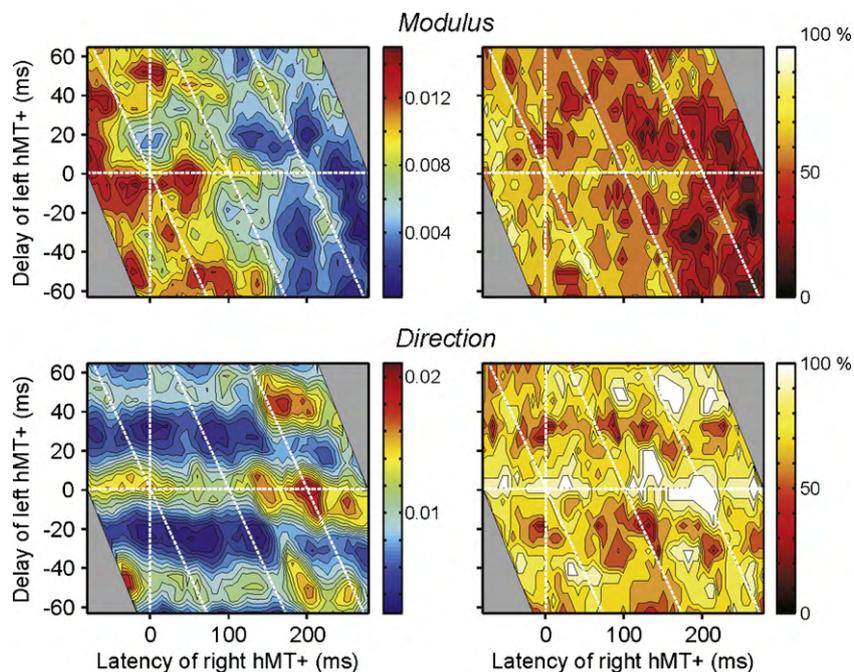


Fig. 7. Temporal pattern of functional connectivity at low contrast. *Top-left panel:* Time courses of MI_{Cor} values based on the current vector moduli. The average over all subjects is shown. Horizontal axes show the latency of the left hMT+ from the onset of stimulus. Vertical axes show the delay of the right hMT+ with respect to the latency of the left hMT+. *Top-right panel:* Percentage of subjects exhibiting significant MI_{Ori} values computed with the moduli ($p < 0.05$). *Bottom-left panel:* Time courses of MI values based on the current vector directions, averaged over all subjects. *Right panel:* Percentage of subjects exhibiting significant MI_{Ori} values computed with the directions ($p < 0.05$).

kernel estimator. However, the temporal patterns took the form of smeared patches in the MI analysis with the histogram method. Moreover, the connectivity was significant at fewer latencies and delays with the histogram method than with the Gaussian kernel method. It was also the case in the correlation coefficient analyses, though their sampling method was identical to the MI analyses. It is possible that the computation of the correlation coefficient with smaller size windows will remove some of the blurring. However the computation of the (linear) correlation coefficient with smaller windows will also emphasize linear relations, both the ones arising through actual linkage between the two time series and ones simply arising by chance. Also, the use of a smaller window is not appropriate for the computation of nonlinear relationships, and it would therefore obscure the comparison between the linear and nonlinear methods. Taking into account the simulation results in which the MI estimation with the histogram method possessed the large SD (see [Supplementary Materials](#)), the smeared patches and the fewer latencies and delays of significant connectivity could reflect a noisy estimation of MI with the histogram method. Overall, we found that MI analysis becomes a better estimator of functional connectivity by using the Gaussian kernel optimized with the LH method. For the details of the methodological comparisons, see [Supplementary Materials](#).

4. Discussion

4.1. Single-trial analysis

Single-trial MI analysis demonstrated that the trial variance of the current vector of the left hMT+ was significantly linked with that of the right hMT+. A statistically significant link was found using both the moduli and directions (Figs. 3 and 6, right panels). We also found that the directions of the most and second-most frequent current vectors were often opposite to each other in the same latency range (Fig. 4). This finding implies that, because of averaging, the event-related magnetic field (ERMF) often underestimates the intensity of the neural response to a stimulus; thus, the

ERMF might not be always an appropriate indicator of response intensity. This limitation of averaging should also be relevant for event-related potential (ERP) analysis.

