A hierarchical Bayesian method to resolve an inverse problem of MEG contaminated with eye movement artifacts

Yusuke Fujiwara [a,b,*], Okito Yamashita [b], Dai Kawawaki [b], Kenji Doya [a,b,c], Mitsuo Kawato [a,b], Keisuke Toyama [b], Masa-aki Sato [b]

[a] *Graduate School of Information Science, Nara Institute of Science and Technology, Nara 630-0192, Japan*
[b] *ATR Computational Neuroscience Laboratories, Kyoto 619-0288, Japan*
[c] *Neural Computation Unit, Okinawa Institute of Science and Technology, Okinawa 904-2234, Japan*

ARTICLE INFO

ABSTRACT

The magnetic fields generated by eye movements are major artifacts in MEG measurements. We propose a hybrid hierarchical variational Bayesian method to remove eye movement artifacts from MEG data. Our method is an extension of the hierarchical variational Bayesian method for MEG source localization proposed by Sato et al. [Sato, M., Yoshioka, T., Kajihara, S., Toyama, K., Goda, N., Doya, K., and Kawato, M., (2004). Hierarchical Bayesian estimation for MEG inverse problem. NeuroImage 23(3), 806–826]. First, we assumed a single dipole at each left and right eyeball as a source of eye artifacts. Second, we constructed an EOG forward model describing the relationship between eye dipoles and electric potentials, i.e., EOG. Based on the Bayesian framework, the proposed method concurrently estimates eye and brain current sources from both MEG and EOG data. Thereby the brain current sources can be isolated from eye artifacts. The new method was tested in two ways. In the simulation experiments, the performance of eye artifact removal was evaluated from various aspects; locations of brain current sources, temporal correlation between eye and brain current sources, the level of MEG observation noise and so on. In real MEG experiments, we measured MEG and EOG data during smooth pursuit eye movements for a horizontally or circularly moving target. Our method successfully removed eye artifacts from the simulated and real MEG data with the estimation of brain current sources that were located in eye movement related areas. Our method should be widely applicable to MEG data obtained in tasks with non-negligible eye movements.

© 2008 Elsevier Inc. All rights reserved.

Introduction

Magnetoencephalography (MEG) is a technique that can measure brain activities with millisecond-temporal resolution, but poor spatial resolution remains one of its disadvantages. Several approaches to resolve this disadvantage based on the variational Bayesian method have been proposed (Sato et al., 2004; Daunizeau and Friston, 2007; Kiebel et al., 2008; Friston et al., 2008; Wipf and Nagarajan, 2009). MEG is also sensitive to eye movement artifacts, which limit the applications of its recording to experiments without eye movements. The eyes, which are the source of the artifacts, are located close to the MEG sensors, and the eye current source is two orders of magnitude larger than the brain current sources. For this reason, there have been few MEG studies associated with eye movements such as saccades or smooth pursuit eye movements. It would be very beneficial to invent a reliable method to remove eye artifacts from MEG in order to investigate the brain activity during tasks that inevitably induce eye movements.

Several methods have been proposed to remove eye artifacts from electroencephalograms (EEGs) or MEGs. Early methods simultaneously measured electrooculogram (EOG) with EEG and subtracted the signals that correlated with the EOG data from the EEG data (Kenemans et al., 1991; Meier et al., 1998; Croft and Barry, 2000). Recently, the independent component analysis (ICA) has been applied to remove eye artifacts (Vigario, 1997; Vigario et al., 2000; Jung et al., 2000; Barbati et al., 2004) by assuming statistical independence between the eye artifacts and brain signals. The moving dipole method (Berg and Scherg, 1994) assumes a single dipole in each eye and a few dipoles in the brain. Those dipoles are estimated to minimize the reconstruction error, and the contributions of the eye dipoles are finally removed from the MEG.

Both the correlation and ICA methods, which in principle subtract all components correlated to eye movement, may possibly even subtract the brain signals related to eye movements (see Fukushima et al., 2002 for correlated brain signals with eye movements) as eye artifacts. The moving dipole method may separate brain signals from correlated eye artifacts, but this method is not generally applicable to estimate distributed current sources since it approximates brain activities with only a limited number of dipoles.

* Corresponding author. ATR Computational Neuroscience Laboratories, 2-2-2 Hikaridai Seika-cho, Soraku-gun, Kyoto, Japan. Fax: +81 774 95 1259.
 E-mail address: yusuke-f@atr.jp (Y. Fujiwara).

We propose a new method to remove the eye artifacts, which we call the hybrid hierarchical Variational Bayesian method with Eye Dipoles (hyVBED method). This method is not based on the assumption of signal independence of brain current sources with eye movements and is applicable to the distributed source model. In the distributed source model, we must estimate one to ten thousand current sources from MEG signals obtained by a few hundred sensors. The hierarchical variational Bayesian method (Sato et al., 2004) resolves the ill-posed problem by imposing functional MRI information as hierarchical prior information. The hyVBED method consists of a MEG forward model containing a model of eye artifacts and a new EOG forward model that describes how EOGs are generated from eye current sources. EOG data are not contaminated with brain current sources, and the signal to noise (S/N) ratio of the EOG is two orders of magnitude higher than the MEG data. The problematic point is that the resistance between EOG and the eye current source is not known exactly although the conductance values of several issues around eyes have been investigated in detail (Malmivuo and Plonsey, 1995). The hyVBED method estimates unknown resistance simultaneously with eye and brain current sources.

We focused on smooth pursuit eye movements as one of the factors of eye movement artifacts and evaluated the performance of our method in both simulated and real MEG data during smooth pursuit eye movements. In simulation studies, our method almost perfectly removed the assumed eye artifacts and estimated brain current sources practically free of eye artifacts. Robustness of our method was also demonstrated by manipulating various simulation settings. Correspondingly, in the real MEG study, it successfully removed the eye artifacts and estimated brain current sources in the cortical areas known to be related to smooth pursuit eye movements whereas no significant activity was observed in other cortical areas including those close to both eyes.

Methods

Forward models for EOG and MEG

To develop a method for eye artifact removal, we must know how eye current sources produce EOG and MEG. The eye has a steady electric polarity with a positive charge at the cornea and a negative charge at the retina (Miles, 1939; Carpenter, 1988). For this reason, the eye can be regarded as an electric dipole (Katila et al., 1981; Malmivuo and Plonsey, 1995). When the eyeball rotates, the rotation induces currents inside it. Consequently, this eye current source produces electric potentials (EOG) and magnetic fields (MEG) around the eye. The magnetic fields produced by the eye current sources contaminate the MEG produced by brain current sources and act as eye artifacts.

The hyVBED method utilizes EOG data as well as MEG and fMRI data to estimate the eye and brain current sources. The advantage of using EOG is that it is practically uncontaminated with brain activities, since electric potentials from the eye current sources are much larger than those from the brain activities, although MEG is strongly contaminated with the eye current sources. This provides the basis for the hyVBED method that cancels the eye artifacts.

EOG forward model

Electric potentials from eye current sources are observed as EOGs through a volume conductor that consists of tissues around an eyeball. Unfortunately, there has been no detailed model to calculate EOG from the eye current source. Nevertheless, we can assume a linear relationship between EOG and the eye current source due to the superposition principle of Maxwell's equation. Therefore, we assume the following parametric linear model to calculate EOG $\mathbf{E}(t)$ from eye current source $\mathbf{J}_{eye}(t)$:

$$\mathbf{E}(t) = \mathbf{F} \cdot \mathbf{J}_{eye}(t) + \mathbf{u} + \varepsilon_{eog}(t), \tag{1}$$

where $\mathbf{E}(t)$, $\mathbf{J}_{eye}(t)$, $\mathbf{u}$, and $\varepsilon_{eog}(t)$ are $K \times 1$ vectors and $\mathbf{F}$ is a $K \times K$ matrix. The transfer matrix $\mathbf{F}$, which represents the sensitivity from the eye current source to EOG, is regarded as the unknown parameters that will be estimated by the hybrid hierarchical Bayesian method. Eq. (1) has the same form as Ohm's law, and $\mathbf{F}$ corresponds to resistance [Ω/m]. To separately measure the left/right and horizontal/vertical components of EOG, we place electrode pairs at the above and below eyes for the vertical EOG and the outer and inner canthi for horizontal EOG like Fig. 2B (Jervis et al., 1998). With this configuration, crosstalk between different components can be suppressed, and we assume that the matrix $\mathbf{F}$ only has nonzero entries in the diagonal part. The EOG system measures the DC component and is modeled by EOG bias components $\mathbf{u}$ in Eq. (1). Observation noise $\varepsilon_{eog}(t)$ is assumed to be a Gaussian white noise with mean zero.

MEG forward model

The MEG forward model is a modified version of the hierarchical variational Bayesian method (Sato et al., 2004) that includes a new term representing the effect of the eye current sources. Observed MEG data $\mathbf{B}(t)$ at time t are assumed to be the sum of three components: the magnetic fields generated from brain current sources $\mathbf{J}_{brain}(t)$, those generated from eye current sources $\mathbf{J}_{eye}(t)$, and observation noise $\varepsilon_{meg}(t)$:

$$\mathbf{B}(t) = \mathbf{G}_{brain} \cdot \mathbf{J}_{brain}(t) + \mathbf{G}_{eye} \cdot \mathbf{J}_{eye}(t) + \varepsilon_{meg}(t), \tag{2}$$

where $\mathbf{B}(t)$ is an $N \times 1$ vector, $\mathbf{G}_{brain}$ is an $N \times L$ matrix, $\mathbf{J}_{brain}(t)$ is an $L \times 1$ vector, $\mathbf{G}_{eye}$ is an $N \times K$ matrix, and $\mathbf{J}_{eye}(t)$ is a $K \times 1$ vector. Constants N, L, and K denote the number of sensors, brain current sources, and eye current sources, respectively. $\mathbf{G}_{brain}$ and $\mathbf{G}_{eye}$ are called the lead field matrices, whose n,l-th and n,k-th elements describe the sensitivity of the n-th sensor (gradiometer) when a unit dipole is set on the l-th and k-th locations. The lead field matrix for brain current sources $\mathbf{G}_{brain}$ is calculated by the Sarvas equation assuming a homogeneous conductor inside a spherical skull (Sarvas, 1987). The lead field matrix for eye current sources $\mathbf{G}_{eye}$ is calculated by the Biot–Savart law according to Katila et al., 1981, who argues that the eye current source can be modeled as a single current dipole and the volume current inside the eyeball can be neglected. Observation noise $\varepsilon_{meg}(t)$ is assumed to obey a Gaussian distribution with zero mean. In contrast to the EOG system, the MEG system does not measure a DC component because it measures the amount of change from a baseline magnetostatic field at the onset of the measurement.

Both EOG and MEG forward models (Eqs. (1) and (2)) are summarized as follows:

$$\begin{bmatrix} \mathbf{B}(t) \\ \mathbf{E}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{brain} & \mathbf{G}_{eye} \\ 0 & \mathbf{F} \end{bmatrix} \begin{bmatrix} \mathbf{J}_{brain}(t) \\ \mathbf{J}_{eye}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{u} \end{bmatrix} + \begin{bmatrix} \varepsilon_{meg}(t) \\ \varepsilon_{eog}(t) \end{bmatrix}. \tag{3}$$

Note that MEG data $\mathbf{B}(t)$ and EOG data $\mathbf{E}(t)$ are only observations in this model, and the other variables, $\mathbf{J}_{brain}(t)$, $\mathbf{J}_{eye}(t)$, $\mathbf{F}$, $\mathbf{u}$, $\varepsilon_{meg}(t)$, and $\varepsilon_{eog}(t)$ are estimated, given $\mathbf{G}_{brain}$ and $\mathbf{G}_{eye}$.

Current source estimation from EOG and MEG

We resolve the current source estimation problem according to a Bayesian perspective. A likelihood function is constructed based on the EOG and MEG forward models (Eqs. (1) and (2)). As prior distributions for the current sources, the resistance, and the EOG bias, the automatic relevance determination (ARD) model (Neal, 1996) is employed (relationship between variables is described in Fig. 1) that consists of hierarchical priors that promote the sparsity of resulting current source configurations. In addition fMRI information can be incorporated as a soft constraint that determines the strength of the brain current sources even if the fMRI experiment does not exactly match the MEG experiment. The efficacy of this approach has been confirmed by both simulation and real experimental studies (Sato et al., 2004;
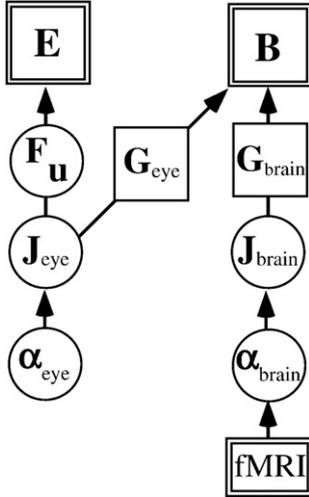
Fig. 1. Graphical representation of the model of our hyVBED method. Variables in circle and in a double-lined square denote variables to be estimated and to be observed, respectively. A variable in a single-lined square is a known constant. The likelihood function for EOG consists of four variables: observed EOG **E**, eye current sources $\mathbf{J}_{\text{eye}}$, resistance **F**, and an EOG bias **u**. The likelihood function for MEG consists of five variables: observed MEG **B**, brain current sources $\mathbf{J}_{\text{brain}}$, eye current sources $\mathbf{J}_{\text{eye}}$, the lead field matrix for brain current sources $\mathbf{G}_{\text{brain}}$ and that for the eye current source $\mathbf{G}_{\text{eye}}$. The Gaussian prior distributions are assumed for brain and eye current sources $\mathbf{J}_{\text{brain}}$, $\mathbf{J}_{\text{eye}}$ given their variances $\boldsymbol{\alpha}_{\text{brain}}$, $\boldsymbol{\alpha}_{\text{eye}}$. As hyper-prior distributions, the gamma distributions are further assumed for the brain and eye current variances $\boldsymbol{\alpha}_{\text{brain}}$, $\boldsymbol{\alpha}_{\text{eye}}$. Spatial information obtained from fMRI data (i.e., T-map) is imposed on mean parameters in the hyper-prior distributions of brain current variance $\boldsymbol{\alpha}_{\text{brain}}$ if available. Under these assumptions, brain and eye current sources $\mathbf{J}_{\text{brain}}$, $\mathbf{J}_{\text{eye}}$, current variances $\boldsymbol{\alpha}_{\text{brain}}$, $\boldsymbol{\alpha}_{\text{eye}}$, resistance **F**, and EOG bias **u** are estimated simultaneously by the variational Bayesian algorithm (see Methods).

Nummenmaa et al., 2007; Yoshioka et al., 2008). Posterior distributions, which are calculated from the likelihood functions and the prior distributions by the variational Bayesian method, provide estimated values of the current sources, the resistance, and the EOG bias. The novelty of the hyVBED method is that we build a likelihood function for EOG data and estimate the source of the eye artifacts not only from MEG data but also from the EOG data whose S/N ratio is much higher than the MEG data.

