ELSEVIER

Journal of Neuroscience Methods 157 (2006) 337–354

JOURNAL OF NEUROSCIENCE METHODS

www.elsevier.com/locate/jneumeth

Magnetoencephalographic artifact identification and automatic removal based on independent component analysis and categorization approaches

Feng Rong [a], José L. Contreras-Vidal [a,b,*]

[a] *Department of Kinesiology and Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD 20742, United States*
[b] *Bioengineering Program, University of Maryland, College Park, MD 20742, United States*

Received 23 June 2005; received in revised form 24 April 2006; accepted 26 April 2006

Abstract

Artifact signals from eye movements, heart beat and muscle activity contaminate magnetoencephalographic (MEG) signals generated from the neural activities inside the brain. Rejection of contaminated trials not only causes data loss, but can also significantly increase the experimental time or even prevent the analysis of highly contaminated or noisy data. We combined the use of independent component analysis (ICA) and clustering methods to isolate the artifacts from MEG signals. Threshold-based clustering analyses based on the topographic pattern, statistical aspects and power spectral patterns of independent components (ICs) successfully identified ICs related to certain types of artifacts. Unsupervised neural network based on the Adaptive Resonance Theory (ART) also categorized the artifact ICs, albeit with lower accuracy. Performance of the identification methods were evaluated with measurements of underestimation and overestimation of the target artifactual ICs. The combination of threshold-based clustering and ART-2 neural network categorization methods demonstrated the best identification performance. Comparison between contaminated and artifact-cleaned MEG signal waveforms showed the efficiency of the proposed methods of artifacts rejection. The analysis of the artifact components suggested the possibility of automatic artifact removal based on general templates.

Keywords: Artifacts; Magnetoencephalography; ART-2; Categorization; Clustering; ICA

1. Introduction

The analysis of magnetoencephalography (MEG) signals always requires the identification and removal of artifacts such as eye movements, heart beat and muscle activity due to head and hand/arm movements produced during the experimental session. These biological events contaminate the MEG signals generated by task-relevant brain activities and thus make direct analysis of the MEG signal with artifacts difficult (Hämäläinen et al., 1993). Regression methods in the time domain (Hillyard and Galambos, 1970) and the frequency domain (Woestenburg et al., 1983) have been proposed for the removal of eye movements and other artifacts in electroencephalographic (EEG) signals. However, the lack of eye-movement reference electrodes in MEG recording makes it difficult to remove artifacts in MEG signals with these methods. Other methods based on signal thresholds for the detection of large artifacts such as eye blinks, need an arbitrarily set magnitude threshold, which makes the identification and rejection of artifacts with smaller magnetic field magnitude such as heart beat difficult (Hämäläinen et al., 1993). Moreover, as this approach requires rejection of data segments with artifacts, valuable brain information contained in the rejected segment represents an unacceptable data loss. Thus, the development of methods to identify the artifacts and remove them without loss of information related to the cerebral activity is of the utmost importance for MEG studies involving eye movements and/or subject motor responses.

Independent component analysis (ICA, Bell and Sejnowski, 1995) has been successfully applied to the processing of electroencephalography signals (Makeig et al., 1996; Delorme and Makeig, 2004). Various methods have been proposed in the identification of components related to artifacts in both EEG (Jung et al., 1998, 2000; Delorme et al., 2001; Joyce et al., 2004) and

* Correspondence to: 2363 HHP Bldg, Department of Kinesiology, University of Maryland, College Park, MD 20742, United States.
Tel.: +1 301 405 2495; fax: +1 301 405 5578.
E-mail addresses: rongfeng@glue.umd.edu (F. Rong), pepeum@umd.edu (J.L. Contreras-Vidal).

0165-0270/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2006.04.024

MEG (Vigario et al., 2000, Tang et al., 2002; James and Gibson, 2003; Barbati et al., 2004). The criteria used by most of these methods focus on a single feature of the artifactual components related to topographic patterns, statistical patterns, or spectral properties. One recent study claimed automatic detection of the artifactual components in MEG signal based on the analysis of the statistical and spectral aspects of the components (Barbati et al., 2004). However, this study did not take into account the scalp map of the components, which represents the fixed spatial feature of the components and has been used successfully in identification of the artifacts in former EEG studies (Delorme and Makeig, 2004).

Here we propose an approach for efficient artifact identification and automatic removal of MEG signals based on iterative clustering of independent components (ICs) features including the scalp map, the statistical and the spectral properties. Both threshold-based clustering in feature space and automatic neural network-based categorization were implemented for discrimination of certain types of artifacts in MEG signals. We report that both approaches can detect and remove artifacts for single-trial and multiple-trials (concatenated) MEG signals from datasets within subject and across subjects. Evaluations of performance showed that threshold-based clustering could identify the artifacts with higher accuracy, whereas the neural network categorization performed with minimal intervention from the researchers. The performance gain obtained by combining these two methods is also demonstrated, namely the improvement in alleviating overestimation and/or underestimation of target artifactual ICs.

We have applied our proposed methods for ICA-based artifactual IC identification and removal on both single-trial based epochs and the concatenated MEG signals (Delorme and Makeig, 2004). The iterative template updating used for the categorization and clustering of the identified components suggests that the same type of independent components with slightly different features can be clustered together. The statistical analysis of the feature values also showed that the components related to artifacts can be significantly discriminated with combination of two or more features. Since the proposed methods are not specific or restricted to any biological event, they may also be useful in the identification and categorization of ICs related to brain signals generated by similar functional brain networks.

2. Methods

2.1. MEG signal collection and pre-whitening

2.1.1. Participants

Five healthy right-handed adults volunteered to participate in the study after giving informed consent. All subjects had normal or corrected-to-normal vision. All procedures were approved by the Institutional Review Board at the University of Maryland, College Park.

2.1.2. Apparatus

Subjects lay supine on a table with their heads inside the recording dewar of the MEG device, which was located in a magnetic-shielded room in the KIT-MEG laboratory at the University of Maryland, College Park. The subjects performed center-out drawing movements with an optic pen on a glass panel in front of them (Fig. 1). During the experiment, the subjects' vision of the pen and their hands was occluded by a black curtain. A video camera placed over the glass panel captured the movement of the pen tip with a 60 Hz sampling frequency. The signal of the movement was transformed onto a screen cursor moving on a monitor screen outside the room and another glass panel over the subject's head, providing the subjects with visual information of the pen tip movement. The subjects' task was to draw a line from a home circle centered on the screen to one of the targets at the four corners of the screen. MEG signals were recorded with the 192-neuro-magnetometer-channel system (Kado et al., 1999). This system uses coaxial type first-order gradiometers which have a magnetic field resolution of 4 ft/Hz$^{1/2}$ or 0.8(ft/cm)/Hz$^{1/2}$ at the white noise region. One hundred and fifty seven channels (channel index 0–156) out of the 192 channels were used to record magnetic signals related to brain activity. The remaining 35 channels are commonly reserved for triggering or other purposes. In this experiment, four of these channels were used for temporal markers of behavioral events (160: target onset; 161: go signal; 162: movement onset; 163: target acquisition), and three channels (channel index 157–159) served as references for noise reduction (Adachi et al., 2001). The continuous MEG signals were sampled at 1 kHz and on-line filtered in the SQUID electric circuit units using three kinds of analog filters: high-pass filter (1st order Butterworth, fl = 1 Hz, attenuation step: 6 dB/octave), low-pass filter (2nd

Fig. 1. Experimental setup. The subjects' view of their hands and arms has been occluded during the experiment. The participants were asked to move an optic pen on a plexiglass panel in front of them. A camera captured the movement of the pen tip and transformed it into cursor movement on a screen over the subjects' heads.

order Butterworth, fh = 100 Hz, attenuation step: 12 dB/octave), and a notch filter (1st order paired band elimination filter, center frequency = 60 Hz, attenuation rate: >30 dB at center frequency).

2.1.3. MEG signal pre-whitening

The non-periodical low-frequency environmental noises (e.g., noise from urban traffic surrounding the MEG facility) recorded by three reference sensors during MEG measurements were pre-whitened using the continuously adjusted least square method (CALM, Adachi et al., 2001), which is implemented in MEG160 software (MEG Laboratory 2.001 M, Yokogawa Corporation, Eagle Technology Corporation, Kanazawa Institute of Technology). The pre-whitened MEG signal was epoched with 10.2 s time length, which corresponded to 10,201 samples per sensor per epoch. Each epoch was viewed as a single trial. Four datasets were constructed with the pre-whitened MEG signal to test the methods of ICA and artifactual IC identification: dataset one was from one randomly picked trial; dataset two consisted of 95 trials from one subject; dataset three contained 123 trials randomly selected across all five subjects, and dataset four contained 21 trials from one subject. Dataset four was used to compare the performance of artifactual IC identification methods when applying ICA on the concatenated continuous MEG signal and on single-trial based epochs. The proposed methodology for artifact identification and removal is illustrated in Fig. 2.

2.2. Independent component analysis

First, we applied Infomax ICA (EEGLAB v4.512, which was downloaded from http://www.sccn.ucsd.edu/eeglab/index.html) on the pre-whitened MEG signals. We did not use the extended ICA algorithm because the main target sources (artifactual sources) of this study were mainly supra-Gaussian distributed (kurtosis > 0) (Delorme and Makeig, 2004). For the first three datasets, ICA was applied on single-trial based epochs. For dataset four, ICA was applied on both a concatenated dataset and in single-trial based epochs.

For the application of ICA, we assumed that the MEG signal $x(t)$ was generated from sources $s(t)$ with a linear mixing

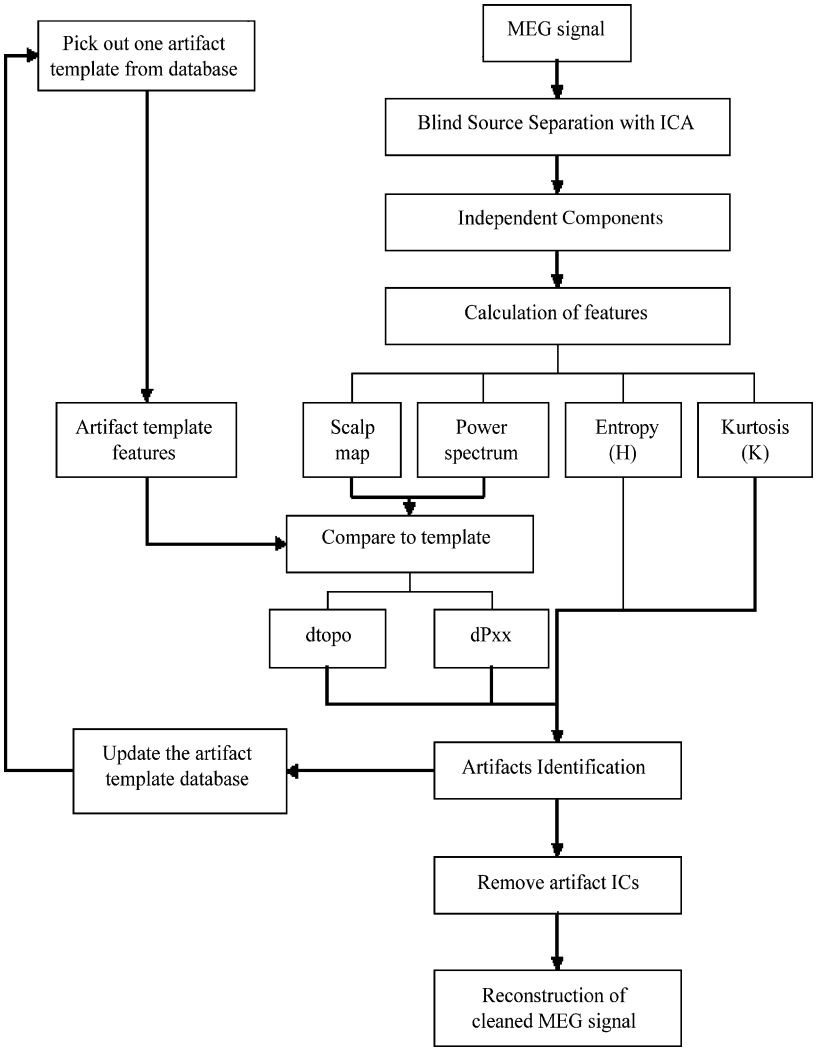

Fig. 2. The illustration of the procedure for independent component calculation, artifactual IC identification, template updating, artifact rejection and reconstruction of cleaned MEG signal.