The influence of sensor noise might be larger when current sources are reconstructed from single-trial MEG signals than when they are reconstructed from averaged MEG signals. The sensor noise near the left hMT+ was independent of that near the right hMT+; thus, the influence of noise on the current estimation in the left hMT+ was also independent of that in the right hMT+. As a consequence, the random nature of sensor noise might lower the total MI values between the left and right hMT+. However, sensor noise would not significantly change the temporal pattern of MI values.

4.2. Current vector direction

The neural current that is responsible for the generation of the MEG signal is generally assumed to flow perpendicularly to the cortical surface of a sulcus. In accordance with this general assumption, we identified the sulcus segments whose orientations could explain our estimation of highly frequent directions (Fig. 5). These findings confirm the presence of a possible neural substrate for our direction results and thus support the reliability of current vector estimation by MFT analysis.

The hMT+ has been reliably colocalized to the ITS, ALITS, PCITS, and LOS (Dumoulin et al., 2000). In the current study, the ROIs covered the various sulci in the left hemispheres of Subs. 1 and 2 (Fig. 5). Among the sulci, only the ALITS segment could explain the direction of the frequent current vector by its orientation. The curvature of that sulcus in the ROI in the right hemisphere of Sub. 1 made it possible to estimate the location of that current source at a higher spatial resolution relative to the size of the ROI (10 mm in diameter).

4.3. Joint probability estimation with an optimized kernel

We used the Gaussian kernel optimized by the LH method to estimate the JPD distribution from a limited sample size. This estimation is a straightforward procedure based on the samples, and

thus its results are objective. On the other hand, the often-used histogram method requires choices for bin size and the origin of the binning. These choices by the user can affect the JDP distribution results and hence limit their objectivity. Frequent current vector directions were clearly identified from the JPD distributions with the optimized kernel, allowing the direction results to be compared with the sulcus orientations. The optimized kernel method produces a smaller SD for the MI estimation than the histogram method (Treves and Panzeri, 1995; Darbellay and Vajda, 1999; Kraskov et al., 2004); this makes it easy to detect the peak of functional connectivity in relation to the latencies and delays (Figs. 3 and 6), whereas MI time courses obtained by the histogram method often take the form of smeared patches (Supplementary Fig. 7A and C) because of the large SD. Taking into account these observations, we deem the optimized kernel method to be appropriate for our study purpose. More detailed simulation-based comparisons with the histogram method can be found in Supplementary Materials.

4.4. Nonlinear relationship in functional connectivity

For a comparison with the MI analysis, which accounts for both linear and nonlinear relationships, we also examined correlation coefficient analyses as representative of methods for detecting only linear relationships. Correlation coefficient analyses obtained a similar temporal pattern of functional connectivity, i.e., the connectivity based on the moduli was relatively strong in the prestimulus period, continuing until a latency of 100 ms (Fig. 3B and 8B), whereas the connectivity measured with the directions became strong after that latency (Fig. 6B and 8D). However, in spite of the same sampling condition, the peak correlation coefficient values computed with the moduli were not as sharp as those of MI, which might be a reflection of the presence of a nonlinear component at peaks of functional connectivity.

4.5. Strength of functional connectivity

The MI_{Cor} values in the current study were always less than 0.1 (Figs. 3 and 6). Two previous MEG studies performed MI analyses comparable to that of the current study in which the MI value was defined in accordance with the standard form (Kwapień et al., 1998) and applied to tomographic estimates of neural currents. In these studies, the MI values obtained were less than 0.2, both between the left and right auditory areas (Kwapień et al., 1998) and between V1/V2 and hMT+ (Maruyama et al., 2009). Here, we discuss the MI values obtained in these two studies and our MEG study relative to those obtained in other previous studies, described below.