Likelihood functions

Assuming that the EOG and MEG observation noises follow different Gaussian distributions with a spherical covariance, the EOG forward model (Eq. (1)) and the MEG forward model (Eq. (2)) lead to likelihood functions observing the time course of electric potentials $\mathbf{E}(t)$ and magnetic fields $\mathbf{B}(t)$ from time $t=1$ to $t=T$:

$$P\left(\mathbf{E}|\mathbf{J}_{\text{eye}},\mathbf{F},\mathbf{u}\right) \propto \exp\left[-\frac{1}{2}\beta_{\text{E}}\sum_{t=1}^{T} ||\mathbf{E}(t) - \mathbf{F}\,\mathbf{J}_{\text{eye}}(t) - \mathbf{u}||^2\right], \quad (4)$$

$$P\left(\mathbf{B}|\mathbf{J}_{\text{brain}},\mathbf{J}_{\text{eye}}\right) \propto \exp\left[-\frac{1}{2}\beta_{\text{B}}\sum_{t=1}^{T} ||\mathbf{B}(t) - \mathbf{G}_{\text{brain}}\cdot\mathbf{J}_{\text{brain}}(t) - \mathbf{G}_{\text{eye}}\cdot\mathbf{J}_{\text{eye}}(t)||^2\right], \quad (5)$$

where β_{E}^{-1} and β_{B}^{-1} are the unknown variances of EOG and MEG observation noises, respectively.

Prior distribution for current sources

The prior distributions of the brain and eye current sources are given by

$$P_0\left(\mathbf{J}_{\text{brain}},\mathbf{J}_{\text{eye}}|\boldsymbol{\alpha}_{\text{brain}},\boldsymbol{\alpha}_{\text{eye}}\right) \propto \exp$$

$$\times\left[-\frac{1}{2}\sum_{t=1}^{T}\left\{\mathbf{J}'_{\text{brain}}(t)\cdot\mathbf{A}_{\text{brain}}\cdot\mathbf{J}_{\text{brain}}(t) + \mathbf{J}'_{\text{eye}}(t)\cdot\mathbf{A}_{\text{eye}}\cdot\mathbf{J}_{\text{eye}}(t)\right\}\right], \quad (6)$$

where $\mathbf{A}_{\text{brain}}=\text{diag}(\boldsymbol{\alpha}_{\text{brain}})$ and $\mathbf{A}_{\text{eye}}=\text{diag}(\boldsymbol{\alpha}_{\text{eye}})$. Vectors $\boldsymbol{\alpha}_{\text{brain}}$ and $\boldsymbol{\alpha}_{\text{eye}}$ are the inverse variances of the brain and eye current sources, respectively. Note that $\boldsymbol{\alpha}_{\text{brain}}$ and $\boldsymbol{\alpha}_{\text{eye}}$, which are assumed to be time invariant, are unknown parameters determined from the observed data. For convenience, we reformulate the prior distribution (Eq. (6)) as

$$P_0\left(\mathbf{J}|\boldsymbol{\alpha}_{\mathbf{J}}\right) \propto \exp\left[-\frac{1}{2}\sum_{t=1}^{T} \mathbf{J}'(t)\cdot\mathbf{A}_{\mathbf{J}}\cdot\mathbf{J}(t)\right], \quad (7)$$

where $\mathbf{J}' = \left[\mathbf{J}'_{\text{brain}}\ \mathbf{J}'_{\text{eye}}\right]$, $\boldsymbol{\alpha}'\mathbf{J} = \left[\boldsymbol{\alpha}'_{\text{brain}}\ \boldsymbol{\alpha}'_{\text{eye}}\right]$, and $\mathbf{A}_{\mathbf{J}}=\text{diag}(\boldsymbol{\alpha}_{\mathbf{J}})$.

Hyper-prior distribution for current variances

Furthermore the hierarchical priors for the inverse variances of current sources $\boldsymbol{\alpha}_{\mathbf{J}}$ are assumed:

$$P_0(\boldsymbol{\alpha}_{\mathbf{J}}) = \prod_i \mathcal{G}\left(\boldsymbol{\alpha}_{\mathbf{J}(i)}|\overline{\boldsymbol{\alpha}}_{\mathbf{J}0(i)},\gamma_{\mathbf{J}0(i)}\right), \quad (8)$$

where $\mathcal{G}(\boldsymbol{\alpha}|\overline{\boldsymbol{\alpha}},\gamma)$ represents the Gamma distribution with mean $\overline{\boldsymbol{\alpha}}$ and degree of freedom γ. $\overline{\boldsymbol{\alpha}}_{\mathbf{J}0(i)}$ is a mean prior of an inverse current variance and $\gamma_{\mathbf{J}0(i)}$ is the confidence parameter of corresponding $\overline{\boldsymbol{\alpha}}_{\mathbf{J}0(i)}$. A prior current variance $\overline{v}_{\mathbf{J}0}\equiv\overline{\boldsymbol{\alpha}}_{\mathbf{J}0}^{-1}$ represents the prior information about a current intensity. For large $\overline{v}_{\mathbf{J}0}$, estimated current $\overline{J}$ could be large. For small $\overline{v}_{\mathbf{J}0}$, estimated current $\overline{J}$ tends to be small. The confidence parameter $\gamma_{\mathbf{J}0}$ controls the spread of the distribution of $\boldsymbol{\alpha}_{\mathbf{J}}$. For **small** $\gamma_{\mathbf{J}0}$, the distribution spreads uniformly, and prior information $\overline{v}_{\mathbf{J}0}$ does not affect the current estimation (noninformative prior). In contrast, for large $\gamma_{\mathbf{J}0}$, since the distribution is concentrated around prior mean $\overline{v}_{\mathbf{J}0}$, prior information $\overline{v}_{\mathbf{J}0}$ influences the current estimation more strongly. When fMRI data are available for which the same task paradigm as a MEG experiment was used, the fMRI information can be imposed on the prior variance parameter for brain current sources $\overline{v}_{\mathbf{J}0}$ (Fig. 1) and the confidence parameter $\gamma_{\mathbf{J}0}$ is set to large (detailed value is described in Yoshioka et al., 2008). When fMRI data are not available, noninformative prior ($\gamma_{\mathbf{J}0}=0$) or uniform spatial prior (the same value for all $\overline{v}_{\mathbf{J}0}$) can be used. The confidence parameter for the eye current source is set to zero because we have no prior information about it.

Prior and hyper-prior distributions for resistance and EOG bias

The prior distributions of the resistance **F** and the EOG bias **u** are given by

$$P_0(\mathbf{F}|\boldsymbol{\alpha}_{\mathbf{F}}) = \exp\left[-\frac{1}{2}\sum_k \boldsymbol{\alpha}_{\mathbf{F}(k)}\mathbf{F}^2_{(k,k)}\right] \quad (9)$$

and

$$P_0(\mathbf{u}|\boldsymbol{\alpha}_{\mathbf{u}}) = \exp\left[-\frac{1}{2}\sum_k \boldsymbol{\alpha}_{\mathbf{u}(k)}\mathbf{u}^2_{(k,k)}\right], \quad (10)$$

where $\boldsymbol{\alpha}_{\mathbf{F}}$ and $\boldsymbol{\alpha}_{\mathbf{u}}$ are the inverse variances of the resistance and the EOG bias, respectively. Noninformative priors are imposed as the hierarchical priors for the inverse variances of resistance $\boldsymbol{\alpha}_{\mathbf{F}}$:

$$P_0(\boldsymbol{\alpha}_{\mathbf{F}}) = \prod_k 1/\boldsymbol{\alpha}_{\mathbf{F}(k)}. \quad (11)$$

The hierarchical priors for $\boldsymbol{\alpha}_{\mathbf{u}}$ have the same form as Eq. (11).

Posterior distribution

The estimation of the current sources is done by calculating a posterior distribution $P(\mathbf{J}|\mathbf{B},\mathbf{E})$ from the likelihood function (Eqs. (4), (5)), the prior distribution (Eqs. (7), (9), (10)) and the hyper-prior distribution (Eqs. (8), (11)). The posterior distribution $P(\mathbf{J}|\mathbf{B},\mathbf{E})$ is obtained from a joint posterior distribution $P(\mathbf{J},\mathbf{F},\mathbf{u},\boldsymbol{\alpha}_{\mathbf{J}},\boldsymbol{\alpha}_{\mathbf{F}},\boldsymbol{\alpha}_{\mathbf{u}}|\mathbf{B},\mathbf{E})$ as follows:

$$P(\mathbf{J},\mathbf{B},\mathbf{E}) = \int d\mathbf{F}d\mathbf{u}d\boldsymbol{\alpha}_{\mathbf{J}}d\boldsymbol{\alpha}_{\mathbf{F}}d\boldsymbol{\alpha}_{\mathbf{u}}P(\mathbf{J},\mathbf{F},\mathbf{u},\boldsymbol{\alpha}_{\mathbf{J}},\boldsymbol{\alpha}_{\mathbf{F}},\boldsymbol{\alpha}_{\mathbf{u}}|\mathbf{B},\mathbf{E}). \quad (12)$$

Since the joint posterior distribution $P(\mathbf{J}, \mathbf{F}, \mathbf{u}, \alpha_J, \alpha_F, \alpha_u | \mathbf{B}, \mathbf{E})$ cannot be calculated analytically, we approximate it using the variational Bayesian method (Attias, 1999; Sato, 2001). The algorithm becomes a three-step iterative procedure of the estimation of $\mathbf{J}$, $\mathbf{F}$ and $\mathbf{u}$, and α. The derivation and details of the algorithm are shown in Appendix A.

Current source estimation without EOG data

Even if EOG data are not available, the eye and brain current sources can be estimated from MEG data alone using the likelihood (Eq. (5)), prior distribution (Eq. (7)), and hyper-prior distribution (Eq. (8)). This method is referred to as "the VBED method" (the hierarchical Variational Bayesian method with Eye Dipoles). The estimation procedure is identical to the original hierarchical variational Bayesian method (Sato et al., 2004). Its performance will be compared with the hyVBED method in both simulation and real experiments.

Real EOG and MEG experiments

Pursuit eye movement task

Two kinds of moving target patterns were used for the smooth pursuit task; horizontal and circular movements (Fig. 2A). The horizontal eye movement task was a main experiment in which the performance of eye artifact removal and spatio-temporal patterns of cortical activities related with the eye movements were investigated in detail. On the other hand, the circular eye movement task was a supplemental experiment to verify the applicability of the proposed method to remove artifacts composed of both horizontal and vertical eye movements.

Three right-handed males 24 or 25 years old with normal vision and no history of neurological injury participated under informed consent. All the three subjects participated in the horizontal eye movement task and one of the three subjects participated in the circular eye movement task in a separate day. They were instructed to pursue a target (a red spot) moving sinusoidally along a horizontal meridian or circularly on a screen (horizontal amplitude and circular radius, 4°; frequency, 0.7 Hz; cycle, 2.5 for MEG, 5.7 for fMRI). The horizontally moving target was a laser spot projected onto the

screens in MEG and fMRI systems from a galvanometer mirror system (Kiyohara Kougaku) controlled by a visual stimulus generator (VSG2/5, Cambridge Research Systems). The circularly moving target was projected by a projector.

EOG and MEG recordings

The horizontal and vertical EOG data were recorded simultaneously with the MEG data through the four pairs of electrodes upper and below, left and right side of each eye (Fig. 2B; sampling rate, 1000 Hz). Before the eye movement task, an EOG calibration task was conducted to obtain transform coefficients from EOG measurements to eye positions (Miles, 1939; Shackel, 1960). The horizontal and vertical EOG were calibrated separately. In the horizontal calibration, a subject was instructed to sequentially fixate at four horizontal targets whose displacements were ±1.34, ±4° from the center of the screen. Then two coefficients (a bias and a weight constant) of a linear transformation from a horizontal EOG measurement (μV) to the horizontal eye position (°) were calculated by the least square method using the four target displacements and the corresponding horizontal EOG measurements. The vertical calibration was done in the same way. By this calibration, EOG measurements during smooth pursuit task were converted to eye positions by which we assessed subjects' task performance.

MEGs for horizontal eye movements were recorded using a whole-head biomagnetic imaging system (Shimadzu Corp.) that consists of 201 gradiometers (sampling rate, 1000 Hz). In the experiments, only 188 sensors were used for the measurements because the remaining 13 sensors failed. MEGs for circular eye movements were recorded using a whole-head biomagnetic imaging system (Yokogawa Corp.) that consists of 210 gradiometers (sampling rate, 1000 Hz). The subjects had to maintain their gaze at the target as it stayed in the center of the screen during the prestimulus interval (300 ms) and to pursue the moving target during the eye movement period (3571 ms, corresponding to a 2.5 cycle). The inter-trial intervals were 2500 ms, and eye blinks were only allowed during this period. MEG recoding of 60 trials were conducted for each subject.

EOG and MEG data were preprocessed as follows. The high frequency components of EOG and MEG signals were removed by a lowpass filter with a 100 Hz cut-off frequency. Drift components
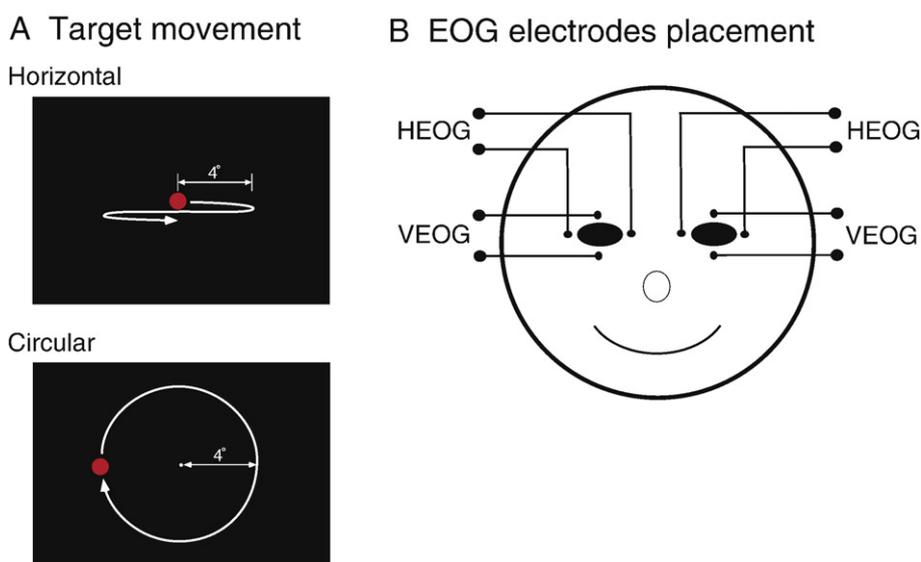
Fig. 2. (A) Target movements and (B) EOG electrodes placement used in the real experiment. Two pairs of electrodes were placed around each eye. One pair was placed on upper and below side and the other was placed on the left and right side of an eye. The horizontal EOGs (HEOG) and vertical EOGs (VEOG) were measured during the real MEG experiment. By this placement, the horizontal and vertical EOG were able to measure horizontal and vertical eye movements respectively while a crosstalk between the horizontal and vertical EOGs was inhibited.

were removed by fitting quadratic functions and subtracting fitted curves from the data. Trials containing eye blinks were discarded (54–60 trials remained). Then the remaining trials of the EOG and MEG signals were averaged, and the baselines of their EOG and MEG signals were corrected by subtracting the mean of the prestimulus intervals from the averaged signals. The data collected during the initial and terminal phases of the smooth pursuit eye movements (duration, 357 ms) were discarded and those during the maintenance phase (358–3214 ms) were used for analysis in the hyVBED and VBED methods. The beginning of the analysis period was defined as the origin of the time axis in Figs. 10–12.