procedure

$$x(t) = A\,s(t) \tag{1}$$

where t was a vector of the sampling time with length T, $x(t) = [x_1(t)\ldots x_n(t)]$ was a $n \times T$ matrix of noise reduced MEG signal recorded from n sensors; $s(t) = [s_1(t)\ldots s_m(t)]$ was a $m \times T$ matrix of m sources' activations; A was an unknown $n \times m$ full-rank mixing matrix. The application of ICA involved processing the MEG signals following the demixing procedure:

$$\hat{s}(t) = W\,x(t) \tag{2}$$

where $\hat{s}(t) = [\hat{s}_1(t)\ldots\hat{s}_m(t)]$ is an m-dimensional matrix of the independent components, which represents the estimation of the sources $s(t)$; W is a demixing matrix

$$W = \hat{A}^{+} \tag{3}$$

where $\hat{A}^{+}$ denotes an estimate of the pseudo-inverse of the mixing matrix A.

After the artifacts had been identified, the MEG signals were reconstructed with the components which have not been labeled as artifacts. Thus, in the *remixing* matrix, the weight vectors of the identified artifactual ICs have been set to zeros, such that

$$x'(t) = W^{+}_{\text{clean}}\hat{s}(t) \tag{4}$$

where $x'(t)$ is the reconstructed artifacts-free MEG signal for n sensors, W^{+} is the pseudo-inverse of the demixing matrix W, and W^{+}_{clean} is the $n \times m$ remixing matrix with its columns corresponding to the artifactual ICs set to zeros, $\hat{s}(t)$ is the activation matrix of ICs as mentioned above.

It is important to note that the application of Infomax ICA on EEG/MEG signals makes at least the following four assumptions (Vigario et al., 2000): (1) the underlying sources are statistically independent[1] and spatially stationary, however, if the total temporal independence cannot be reached, the sources are nevertheless computed with maximal independence to each other; (2) the sources do not have a Gaussian distribution[2]; (3) the components are characterized by a fixed, but possibly overlapping, spatial topography; and (4) the mixing process is instantaneous and stationary. The instantaneous mixing assumption can be met if we accept the quasi-static approximation of Maxwell equation (Jackson, 1999). The widely accepted current dipole model of the brain sources of MEG signals is congruent with the assumed stationarity of the ICs (Hämäläinen et al., 1993).

[1] Independence is not a strong constraint of ICA as ICA can be viewed as a maximum projection algorithm so it will find the sources which are maximally independent. This makes sense when studying brain processes as it is unreasonable to assume that the activity in one part of the brain is completely independent of the activity in another part of the brain.

[2] Strictly speaking, it is not necessary to assume for the information maximization in ICA that the sources not have a Gaussian distribution. In the case of Gaussian distributed signals, ICA merely degenerates to regular old PCA with summarized variance equivalent to information and uncorrelated components also being independent (Vigario et al., 2000).

2.3. Component feature calculation and clustering identification

Four feature values were calculated for identification of the artifactual ICs: the pairwise distance between topographic patterns (dtopo); the entropy of the component activations (H); the pairwise distance between the power spectral density patterns of the component activations (dPxx), and the global kurtosis coefficient of the component activations (K). The ICs were visualized as scatter points in feature space for threshold-based clustering of the artifactual ICs. For dataset one, the EKG-related artifactual ICs were clustered with criteria of dtopo ≤ 0.2, dPxx ≤ 0.5, $H \leq 2.78$ and $K > 0$, the ocular movement-related artifactual ICs were clustered with criteria of dtopo ≤ 0.2, dPxx ≤ 0.5, $H \leq 3.0$ and $K > 20$. For dataset two and three, the EKG related artifactual ICs were clustered with criteria of dtopo ≤ 0.2, dPxx ≤ 0.5, $H \leq 2.8$ and $K > 0$, the EOG1-related artifactual ICs were clustered with criteria of dtopo < 0.2, dPxx ≤ 0.5, $H \leq 3.0$ and $K > 10$. The EOG2-related ICs were clustered with criteria of dtopo ≤ 0.2, dPxx ≤ 0.5, $H \leq 3.0$ and $K > 10$. For dataset four, the threshold-based clustering methods used two sets of thresholds for each type of artifactual ICs, which are provided in Table 3.

2.3.1. Distance between IC topographic patterns

The ICs are commonly viewed as related to field activity with fixed spatially distributed sources (Makeig et al., 2002, 2004). In applications of ICA, the spatial distribution of sources is usually represented as scalp maps (Makeig et al., 2002, 2004) or field maps (Tang et al., 2002), which are represented mathematically as projection of the ICs on sensor space. This measurement represents the estimate of the mixing matrix $\hat{A}$, which is denoted as the pseudo-inverse of the demixing matrix (that is, W^{+}). The pairwise distance of these topographic patterns can give us information how close the spatial distributions of these sources are. In our study we used the correlation method to calculate this distance,

$$\text{dtopo} = 1 - \frac{(\text{Topo}_c - \overline{\text{Topo}_c})(\text{Topo}_t - \overline{\text{Topo}_t})'}{\sqrt{\begin{array}{c}(\text{Topo}_c - \overline{\text{Topo}_c})(\text{Topo}_c - \overline{\text{Topo}_c})' \\ \times\,(\text{Topo}_t - \overline{\text{Topo}_t})(\text{Topo}_t - \overline{\text{Topo}_t})'\end{array}}} \tag{5}$$

where Topo_c was the topographic pattern of the compared IC and Topo_t was the topographic pattern of the template IC. If this value is close to zero then the topographic pattern of the current IC is a close match to that of the template IC, which represents a certain type of artifacts. In this study we used the built-in *pdist.m* function in statistical toolbox v4.0 and MATLAB 6.5 (The Math Works Inc., MA, USA) to compute dtopo. Note that because the exact polarity of the components separated by ICA is unknown, and because eye movement-related IC's of opposite polarity had slightly different reversing projection patterns to each other, we allowed for two different templates rather than use the absolute value of the fraction above (Eq. (5)) to increase the sensitivity of the proposed method.

2.3.2. Entropy of the IC activation

The entropy of the IC was defined as:

$$H = -\sum_{t=1}^{T} p(\hat{s}(t) = y) \log(p(\hat{s}(t) = y)) \quad (6)$$

where $p(\hat{s}(t) = y)$ was the probability density of observing the activity value y in IC activation $\hat{s}(t)$. Higher entropy values corresponded to more 'random' source activities. ICs with lower entropy than other ICs can be viewed as artifactual ICs with characteristic probability densities (Barbati et al., 2004).

2.3.3. Kurtosis of the activation

The global kurtosis coefficient of one IC was calculated as

$$K = d_4 - 3 \times (d_2^2) \quad (7)$$

where $d_n = E\{(\hat{s}(t) - d_1)^n\}$ was the nth central moment of the IC component activity, and E, or the mean, was the expectancy function. With this equation, ICs with $K = 0$ represented components with Gaussian distributed activations. ICs with much higher kurtosis values than the average were viewed as artifacts which have supra-Gaussian distributed activities. Here we used the built-in *kurt.m* function provided by the EEGLAB software package (Delorme and Makeig, 2004).

2.3.4. Correlation between power spectral densities

We used the built-in *pmtm.m* function of signal processing toolbox v6.0 and MATLAB 6.5 to compute the power spectral density (PSD) of the IC activities with multi-taper spheroidal methods (nw = 3.5, which corresponded to the use of seven discrete prolate spheroidal sequences as data tapers for the multi-taper estimation method). The sampling frequency was 1000 Hz, and the window size of the fast Fourier transform (fft) was 512. Their distances to the chosen template PSD were computed with the same equation used in the computation of topographic pattern distances with substitution of topographic pattern vectors with power spectral pattern vectors. The higher the correlation between IC power spectrums the lower their pairwise distance, which means the IC's PSD represents a close match to the template.

2.4. Categorization of ICs with ART-2 network

Another method we used to identify artifactual ICs consisted of applying the analog adaptive resonance theory (so-called ART-2) neural network to categorize IC features. ART networks are self-organizing, dynamic neural networks that recognize patterns with stable structures (Carpenter and Grossberg, 1987a). The two-layer ART-2 network architecture and equations, as originally proposed by Carpenter and Grossberg (1987b), are reproduced with permission in Appendix A for the convenience of the reader. The inputs to the network were vectors composed of four normalized feature values of each IC in the order of [dtopo dPxx H K]. dtopo and dPxx values ranged from 0 to 2 and represented the outcome of the correlation-based distance metric described above. H values were normalized between 0 and 1 based on the minimum and maximum H values in each dataset. Normalized K values were obtained by dividing the computed global kurtosis coefficient by 50, which was arbitrarily selected based on experience. The ART-2 output consisted of two nodes or categories [target IC artifact, IC non-artifact]. Categorization results were analyzed after the ART-2 system had stabilized for a given pattern mismatch threshold ρ (so-called the 'vigilance' parameter by Carpenter & Grossberg). All other parameters in the ART-2 network were set as in the original publication and no parameter optimization was performed (see Appendix A for the model parameters).

2.5. Artifact identification and updating of the templates

All datasets were processed with the proposed artifact identification and removal procedures. As illustrated in Fig. 2, four feature attributes {dtopo, dPxx, H, K} were extracted from the ICs. The initial artifact template was chosen based on topographic and/or activation patterns previously characterized as certain type of artifacts (e.g., eye movement) in former studies (Tang et al., 2002; Vigario et al., 2000). Artifactual ICs have been identified through iterations with threshold-based clustering, neural network categorization or the combination of these two methods. After each iteration the template was updated with the average of the target artifactual ICs among the identified ICs.