Chen et al. (2008) applied an MI analysis to magnetic field signals of sensor channels in which the largest MI values exceeded 0.9. The signal of each channel contains contributions from many regional neural generators. Equivalently, each of these generators contributes to many channels. The high MI values obtained by Chen et al. (2008) might reflect contributions to different channels from the same generators. In EEG studies, Jeong et al. (2001) and Na et al. (2002) applied MI analysis to signals from many pairs of electrodes. In their studies, base-two logarithms were used in the definition of entropy, whereas our MEG studies used natural logarithms (Eqs. (8) and (9)). This produces a 1.44-fold difference ($=\log_2 P/\ln P$) in the MI value. Some MI values that they obtained when converted to the natural logarithm form exceeded 0.4, which is also higher than our MI results. However, we note that, in comparison with the MI computed between regional estimations of the current vector, as in the present study, the MI computed between electrodes can be more strongly influenced by correlations due to common sensitivity to the same generator being mixed with correlations due to interactions between distinct neural generators.

Previous fMRI studies have rarely performed a comparable MI analysis. However, the range of correlation coefficients (R) between -0.8 and 0.8 , provided by Fox et al. (2005), may allow the typical range of MI values to be inferred, since in a bivariate normal distribution [see Eq. (S5) in Supplementary Materials], R values can be converted to MI values by the equation $MI = 0.5 \times \ln(1 - R^2)$. By using this equation, the range of MI values in the fMRI study of Fox et al. (2005) can be estimated to be up to around 0.5, which is also larger than our MI results. The large difference in the connectivity strength requires further investigation to pinpoint how the very different temporal sensitivities and underlying mechanisms between MEG and fMRI signals contribute to the connectivity estimates.

4.6. Time courses of functional connectivity

The previous fMRI study found a positive correlation between the left and right hMT+ under various resting conditions, including while the subjects fixated on a crosshair without any other visual stimuli (Fox et al., 2005). Similarly, the subjects in the current study fixated on a central black dot without any other stimuli during the prestimulus period. Thus, the relatively high MI values of the current vector moduli in the prestimulus period (Fig. 3) agree with the fMRI study results of Fox et al. (2005). The intensity of neural responses can be detected by current vector moduli as well as by fMRI signals, in spite of the large modality difference of signals between the methods.

Under the high-contrast condition, the relatively strong connectivity computed with the current vector moduli continued until a latency of around 100 ms after the onset of the stimulus. The delay of the ipsilateral hMT+ relative to the contralateral one could be both positive and negative (Fig. 3B), which may be a reflection of bidirectional information exchange between the areas related to the appearance of the visual stimulus. With regard to the strongest modulus connectivity, obtained in most of the subjects at a latency of around 100 ms, the delay was negative, suggesting feedback processing from the ipsilateral hMT+ (right hMT+ relative to the bottom-right visual field) to the contralateral one. The long delay of 19.2 ms of the strongest modulus connectivity allows the possibility of an indirect pathway, as implied by ffytche et al. (2000). The link between the hMT+ areas in the two hemispheres might be unspecific before the stimulus is adequately processed (i.e., in the prestimulus period and until about 100 ms after onset). In the early period the ipsilateral hMT+ might send signals to higher frontal cortical regions, such as the frontal eye fields that could input signals to the contralateral hMT+ (Silvanto et al., 2006). Such pathway might be beneficial when subjects' attention can facilitate visual motion processing. After peaking at around 100 ms, the modulus-based MI values declined.

Alternatively, high MI values appeared for the current vector directions, beginning around 100 ms and continuing until 200 ms. The delays of the major peak of MI values were centered around zero at high contrast until it vanished at the latency of 200 ms (Fig. 6B), but not at low contrast (Fig. 7B). With long delays in positive (+50 ms) and negative (−60 ms) two other peaks were observed at latencies of 100 and 200 ms, but their magnitude was lower than the major peak. While the longer delays of high MI values allow for the possibility of indirect connectivity, the short delays observed in the direction-based MI values are more consistent with direct communication between the two hMT+ areas (Van Essen et al., 1982; ffytche et al., 2000). At high contrast a focused link between the hMT+ areas might come out after the appropriate stimulus features are selected (i.e., latencies around and after 100 ms). The later stage results could then be a signature of the integration of stimulus representations in the left and right hMT+ areas into coherent representational states, most likely through direct communication

via the corpus callosum, although common input from a third area remains a possibility (Singer, 1993).