MRI and fMRI recordings

Structural MRI and fMRI data were obtained by a 1.5-T MRI scanner (Shimadzu-Marconi, Magnex Eclipse). Subject head motions were sustained by a bite bar. Whole brain functional images were obtained with an echo-planar imaging sequence (repetition time = 3 s, echo time = 49 ms, field of view = 192 mm × 192 mm, matrix = 64 × 64, voxel size = 3 mm × 3 mm × 5 mm). T1-weighted images (repetition time = 20 ms, echo time = 2.26 ms, field of view = 256 mm × 256 mm, matrix = 256 × 256, voxel size = 1 mm × 1 mm × 1 mm) were also acquired for the construction of a cortical surface model for current source localization. During the fMRI measurements, eye movements were recorded with the MR-Eyetracker (Cambridge Research Systems). A block design was used for the fMRI experiments. One session consisted of four repetitions of a task block and a rest block. In the task block, a period of 4 s, where the subject pursued the moving target, was followed by a period of 2 s, where the subject was allowed to blink. These periods were repeated alternately three times (total 18 s). In the rest block, the target was presented at the center of the screen for 18 s. Five sessions were conducted.

Cortical model

We constructed a cortical model for source localization based on Kajihara et al. (2004). A polygon cerebral cortex model (about 20,000 vertex points; intervertex distance, 1.5 ± 0.5 mm) was constructed using BrainVoyager software (Brain innovation) from the T1 structural image. Brain current sources were located on the vertex points of the cortical surface model with the orientation of the current sources perpendicular to the cortical surface. The positions of each eyeball were obtained by visual inspection from the T1 structural image, and a single dipole was placed at the center of each eyeball. MEG and MRI coordinates were co-registered according to Kajihara et al. (2004). First, the 3D shape of a face measured by a 3D scanner (VIVID 700, Konica Minolta, Japan) was registered to the MRI image. Then, four calibration marker coils on the face were measured by a 3D scanner and a MEG system, and marker positions were registered both on MEG and MRI. Finally, MEG and MRI coordinates were co-registered through the registered marker positions.

FMRI analysis

The fMRI data were analyzed using the SPM99 (Wellcome Department of Cognitive Neurology) to restrict locations of current sources and obtain prior information of the variances of the brain current sources. All EPI images of each session were co-registered to the first EPI image of the session to the correct head movements. Then the images were smoothed with a Gaussian filter of 6 × 6 × 10 mm FWHM. We defined a boxcar function for the task blocks and investigated the brain regions that significantly increased their activities ($p < 0.01$ uncorrected). Precentral sulcus, lateral occipito–temporal cortex, and intraparietal cortex were activated. The T-values obtained by the fMRI analysis were used for the prior information of the variance of brain currents for the hyVBED and VBED method. Possible locations of the brain current sources were restricted to the vertex points in the brain regions where statistically significant fMRI activities were found (around 2000 current sources for the horizontal eye movements, 7985 current sources for the circular eye movements).

Parameter settings for the hierarchical Bayesian method in real experiments

The hyVBED method was applied to the trial-averaged EOG and MEG data using an individual's cortical surface model and fMRI results. The hyVBED method has six types of parameters: the prior current variances for the brain and eye current sources $\bar{v}_{\mathrm{brain}0(l)}$, $\bar{v}_{\mathrm{eye}0(k)}$, the confidence parameters corresponding to these prior current variances $\gamma_{\mathrm{brain}0(l)}$, $\gamma_{\mathrm{eye}0(k)}$, and the initial values of $\mathbf{F}$ and $\mathbf{u}$. For each vertex point, we assumed the following relations between the prior variance of a brain current source $\bar{v}_{\mathrm{brain}0(l)}$ and a T-value obtained from fMRI data $t_{(l)}$ (Yoshioka et al., 2008):

$$\bar{v}_{\mathrm{brain}0(l)} = v_0 + (m_0 - 1)v_0 t_{(l)}^2, \tag{13}$$

where $t_{(l)}$ was a normalized T-value on the l-th vertex. Normalized T-values were computed by dividing the original T-values by the maximum of those T-values (thus ranging from 0 to 1). v_0 provides a baseline of the current variance, which was estimated from the prestimulus interval of the MEG data by the minimum norm estimation (Wang et al., 1992). m_0 is a variance magnification parameter that specifies the scaling between the current variances in the baseline and task periods. We used $m_0 = 100$. The confidence parameters for brain current sources $\gamma_{\mathrm{brain}0}$ were set to 10 for all vertex points. The prior variances of eye current sources $\bar{v}_{\mathrm{eye}0}$ were set to the maximum of $\bar{v}_{\mathrm{brain}0}$. For the horizontal eye movements, only horizontal EOGs were used and the initial values of resistance $\mathbf{F}$ to the horizontal EOGs were set to 1000 Ω·m. For the circular eye movements, both horizontal and vertical EOGs were used and the initial values of the corresponding elements of $\mathbf{F}$ were set to 1000 Ω·m. The initial values of the corresponding elements of EOG bias $\mathbf{u}$ were set to 0 μV.

Simulation experiments

We conducted two kinds of simulation experiments using artificial EOG and MEG data to evaluate the performance of the hyVBED method. In simulation experiment 1, we evaluated how a position of brain current sources affects the performance of eye artifact removal. In simulation experiment 2, we generated artificial data mimicking real experiments during smooth pursuit eye movements and tested the hyVBED method under various conditions.

Analysis settings for hyVBED and VBED methods

We used the following parameters and initial values for the hyVBED and VBED methods. These parameters were common to all simulation experiments described below. For the hyVBED method, the initial values of resistance $\mathbf{F}$ and EOG bias $\mathbf{u}$ were set to the same values used in the real data analysis (resistance, 1000 Ω·m; bias 0 μV). As parameters in the hyper-prior distribution of brain and eye current variances (see Eq. (8)), we used different parameters from real data analysis, i.e., noninformative priors ($\gamma_{\mathrm{brain}} = 0$, $\gamma_{\mathrm{eye}} = 0$, and $\bar{v}_{\mathrm{brain}0(l)}$, $\bar{v}_{\mathrm{eye}0(l)}$ were arbitrary) were used, because no fMRI information was assumed in this simulation study. Four types of parameters $\bar{v}_{\mathrm{brain}0(l)}$, $\bar{v}_{\mathrm{eye}0(l)}$, $\gamma_{\mathrm{brain}0(l)}$, and $\gamma_{\mathrm{eye}0(k)}$ for the VBED method were set to the same values as those for the hyVBED method.

Performance evaluation measure

Since the true values are known in the simulation experiment, we used the normalized root mean square errors (nRMSE)

between the true and estimated values as a performance evaluation measure:

$$nRMSE = \sqrt{\sum_{t=1}^{T} ||\mathbf{X}_{true}(t) - \overline{\mathbf{X}}(t)||^2 \Big/ \sum_{t=1}^{T} ||\mathbf{X}_{true}(t)||^2} \times 100, \qquad (14)$$

where $\mathbf{X}_{true}(t)$ and $\overline{\mathbf{X}}(t)$ are the true and estimated values, respectively. This measure expresses an estimation error normalized by the norm of the true value in percent. The nRMSE of 0 means that the estimated value is equal to the true value. The nRMSE of 100 means that the difference between estimated and true values is the same level as the norm of the true value.

The hyVBED method does not directly estimate the eye artifact component in the MEG sensor space. It estimates eye current sources simultaneously with brain current sources. To evaluate the artifact removal performance in the sensor space, we reconstructed an eye artifact component in the MEG sensor space from estimated eye current sources $\overline{\mathbf{J}}_{eye}(t)$ as

$$\overline{\mathbf{B}}_{artifact}(t) = \mathbf{G}_{eye} \cdot \overline{\mathbf{J}}_{eye}(t). \qquad (15)$$
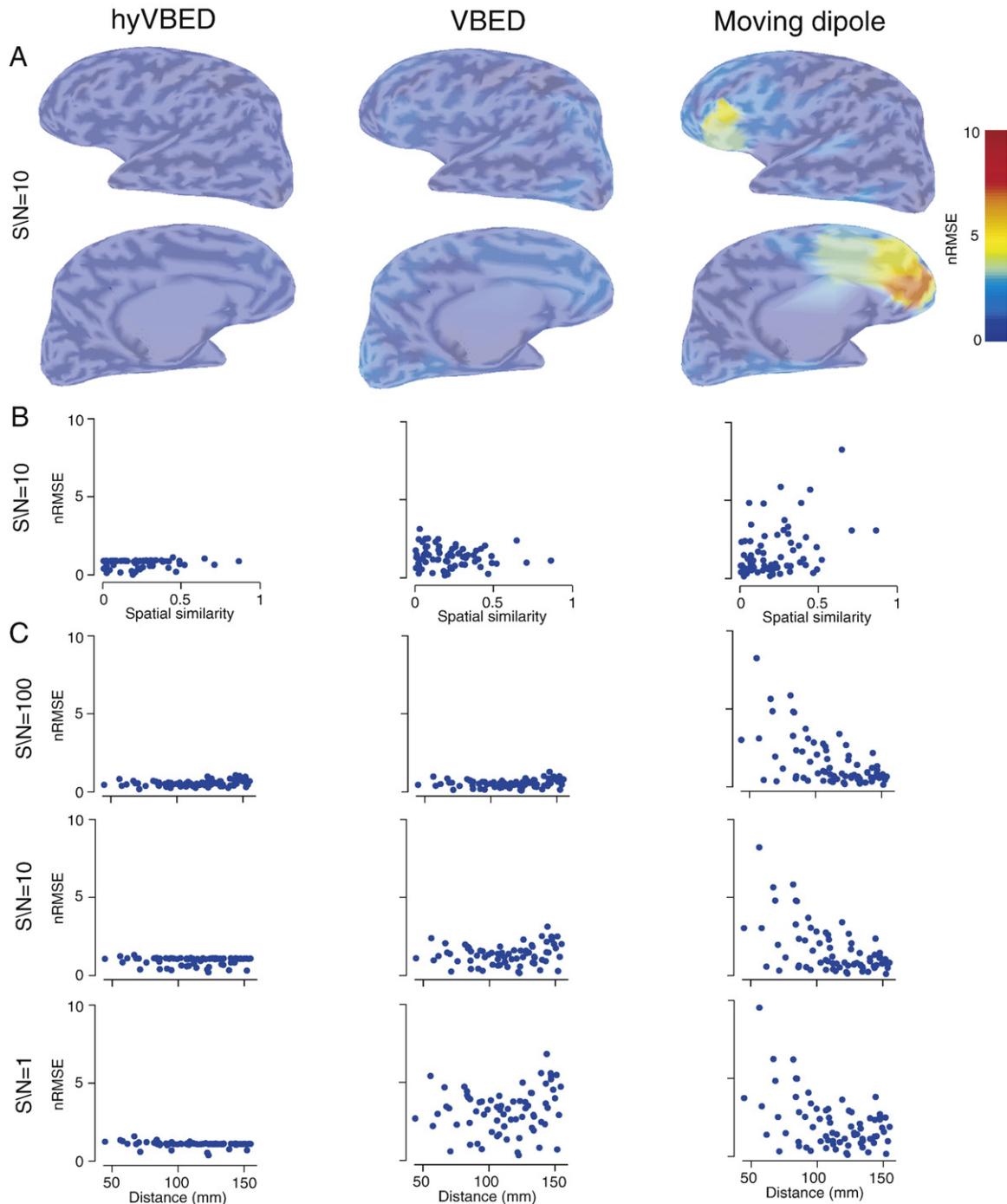
Fig. 3. Normalized root mean square errors (nRMSE) of eye artifact removal for various locations of brain current sources. (A) The nRMSEs for MEG S/N = 10 were mapped to the area where brain current sources were assumed on inflated brain surface models. (B) NRMSE versus spatial similarity between lead fields for eye and brain current sources. (C)The nRMSE of all the three levels of MEG S/N ratio were plotted versus the mean distance between eye and the brain area.

The estimation error of eye artifacts $\overline{\mathbf{B}}_{\text{artifact}}(t)$ was evaluated by the nRMSE measure (Eq. (14)).

Simulation experiment 1: brain current sources in various locations

The purpose of this simulation experiment is to evaluate how a location of brain current sources affects the performance of the eye artifact removal.

Setting

We generated multiple simulation data sets, each of them consisted of a single dipole at each eyeball and brain current sources distributed around a single peak dipole. A low-spatial-resolution cortical model (8036 vertex points; inter-vertex distance, 4.9±1.6 mm) was constructed from subject 1's T1 image in the real experiment and the eye dipoles were placed at the same positions as the real experiment. The location of the brain current sources was changed systematically as follows. The left hemisphere was divided into 72 areas that have almost equal area and the inter-area distance (25±7 mm). A peak dipole (amplitude, 0.3 nAm) was placed at the center of each area and a Gaussian spatial filter with standard deviation 10 mm was applied to generate distributed brain current sources around the peak position (about 50 dipoles were contained). The waveforms of the eye and brain current source were assumed to be the same form. These waveforms were one cycle of sinusoidal waveform at the frequency of 1 Hz. Only the horizontal component of the eye dipoles was considered (corresponding to the horizontal eye movement task). The amplitude of each eye dipole was 20 nAm (the average for the three subjects estimated from the real MEG experiments). Magnetic fields were generated based on the MEG forward model (Eq. (2)) using the simulated eye and brain current sources described above. We assumed 188 MEG sensors with the same positions as the real experiments for subject 1. The observation noise generated from the Gaussian distribution with spherical covariance (i.e., independent among MEG sensors) was added. We prepared three levels of the MEG observation noise (S/N = 1, 10, 100) to test the influence of the observation noise. Artificial EOG data were generated based on the EOG forward model (Eq. (1)) where all the elements of the resistance were 1250 Ω·m, the left horizontal EOG bias, the right horizontal EOG bias and the both vertical horizontal EOG biases were 5 μV, −5 μV, and 0 μV, respectively. The S/N ratio of the EOG observation noise was 5000. These resistances and the EOG biases were in a similar range of estimated values from the real experimental data.