2.6. Independent component datasets

We set the IC number of each trial to 157, then dataset one contained 157 ICs, dataset two contained 14,915 ICs (95 trials × 157 ICs/trial = 14,915 ICs) and dataset three contained 19,311 ICs (123 trials × 157 ICs/trial = 19,311 ICs). For dataset four, there were 157 ICs for the concatenated ICA analysis, and 3297 ICs (21 trials × 157 ICs/trial = 3297 ICs) for the ICA analysis on single-trial based epochs. Artifactual IC templates related to two types of ocular movements (EOG1 and EOG2), heart beat (EKG) and muscle activity were found in the datasets (Fig. 7). Three types of components related to muscle activity were displayed. EKG, EOG1 and EOG2 IC templates were applied in artifactual IC identification. Both identification procedures were employed for all four datasets. For dataset one and four, the results were calculated with one repetition of the identification loop, whereas for datasets two and three, this loop was iterated three times with template updating after each round. Further steps of removing the artifacts and reconstructing the cleaned MEG signal were performed on single trial basis for all four datasets.

2.7. Evaluation of performance

2.7.1. Comparison of performance between threshold-based clustering methods and ART-2 categorization neural network

We then compared the performance of the two methods in identifying the artifactual ICs. We picked dataset three for this purpose since dataset three contained MEG signal from all subjects, which was closer to most of the empirical conditions

than dataset one and dataset two. The artifactual ICs were identified with template EKG and EOG1 since they were the most prominent artifactual ICs found in the datasets. We applied both methods with three iterations. For the ART-2 categorization neural network, we used five vigilance levels (0.95, 0.96, 0.97, 0.98 and 0.99). The criteria for the threshold-based clustering method were the same as described in Section 2.3 for the corresponding type of artifactual ICs. The target artifactual ICs among the identified ICs were selected via visual inspection. The number of identified ICs and selected target artifactual ICs were

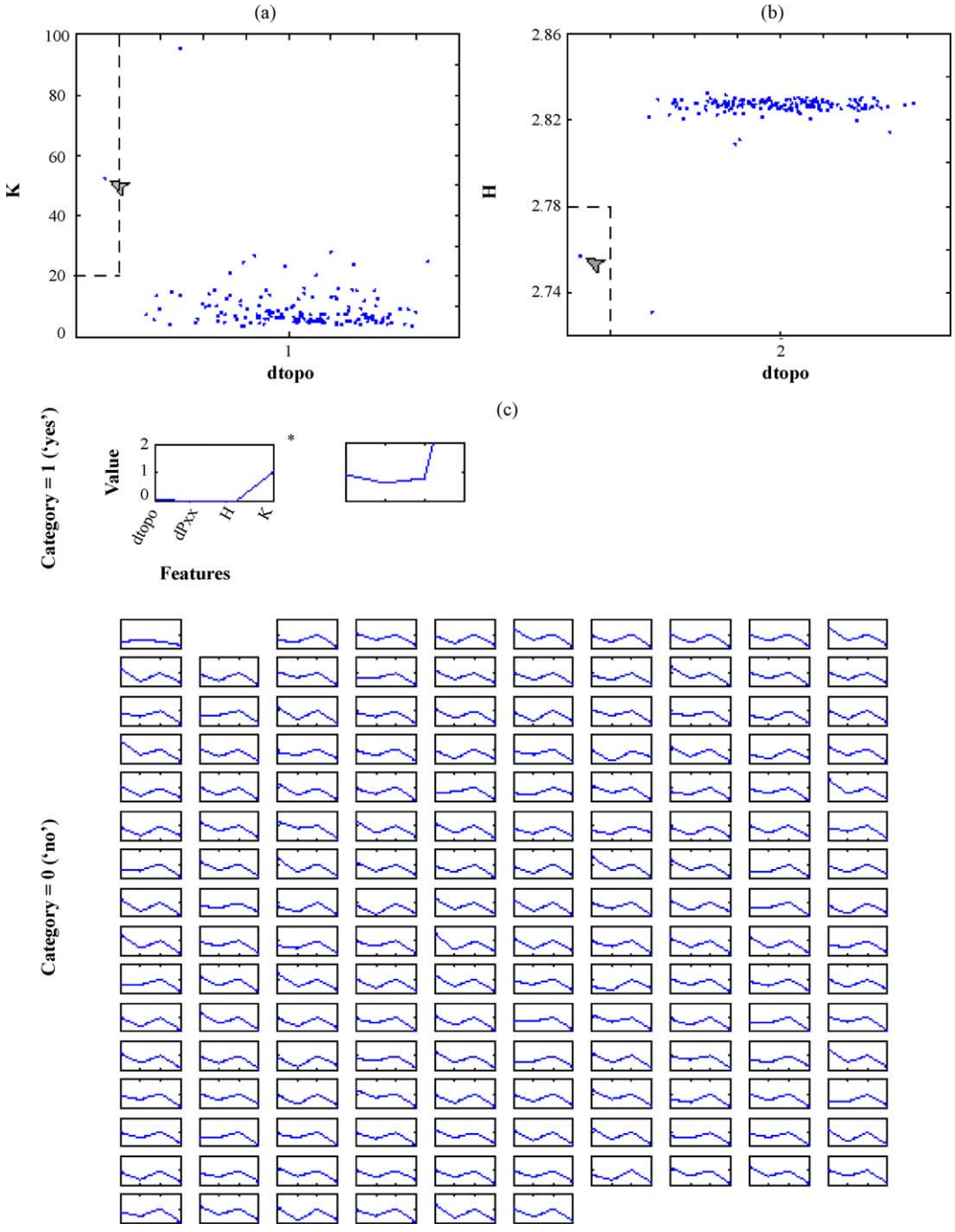

Fig. 3. Threshold-based clustering of ICs for artifactual IC identification of single trial MEG signals. (a) The identification of the IC related to EOG1 artifact with threshold-based clustering method (indicated by arrowhead), feature values of dtopo and *K* were displayed. (b) The identification of the IC related to EKG artifact with threshold-based clustering method (indicated by arrowhead), feature values of dtopo and *H* were displayed. The threshold criteria for clustering were indicated with gray dash lines. (c) The categorization of the ICs with ART-2 network. The inputs to the network were the vectors composed of four normalized feature values of each IC. The IC feature values of dtopo and dPxx (see text for the description of the calculation) were calculated from template related to EOG1 artifactual ICs. The ICs were categorized into 'yes' and 'no' groups. 'Yes' group contained ICs identified as EOG1 artifacts, 'no' group contained ICs identified as non-EOG1. Two ICs were identified as EOG1 artifactual ICs with the network method, while only one was target EOG1 artifactual IC. The target IC was selected via visual inspection and was labeled with (*).

counted. The correctness of the identification was calculated as the proportion of target artifactual ICs in the identified ICs.

$$\text{Correctness} = \frac{\text{Number of target ICs}}{\text{Number of identified ICs}} \times 100\% \quad (8)$$

The number of identified ICs and the correctness of identification are plotted in Fig. 9 to compare the performance between the methods. The performance of ART-2 categorization neural network was displayed with five vigilance levels.

2.7.2. Comparison of performance on continuous data and single-trial based epochs

We applied ICA and artifactual IC identification methods on dataset four to compare the performance between the ICA

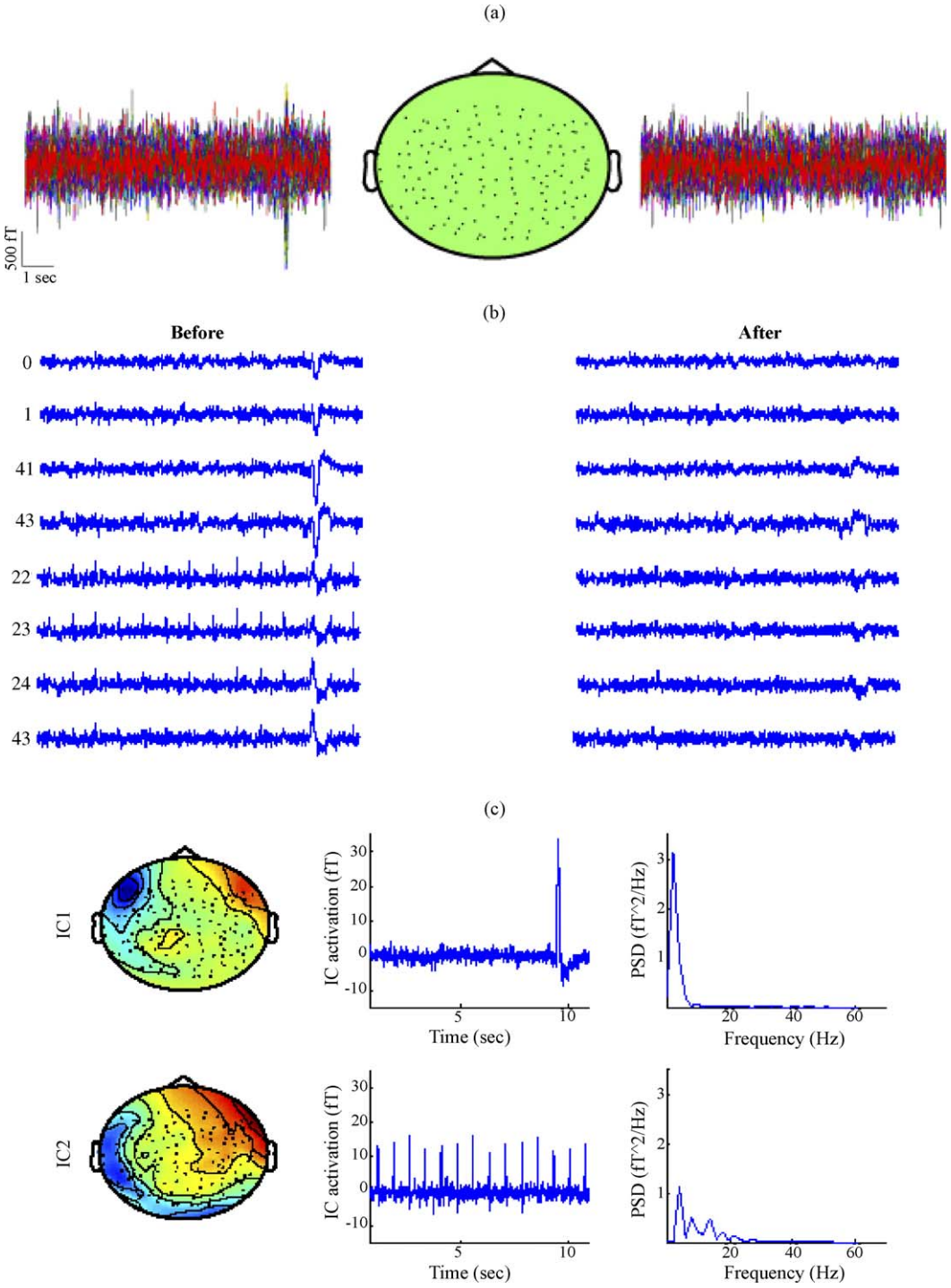

Fig. 4. Single trial artifactual component identification and rejection. (a) Whole channel noise-reduced raw MEG signal (left); 2D top view of channel position (center), and whole channel artifact-cleaned MEG signal (right). (b) Bilateral frontal channel recording before removing artifactual components (left panel), and MEG signals on same channels after rejection of artifactual components (right panel). The scales are the same as in (a). (c) The components found in this trial. (Upper) topographic pattern, activity and power spectrum of the component identified as EOG1; (lower) component identified related to heart beat (EKG).

analysis of the continuous (concatenated) MEG signals and the single-trial based epochs. The concatenated dataset contained 21 trials of MEG signals since it was the highest amount of continuous data our data analysis workstation (Gateway 6100 series with 3.20 GHz CPU and 1.0 GB RAM) could process using ICA without memory problems. Percentage of variance accounted for (pvaf) of the ICs for each sensor were calculated for each trial.

$$\text{pvaf} = \frac{\text{Var}(x(t) - x'(t))}{\text{Var}(x(t))} \tag{9}$$

where $x(t)$ and $x'(t)$ had the same mathematical denotation to the ones in Eqs. (1) and (4). Var () operation was to compute the sample variance of the signal.