The modulus- and direction-based connectivity appeared consistent with two distinct periods of transcranial magnetic stimulation (TMS) studies, in which subjects' performance of motion-direction detection was reduced by stimulating the hMT+ areas before 40 ms or between 80 and 210 ms relative to motion onset (d'Alfonso et al., 2002; Sack et al., 2006; Laycock et al., 2007). The functional distinction of neural processing in the separated periods has not been fully understood yet. Our results suggest a possibility that the two hMT+ areas exchange information indirectly with contributions of other areas in order to interpret better the visual inputs. Visual features computed in the separate hMT+ areas are then integrated into a coherent perception, possibly via direct communication.

In summary, we showed that the MI results based on the modulus and direction were sensitive to different aspects of connectivity, and they were consistently so across subjects. The use of information about the direction of the current adds specificity, because one of the nearby sulci can be selected as the most likely neural generator on the basis of how well its orientation matches the current density direction corresponding to a peak in the MI. In conclusion, this study demonstrated that MI analyses of current vector moduli and directions can disentangle different types of functional connectivity into different dynamics, thanks to the ability of current vector direction analysis to reveal detailed aspects of functional connectivity.

Acknowledgements

We thank all the staff of the Laboratory for Human Brain Dynamics (1999–2009) at the Brain Science Institute, RIKEN for assistance with the experiment. The experiment and data analysis was performed at RIKEN and the revisions of the work continued by the authors at their new posts in France and Cyprus. Following the closing of the MEG Laboratory at RIKEN in September 2009, the contribution from AAI was supported by Cyprus Research Promotion Foundation grants (upgrade/info/0308/02 and human/sociol/0308/BE). The authors thank Dr. Peter Fenwick and Elizabeth Fenwick for revising the English used in this article.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jneumeth.2010.07.010.

References

- Astolfi L, Cincotti F, Mattia D, Marciani MG, Baccala LA, Fallani FD, et al. Comparison of different cortical connectivity estimators for high-resolution EEG recordings. *Hum Brain Mapp* 2007;28:143–57.
- Batschelet E. *Circular statistics in biology*. London: Academic Press; 1981.
- Buchel C, Friston KJ. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cereb Cortex* 1997;7:768–78.
- Chen CC, Hsieh JC, Wu YZ, Lee PL, Chen SS, Niddam DM, et al. Mutual-information-based approach for neural connectivity during self-paced finger lifting task. *Hum Brain Mapp* 2008;29:265–80.
- d'Alfonso AA, van Honk J, Schutter DJ, Caffé AR, Postma A, de Haan EH. Spatial and temporal characteristics of visual motion perception involving vs visual cortex. *Neurol Res* 2002;24:266–70.
- Damasio AR. The brain binds entities and events by multiregional activation from convergence zones. *Neural Comput* 1989;1:123–32.
- Darbellay GA, Vajda I. Estimation of the information by an adaptive partitioning of the observation space. *IEEE Trans Inform Theory* 1999;45:1315–21.
- David O, Cosmelli D, Friston KJ. Evaluation of different measures of functional connectivity using a neural mass model. *Neuroimage* 2004;21:659–73.
- Dumoulin SO, Bittar RC, Kabani NJ, Baker CL, Le Goualher G, Pike GB, et al. A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cereb Cortex* 2000;10:454–63.