Result

We explain the estimation results in the case of MEG S/N = 10 at first. The nRMSEs of the eye artifacts were mapped onto the area where the brain current sources were assumed (Fig. 3A). The nRMSEs
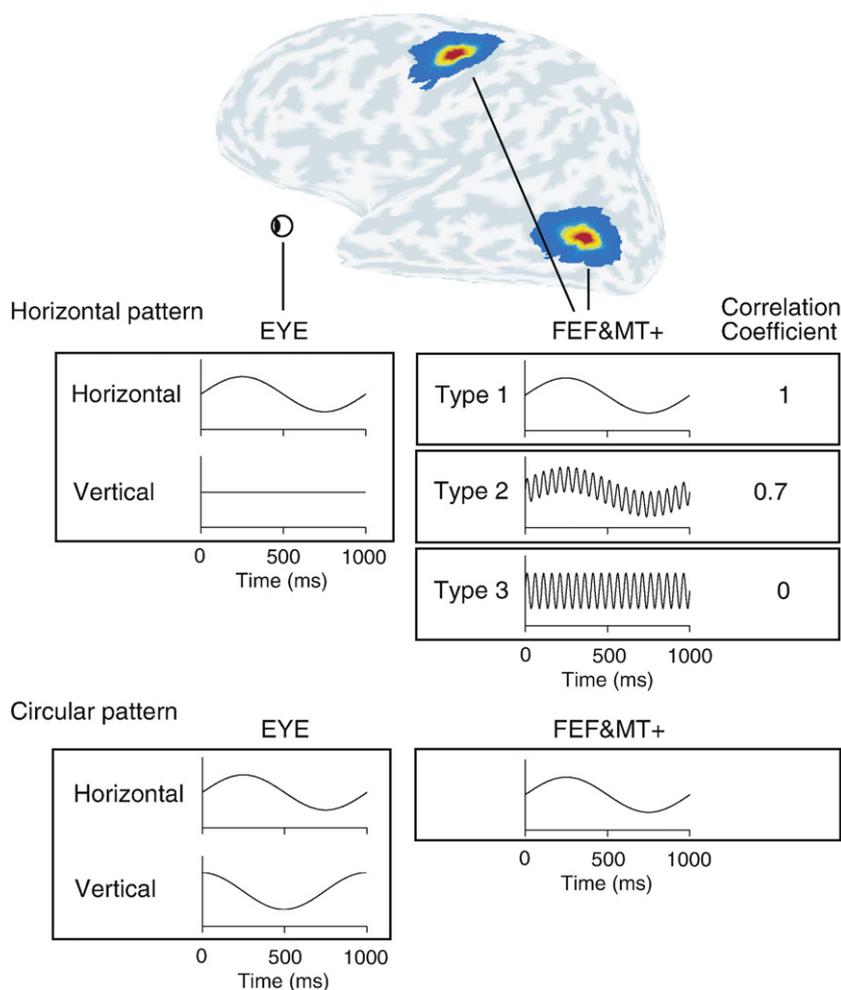
Fig. 4. Settings of the simulation experiment 2. We assumed the brain activities in FEF and MT+ of both hemispheres as shown in an inflated brain surface model. Color scale changes from blue to red as the amplitude changes from low to high. Corresponding to horizontal and circular eye movements, the waveforms shown in 'horizontal pattern' and 'circular pattern' were assumed for eye current sources, respectively. The horizontal and vertical components of eye current sources are shown for each pattern. We prepared three different types of brain current sources for the 'horizontal pattern'. Type 1 brain current sources were completely correlated with eye current sources (correlation coefficient = 1). Type 2 brain current sources were highly correlated with eye current sources (correlation coefficient = 0.7). Type 3 brain current sources were uncorrelated with eye current sources (correlation coefficient = 0).

were also plotted versus the spatial similarity between lead fields of the eye and brain current sources (Fig. 3B) and the mean distance between the eyes and brain area (Fig. 3C, middle row). The spatial similarity was defined as the absolute correlation of these lead fields. The nRMSEs of the hyVBED method were lower than 1.2 in all assumed locations of brain current sources, while that of the moving dipole method (see Appendix B for methodological details) became larger when the mean distance was smaller or the spatial similarity was higher. The errors of the VBED method became larger in frontal and occipital areas. In particular, in the occipital areas it was more than twice the errors of the hyVBED method.

We also plotted the nRMSEs versus the mean distance between the eye and brain area in the case of MEG S/N=1 and 100 (Fig. 3C, top and bottom rows). Although the nRMSEs of the hyVBED and VBED

methods were similar level in the case of the highest S/N ratio (=100), those of the VBED method became larger as the S/N ratio was getting lower (Fig. 3C, middle column). On the other hand, the nRMSEs of the hyVBED method were almost the same for the different noise level. In S/N=1, the highest nRMSE of the hyVBED, VBED and moving dipole methods were 1.6, 6.9 and 9.6, respectively. The results of the moving dipole method were not changed regardless of noise levels. However, the nRMSE of the moving dipole method increased as the mean distance between the eye and brain area decreased.

Simulation experiment 2: brain current sources in FEF and MT+

We constructed similar simulation settings to the real experiment during smooth pursuit eye movements and manipulated various
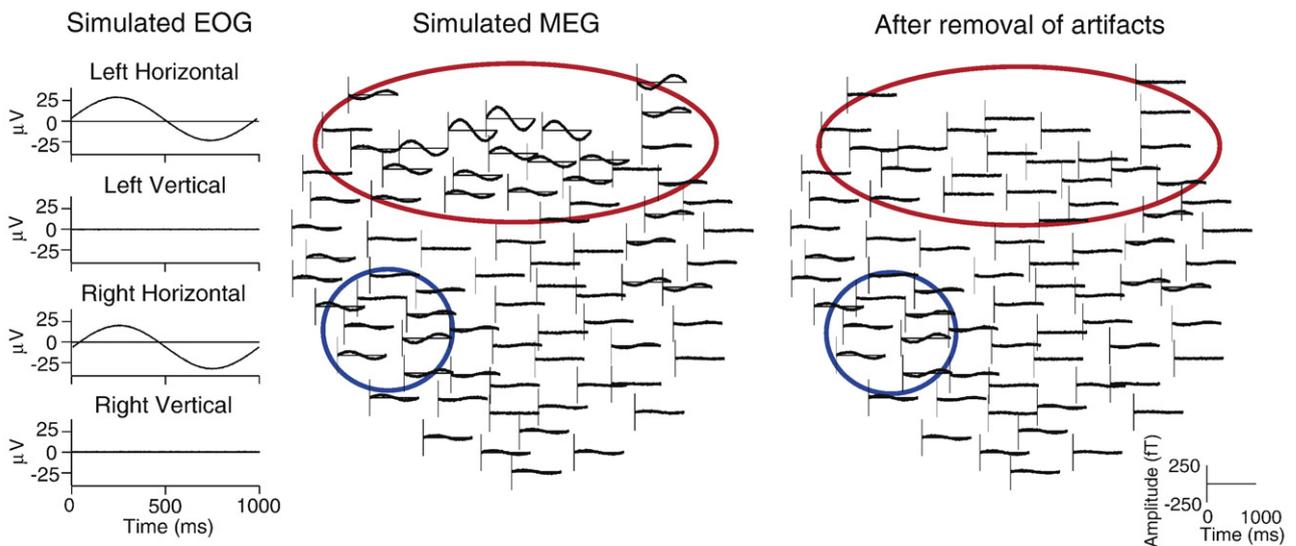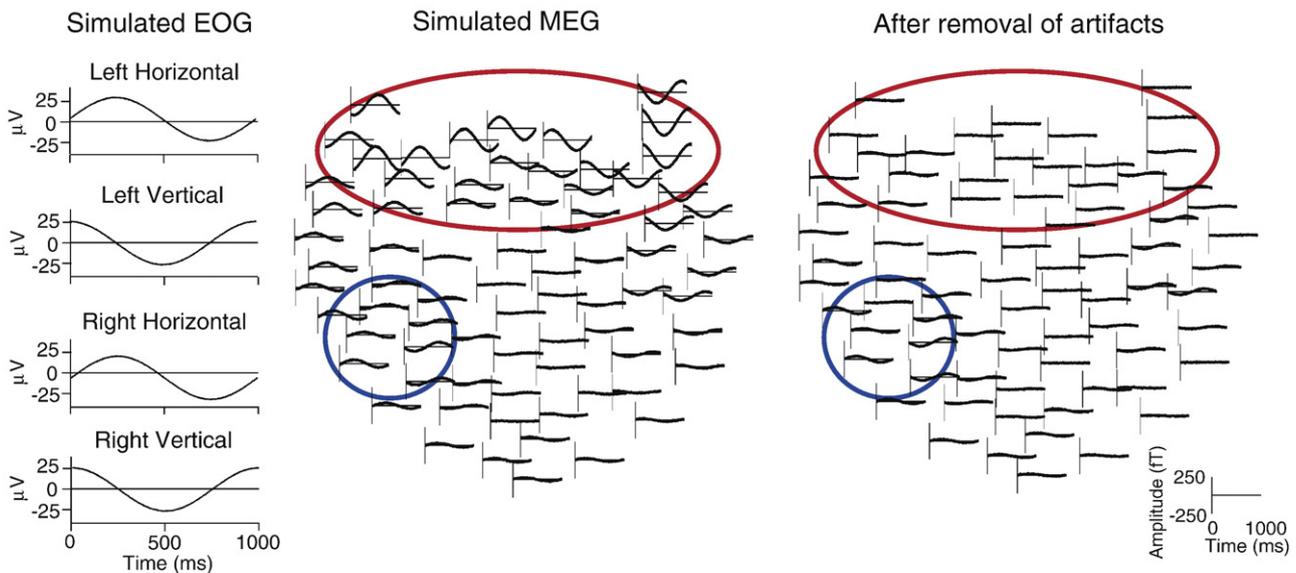
Fig. 5. Simulated EOGs (left), simulated MEGs from eye and type 1 brain current sources (middle) and MEGs after removing eye artifacts by hyVBED method (right) in the simulation experiment 2. (A) Horizontal eye movements. (B) Circular eye movements. Eye artifacts generated from horizontal and circular eye movements were successfully removed by the hyVBED method on sensors located in frontal parts. These sensors were strongly contaminated by eye artifacts in simulated MEG data (emphasized by a red circle). In contrast, the magnetic fields were not changed on sensors in occipital parts. It shows that the magnetic fields from simulated brain current sources (emphasized by blue circles) were conserved. 75 sensors from total 188 sensors were selected for the visualization purpose, which were distributed in almost equal distance.
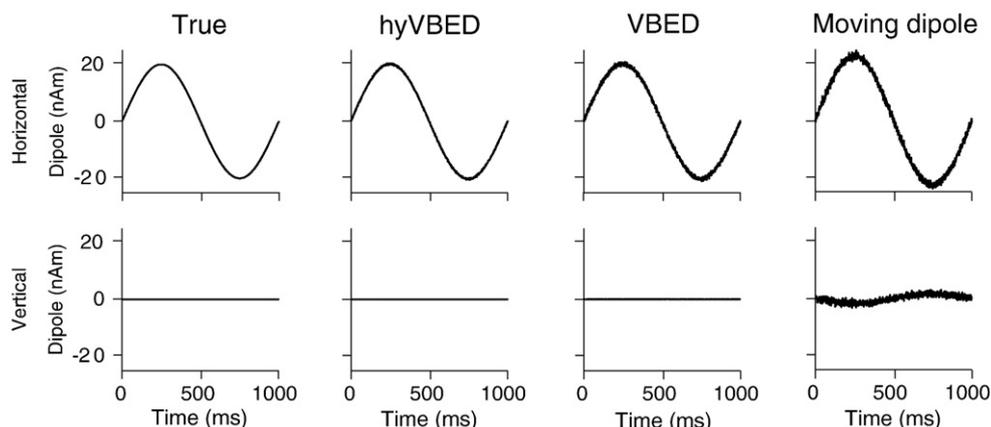
Fig. 6. Estimated eye current sources for horizontal eye movements by the hyVBED, VBED and moving dipole methods from the baseline setting. Upper and lower panels show horizontal and vertical components of estimated eye current source, respectively. Simulated eye current sources are shown in the leftmost panels.

aspects of the simulation settings to validate the applicability of the hyVBED method for real data.

Baseline simulation setting

In this simulation experiment, distributed brain current sources were placed in four areas: the left frontal eye field (FEF), the right FEF, the left middle temporal + area (MT+), and the right MT+ (Fig. 4). To locate these current sources, we employed the same high-resolution cortical model of subject 1 in the real experiment (23,623 vertex points; intervertex distance, 1.5±0.5 mm). FEF and MT+ were reported to be involved in smooth pursuit eye movements (Newsome et al. 1988; MacAvoy et al., 1991; Fukushima et al., 2002; Petit and Haxby, 1999; Krauzlis, 2004; Kawawaki et al., 2006). The same amplitudes, the same spatial filter and the same waveforms of current sources as described in the simulation experiment 1 were used. Note that the eye and brain current sources have equal waveforms and only the horizontal component of eye dipoles was non-zero whereas the vertical component was zero for all time points (see the Type 1 setting in Fig. 4). These waveforms of the eye dipoles correspond to a real experiment when a subject pursuits a horizontally moving target. Magnetic fields and EOGs were generated using the same parameters as described in the simulation experiment 1. In this setting, MEG observation noise of S/N ratio=10 was added (Fig. 5A). When estimating current sources, we used the restricted brain model in which the locations of brain current sources are restricted a priori. This model is available when we have sufficiently reliable spatial information about brain activities from fMRI experiment (employed in the real experiment analysis, see Methods). In this simulation experiment, we restricted the source locations around FEF and MT+ in both hemispheres (1991 dipoles).

Results of the baseline setting

The estimation results for the baseline setting (type 1 brain current, horizontal eye movement and the restricted brain model) were evaluated in terms of eye current sources, MEG sensor level and brain current sources. In Fig. 6, the eye current sources obtained by the hyVBED method, the VBED method and the moving dipole method are shown. The hyVBED method estimated the eye current sources the best. The estimation error of the hyVBED method was smaller than the VBED method. In particular the eye current sources estimated by the VBED method were contaminated by high frequency noise that was observed in magnetic fields. The moving dipole method failed to estimate the vertical component, and the amplitude of the horizontal component was bigger than the true data. The estimation error of the eye current sources was calculated using the nRMSE measure (Eq.

(14)). The hyVBED method showed the smallest estimation error (nRMSE=1.4). The estimation error of the VBED method (nRMSE=3.1) was twice as high as the hyVBED method. That of the moving dipole method was the highest (nRMSE=17).

In Fig. 5A, the magnetic fields, after the eye artifacts were removed using the hyVBED method, are shown. The large sinusoidal waveforms on the sensors in the anterior part (red circle) resulting from eye current sources were removed while the sinusoidal waveforms on the sensors above the left FEF (blue circle) still remained. About the nRMSEs for eye artifacts (Fig. 7), the hyVBED method showed the smallest error (nRMSE=1.0). The nRMSEs of the VBED and moving dipole methods were 1.7 and 4.5 times higher, respectively, than the hyVBED method.

The nRMSEs for the brain current sources are summarized in Fig. 8. The error bars were standard errors calculated from 100 times simulations. An analysis of variance (ANOVA) test and a *t*-test were performed to compare the estimation errors between methods. The estimation errors of the hyVBED method in all four areas are the lowest. In the left and right FEF, the errors of the hyVBED method are
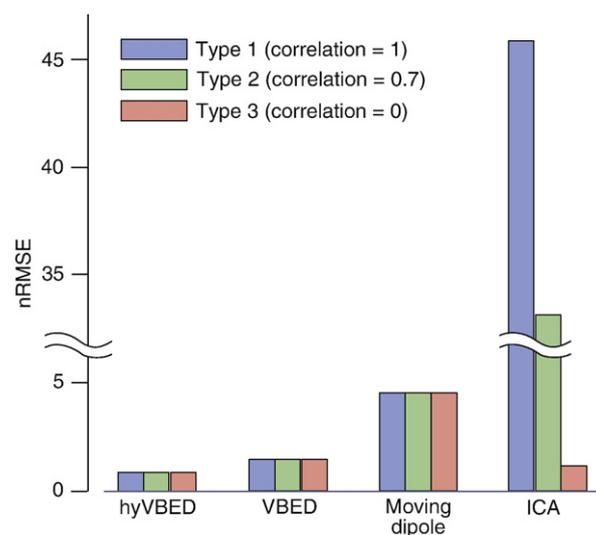
Fig. 7. Normalized root mean square errors (nRMSE) of artifact removal from the simulation experiment 2. The performance of four methods, the hyVBED, VBED, moving dipole method and ICA were compared using the three types of brain current sources (see 'horizontal pattern' of Fig. 4) as indicated by legend in upper right. As the ICA algorithm, the joint diagonalization of cumulant matrices (JADE) was used.
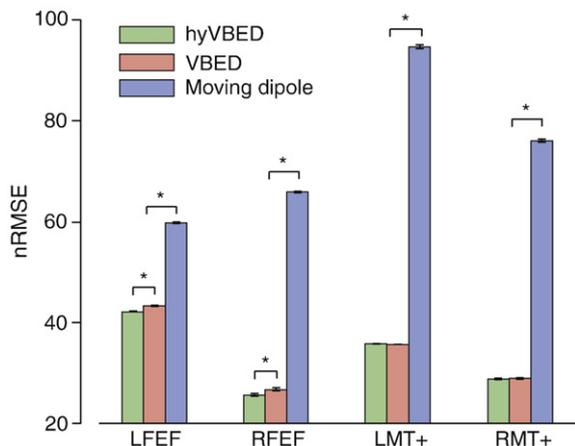
Fig. 8. Normalized root mean square errors (nRMSE) between simulated brain and estimated current sources for the baseline setting (the simulation experiment 2). The hyVBED, VBED and moving dipole methods were compared. Error bars are standard error calculated from 100 Monte Carlo simulations. In all four areas (left and right FEF and MT+), estimation errors of both the hyVBED and VBED methods were lower than those of the moving dipole method ($p<0.05$). In left and right FEFs, the estimation errors of hyVBED method were lower than VBED method ($p<0.05$).

significantly lower than those of the VBED method. There is no significant difference in the left and right MT+ between the hyVBED and VBED methods. This result suggests that accurate estimation of the eye current sources improves the estimation of brain current sources in the anterior areas. The error of the moving dipole method was high, especially in the left MT+. Note that the initial positions of the brain dipoles in the moving dipole method were assumed to be the true positions. Even in this optimal setting, the estimation of the moving dipole method was worse than the hyVBED and VBED methods, because the spatially distributed currents were approximated by a single dipole.