The 10 sensors with the highest absolute values in the topographic component of the target artifactual IC templates were selected as the most influenced sensors, while the 10 sensors with lowest absolute values were selected as the least influenced sensors. The pvaf values calculated on most influenced sensors and least influenced sensors were statistically analyzed with repeated measures ANOVA. The mean pvaf values between ICA on concatenated signal and single-trial epochs were pair-wisely compared. Significant differences were the comparisons with probability value $p < 0.05$. The statistical analyses were performed with SAS v9.1 (SAS Institute Inc., Cary, NC, USA.).

2.7.3. Evaluation of overestimation and underestimation of identification methods

With dataset four, EKG, EOG1 and EOG2 artifactual ICs were identified with the ART-2 categorization neural network, higher threshold-based clustering, lower threshold-based clustering, and the combined method involving the neural network categorization on the ICs which have been identified with higher threshold-based clustering method. These identification methods were applied on both concatenated data and single-trial epochs. Numbers of target artifactual ICs in identified ICs were visually examined. Underestimation was evaluated with a measurement of coverage, which was computed as the ratio of the number of target artifactual ICs within the identified ICs divided by the number of target ICs in the dataset; if coverage value was less than 100%, this type of artifactual ICs was underestimated

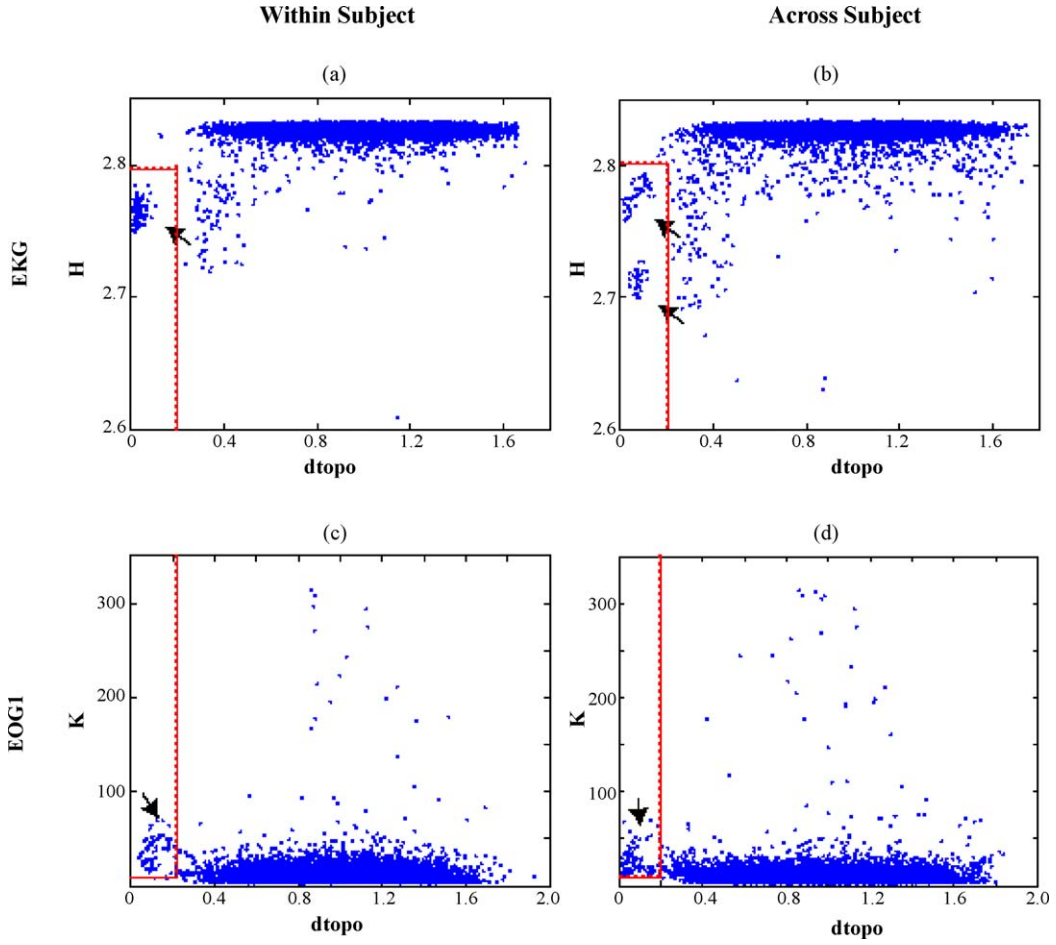

Fig. 5. Threshold-based clustering of artifactual ICs across multiple trials, the threshold criterion for clustering was indicated with red dash lines, the clustered ICs were labeled with arrowheads. (a) Clustering of EKG-related artifactual ICs in dataset two, dtopo and *H* values were displayed. (b) Clustering of EKG-related artifactual ICs in dataset three, dtopo and *H* values were displayed. (c) Clustering of EOG1-related artifactual ICs in dataset two, dtopo and *K* values were displayed. (d) Clustering of EOG1-related artifactual ICs in dataset three, dtopo and *K* values were displayed.

with this method. Overestimation was evaluated with measurement of correctness, which was computed using Eq. (8). If the percentage correctness value was less than 100%, then this type of artifactual ICs was overestimated with this method.

3. Results

3.1. Identification of artifacts and effect of removal in single trials

In general, scalp map distances (dtopo) and global kurtosis coefficients (K) were adequate for discriminating EOG artifacts, whereas scalp map distances and entropy (H) values were sufficient in identifying EKG artifacts. For dataset one, threshold-based clustering methods successfully identified components related to EKG and EOG1, as displayed in scatter plot of dtopo versus H and in scatter plot of dtopo versus K (Fig. 3(a) and (b)), in which gray dash lines indicated the threshold used for clustering. No EOG2 artifactual IC was identified in dataset one.

Fig. 3(c) shows the inputs to ART-2 network for identification of type 1 EOG (EOG1) artifactual ICs and the identification results. Two ICs were identified as EOG1 artifacts—one more than EOG1 ICs identified through threshold-based clustering process. The same IC identified through network method and threshold-based clustering method were labeled with *. The vigilance level of the reset system in the ART-2 network was set to 0.96.

Fig. 4(a) depicts the whole-channel pre-whitened MEG signals before and after artifact removal (left and right panels, respectively), and the 2D view of the sensor positions (center panel). Fig. 4(b) shows the bilateral frontal channel recordings with synchronous activity caused by eye movement and heart beat related artifactual components. Comparison between the MEG signal before and after removal of the artifactual components demonstrated successful decontamination of the MEG signals by removing the identified artifactual ICs. The scalp map pattern, activity and power spectral density of the artifactual ICs identified in this trial are shown in Fig. 4(c). These artifactual ICs are consistent with the artifactual components identified in former studies (Vigario et al., 2000).

3.2. Identification of artifacts in multiple trials

Fig. 5 shows threshold-based clustering of the ICs for the multiple trial datasets after three iterations. Red dash lines indicate the threshold used for clustering. Fig. 5(a) and (c) are scatter plots of EKG and EOG1 feature values for dataset two (within subjects); Fig. 5(b) and (d) correspond to scatter plots for dataset three (across subjects). For EKGs, the features displayed are dtopo and H. For EOG1s, the features displayed are dtopo and K. Clusters marked with arrows were selected, and corresponding ICs were removed on a single trial basis. We also compared mean values of the IC features between the identified ICs and the un-identified ICs with Kruskal–Wallis non-parametric statistical analyses. Both dtopo ($\chi^2 = 307.33$, $p < 0.001$), and K value ($\chi^2 = 251.23$, $p < 0.001$) between EOG1 and non-EOG1 ICs, and dtopo ($\chi^2 = 274.32$, $p < 0.001$), and H value ($\chi^2 = 277.24$,

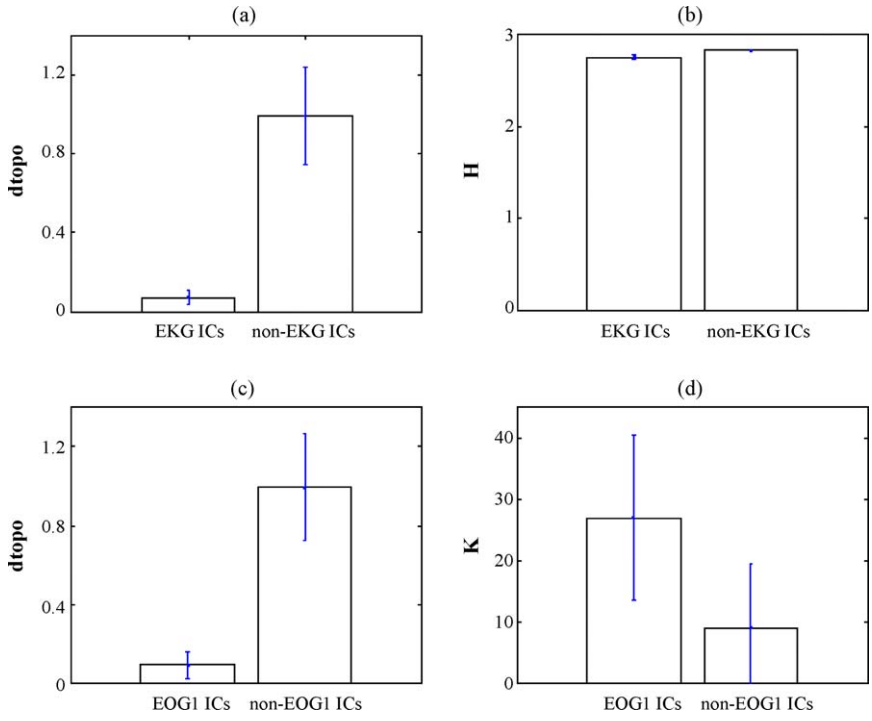

Fig. 6. Comparison of feature values between artifactual ICs and non-artifactual ICs, independent components were categorized from dataset three. (a) Comparison of dtopo values between EKG-related ICs and non-EKG ICs. (b) Comparison of H values between EKG-related ICs and non-EKG ICs. (c) Comparison of dtopo values between EOG1-related ICs and non-EOG1 ICs. (d) Comparison of K values between EOG1-related ICs and non-EOG1 ICs.