- ffytche DH, Howseman A, Edwards R, Sandeman DR, Zeki S. Human area V5 and motion in the ipsilateral visual field. *Eur J Neurosci* 2000;12:3015–25.
- Fox MD, Snyder AZ, Vincent JL, Corbetta M, van Essen DC, Raichle ME. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci U S A* 2005;102:9673–8.
- Friston KJ. Functional and effective connectivity in neuroimaging: a synthesis. *Hum Brain Mapp* 1994;2:56–78.
- Friston KJ. Another neural code? *Neuroimage* 1997;5:213–20.
- Friston KJ. The labile brain. I. Neuronal transients and nonlinear coupling. *Philos Trans R Soc Lond Ser B: Biol Sci* 2000;355:215–36.
- Friston K. Functional integration and inference in the brain. *Prog Neurobiol* 2002;68:113–43.
- Friston KJ, Ungerleider LG, Jezzard P, Turner R. Characterizing modulatory interactions between areas V1 and V2 in human cortex: a new treatment of functional MRI data. *Hum Brain Mapp* 1995;2:211–24.
- Gevins AS, Doyle JC, Cuttillo BA, Schaffer RE, Tannehill RS, Ghannam JH, et al. Electrical potentials in human brain during cognition: new method reveals dynamic patterns of correlation. *Science* 1981;213:918–22.
- Gross J, Kujala J, Hamalainen M, Timmermann L, Schnitzler A, Salmelin R. Dynamic imaging of coherent sources: studying neural interactions in the human brain. *Proc Natl Acad Sci U S A* 2001;98:694–9.
- Grynspan F, Geselowitz DB. Model studies of the magnetocardiogram. *Biophys J* 1973;13:911–25.
- Hamalainen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev Mod Phys* 1993;65:413–97.
- Huk AC, Dougherty RF, Heeger DJ. Retinotopy and functional subdivision of human areas MT and MST. *J Neurosci* 2002;22:7195–205.
- Ioannides AA. Magnetic field tomography: theory, applications and examples from the visual system. In: *Recent Advances in Human Brain Mapping*. Excerpta Medica, International Congress Series, vol. 1232; 2002. p. 261–70.
- Ioannides AA, Bolton JPR, Clarke CJS. Continuous probabilistic solutions to the bi-magnetic inverse problem. *Inverse Probl* 1990;6:523–42.
- Ioannides AA, Corsi-Cabrera M, Fenwick PB, del Rio PY, Laskaris NA, Khurshudyan A, et al. MEG tomography of human cortex and brainstem activity in waking and REM sleep saccades. *Cereb Cortex* 2004;14:56–72.
- Ioannides AA. Estimation of brain activity using magnetic field tomography and large scale communication within the brain. In: Ho MW, Popp FA, Warnke U, editors. *Bioelectronics and biocommunication*. Singapore: World Scientific; 1994. p. 319–53.
- Ioannides AA, Fenwick PB, Liu L. Widely distributed magnetoencephalography spikes related to the planning and execution of human saccades. *J Neurosci* 2005;25:7950–67.
- Ioannides AA, Liu LC, Kwapien J, Drozd S, Streit M. Coupling of regional activations in a human brain during an object and face affect recognition task. *Human Brain Mapp* 2000;11:77–92.
- Jahn O, Cichocki A, Ioannides AA, Amari S. Identification and elimination of artifacts from MEG signals using extended independent component analysis. In: Yoshimoto T, Kotani M, Kuriki S, Karibe H, Nakasato N, editors. *Recent Advances in Biomagnetism*. Sendai: Tohoku Univ; 1999. p. 224–8.
- Jeong J, Gore JC, Peterson BS. Mutual information analysis of the EEG in patients with Alzheimer's disease. *Clin Neurophysiol* 2001;112:827–35.
- Kandel ER, Schwartz JH, Jessell TM. *Principles of neural science*. New York: McGraw-Hill; 2000.
- Kraskov A, Stogbauer H, Grassberger P. Estimating mutual information. *Phys Rev E* 2004;69:066138.
- Kwapien J, Drozd S, Liu LC, Ioannides AA. Cooperative dynamics in auditory brain response. *Phys Rev E* 1998;58:6359–67.
- Laycock R, Crewther DP, Fitzgerald PB, Crewther SG. Evidence for fast signals and later processing in human V1/V2 and V5/MT+: a TMS study of motion perception. *J Neurophysiol* 2007;98:1253–62.
- Lin FH, Belliveau JW, Dale AM, Hamalainen MS. Distributed current estimates using cortical orientation constraints. *Hum Brain Mapp* 2006;27:1–13.
- Liu L, Ioannides AA. Spatiotemporal dynamics and connectivity pattern differences between centrally and peripherally presented faces. *Neuroimage* 2006;31:1726–40.
- Llinás R, Nicholson C. Analysis of field potential in the central nervous system. In: Stevens CF, editor. *Handbook of electroencephalography and clinical neurophysiology*. Amsterdam: Elsevier; 1974. p. 61–83.
- Mardia KV, Jupp PE. *Directional statistics*. West Sussex: Wiley; 2000.
- Maruyama M, Palomo DD, Ioannides AA. Stimulus-contrast-induced biases in activation order reveal interaction between V1/V2 and human MT+. *Hum Brain Mapp* 2009;30:147–62.
- Moon YI, Rajagopalan B, Lall U. Estimation of mutual information using kernel density estimators. *Phys Rev E* 1995;52:2318–21.
- Moradi F, Liu LC, Cheng K, Waggoner RA, Tanaka K, Ioannides AA. Consistent and precise localization of brain activity in human primary visual cortex by MEG and fMRI. *Neuroimage* 2003;18:595–609.
- Murakami S, Zhang T, Hirose A, Okada YC. Physiological origins of evoked magnetic fields and extracellular field potentials produced by guinea-pig CA3 hippocampal slices. *J Physiol* 2002;544:237–51.
- Na SH, Jin SH, Kim SY, Ham BJ. EEG in schizophrenic patients: mutual information analysis. *Clin Neurophysiol* 2002;113:1954–60.