Manipulations of various simulation settings

We manipulated various aspects of the simulation setting from the baseline setting described in the previous paragraph; temporal correlation between eye and brain current sources, waveforms of the eye current sources (circular pattern), the levels of MEG observation noise and the source locations (whole brain) used in estimation.

Influence of temporal correlation between the eye and brain current sources

To evaluate how temporal correlation between the eye and brain current sources affects the performance of artifact removal, the waveform of the brain current sources were changed while the waveform of the eye current source was fixed to the sinusoidal wave of 1 Hz (only horizontal component was considered):

Type 1: 1 Hz sine wave (correlation coefficient = 1)
Type 2: sum of 1 Hz and 20 Hz sine waves (correlation coefficient = 0.7)
Type 3: 20 Hz sine wave (correlation coefficient = 0)

We assumed the same waveforms over four brain areas (MT+ and FEF of both hemispheres).

The performances of the eye artifact removal were compared between methods including ICA. The hyVBED method showed the smallest error (nRMSE = 1.0) in all correlation cases (Fig. 7). There was no resultant difference of reconstruction between the three types of simulated data except ICA. For the correlation coefficient of 0, ICA showed the second smallest error (nRMSE = 1.2) after the hyVBED method. By contrast, if there are correlations between the eye and brain current sources, ICA showed the highest error (nRMSE = 46 for correlation coefficient = 1; nRMSE = 32 for correlation coefficient = 0.7).

Circular eye current pattern

The circular eye current pattern was tested in addition to the horizontal one. The circular pattern consisted of one cycle of sine waveform for the horizontal component and one cycle of cosine wave for vertical component at each eye (Fig. 4). The magnetic fields, after the eye artifacts were removed using the hyVBED method, are shown in Fig. 5B. The large sinusoidal waveforms on the sensors in the anterior part were diminished (circled with red line in Fig. 5B), whereas the magnetic fields from the brain current sources still remained (circled with blue line in Fig. 5B). The errors of artifact removal of the hyVBED method for both circular and horizontal patterns were almost the same (solid lines in Fig. 9).

Influence of the level of MEG observation noise

Even when the MEG S/N ratio was changed to one tenth (S/N = 1) or ten times (S/N = 100), the hyVBED showed the smallest nRMSEs in the all conditions (solid lines in Figs. 9A and B). In particular if the errors in the case S/N = 1 were compared with those in the case S/N = 10, the nRMSE of the hyVBED method did not differ while the nRMSE of the
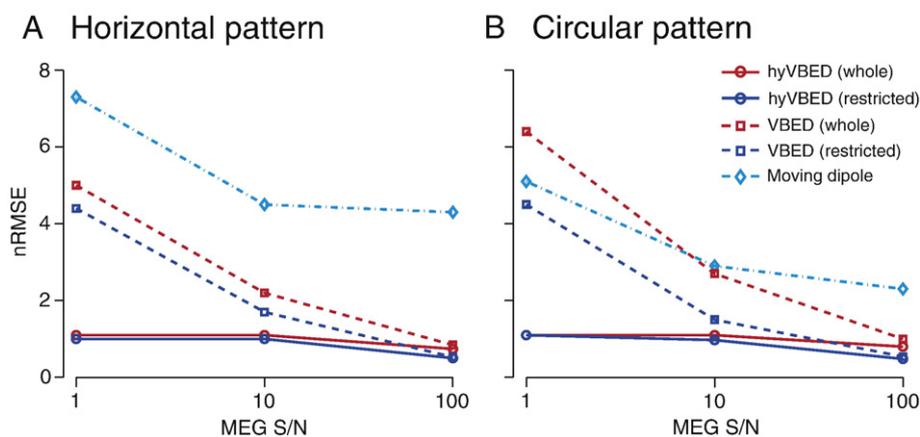
Fig. 9. Normalized root mean square errors (nRMSE) of artifact removal under various simulation settings from the simulation experiment 2. The level of MEG observation noise (S/N = 1, 10, 100), the source location (the whole or restricted brain model) and eye current patterns (horizontal and circular patterns) were manipulated and the performances of artifact removal were compared among hyVBED (solid lines), VBED (dashed lines) and moving dipole method (chain cyan line). The results corresponding to the horizontal eye movement pattern and the circular eye movement pattern are shown in panel (A) and (B), respectively. The whole brain model (red lines) and the restricted brain model (blue lines) were applied for the hyVBED and VBED methods. The whole brain model assumed brain current sources over the whole brain whereas the restricted brain model assumed brain current sources only around FEF and MT+ in both hemispheres a priori.

VBED (dashed lines) and the moving dipole methods (chain line) significantly increased. The nRMSEs of the VBED and moving dipole method in the case S/N=1 were about 2.5 and 1.6 times as high as those in the case of S/N=10, respectively.

Whole brain model

We also applied the whole brain model (red lines in Figs. 9A and B), possible current locations were over a whole brain instead of the restricted brain model (blue lines) to see the difference between the nRMSE of the whole brain model and the restricted brain model. The restricted model was superior to the whole brain model in all the cases.

Application to real data

Performance of eye movement task

The performance of the pursuit eye movement during the MEG measurements was estimated as the pursuit gain, which is the ratio of the eye velocity to the target velocity, and the phase lead of the eye movement velocity to the target velocity. For the horizontal eye movement, the pursuit gain and phase lead were 0.81–0.87 and 4–18 ms. For the circular eye movement, the horizontal gain and phase were 1.14 and −17 ms and the vertical gain and phase were 1.36 and 32 ms. These results indicate that each subject performed the task with reasonable precision and in a predictive way.

Eye current sources and resistances

Figs. 10A and B show the estimated eye current sources for subject 1 during the horizontal eye movements by the hyVBED and VBED methods, respectively. The eye current sources estimated by both methods did not differ significantly except that those for the VBED method were noisier than the hyVBED method. The mean amplitudes between subjects had similar levels (hyVBED, left, 21.8 ±3.2 nAm, right, 20.0±10.6 nAm; VBED, left 25.0±3.5 nAm, right, 20.2±9.3 nAm). The S/N ratios of the eye current sources, which were defined as the ratio of the power of 0.7 Hz to that of all other frequencies, were several times greater for the hyVBED (S/N=70 for subject 1, S/N=131 for subject 2, S/N=71 for subject 3) than the VBED method (S/N=24 for subject 1, S/N=18 for subject 2, S/N=30 for subject 3). This is because the hyVBED method estimates eye current sources not only from MEG but also EOG whose S/N ratio is much higher than MEG. To the best of our knowledge, resistance and EOG bias were estimated for the first time by the hyVBED method. The subject average of the left- and right-side resistances were 1486±291 Ω·m and 1821±919 Ω·m, and the left- and right-side EOG biases were 3.5±1.23 μV and −3.43±2.86 μV.

For the circular eye movement, the amplitude of eye current sources, the resistance values and the EOG biases were estimated in similar range to those for the horizontal eye movements. The amplitudes of eye current sources of left horizontal, left vertical, right horizontal and right vertical component, were 21.7, 34.5, 36.0 and 36.3 nAm, respectively. The resistance and the EOG bias for these four components were 3005, 1593, 1949, 1351 Ω·m and 0, −7.89, 0.61, 0.60 μV, respectively.

MEG signals after artifact removal

We also confirmed the performance of the hyVBED method to remove the eye artifacts at the MEG level. The raw MEGs of the horizontal eye movement task by subject 1 contain big eye artifacts highly correlated with eye movement for sensors in the frontal region close to the eye (Fig. 11A middle). After the removal of eye artifacts by the hyVBED method, the eye artifacts could not be visible (Fig. 11A right). A similar dramatic effect of eye artifact removal was confirmed for the other two subjects and the circular eye movements (Fig. 11B). The results are summarized in Table 1. More than half of the global magnetic field power (60–80%) was removed from the whole MEG sensors (left column in Table 1). In particular, most of the power at the frontal sensors near the eyes (≤120 mm) was diminished significantly (middle column in Table 1). About 95% of the power of the 0.7 Hz component in the frontal sensors was removed (right column in Table 1).

Brain current sources

For the horizontal eye movement experiment, we investigated spatial and temporal patterns of brain current sources that were estimated with the eye current source simultaneously. Figs. 12A and B show the spatial patterns of the brain current sources of subject 1, estimated by the hyVBED method. The root mean square of each current source was integrated across the entire analysis period (0–2856 ms) and mapped on an inflated cortical surface model. Strong activities were found in several areas of the prefrontal and temporal cortices of both hemispheres, corresponding to those identified as human FEF and MT+ by the imaging study (Petit and Haxby, 1999; Kawawaki et al., 2006). Additionally, weaker activities were found in areas corresponding to the intraparietal and occipital cortices.

The temporal patterns of the current sources in FEF and MT+ are shown in the insets of Figs. 12C, D and F–H. Firstly, we manually defined the area that formed a cluster around a peak current source,
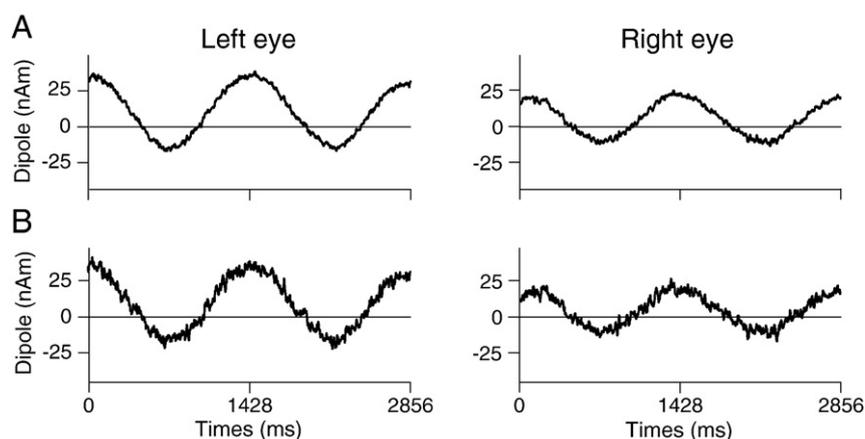
Fig. 10. (A) Eye current sources estimated by hyVBED method and (B) those estimated by the VBED method (from the real MEG experiment). Time courses of the estimated eye current source of subject 1 are plotted during maintenance phase.

A Horizontal Eye Movement
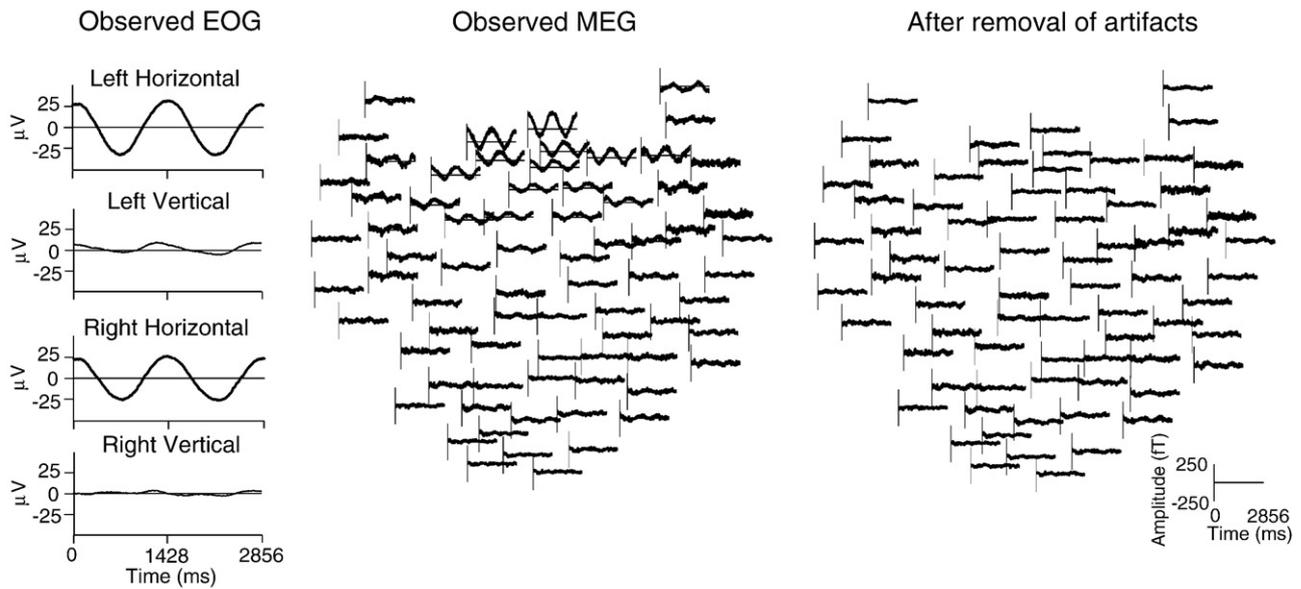
B Circular Eye Movement
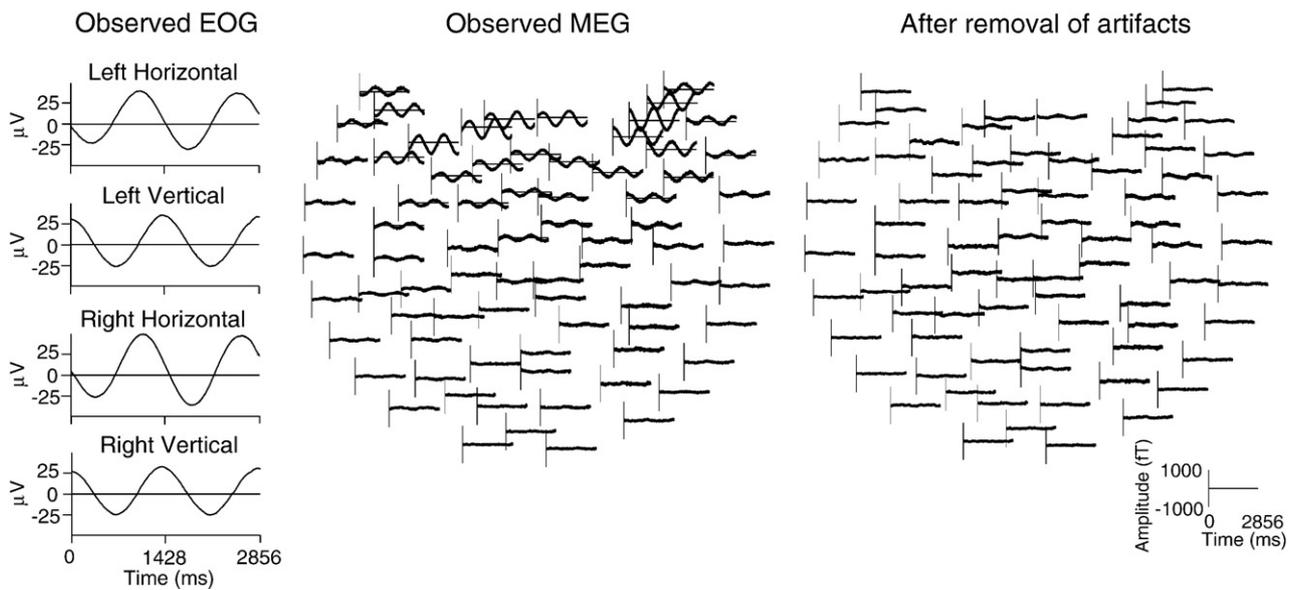
Fig. 11. Trial averaged EOGs (left), trial averaged MEGs (middle) and MEGs after eye artifact removed by hyVBED method (right). (A) Horizontal eye movement. (B) Circular eye movement. Each panel in middle and right columns shows a time course of magnetic fields on each MEG sensor. 75 sensors from total 188 (A) and 210 (B) sensors were selected for the visualization purpose, which were distributed in almost equal distance.