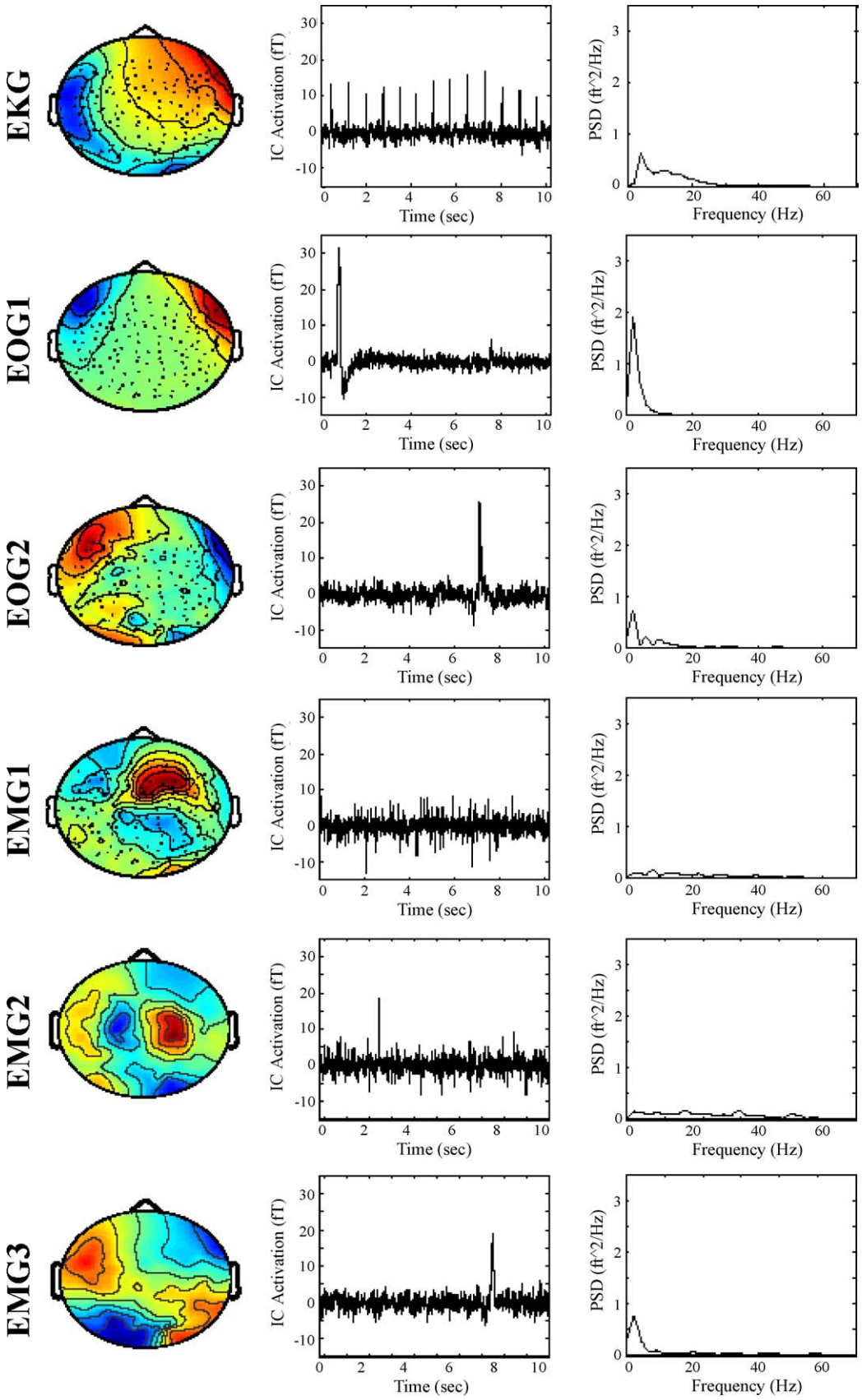

Fig. 7. The artifactual IC templates. Left column: topographic patterns of IC templates; center column: activations of IC templates; right column: power spectral patterns of IC templates. First row: the features of the template IC related to EKG; second row: features of template IC related to EOG1; third row: features of template IC related to EOG2; fourth row to sixth row: features of template ICs related to three instances of muscle activities.

$p < 0.001$) between EKG and non-EKG ICs showed statistically significant differences (Fig. 6).

3.3. The artifactual IC templates and update

Fig. 7 depicts the initial templates for the artifactual IC scalp maps (left panel), activations (center panel) and power spectral density plots (right panel). The six templates represent the artifactual ICs related to EKG, EOG1, EOG2 and three instances of muscle activities, respectively. EKG template IC showed the characteristic activation pattern. EOG1 and EOG2 template ICs represented the types of eye movements with opposite direction of electric sources. The three muscle activity-related IC templates were selected with the similar topographic patterns as the ones showed in a former study (Vigario et al., 2000) and with the characteristic bursting pattern in IC activations. Fig. 8 shows the comparison between the initial template and the final identified IC features in the identification of EKG artifactual ICs with dataset two. The similarity between the initial template topographic pattern and the mean topographic pattern of identified ICs indicated the closeness of the identified EKG ICs. The similarity of the characteristics between same types of ICs also can be seen in the power spectral patterns. The typical 4 and 8 Hz peaks in EKG ICs are clearly seen in the averaged EKG IC power spectrum.

3.4. Performance of identification methods

As shown in Fig. 9, the threshold-based clustering method identified more ICs and showed higher correctness than the ART-2 network categorization method although the threshold-based clustering method for EOG1 artifactual ICs did not reach 100% correctness. As expected, different vigilance levels showed different performance levels, so that increasing the vigilance level resulted in an increased number of identified ICs. However, the correctness of identification decreased when $\rho > 0.97$.

3.5. Continuous signal versus single-trial based identification

Table 1 and Fig. 10 illustrate the identification and clustering of artifactual ICs in dataset 4. The scatter plots in Fig. 10 illustrate the ICs in feature space constructed with the two features with the highest discriminatory power for the target artifactual ICs.

As shown in Table 1 and Fig. 10(a) and (c), two EKG and two EOG1 artifactual ICs were identified within 157 ICs when applying ICA on the concatenated MEG signal. EOG2 artifactual ICs were not found in these 157 ICs (Fig. 10 (e)), so we only compared pvaf (percentage variance accounted for) values for the identified EKG and EOG1 artifactual ICs (Table 2).

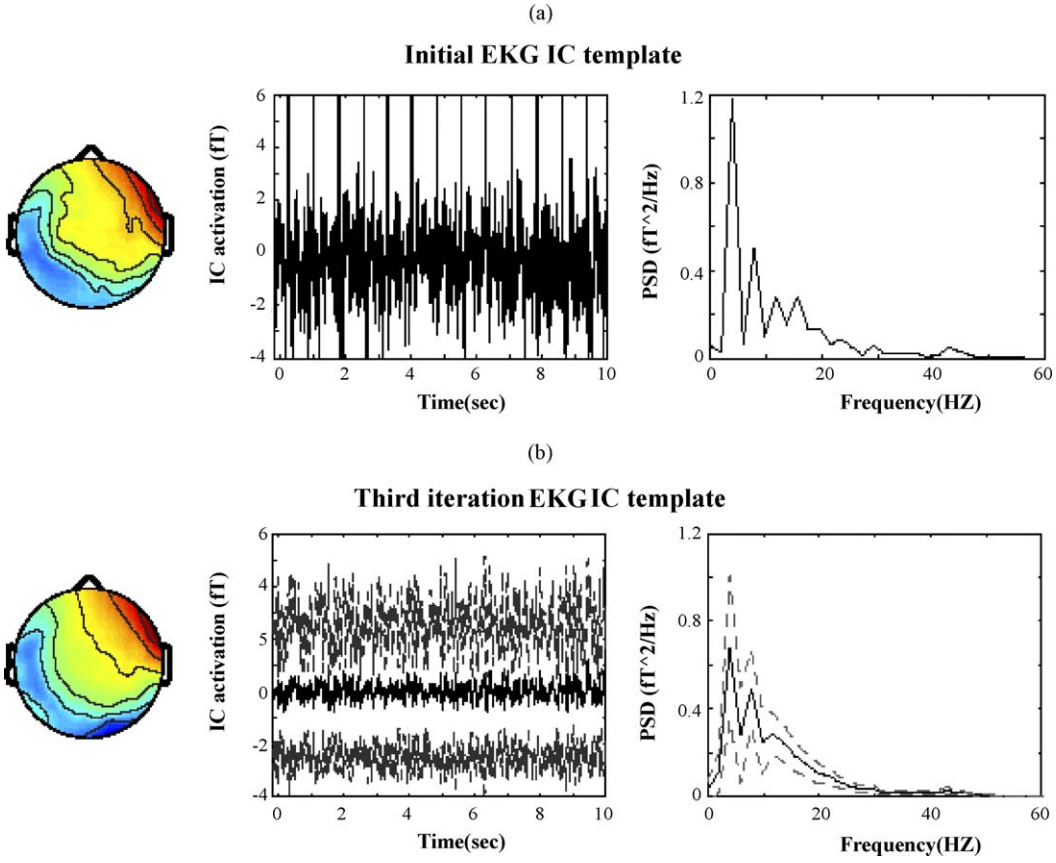

Fig. 8. The illustration of iterative identification processes of artifactual ICs. (a) The initial template of EKG-related IC for identification of EKG in dataset two. Topographic pattern, activation and power spectral pattern of the template IC were showed. (b) The identification results after three iterations. Left: mean topographic pattern of identified ICs; center: mean activation and 95% confidence interval (light gray dash lines) of identified ICs, and; right: mean power spectrum and 95% confidence interval (light gray dash lines) of identified ICs.

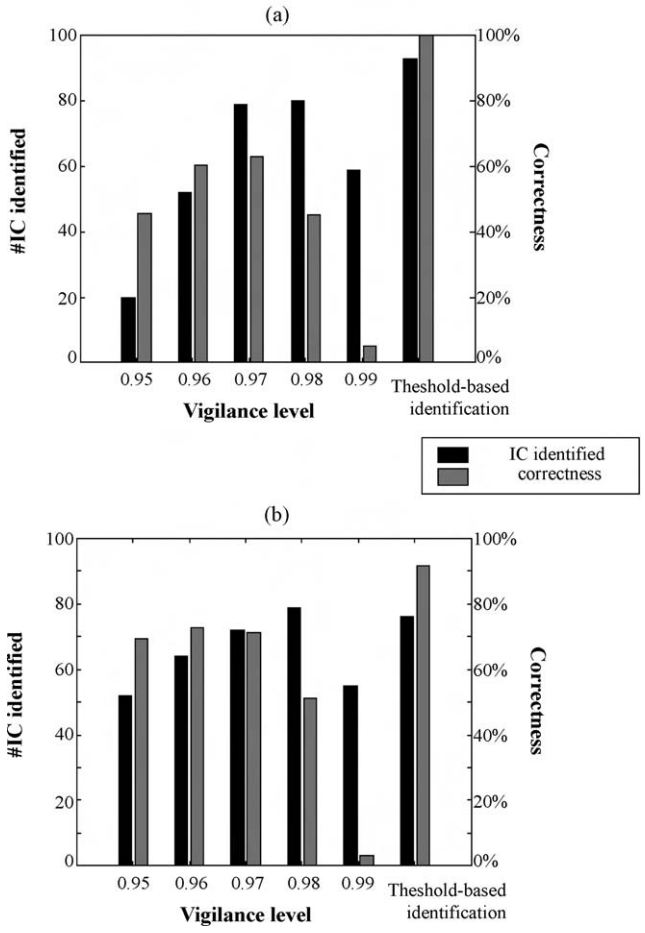

Fig. 9. Performance evaluation of artifactual IC identification with ART-2 network and threshold-based clustering. Data from dataset two. Vigilance levels used in ART-2 network were 0.95–0.99. (a) Identification of EKG-related ICs. Left *y*-axis indicates the number of IC identified with each method as indicated in *x*-axis, right *y*-axis indicates the correctness values. (b) Identification of EOG1-related ICs.