- Perfetti B, Franciotti R, Della Penna S, Ferretti A, Caulo M, Romani GL, et al. Low- and high-frequency evoked responses following pattern reversal stimuli: a MEG study supported by fMRI constraint. *Neuroimage* 2007;35:1152–67.
- Poghosyan V, Ioannides AA. Attention modulates earliest responses in the primary auditory and visual cortices. *Neuron* 2008;58:802–13.
- Rodriguez E, George N, Lachaux JP, Martinerie J, Renault B, Varela FJ. Perception's shadow: long-distance synchronization of human brain activity. *Nature* 1999;397:430–3.
- Sack AT, Kohler A, Linden DEJ, Goebel R, Muckli L. The temporal characteristics of motion processing in hMT/V5+: combining fMRI and neuronavigated TMS. *Neuroimage* 2006;29:1326–35.
- Samengo I. Estimating probabilities from experimental frequencies. *Phys Rev E: Stat Nonlin Soft Matter Phys* 2002;65:046124.
- Shannon CE. The mathematical theory of communication. *Bell Syst Tech J* 1948;27:379–423.
- Silvanto J, Lavie N, Walsh V. Stimulation of the human frontal eye fields modulates sensitivity of extrastriate visual cortex. *J Neurophysiol* 2006;96:941–5.
- Silverman BW. Density estimation for statistics and data analysis. London: Chapman & Hall; 1986.
- Singer W. Synchronization of cortical activity and its putative role in information processing and learning. *Annu Rev Physiol* 1993;55:349–74.
- Smith AT, Wall MB, Williams AL, Singh KD. Sensitivity to optic flow in human cortical areas MT and MST. *Eur J Neurosci* 2006;23:561–9.
- Taylor JG, Ioannides AA, Muller-Gartner HW. Mathematical analysis of lead field expansions. *IEEE Trans Med Imaging* 1999;18:151–63.
- Treves A, Panzeri S. The upward bias in measures of information derived from limited data samples. *Neural Comput* 1995;7:399–407.
- Van Essen DC, Newsome WT, Bixby JL. The pattern of interhemispheric connections and its relationship to extrastriate visual areas in the macaque monkey. *J Neurosci* 1982;2:265–83.
- Varela F, Lachaux JP, Rodriguez E, Martinerie J. The brainweb: phase synchronization and large-scale integration. *Nat Rev Neurosci* 2001;2:229–39.