keeping from overlapping with the other area. And then the spatially cumulative value (Fig. 12, blue trace in the insets) was calculated by summing up the current sources more than 50% of the peak in each area. For comparison, the eye position converted from the EOG data was overwritten (red trace in the insets). To estimate a time lag between the current sources and the eye position, cross correlations with different time delays between the cumulative current sources and the eye position time course were calculated. We found two active loci in the left FEF that are supposed to be the major cortical area to drive the pursuit eye movement (Krauzlis 2004). The one located in the posterior part of the left FEF exhibited activity with an amplitude of 0.7 nAm, whose frequency (0.7 Hz) and phase lead (5°) roughly matched the eye position (Fig. 12D). The other locus located in the anterior part of the left FEF showed activity with an amplitude of

0.5 nAm, and its frequency was the same as the eye position and its phase lead was almost opposite (174°) (Fig. 12C). It might be considered that a single current source at one side of the sulcus was estimated as two current sources with opposite phases at both sides of the sulcus when the normal direction of both sides had the opposite direction, due to the ill-posed nature of the MEG inverse problem. However, the two peak positions were not close to each other (distance of the two peak positions, 11 mm; relative angle, 156°). These facts indicate that the estimated two current loci were inherent. The amplitude (0.5 nAm) and frequency (0.7 Hz) of the right FEF were identical to the anterior part of the left FEF, but the phase lead was different (44°) (Fig. 12G). Stronger activities (2–8 times bigger than the FEF amplitude) were found in the MT+, which is supposed to be the major cortical area for receiving such visual information as a

Table 1
Amount of magnetic fields removed by hyVBED method

		Whole sensors (%)	Frontal sensors (%)	0.7 Hz power of frontal sensors (%)
Horizontal	Subject 1	82.1	94.1	98.2
	Subject 2	69.9	91.5	96.3
	Subject 3	62.4	83.3	91.8
Circular		87.5	94.1	96.0

Ratios of root mean squared values of removed magnetic fields to those of original magnetic fields (before removing eye artifact) are shown by percentage. Left and middle columns represent ratios for whole and frontal sensors within 120 mm from both eyes (For horizontal, subject 1, 51 sensors; subject 2, 48 sensors; subject 3, 47 sensors; for circular, 72 sensors), respectively. Right column represents a ratio of 0.7 Hz powers of removed magnetic fields in frontal sensors to those of original magnetic fields.

retinal slip during smooth pursuit eye movements (Figs. 12F and H). The left MT exhibited activity with a frequency of 0.7 Hz and phase lead (−148°) to the eye position, while the main frequency (1.4 Hz) of the right MT+ was twice that of the eye position. We further investigated whether activities in the frontal region anterior to FEF might represent the remainder of the eye artifacts, since that region is close to the eyeball that is the source of the eye artifact. The main frequency of the waveform (1.05 Hz) in the left inferior frontal gyrus (Fig. 12E) was higher than the eye position. This result suggests that the estimated current sources in the left FEF might not be due merely to the eye artifacts.

The same spatiotemporal patterns as for subject 1 were found in other two subjects. We found strong activities in FEF and MT+ of both hemispheres; in particular, the waveforms of the left FEF and the right MT+ resembled subject 1. For subject 2, both the posterior and the anterior parts of the left FEF showed similar activities to subject 1, i.e., both frequencies were 0.7 Hz and the phase leads of the posterior and anterior parts were 5° and 154°, respectively (distance of the two peak positions, 5 mm; relative angle, 108°). For subject 3, two active peaks were placed in the dorsal and ventral parts of the left FEF. The frequencies of both activities (0.7 Hz) were the same as the other two subjects, and the phase lead of the dorsal part (−1°) and the ventral part (144°) were roughly matched with those of the posterior and anterior parts of the other two subjects (distance of the two peak positions, 10 mm; relative angle, 153°). The right MT+ of both subjects 2 and 3 showed that the main frequencies (1.4 Hz) were twice as high as the eye position, which was identical with subject 1. In the right FEF and the left MT+, we did not observe consistent results among subjects.

As shown in the last results for subject 1, we also investigated the waveforms of the current sources located nearer an eyeball than FEF for subjects 2 and 3 to check whether the current sources in the left FEF were the remainders of the eye artifacts. We chose a current source located in the Sylvian fissure for subject 2 (it was the only current source anterior to FEF) and a left inferior frontal gyrus for subject 3. The waveforms of the current sources were different from those of the eye position (main frequency for subject 2, 9.8 Hz; subject 3, 2.8 Hz). Finally the Talairach coordinates of the peak vertices in FEF
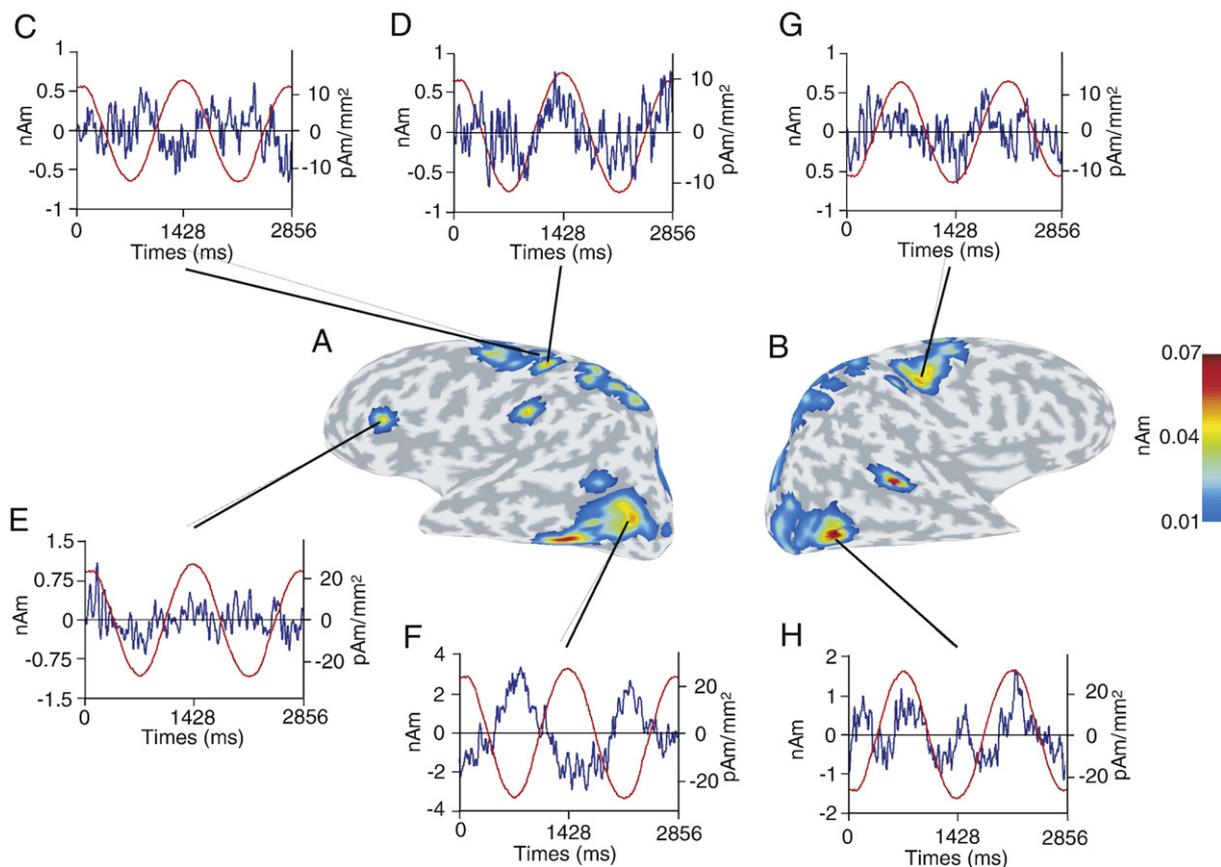
Fig. 12. Spatiotemporal patterns of brain current sources estimated by hyVBED method. (A) and (B), spatial patterns of brain current sources in left and right hemispheres of subject 1, respectively. Mean root square intensity of current sources is shown on inflated brain surface models of the subject by pseudocolor representation. Blue traces of insets (C–H) represent time course of cumulative current sources in loci indicated on the maps, which is the sum of current sources more than half of the peak dipole activity for each loci. The scale on left is a dipole moment and the scale on right is a dipole current density obtained from dipole moment divided by area corresponding to it. Those for red traces, eye position expressed such that leftward deviation is positive in a left hemisphere and vice versa in a right hemisphere. Amplitude of eye position is 4° in all insets. Inset C–H, anterior and posterior parts of left FEF, left inferior frontal gyrus, a left MT+, right FEF, and right MT+, respectively.

and MT+ were computed and compared with those obtained from a previous fMRI study (Petit and Haxby, 1999). The distance discrepancies ranged from 5 to 17 mm.

For two subjects, the VBED method showed almost the same brain current sources as the hyVBED method. However, for one subject, the VBED method showed different results from the hyVBED method. In particular, the left FEF exhibited a different phase lead (52°) to the eye position while the other two subjects showed the same phase lead (5°).

Discussion

In this study, we proposed a novel method for the removal of eye artifacts to investigate brain activities during eye movements using MEG. Our method is an extension of the hierarchical variational Bayesian method for source localization proposed by Sato et al. (2004). The key idea of our method is construction of two forward model; a forward model for MEG consisting of both the brain and eye current sources and that for EOG consisting of eye current sources. According to these forward models, the brain and eye current sources are estimated from EOG and MEG data by the variational Bayesian method. In particular, its novel point is that resistance and EOG bias are estimated simultaneously with current sources, improving the estimation of eye current sources. Using simulated data, we have confirmed that our method is capable of removing eye artifacts wherever brain current sources are located. We have also demonstrated that the performance of our method was significantly better than the independent component analysis when eye and brain current sources were correlated. Furthermore we applied it to real experimental data during smooth pursuit eye movements and showed that most eye artifacts were successfully removed while brain activities related to the task remained.

In the simulation experiment 1, we investigated positional influence of brain current sources. The hyVBED and VBED method robustly removed artifacts regardless of the distance between the eye position and the position where a brain current source was placed. On the other hand, the estimation error of the moving dipole method increased as the positions of eye and brain current source got close. Moreover the hyVBED and VBED methods were little affected by the spatial similarity that was defined by absolute correlation of the lead field matrices generated from eye dipole and brain current sources. We further observed that the amount of MEG observation noise degrades the performance of the VBED method while the performance of the hyVBED method was not affected. This result shows the clear benefit of using EOG data for artifact removal. The applicability of the hyVBED method to low S/N ratio data indicates that the hyVBED method may be able to remove eye artifacts from single trial MEG data whose S/N ratio is very low.

In the simulation experiment 2, we generated the artificial data that simulated the real smooth pursuit eye movement experiment and evaluated our method under various simulation settings. When temporal correlation between eye current sources and brain current sources was manipulated, the performance of ICA decreased significantly as the correlation increased while the performances of the hyVBED, VBED, moving dipole, methods were not influenced. This is because ICA relies on the statistical independence of MEG while the latter three approaches rely on different spatial patterns of magnetic fields generated from the brain and eye current sources. The moving dipole method assumes a few dipoles to estimate brain current sources. On the other hand, the hyVBED and VBED methods assume distributed current sources over a cortex. Accordingly the moving dipole method did not sufficiently approximate spatially spreading brain current sources. If brain current sources can be approximated by a few dipoles, the moving dipole method would not be much different from the hyVBED and VBED method.

We also compared the whole brain model with the restricted brain model. In the restricted model, possible locations of brain current sources were restricted around true source locations a priori. The restricted model was superior to the whole brain model. However the restricted model could degrade performance of artifact removal if restricted areas neglect true source locations. In real data, it is generally unknown whether restricted areas ignore important brain sources or not. Therefore it is desirable to determine the brain model (whole or restricted) using some objective criteria such as the Bayes' factor (Daunizeau et al., 2005) or the free energy (Nummenmaa et al., 2007). The free energy can be used for this purpose in our model, but the analytical form is very complicated and thus the model comparison using the free energy will remain as a future work.

In the real MEG experiments, it is fundamentally difficult to confirm that eye artifacts are completely removed while brain activities remain since true brain activities are unknown. The performance of eye artifact removal was evaluated from the following viewpoints. First, the signal level of the MEG sensors was checked. The big sinusoidal eye artifacts near the face were removed whereas MEG signals in the occipital region remained. Second, brain activities in the areas related to smooth pursuit eye movements were investigated. The similar features of the waveforms in the left FEF and right MT+ were observed between three subjects. Third, the influence of eye artifacts in the frontal cortex was examined. If the current sources in the frontal cortex were not affected by eye artifacts, current sources in the other areas might be less affected. The rationale behind this idea is that the eye artifact is more likely to be localized in the frontal cortex since the area is very close to eye. We observed that activities in frontal cortex were not correlated with eye movements. This suggests that FEF activities were not merely the remainder of eye artifacts. These evidences support that the proposed method successfully removed eye artifacts. Although some common features of the brain current sources were observed among three subjects, it was difficult to conclude physiological findings of these brain current sources. This is beyond the aim of this article.

The selection of hyper-parameters is one of key issues for success of our method. We have adopted the same hyper-parameters as described in Yoshioka et al. 2008 in which systematic exploration of the hyper-parameters was discussed. The similar analysis should be desired for our method but this is out of a scope of the current study. Thus we briefly investigated several settings of hyper-parameters; three kinds of the magnification parameter $m_0 = 10, 100, 1000$ and three kinds of confidence parameters $\gamma_{brain0} = 1, 10, 100$. The magnification parameter determines scale of brain current variances and the confidence parameters determines confidence of fMRI information (freedom of the Gamma distribution). When the magnification parameter was varied while the confidence parameter was fixed to 10, the results were not basically changed except one case ($m_0 = 1000$, for one subject). When the confidence parameter was varied, the results was unchanged as long as $m_0 = 10$ or 100. These observations demonstrated that the hyper-parameter setting adopted in this article was in rather robust range. It should be noted that the principled way of choosing hyper-parameter settings may be to maximize the free energy. But Nummenmaa et al., 2007 discussed the difficulty of choosing the hyper-parameters merely based on the free energy.