Fig. 11(a) and (b) shows the 2D projection (location) of the least (denoted by the filled dots) and the most (denoted by the white dots) influenced sensors. As shown in Fig. 11(c) and (d), for the most influenced sensors, the mean pvaf values of the identified artifactual ICs for the ICA processing of single-trial based epochs were similar to the identified artifactual ICs using ICA on the concatenated MEG signals (see Table 2). For the least influenced sensors, the mean pvaf values of the identified artifactual ICs from the single-trial based epochs were significantly smaller statistically for the EKG ICs ($p < .01$; Table 2). No statistically significant difference was found for EOG1 ICs in the least influenced sensors.

3.6. Overestimation and underestimation in identification of artifactual ICs.

Table 1 shows the evaluation of overestimation and underestimation of the identification methods. The results were computed from dataset four. For EKG ICs, the higher threshold-based clustering method performed best for the concatenated ICA procedure (coverage = 100%, correctness = 100%), whereas the lower threshold-based clustering method performed best for the single-trial based epochs (coverage = 100%, correctness = 100%). Applying the ART-2 neural network categorization on ICs identified by the higher threshold-based clustering method alleviated the overestimation caused by the lower clustering threshold with slight underestimation of the EKG-related ICs computed from single-trial based epochs.

For EOG1 and EOG2 ICs, both coverage and correctness performance based on the neural network categorization were not as good as those seen in the threshold-based clustering method. The higher threshold-based method captured all artifactual ICs at expense of correctness because of the broader threshold criteria used, whereas the lower threshold-based method achieved 100% correctness, but the coverage was low. Applying the ART-2 method on the ICs identified with the higher threshold-based method increased the correctness ratio.

No EOG2 ICs were identified or visually found in ICs computed from the concatenated MEG signal. However, for the single-trial based epoch IC dataset, seven EOG2 ICs were identified. The lower threshold-based clustering method has identified only one IC among the seven ICs, whereas the neural network categorization identified six of them with very low correctness. The higher threshold-based clustering method identified all seven ICs with less than 50% correctness. Applying network categorization on ICs identified with the higher threshold-based clustering method increased correctness to 54%.

4. Discussion

In this paper we proposed a method for artifact identification and removal in MEG signals based on independent component analysis and clustering methods. The procedure consisted of four iterated consecutive steps: (1) ICA computation, (2) calculation of IC features, (3) identification of artifact ICs with iterative procedures and template updating, and (4) removal of artifactual ICs and reconstruction of artifact-cleaned MEG signals.

In this study we tested two IC identification methods. Both methods identified the artifactual ICs based on feature values computed from the scalp map, power spectral density, kurtosis and entropy. The threshold-based clustering method showed effective identification and higher accuracy; whereas the automatic ART-2 network categorization method had the highest identification rates and correctness at vigilance level of 0.97. Combination of these two methods increased the accuracy of identification with no or very small decrease of coverage for target artifactual ICs. To our knowledge, this is the first attempt to use an ART-2 categorization neural network in artifact identification and automatic removal for MEG signals.

For artifactual IC identification and removal, both overestimation and underestimation should be avoided since overestimation may cause data loss related to non-artifactual sources, where underestimation may cause insufficient removal of artifacts in the MEG signals. We evaluated the performance of each method using measurements of coverage and correctness of identification. The results showed that for EKG ICs, the threshold-based method was better in avoiding both overestimation and underestimation. This may be because of the amount of information

Table 1
Evaluation of performance—over estimation and underestimation

	ICA processes	Methods[a]	Parameters[b]	Number of artifactual ICs in dataset	Number of ICs identified	Number of target artifactual ICs among identified ICs	Coverage (%)	Correctness (%)
EKG	CAT[c]	nw	0.97	2	2	1	50.00	50.00
		Ltb	[0.4 0.5 0 2.84]	2	1	1	00.00	100.00
		Htb	[0.8 0.5 0 2.84]	2	2	2	100.00	100.00
		comb	nw + Htb	2	2	2	100.00	100.00
	EPOCH	nw	0.97	21	37	18	85.71	48.65
		Ltb	[0.2 0.5 0 2.8]	21	21	21	100.00	100.00
		Htb	[0.4 0.5 0 2.84]	21	35	21	100.00	60.00
		comb	nw + Htb	21	21	18	85.71	85.71
EOG1	CAT	nw	0.97	2	3	1	50.00	33.33
		Ltb	[0.5 0.5 10 2.84]	2	2	2	100.00	100.00
		Htb	[0.8 0.5 10 2.84]	2	2	2	100.00	100.00
		comb	nw + Htb	2	2	2	100.00	100.00
	EPOCH	nw	0.97	19	48	16	84.21	33.33
		Ltb	[0.2 0.5 10 3.0]	19	7	7	36.84	100.00
		Htb	[0.5 0.5 10 3.0]	19	23	19	100.00	82.61
		comb	nw + Htb	19	21	18	94.74	85.71
EOG2	CAT[d]	nw	0.97	0	3	0	n/a	n/a
		Ltb	[0.5 0.5 10 3.0]	0	0	0	n/a	n/a
		Htb	[0.8 0.5 10 3.0]	0	0	0	n/a	n/a
		comb	nw + Htb	0	n/a	n/a	n/a	n/a
	EPOCH	nw	0.97	7	51	6	85.71	11.76
		Ltb	[0.2 0.5 10 3.0]	7	1	1	14.29	100.00
		Htb	[0.5 0.5 10 3.0]	7	15	7	100.00	46.67
		comb	nw + Htb	7	13	7	100.00	53.85

[a] Methods: nw, ART-2 categorization neural network; Ltb, lower threshold-based clustering; Htb, higher threshold-based clustering; nw + Htb, ICs were identified with Htb method first, the identified ICs were then processed again with neural network method.
[b] Parameters: the parameters for nw method were the vigilance levels. The parameters for threshold-based clustering methods were vectors composed of threshold values, the values were $[T_{\text{dtopo}}\ T_{\text{dPxx}}\ T_K\ T_H]$. The identification criteria were dtopo $\leq T_{\text{dtopo}}$, dPxx $\leq T_{\text{dPxx}}$, $H \leq T_H$ and $K > T_K$.
[c] ICA process: CAT, MEG signal was concatenated from 21 trials, 157 ICs were computed from whole dataset; EPOCH, ICs were computed on single trial basis, each trial has 157 ICs.
[d] Since no EOG2 ICs were identified from CAT IC dataset, coverage and correctness values have not been computed. Hence, no evaluation of overestimation and underestimation was performed on this dataset.

carried in EKG ICs, which can be separated from other ICs by selecting an appropriate threshold. This effect has been demonstrated in Fig. 3(b) and Fig. 5(a) and (b). For EOG1 and EOG2 ICs, the threshold-based methods performed better than the neural network, albeit with imperfect performance—the higher threshold-based method demonstrated overestimation of the artifactual ICs whereas the lower threshold-based method demonstrated underestimation of the artifactual ICs. Combination of these two methods with the ART-2 method on ICs identified by higher threshold-based method alleviated the overestimation while keeping the benefit of coverage; the only exception was the failure to identify one of the EOG1 ICs. As a conclusion, the combined method appears to be the best approach showing higher accuracy and automatization.

Table 2
Evaluation of performance—ICA on continuous MEG signal (CAT) vs. ICA on single-trial epochs (EPOCH)

	ICA process[a]	Most influenced sensors[b]				Least influenced sensors[b]			
		Mean pvaf	S.E.M.[c]	*F* value	Probability	Mean pvaf	S.E.M.	*F* value	Probability
EKG	CAT	18.6	3.98	0.1	$p > 0.05$	2.9	0.64	8.42	$p < 0.01$
	EPOCH	16.8	3.98			0.2	0.64		
EOG1	CAT	10.1	2.26	0.87	$p > 0.05$	1.1	0.35	1.53	$p > 0.05$
	EPOCH	7.1	2.26			0.4	0.35		

[a] ICA process: CAT, MEG signal was concatenated from 21 trials, 157 ICs were computed from whole dataset; EPOCH, ICs were computed on single trial basis, each trial has 157 ICs.
[b] Sensors were grouped as most influenced sensors and least influenced sensors. Most influenced sensors were the sensors with highest absolute values in the remixing vectors of the target artifactual IC templates. Least influenced sensors were the sensors with lowest absolute values in the remixing vectors of the target artifactual IC templates.
[c] S.E.M.: standard error of means.