The choice of initials values is also important. In particular, we found in the simulation study that the result was to some extent sensitive to an initial value of the resistance **F**. If the initial value 0 was used, brain current sources were not estimated correctly although MEG after eye artifact removal seemed to be correct. As the initial value was over 1000, the robust estimation of the brain and eye current estimation was obtained. Therefore 1000 Ω·m was adopted as an initial value for the real experimental data. This choice of the initial value could be reasonable because this value is due to physiological substances consisting of skin and an eyeball thus the order of **F** should not differ among subjects.

We assumed an eye current source as a single dipole placed in the center of an eyeball based on Katila et al., 1981. Instead of a single dipole model, we also tried a seven-dipole model in which six dipoles were placed around the center of an eyeball at a distance of 10 mm.

The result of the seven-dipole model did not differ from that of the single-dipole model in the real experiment. This showed the single dipole model was capable of approximating an eye current source adequately

Our method provides the estimate of resistance as well as that of eye current sources through the MEG and EOG forward models. To the best of our knowledge, this is for the first time that the resistance is estimated from EOG data. If we estimate the resistance value between two electrodes around an eye at an inter-electrode distance of 0.04 m using the skin conductance level 0.3 μS (Malmivuo and Plonsey, 1995), it will be 13,300 Ω·m (=0.04/(0.3 × 10^{-6})). This value is about ten times of our estimates (left, 1486±291; right, 1821±919 Ω·m). This discrepancy is mainly due to overestimation of resistance calculated from the skin conductance level, because the resistance consists of several tissues such as an eyeball whose conductance is much higher than the skin (Malmivuo and Plonsey, 1995).

Once the resistance and EOG bias are determined by the hyVBED method, the eye artifacts can be removed online from only EOG data. Firstly eye current sources are estimated from the EOG data with the determined resistance and EOG bias. Then eye artifacts are calculated by multiplying a lead field matrix by the estimated eye current sources. Finally the eye artifacts are removed from the observed MEG data. This procedure can be conducted online because these calculations are very fast. This method can be applied to MEG data during different eye movement tasks because the resistance and EOG bias do not depend on patterns of eye movements.

Acknowledgments

Appendix A

The derivation and a detailed algorithm for the hyVBED method are explained here. Likelihood functions (Eqs. (4), (5)), prior distributions (Eqs. (7), (9), (10)), and hyper-prior distributions (Eqs. (8), (11)) are summarized below:

[likelihood]

$$P\left(\mathbf{B}|\mathbf{J}_{\text{brain}}, \mathbf{J}_{\text{eye}}, \beta_{\text{B}}\right) \propto \exp$$

$$\times \left[-\frac{1}{2}\beta_{\text{B}} \sum_{t=1}^{T} ||\mathbf{B}(t) - \mathbf{G}_{\text{brain}} \cdot \mathbf{J}_{\text{brain}}(t) - \mathbf{G}_{\text{eye}} \cdot \mathbf{J}_{\text{eye}}(t)||^2\right] \quad (16)$$

$$P\left(\mathbf{E}|\mathbf{J}_{\text{eye}}, \mathbf{F}, \mathbf{u}, \beta_{\text{E}}\right) \propto \exp\left[-\frac{1}{2}\beta_{\text{E}} \sum_{t=1}^{T} ||\mathbf{E}(t) - \mathbf{F} \cdot \mathbf{J}_{\text{eye}}(t) - \mathbf{u}||^2\right] \quad (17)$$

[prior]

$$P_0(\mathbf{J}|\boldsymbol{\alpha}_{\mathbf{J}}) \propto \exp\left[-\frac{1}{2} \sum_{t=1}^{T} \mathbf{J}'(t) \cdot \mathbf{A}_{\text{J}} \cdot \mathbf{J}(t)\right] \quad (18)$$

$$P_0(\mathbf{F}|\boldsymbol{\alpha}_{\mathbf{F}}) = \exp\left[-\frac{1}{2} \sum_k \boldsymbol{\alpha}_{\text{F}(k)} \mathbf{F}^2_{(k,k)}\right] \quad (19)$$

$$P_0(\mathbf{u}|\boldsymbol{\alpha}_{\mathbf{u}}) = \exp\left[-\frac{1}{2} \sum_k \boldsymbol{\alpha}_{\text{u}(k)} \mathbf{u}^2_{(k,k)}\right] \quad (20)$$

$$P_0(\beta_{\text{B}}) = 1/\beta_{\text{B}} \quad (21)$$

$$P_0(\beta_{\text{E}}) = 1/\beta_{\text{E}} \quad (22)$$

[hyper−prior]

$$P_0(\boldsymbol{\alpha}_{\mathbf{J}}) = \prod_i \mathcal{G}\left(\boldsymbol{\alpha}_{\text{J}(i)}|\overline{\boldsymbol{\alpha}}_{\text{J0}(i)}, \gamma_{\text{J0}(i)}\right) \quad (23)$$

$$P_0(\boldsymbol{\alpha}_{\mathbf{F}}) = \prod_k 1/\boldsymbol{\alpha}_{\text{F}(k)} \quad (24)$$

$$P_0(\boldsymbol{\alpha}_{\mathbf{u}}) = \prod_k 1/\boldsymbol{\alpha}_{\text{u}(k)} \quad (25)$$

where $\mathbf{J}' = [\mathbf{J}'_{\text{brain}} \ \mathbf{J}'_{\text{eye}}]$. We estimate a posterior distribution $P(\mathbf{J}|\mathbf{B},\mathbf{E})$ from the distributions listed above when MEG $\mathbf{B}$ and EOG $\mathbf{E}$ are observed and lead field matrices $\mathbf{G}_{\text{brain}}$ and $\mathbf{G}_{\text{eye}}$ are given:

$$P(\mathbf{J}|\mathbf{B},\mathbf{E}) = \int d\mathbf{F_u} d\boldsymbol{\alpha} P(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha}|\mathbf{B},\mathbf{E})$$

$$P(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha}|\mathbf{B},\mathbf{E}) = \frac{P(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha},\mathbf{B},\mathbf{E})}{P(\mathbf{B},\mathbf{E})}$$

$$P(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha},\mathbf{B},\mathbf{E}) = P\left(\mathbf{B}|\mathbf{J}_{\text{brain}},\mathbf{J}_{\text{eye}},\beta_{\text{B}}\right)$$
$$\times P\left(\mathbf{E}|\mathbf{J}_{\text{eye}},\mathbf{F},\mathbf{u},\beta_{\text{E}}\right) P_0(\mathbf{J}|\boldsymbol{\alpha}_{\mathbf{J}}) P_0(\mathbf{F}|\boldsymbol{\alpha}_{\mathbf{F}}) P_0(\mathbf{u}|\boldsymbol{\alpha}_{\mathbf{u}}) P_0(\boldsymbol{\alpha})$$

$$P(\mathbf{B},\mathbf{E}) = \int d\mathbf{J} d\mathbf{F_u} d\boldsymbol{\alpha} P(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha},\mathbf{B},\mathbf{E})$$

where the variables are summarized in $\mathbf{F_u} = \{\mathbf{F},\mathbf{u}\}$ and $\boldsymbol{\alpha} = \{\boldsymbol{\alpha}_{\mathbf{J}},\boldsymbol{\alpha}_{\mathbf{F}},\boldsymbol{\alpha}_{\mathbf{u}},\beta_{\text{B}},\beta_{\text{E}}\}$. Since the calculation of the marginal distribution $P(\mathbf{B},\mathbf{E})$ contains a nonlinearity about the integration of $\boldsymbol{\alpha}$, the posterior cannot be calculated analytically. We employed the variational Bayesian (VB) method for posterior approximation. In this method, a trial distribution $Q(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha})$ is introduced to maximize the free energy defined by the following equation:

$$F(Q) = \int d\mathbf{J} d\mathbf{F_u} d\boldsymbol{\alpha} Q(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha}) \log\left[\frac{P(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha},\mathbf{B},\mathbf{E})}{Q(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha})}\right]$$

$$= \log P(\mathbf{B},\mathbf{E}) - KL[Q(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha})||P(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha}|\mathbf{B},\mathbf{E})], \quad (26)$$

where KL means the Kullback–Leibler distance. We assume the following factorized approximation:

$$Q(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha}) = Q_{\mathbf{J}}(\mathbf{J}) Q_{\mathbf{F}}(\mathbf{F_u}) Q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}). \quad (27)$$

Under this assumption, the free energy is rewritten as

$$F(Q) = \int d\mathbf{J} d\mathbf{F_u} d\boldsymbol{\alpha} Q_{\mathbf{J}}(\mathbf{J}) Q_{\mathbf{F}}(\mathbf{F_u}) Q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) \log P(\mathbf{B},\mathbf{E}|\mathbf{J},\mathbf{F_u})$$
$$- KL\left[Q_{\mathbf{J}}(\mathbf{J}) Q_{\mathbf{F}}(\mathbf{F_u}) Q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})||P_0(\mathbf{J}|\boldsymbol{\alpha}_{\mathbf{J}}) P_0(\mathbf{F_u}|\boldsymbol{\alpha}_{\mathbf{F}},\boldsymbol{\alpha}_{\mathbf{u}}) P_0(\boldsymbol{\alpha})\right]. \quad (28)$$

A maximization of the free energy is done by iterative maximization of $Q_{\mathbf{J}}(\mathbf{J})$, $Q_{\mathbf{F}}(\mathbf{F_u})$, and $Q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})$. In the first step (J-step), the free energy $F(Q)$ is maximized with respect to $Q_{\mathbf{J}}(\mathbf{F_J})$, while $Q_{\mathbf{F}}(\mathbf{F_u})$ and $Q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})$ are fixed:

$$Q_{\mathbf{J}}(\mathbf{J}) \propto \exp\left[\int d\mathbf{F_u} d\boldsymbol{\alpha} Q_{\mathbf{F}}(\mathbf{F_u}) Q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) \log P(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha},\mathbf{B},\mathbf{E})\right]. \quad (29)$$

In the second step (F-step), the free energy $F(Q)$ is maximized with respect to $Q_{\mathbf{F}}(\mathbf{F_u})$, while $Q_{\mathbf{J}}(\mathbf{J})$ and $Q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})$ are fixed:

$$Q_{\mathbf{F}}(\mathbf{F_u}) \propto \exp\left[\int d\mathbf{J} d\boldsymbol{\alpha} Q_{\mathbf{J}}(\mathbf{J}) Q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) \log P(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha},\mathbf{B},\mathbf{E})\right]. \quad (30)$$

In the third step (α-step), the free energy $F(Q)$ is maximized with respect to $Q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})$, while $Q_{\mathbf{J}}(\mathbf{J})$ and $Q_{\mathbf{F}}(\mathbf{F_u})$ are fixed:

$$Q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) \propto \exp\left[\int d\mathbf{J} d\mathbf{F} Q_{\mathbf{J}}(\mathbf{J}) Q_{\mathbf{F}}(\mathbf{F_u}) \log P(\mathbf{J},\mathbf{F_u},\boldsymbol{\alpha},\mathbf{B},\mathbf{E})\right]. \quad (31)$$

The J-, F-, and α-steps are repeated until the free energy converges.

The detailed algorithm is described below. In the J-step, the distribution $Q_J(\mathbf{J})$ is given by $Q_J(\mathbf{J}) = \mathcal{N}\left(\mathbf{J}(t)|\bar{\mathbf{J}}(t),\Sigma_J^{-1}\right)$:

$$\bar{\mathbf{J}}(t) \equiv \begin{bmatrix} \bar{\mathbf{J}}_{\text{brain}}(t) \\ \bar{\mathbf{J}}_{\text{eye}}(t) \end{bmatrix}$$

$$= \Sigma_J^{-1}\left(\begin{bmatrix} \mathbf{G}'_{\text{brain}} & 0 \\ \mathbf{G}'_{\text{eye}} & \bar{\mathbf{F}}' \end{bmatrix}\begin{bmatrix} \bar{\beta}_B\mathbf{I} & 0 \\ 0 & \bar{\beta}_E\mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{B}(t) \\ \mathbf{E}(t) \end{bmatrix} - \begin{bmatrix} 0 \\ \bar{\beta}_E(\bar{\mathbf{F}}'\bar{\mathbf{u}} + \mathbf{S}_{\text{Fu}}) \end{bmatrix}\right)$$

$$\Sigma_J = \begin{bmatrix} \mathbf{G}'_{\text{brain}} & 0 \\ \mathbf{G}'_{\text{eye}} & \bar{\mathbf{F}}' \end{bmatrix}\begin{bmatrix} \bar{\beta}_B\mathbf{I} & 0 \\ 0 & \bar{\beta}_E\mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{G}_{\text{brain}} & \mathbf{G}_{\text{eye}} \\ 0 & \bar{\mathbf{F}} \end{bmatrix} + \begin{bmatrix} \bar{\mathbf{A}}_{\text{brain}} & 0 \\ 0 & \bar{\mathbf{A}}_{\text{eye}} \end{bmatrix}$$

$$+ \begin{bmatrix} 0 & 0 \\ 0 & \bar{\beta}_E\mathbf{S}_{\text{FF}} \end{bmatrix} \tag{32}$$

where $\mathcal{N}\left(\mathbf{J}(t)|\bar{\mathbf{J}}(t),\Sigma_J^{-1}\right)$ represents a Gaussian distribution with mean $\bar{\mathbf{J}}(t)$ and covariance matrix Σ_J^{-1}. The vector $\mathbf{S}_{\text{Fu}}$ and the matrix $\mathbf{S}_{\text{FF}}$ represent the uncertainty of the estimation of the EOG forward model. These variables are calculated in the F-step. For implementation, to avoid an inverse matrix calculation of a large matrix size, we use the following matrix inversion lemma:

$$\Sigma_J^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\cdot\mathbf{G}'_F\cdot\left(\Sigma_0^{-1} + \mathbf{G}_F\cdot\mathbf{A}^{-1}\cdot\mathbf{G}'_F\right)^{-1}\cdot\mathbf{G}_F\cdot\mathbf{A}^{-1},$$

where

$$\mathbf{G}_F = \begin{bmatrix} \mathbf{G}_{\text{brain}} & \mathbf{G}_{\text{eye}} \\ 0 & \bar{\mathbf{F}} \end{bmatrix}$$

$$\Sigma_0 = \begin{bmatrix} \bar{\beta}_B\mathbf{I} & 0 \\ 0 & \bar{\beta}_E\mathbf{I} \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} \bar{\mathbf{A}}_{\text{brain}} & 0 \\ 0 & \bar{\mathbf{A}}_{\text{eye}} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \bar{\beta}_E\mathbf{S}_{\text{FF}} \end{bmatrix}.$$

This reduces the computational expense $\left(O\left((L+K)^3\right)\to O\left((N+K)^3\right), L\gg N\right)$. The matrix Σ_J^{-1} is partitioned into

$$\begin{bmatrix} \mathbf{S}_{\text{BB}} & \mathbf{S}_{\text{BE}} \\ \mathbf{S}_{\text{EB}} & \mathbf{S}_{\text{EE}} \end{bmatrix} = \Sigma_J^{-1}, \tag{33}$$

where the submatrix $\mathbf{S}_{\text{EE}}$ resents an estimated covariance matrix for $\bar{\mathbf{J}}_{\text{eye}}(t)$. The submatrix $\mathbf{S}_{\text{EE}}$ is used in the F- and α-steps.