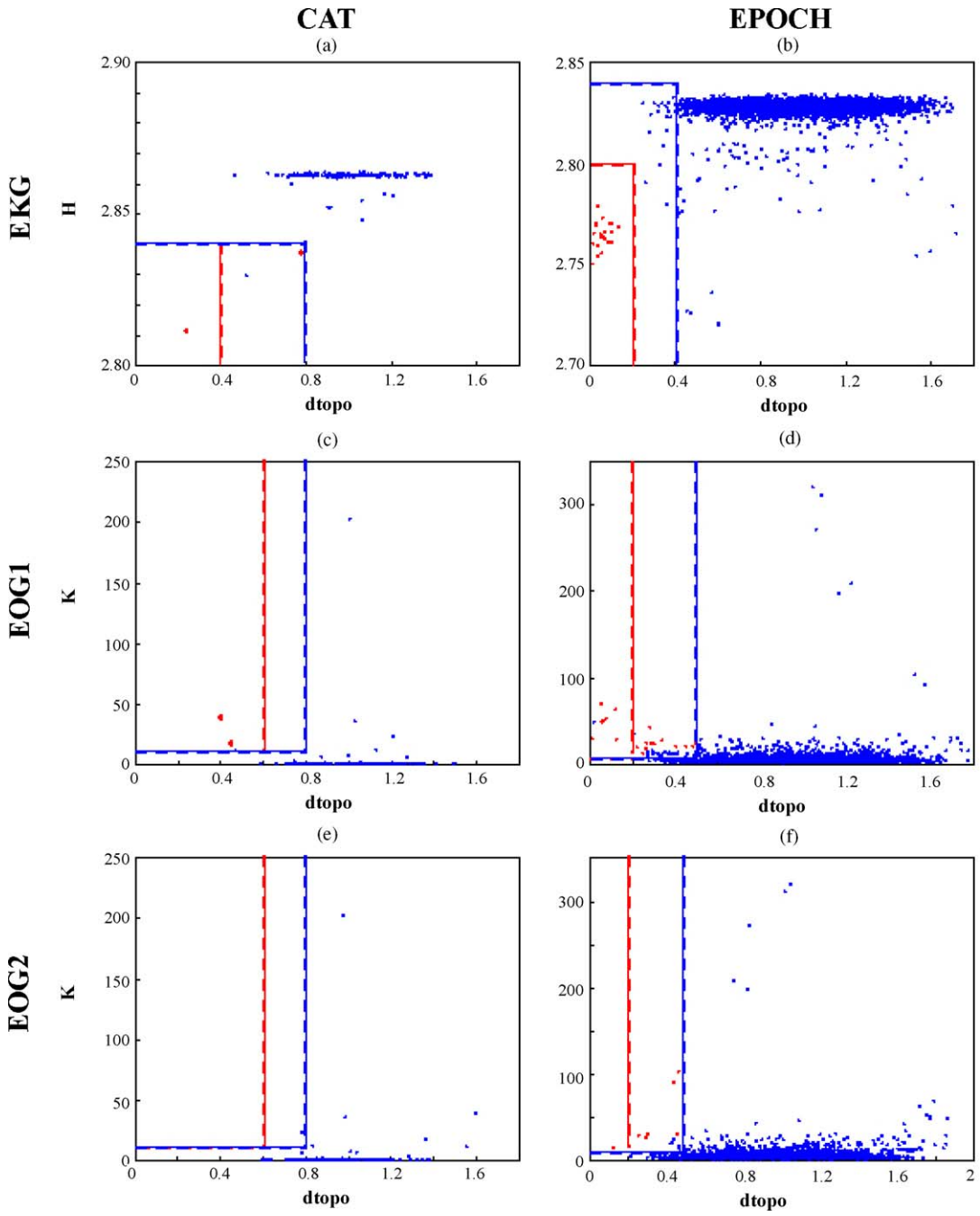

Fig. 10. Threshold-based clustering of artifactual ICs in dataset 4. The threshold criteria for clustering were indicated with red dash lines (lower threshold-based clustering) and blue dash lines (higher threshold-based clustering), which corresponded to the values shown in Table 3. The target artifactual ICs within the clusters were indicated with red dots. Subplots (a), (c) and (e) were scatter plots of the ICs computed from concatenated MEG signals; subplots (b), (d) and (f) were scatter plots of the ICs computed from single trial based epochs. Subplots (a) and (b) illustrate the clustering of EKG-related artifactual ICs in dtopo-*H* space, subplots (c) and (d) illustrate the clustering of EOG1-related artifactual ICs, and subplots (e) and (f) illustrate the clustering of EOG2-related artifactual ICs in dtopo-*K* space.

One issue regarding the application of ICA to blindly separate sources is whether it is appropriate to apply ICA on continuous (concatenated) data rather than on single-trial based epochs. To address this issue, we tested the percentage of variance accounted for (pvaf) by the identified ICs using concatenated data or epochs with 10,201 samples per sensor per trial. The significantly higher pvaf values of the least influenced sensors of EKG ICs indicated that the ICs computed with concatenated ICA contained information from other events which had been mixed with the heart beats since only 157 ICs had been computed with this dataset. As a comparison, applying ICA on single-trial based epochs produced 157 ICs for each epoch, which made separation of the sources relatively easier. Moreover, no EOG2 ICs was identified from the ICs extracted from the concatenated ICA processing, though seven EOG2 ICs have been identified with single-trial based epochs (as demonstrated

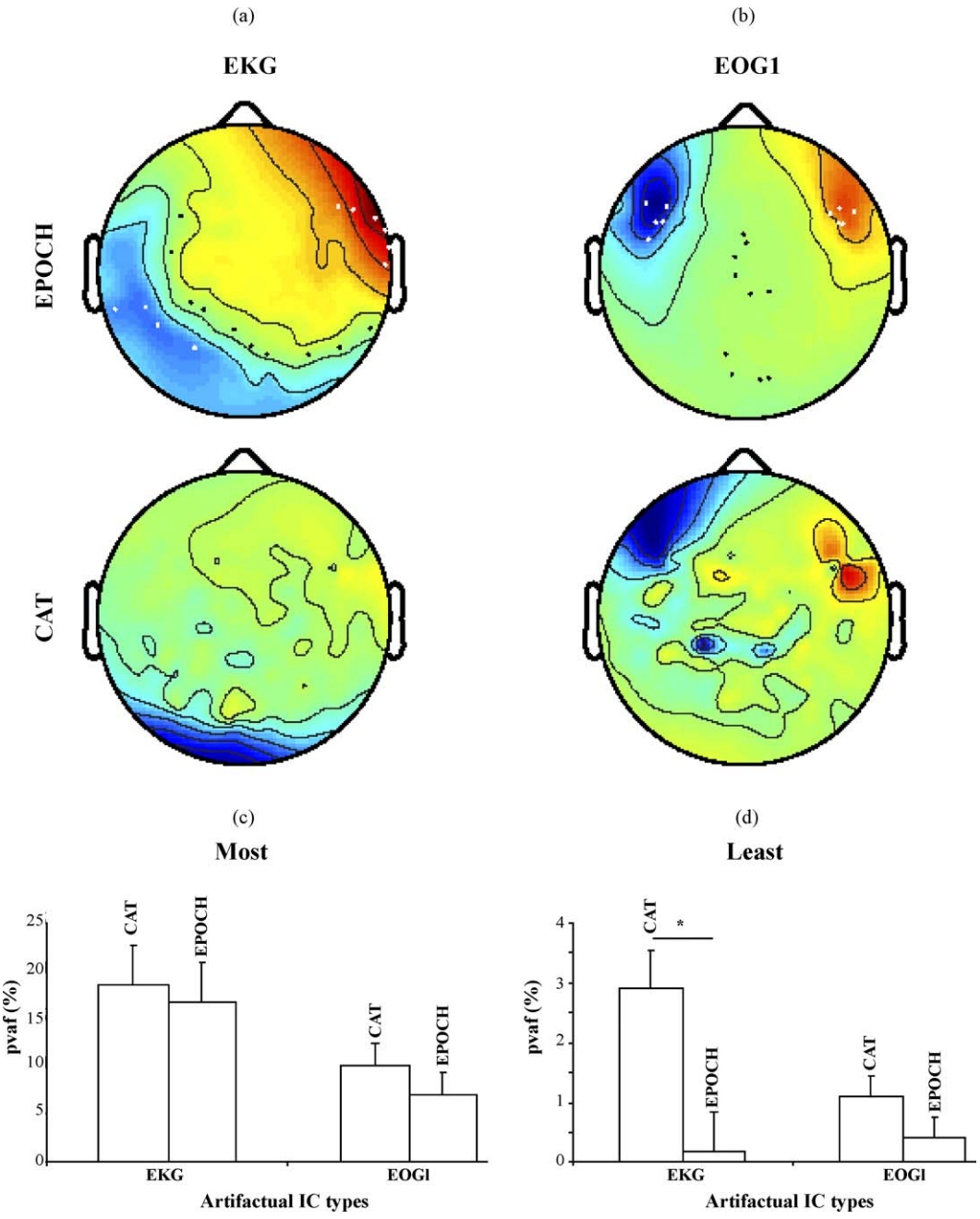

Fig. 11. Comparison of artifactual IC identification and rejection between concatenated vs. single-trial epochs ICA processing. (a) Topographic patterns of EKG-related ICs computed from concatenated MEG signal (CAT) and single-trial based epochs (EPOCH). In EPOCH pattern, sensors selected to compare pvaf values were labeled, with white circles indicating the most influenced sensors, and black circles representing the least influenced sensors. (b) Topographic patterns of EOG1-related ICs and demonstration of the sensors selected to compare pvaf values. (c) Pair-wise comparison of pvaf values on most influenced sensors (most). Height of bars indicated the mean pvaf values of identified target artifactual ICs. Error bars are standard error of means (S.E.M.). (d) Pair-wise comparison of pvaf values on least influenced sensors. (*) Indicates a statistically significant difference with $p < 0.05$.

in Table 2). Thus, we would like to argue that this epoch-based IC identification method works better than the concatenated approach provided there is enough number of data points (samples).

ICA has been shown to be an efficient tool for separating artifactual sources from functional brain signals (Jung et al., 2000) as well as to identify functional sources underlying similar neuronal mechanisms (Contreras-Vidal and Kerick, 2004; Makeig et al., 1999; Makeig et al., 2002; for review, see Delorme and Makeig, 2004) in EEG/MEG signal processing. As compared to segment-rejection methods, artifactual IC identification and rejection can save functional information from mass data loss as well as discover the functional brain activities masked by artifactual signals. Artifactual IC identification and rejection with computation of the features has been proved in this study to be less time and labor consuming for analysis of MEG signals, especially for datasets contained multiple subjects and multiple trials.

In this paper, we have combined typical components features to describe each IC, which have been used separately and emphasized individually in prior studies (Barbati et al., 2004; Delorme and Makeig, 2004). We also used data-driven clustering and categorization methods to identify the artifactual ICs. The results for single trial and multiple trials datasets confirmed that this approach can correctly identify the artifactual ICs efficiently while the best performance was obtained from combination of these two methods. However, the use of independent components rises the question of what their functional significance may be, that is, what is their functional relationship to the brain activities. Although prior ICA studies have shown ICs related to certain types of artifacts or event-related brain dynamics (Vigario et al., 2000), it is still a matter of debate whether independent components with small differences (e.g., from trial to trial or across subjects) in spatial and spectral features represent the same type of biological events. An important finding in the present study is that artifactual ICs have feature values close to each other albeit the variability displayed across subjects and trials (Fig. 6(b) and (d); see Fig. 7 for statistical analysis). The successful identification with template updating also proves that it is possible to describe artifactual ICs using template features with small distributions. As these templates can be used as initial templates for forthcoming MEG studies, it should be possible to create and update a general database of artifact and non-artifactual component templates. Thus, this method is applicable to ICs related to functional brain activities.

Acknowledgements

The authors would like to acknowledge the diligence and insight of the anonymous reviewers who contributed significantly to the revision of this manuscript.

Appendix A. Categorization of ICs with ART-2 neural networks

Based on the Adaptive Resonance Theory developed by Carpenter and Grossberg (1987a), ART networks categorize arbitrary sequences of input patterns with self-organized stable pattern recognition architectures. Implementations have been reported including a digital (binary) system (ART-1; Carpenter and Grossberg, 1987a), a system based on fuzzy logic (fuzzy ART; Carpenter et al., 1991), a system for categorization of analog inputs (ART-2; Carpenter and Grossberg, 1987b), an implementation based on a system of ordinary differential equations capable of stand-alone running in real time (Molenaar and Raijmakers, 1997), and a fast VLSI asynchronous system (ART-1m; Serrano-Gotarredona and Linares-Barranco, 1996). In the present study, we employed the analog implementation of the ART (that is, ART-2) network to categorize the IC features (see Fig. 12). The following description of the architecture and the equations defining the ART-2 network implemented in this paper is based on the original mathematical description (their Eqs. (1)–(20)) in Carpenter and Grossberg (1987a), which are reproduced (with the author's permission) herein as an Appendix A for the convenience of the reader. The network contains two competitive, short-term memory (STM) layers F1 and F2, and two long-term memory (LTM) traces z_{ij} and z_{ji}, which denote the synaptic strength from F1 layer to F2 layer and synaptic strength from F2 layer to F1 layer, respectively. These bottom-up and top-down pathways can be seen as adaptive filters.