In the F-step, the distribution $Q_F(\mathbf{F_u})$ is calculated. We assumed that an observation of EOG $E_{(k)}$ is dependent only on a pair of $F_{(k)}$ and $u_{(k)}$. Therefore $F_{(k)}$ and $u_{(k)}$ are calculated from a pair of $\bar{J}_{\text{eye}(k)}$ and $E_{(k)}$,

$$Q_F\left(F_{(k)},u_{(k)}\right) = \mathcal{N}\left(\begin{bmatrix} F_{(k)} \\ u_{(k)} \end{bmatrix}\middle|\begin{bmatrix} \bar{F}_{(k)} \\ \bar{u}_{(k)} \end{bmatrix},\Sigma_{F(k)}^{-1}\right)$$

$$\begin{bmatrix} \bar{F}_{(k)} \\ \bar{u}_{(k)} \end{bmatrix} = \bar{\beta}_E\cdot\Sigma_{F(k)}^{-1}\cdot\sum_{t=1}^{T}\begin{bmatrix} \bar{J}_{\text{eye}(k)}(t) \\ 1 \end{bmatrix}\cdot E_{(k)}(t)$$

$$\Sigma_{F(k)} = \bar{\beta}_E\sum_{t=1}^{T}\begin{bmatrix} \bar{J}_{\text{eye}(k)}^2(t) + S_{\text{EE}(k,k)} & \bar{J}_{\text{eye}(k)}(t) \\ \bar{J}_{\text{eye}(k)}(t) & 1 \end{bmatrix} + \begin{bmatrix} \bar{\alpha}_{F(k)} & 0 \\ 0 & \bar{\alpha}_{u(k)} \end{bmatrix}. \tag{34}$$

The matrix is partitioned into

$$\begin{bmatrix} S_{\text{FF}(k)} & S_{\text{Fu}(k)} \\ S_{\text{Fu}(k)} & S_{\text{uu}(k)} \end{bmatrix} = \Sigma_{F(k)}^{-1}. \tag{35}$$

For the J-step, $S_{\text{FF}(k)}$ is summarized in a diagonal matrix

$$S_{\text{FF}} = \begin{bmatrix} S_{\text{FF}(1)} & & 0 \\ & \ddots & \\ 0 & & S_{\text{FF}(K)} \end{bmatrix}, \tag{36}$$

and $S_{\text{FU}(k)}$ are summarized in a vector

$$\mathbf{S}_{\text{Fu}} = \begin{bmatrix} S_{\text{Fu}(1)} \\ \vdots \\ S_{\text{Fu}(K)} \end{bmatrix}. \tag{37}$$

In the α-step, $Q_\alpha(\alpha)$ is calculated. The trial distributions for α_J, α_F, α_u, β_B, and β_E are calculated independently. The trial distribution for α_J is given by

$$Q_\alpha\left(\alpha_{J(i)}\right) = \mathcal{G}\left(\alpha_{J(i)}|\bar{\alpha}_{J(i)},\gamma_{J(i)}\right)$$

$$\bar{\alpha}_{J(i)}^{-1} = \gamma_{J(i)}^{-1}\left(\frac{1}{2}\sum_{t=1}^{T}\bar{J}_{(i)}^2(t) + \frac{1}{2}T\Sigma_{J(i,i)}^{-1} + \gamma_{J0(i)}\alpha_{J0(i)}^{-1}\right)$$

$$\gamma_{J(i)} = \frac{1}{2}T + \gamma_{J0(i)}, \tag{38}$$

where $\mathcal{G}\left(\alpha_J|\bar{\alpha}_J,\gamma_J\right)$ represents the Gamma distribution with mean $\bar{\alpha}_J$ and degree of freedom γ_J. When α_J is used in the J-step, α_J is transformed into a diagonal matrix:

$$\begin{bmatrix} \bar{\mathbf{A}}_{\text{brain}} & \\ & \bar{\mathbf{A}}_{\text{eye}} \end{bmatrix} = \begin{bmatrix} \text{diag}(\bar{\alpha}_{\text{brain}}) & 0 \\ 0 & \text{diag}(\bar{\alpha}_{\text{eye}}) \end{bmatrix} = \text{diag}(\bar{\alpha}_J).$$

The trial distributions for α_F and α_u are given by

$$Q_\alpha\left(\alpha_{F(k)}\right) = \mathcal{G}\left(\alpha_{F(k)}|\bar{\alpha}_{F(i)},\gamma_{F(i)}\right)$$

$$\bar{\alpha}_{F(k)}^{-1} = \gamma_{F(k)}^{-1}\left(\frac{1}{2}\bar{F}_{(k)}^2 + \frac{1}{2}S_{\text{FF}(k)}\right)$$

$$\gamma_{F(k)} = \frac{1}{2} \tag{39}$$

and

$$Q_\alpha\left(\alpha_{u(k)}\right) = \mathcal{G}\left(\alpha_{u(k)}|\bar{\alpha}_{u(k)},\gamma_{u(k)}\right)$$

$$\bar{\alpha}_{u(k)}^{-1} = \gamma_{u(k)}^{-1}\left(\frac{1}{2}\bar{u}_{(k)}^2 + \frac{1}{2}S_{\text{uu}(k)}\right)$$

$$\gamma_{u(k)} = \frac{1}{2}. \tag{40}$$

The trial distributions for β_B and β_E are given by

$$Q_\alpha(\beta_B) = \mathcal{G}\left(\beta_B|\bar{\beta}_B,\gamma_{\beta_B}\right)$$

$$\bar{\beta}_B^{-1} = \frac{1}{2}\gamma_{\beta_B}^{-1}\left\{\sum_{t=1}^{T}||\mathbf{B}(t) - \mathbf{G}\cdot\bar{\mathbf{J}}(t)||^2 + T\,\text{Tr}\left(\mathbf{G}\cdot\Sigma_J^{-1}\cdot\mathbf{G}'\right)\right\}$$

$$\gamma_{\beta_B} = \frac{1}{2}TN, \tag{41}$$

where $\mathbf{G} = [\mathbf{G}_{\text{brain}} \quad \mathbf{G}_{\text{eye}}]$ and

$$Q_\alpha(\beta_E) = \mathcal{G}\left(\beta_E|\bar{\beta}_E,\gamma_{\beta_E}\right)$$

$$\bar{\beta}_E^{-1} = \frac{1}{2}\gamma_{\beta_E}^{-1}\left\{\sum_{t=1}^{T}||\mathbf{E}(t) - \bar{\mathbf{F}}\cdot\bar{\mathbf{J}}_{\text{eye}}(t) - \bar{\mathbf{u}}||^2\right.$$

$$+ \sum_{k=1}^{K}\sum_{t=1}^{T}\left[\bar{J}_{\text{eye}(k)}^2(t)S_{\text{FF}(k)} + 2\bar{J}_{\text{eye}(k)}(t)S_{\text{Fu}(k)} + S_{\text{uu}(k)}\right]$$

$$\left. + T\,\text{Tr}\left(\bar{\mathbf{F}}\cdot\Sigma_J^{-1}\cdot\bar{\mathbf{F}}' + \mathbf{S}_{\text{FF}}\cdot\mathbf{S}_{\text{EE}}\right)\right\}$$

$$\gamma_{\beta_E}^{-1} = \frac{1}{2}TK. \tag{42}$$

Appendix B

The previous methods used in the simulation study are briefly explained here.

The moving dipole method

The moving dipole method (Berg and Scherg, 1994) estimates eye current sources at fixed locations in the eyeballs and a few brain current sources. This forward model is

$$\mathbf{B}(t) = \mathbf{G}_{\text{brain}}(\mathbf{r}) \cdot \mathbf{J}_{\text{brain}}(\mathbf{r}, t) + \mathbf{G}_{\text{eye}} \cdot \mathbf{J}_{\text{eye}}(t) + \varepsilon_{\text{meg}}.$$

We explore dipoles $\mathbf{J}_{\text{brain}}$ ($\mathbf{r,t}$), minimizing the square error of reconstruction magnetic fields and moving dipole location $\mathbf{r}$. In simulation experiment 1, we placed single dipoles as initial positions at the center of each brain area in which true peak current sources exist. In simulation experiment 2, each single dipole was placed at the center of left and right FEF and MT+. The estimation errors of eye artifact and eye current source were calculated in the same way as the hyVBED method from Eq. (14). For the estimation error of the brain current source, the nRMSE between the estimated single dipole and the spatially summed simulated current sources was calculated in each area.

ICA

ICA (Vigario, 1997; Vigario et al., 2000; Jung et al., 2000; Barbati et al., 2004) does not use spatial information due to the assumption of a particular current source model. ICA assumes that the eye artifacts and the brain signals are statistically independent. An observed MEG $\mathbf{B}(t)$ is represented as a linear summation of independent components $\mathbf{Z}$:

$$\mathbf{B}(t) = \mathbf{W} \cdot \mathbf{Z}(t),$$

where $\dim \mathbf{B}(t) \geq \dim \mathbf{Z}(t)$. After transforming $\mathbf{B}$ to $\mathbf{Z}$, we find the component of $\mathbf{Z}_{\text{eye}}$ most correlated to the eye movements. Then eye artifacts are calculated as

$$\overline{\mathbf{B}}_{\text{artifact}}(t) = \mathbf{W}_{\text{eye}} \cdot \mathbf{Z}_{\text{eye}}(t),$$

where $\mathbf{W}_{\text{eye}}$ is a mixture matrix corresponding to independent components $\mathbf{Z}_{\text{eye}}$. The error of artifact removal was calculated from Eq. (14). In this paper, we use the joint diagonalization of cumulant matrices (JADE) (Cardoso and Souloumiac, 1993) as ICA.

References

Attias, H., 1999. Inferring parameters and structure of latent variable models by variational Bayes. Proc. 15th Conf. Uncertain. Artif. Intell. 21–30.

Barbati, G., Porcaro, C., Zappasodi, F., Rossini, P.M., Tecchio, F., 2004. Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals. Clin. Neurophysiol. 115 (5), 1220–1232.

Berg, P., Scherg, M., 1994. A multiple source approach to the correction of eye artifacts. Electroencephalogr. Clin. Neurophysiol. 90 (3), 229–241.

Cardoso, J.F., Souloumiac, A., 1993. Blind beamforming for non Gaussian signals. IEE Proc. F 140, 362–370.

Carpenter, R., 1988. Movements of the Eyes, 2nd ed. Pion London.

Croft, R.J., Barry, R.J., 2000. Removal of ocular artifact from the EEG: a review. Neurophysiol. Clin. 30, 5–19.

Daunizeau, J., Friston, K.J., 2007. A mesostate-space model for EEG and MEG. NeuroImage 38, 67–81.

Daunizeau, J., Grova, C., Mattout, J., Marrelec, G., Clonda, D., Goulard, B., Lina, J.M., Benali, H., 2005. Assessing the relevance of fMRI-based prior in the EEG inverse problem: a Bayesian model comparison approach. IEEE Trans. Signal Process. 53, 3461–3472.

Friston, K., Harrison, L., Daunizeau, J., Kiebel, S., Phillips, C., Trujillo-Barreto, N., Henson, R., Flandin, G., Mattout, J., 2008. Multiple sparse priors for the M/EEG inverse problem. NeuroImage 39, 1104–1120.

Fukushima, K., Yamanobe, T., Shinmei, Y., Fukushima, J., 2002. Predictive responses of periarcuate pursuit neurons to visual target motion. Exp. Brain Res. 145 (1), 104–120.

Jervis, B.W., Ifeachor, E.C., Allen, E.M., 1998. The removal of ocular artefacts from the electroencephalogram: a review. Med. Biol. Eng. Comput. 26, 2–12.

Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., Sejnowski, T.J., 2000. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37 (2), 163–178.

Kajihara, S., Ohtani, Y., Goda, N., Tanigawa, M., Toyama, K., 2004. Wiener filter magnetoencephalography of visual cortical activities. Brain Topogr. 17 (1), 13–25.

Katila, T., Maniewski, R., Poutanen, T., Varpula, T., 1981. Magnetic fields produced by the human eye. J. Appl. Phys. 52 (2), 2565–2571.

Kawawaki, D., Shibata, T., Goda, N., Doya, K., Kawato, M., 2006. Anterior and superior lateral occipito–temporal cortex responsible for target motion prediction during overt and covert visual pursuit. Neurosci. Res. 54 (2), 112–123.

Kenemans, J.L., Molenaar, P.C., Verbaten, M.N., Slangen, J.L., 1991. Removal of the ocular artifact from the EEG: a comparison of time and frequency domain methods with simulated and real data. Psychophysiology 28 (1), 114–121.

Kiebel, S.J., Daunizeau, J., Phillips, C., Friston, K.J., 2008. Variational Bayesian inversion of the equivalent current dipole model in EEG/MEG. NeuroImage 39, 728–741.

Krauzlis, R.J., 2004. Recasting the smooth pursuit eye movement system. J. Neurophysiol. 91, 591–603.

MacAvoy, M.G., Gottlieb, J.P., Bruce, C.J., 1991. Smooth-pursuit eye movement representation in the primate frontal eye field. Cereb. Cortex 1 (1), 95–102.

Malmivuo, J., Plonsey, R., 1995. Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields. Oxford Univ Pr.

Meier, T., Rosburg, T., Arnold, R., Kreitschmann-Andermahr, I., Sauer, H., Nowak, H., Witte, H., 1998. Quantification and rejection of ocular artifacts in auditory evoked fields in schizophrenics. Electroencephalogr. Clin. Neurophysiol. 108 (6), 526–535.

Miles, W.R., 1939. The steady polarity potential of the human eye. Proc. Natl. Acad. Sci. U. S. A. 25, 25–36.

Neal, R.M., 1996. Bayesian Learning for Neural Networks. Springer-Verlag.

Newsome, W.T., Wurtz, H., Komatsu, R.H., 1988. Relation of cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. J. Neurophysiol. 60 (2), 604–620.

Nummenmaa, A., Auranen, T., Hämäläinen, M.S., Jääskeläinen, I.P., Sams, M., Vehtari, A., Lampinen, J., 2007. Automatic relevance determination based hierarchical Bayesian MEG inversion in practice. NeuroImage 37, 876–889.

Petit, L., Haxby, J.V., 1999. Functional anatomy of pursuit eye movements in humans as revealed by fMRI. J. Neurophysiol. 82 (1), 463–471.

Sarvas, J., 1987. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Phys. Med. Biol. 32, 11–22.

Sato, M., 2001. On-line model selection based on the variational Bayes. Neural Comput. 13, 1649–1681.

Sato, M., Yoshioka, T., Kajihara, S., Toyama, K., Goda, N., Doya, K., Kawato, M., 2004. Hierarchical Bayesian estimation for MEG inverse problem. NeuroImage 23 (3), 806–826.

Shackel, B., 1960. Pilot study in electro-oculography. Br. J. Ophthal. 44, 89–113.

Vigario, R., Särelä, J., Jousmäki, V., Hämäläinen, M.S., Oja, E., 2000. Independent component approach to the analysis of EEG and MEG recordings. IEEE Trans. Biomed. Eng. 47 (5), 589–593.

Vigario, R.N., 1997. Extraction of ocular artefacts from EEG using independent component analysis. Electroencephalogr. Clin. Neurophysiol. 103, 395–404.

Wang, J.Z., Williamson, S.J., Kaufman, L., 1992. Magnetic source images determined by a lead-field analysis: the unique minimum-norm least-squares estimation. IEEE Trans. Biomed. Eng. 39 (7), 665–675.

Wipf, D. and Nagarajan, S., 2009. A unified Bayesian framework for MEG/EEG source imaging. Neuroimage 44, 947–966.

Yoshioka, T., Toyama, K., Kawato, M., Yamashita, O., Nishina, S., Yamagishi, N., Sato, M., 2008. Evaluation of hierarchical Bayesian method through retinotopic brain activities from fMRI and MEG signals. NeuroImage 42, 1397–1413.