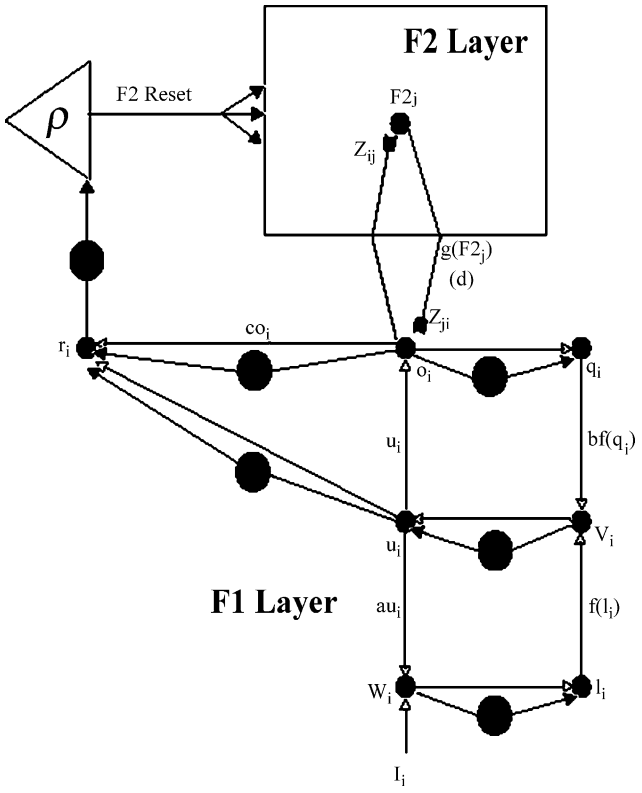

Fig. 12. A typical ART 2 architecture. Open arrows indicate specific patterned inputs to target nodes. Filled arrows indicate nonspecific gain control inputs. The gain control nuclei (large filled circles) nonspecifically inhibit target nodes in proportion to the L_2-norm of STM activity in their source fields (Eqs. (13), (14), (17), (23) and (24)). When F2 makes a choice, $g(F2_j) = d$ if the jth F2 node is active and $g(F2_j) = 0$ otherwise. (Figure and caption are adapted and reproduced with permission from Fig. 6 of Carpenter & Grossberg (1987b, p. 406).)

The F1 layer consists of six processing stages. Each stage has the same dimensionality as the input vector (I_i), which herein is composed of four feature values computed from each IC. The potential V_i, or the STM activity, of the ith node in each of the F1 layer processing stages obeys the membrane equation

$$\varepsilon \frac{\mathrm{d}}{\mathrm{d}t} V_i = -\alpha V_i + (1 - BV_i)J_i^+ - (C + DV_i)J_i^- \tag{10}$$

where J_i^+ denotes the total excitatory input, and J_i^- denotes the total inhibitory input to the ith node. ε represents the ratio between the STM membrane time constant and the LTM membrane time constant. With the constraints $0 < \varepsilon \ll 1$, $B = 0$ and $C = 0$, Eq. (10) reduced to (at steady state)

$$V_i = \frac{J_i^+}{\alpha + DJ_i^-} \tag{11}$$

in the singular form as $\varepsilon \to 0$. With Eq. (11), the STM activity of the ith node in each stage in F1 layer, represented as o_i, q_i, u_i,

v_i, l_i and w_i, becomes:

$$o_i = u_i + \sum_j g(\mathrm{F2}_j) z_{ji} \tag{12}$$

$$q_i = \frac{o_i}{e + ||\boldsymbol{o}||} \tag{13}$$

$$u_i = \frac{v_i}{e + ||\boldsymbol{v}||} \tag{14}$$

$$v_i = f(l_i) + bf(q_i) \tag{15}$$

$$w_i = I_i + au_i \tag{16}$$

$$l_i = \frac{w_i}{e + ||w||} \tag{17}$$

where the expression $||\boldsymbol{x}||$ denotes the L2-norm of a vector $\boldsymbol{x}$; $\mathrm{F2}_j$ is the STM activity of the *jth* F2 node. Thus, z_{ji} is the weight of the projection from the *j*th F2 node to the *i*th node of the output stage O in the F1 layer. Function *f* is the piecewise linear function

$$f(\gamma) = \begin{cases} 0 & \text{if} \quad 0 \le \gamma \le \theta \\ \gamma & \text{if} \quad \gamma \ge \theta \end{cases} \tag{18}$$

where θ is the threshold.

The STM activity pattern in F2 layer represents the categorization output, which is computed in accordance with the 'winner-take-all' law based on the filtered input from the F1 layer. The total input to *j*th F2 node is calculated by

$$\mathrm{F2}_j = o_i \times z_{ij} \tag{19}$$

where o_i is the *i*th output STM activity of F1 layer, and z_{ij} is the weight of the projection from the *i*th output node of F1 layer to the *j*th node in F2 layer. The winner node in F2 layer is then identified as the node that receives the maximal input, and the other nodes are inhibited. The selection is performed through the function $g(\mathrm{F2}_j)$:

$$g(\mathrm{F2}_j) = \begin{cases} d & \text{active node} \\ 0 & \text{inactive node} \end{cases} \tag{20}$$

The top-down and bottom-up LTM trace equations defining the adaptive filters are given by

$$\frac{d}{\mathrm{d}t} z_{ij} = g(\mathrm{F2}_j)[o_i - z_{ij}] \tag{21}$$

$$\frac{d}{\mathrm{d}t} z_{ji} = g(\mathrm{F2}_j)[o_i - z_{ji}] \tag{22}$$

The F2 layer will be reset when a mismatch is detected (for a given vigilance parameter value) between the F1 STM pattern and an active LTM pattern. The degree of match is determined by the activity pattern in node *R* as a vector $\boldsymbol{r}$, in which the *i*th element r_i is determined by

$$r_i = \frac{u_i + co_i}{e + ||\boldsymbol{u}|| + ||c\boldsymbol{o}||} \tag{23}$$

Table 3
Parameter values used in the ART-2 network simulations

a	10
b	10
c	0.1
d	0.9
e	0
θ	0.2

A mismatch will be detected when

$$\frac{\rho}{e + ||\boldsymbol{r}||} > 1 \tag{24}$$

and the F2 layer will be reset. The vigilance parameter ρ is set between 0 and 1, and this is the only parameter that was parametrically varied in the present study. A categorization result will be accepted when both layers are stable and no mismatch has been detected. Table 3 shows the parameters of the network used in the present study.

References

Adachi Y, Shimogawara M, Haruta Y, Ochiai M. Reduction of non-periodic environmental magnetic noise in MEG measurement by continuously adjusted least squares method. IEEE Trans Appl Supercond 2001;11(1):669–72.

Barbati G, Porcaro C, Zappasodi F, Rossini PM, Tecchio F. Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals. Clin Neurophysiol 2004;115:1220–32.

Bell AJ, Sejnowski TJ. An information-maximazation approach to blind separation and blind deconvolution. Neural comput 1995;7:1129–59.

Carpenter G, Grossberg S. A massively parallel architecture for a self organizing neural pattern recognition machine. Comput Vis Graph Image Process 1987a;37:54–115.

Carpenter GA, Grossberg S. ART2: stable self-organization of category recognition codes for analog input patterns. Appl Opt 1987b;26(23):4919–30.

Carpenter G, Grossberg S, Rosen DB. Fuzzy ART: fast stable learning and categorization of analog patterns by an adaptive resonance system. Neural Netw 1991;4:759–71.

Contreras-Vidal JL, Kerick SE. Independent component analysis of dynamic brain responses during visuomotor adaptation. NeuroImage 2004;21:936–45.

Delorme A, Makeig S, Sejnowski TJ. Automatic artifact rejection for EEG data using high-order statistics and independent component analysis. In: Proceedings of the 3rd International Workshop on ICA, San Diego; 2001. p. 457–62.

Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 2004;134:9–21.

Hämäläinen M, Hari R, Ilmoniemi R, Knuutila J, Lounasmaa OV. Magnetoencephalography—theory, instrumentation and applications to noninvasive studies of the working human brain. Rev Mod Phys 1993;65(2):413–97.

Hillyard SA, Galambos R. Eye-movement artifact in the CNV. Electroencephalogr Clin Neurophysiol 1970;28:173–82.

Jackson JD. Classical electrodynamics. 3rd ed. New York: John Wiley & Sons Inc.; 1999. p. 218–23.

James CJ, Gibson OJ. Temporally constrainted ICA: an application to artifact rejection in electromagnetic brain signal analysis. IEEE Trans Biomed Eng 2003;50(9):1108–16.

Joyce CA, Gorodnitsky IF, Kutas M. Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. Psychophysiology 2004;41:313–25.

Jung TP, Humphries C, Lee TW, Makeig S, Mckeown MJ, Iragui V, Sejnowski TJ. Extended ICA removes artifacts from electroencephalographic recordings. Adv Neural Inform Process Syst 1998;10:894–900.

Jung TP, Makeig S, Humphries C, Lee TW, Mckeown MJ, Iragui V, Sejnowski TJ. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 2000;37:163–78.

Kado H, Higuchi M, Shimogawara M, Haruta Y, Adachi Y, Kawai J, Ogata H, Uehara G. Magnetoencephalogram system developed in KIT. IEEE Trans Appl Supercond 1999;9(2):4057–62.

Makeig S, Bell AJ, Jung T-P, Sejnowski TJ. Independent component analysis of electroencephalographic data. In: Touretzky D, Mozer M, Hasselmo M, editors. Advances in neural information processing systems, vol. 8. Cambridge, MA: MIT Press; 1996. p. 145–51.

Makeig S, Westerfield M, Jung TP, Covington J, Townsend J, Sejnowski TJ, Courchesne E. Functionally independent components of the late positive event-related potential during visual spatial attention. J Neurosci 1999;19(7):2665–80.

Makeig S, Westerfield M, Jung TP, Enghoff S, Townsend J, Courchesne E, Sejnowski TJ. Dynamic brain sources of visual evoked responses. Science 2002;295:690–4.

Makeig S, Debener S, Onton J, Delorme A. Mining event-related brain dynamics. Trends Cogn Sci 2004;8(5):204–10.

Molenaar PC, Raijmakers ME. Exact ART: a complete implementation of an ART network. Neural Netw 1997;10(4):649–69.

Serrano-Gotarredona T, Linares-Barranco B. A real-time clustering microchip neural engine. IEEE Trans VLSI Syst 1996;4:195–209.

Tang A, Pearlmutter B, Phung D, Reeb B. Independent component of magnetoencephalography:localization. Neural Comput 2002;14:1828–57.

Vigario R, Sarela J, Jousmaki V, Hamalainen M, Oja E. Independent component approach to the analysis of EEG and MEG recordings. IEEE Trans Biomed Eng 2000;47:587–93.

Woestenburg JC, Verbaten MN, Slangen JL. The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain. Biol Psychol 1983;16:127–47.