

# Linear and Nonlinear Features for Automatic Artifacts Removal from MEG Data Based on ICA

Montri Phothisonothai\*, Hiroyuki Tsubomi\*, Aki Kondo\*, Mitsuru Kikuchi†, Yuko Yoshimura†, Yoshio Minabe†, and Kastumi Watanabe\*

\*Research Center for Advanced Science and Technology

The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904 Japan

E-mail: {montri, htsubomi, kondo, kw}@fennel.rcast.u-tokyo.ac.jp Tel/Fax: +81-3-5453-5249

†Research Center for Child Mental Development, Graduate School of Medical Science, Kanazawa University, Kanazawa 920-8641 Japan

E-mail: minabe@med.m.kanazawa-u.ac.jp

**Abstract**—This paper presents an automatic method to remove physiological artifacts from magnetoencephalogram (MEG) data based on independent component analysis (ICA). The proposed features including kurtosis (K), probability density (PD), central moment of frequency (CMoF), spectral entropy (SpecEn), and fractal dimension (FD) were used to identify the artifactual components such as cardiac, ocular, muscular, and sudden high-amplitude changes. For an ocular artifact, the frontal head region (FHR) thresholding was proposed. In this paper, ICA method was on the basis of FastICA algorithm to decompose the underlying sources in MEG data. Then, the corresponding ICs responsible for artifacts were identified by means of appropriate parameters. Comparison between MEG and artifactual components showed the statistical significant results at  $p < 0.001$  for all features. The output artifact-free MEG waveforms showed the applicability of the proposed method in removing artifactual components.

## I. INTRODUCTION

Magnetoencephalography (MEG) recording is a noninvasive technique in which measures electromagnetic field of brain activities with a good temporal resolution and moderate spatial resolution. MEG data are associated with electrical currents produced by a large number of neurons occurring naturally in the brain in terms of electromagnetic response. The application of using MEG data has been widely succeeded in many fields such as clinical research, biomedical engineering, physiological decoding, and cognitive science [1]–[5]. However, the presence of physiological artifacts could interfere with the measurement during recording, including: eye movements, muscular contractions, cardiac signals, sudden high-amplitude changes, and environmental noise. Generally, cerebral activities are spontaneously masked by noise from biological or technical origins. The distinction of actual MEG signal and noise is common in the experimental literature and is usually made and understood with respect to the experimental protocol in which neural potentials that are related to the experiment and regarded as brain signal or those are not regarded as noise [6]. Moreover, both noise and artifacts make analysis much more difficult and can even be mistaken for interpretation.

Recently, independent component analysis (ICA) has been proven to be an effective and applicable method for removing artifacts and noise in multi-channel physiological measure, e.g., electroencephalography (EEG) and MEG data [7]–[10], [22]. However, a few assumptions are made when using. And, ICA is unable to automatically determine the correct order of independent components (ICs), scale, or polarity of the sources, which make artifact identification difficult. Due to these limitations, the method for selecting artifactual ICs still remains challenging. Therefore, the availability of a standardized procedure for eliminating such problem would be extremely useful for the study of the human brain function [11], [12].

For the related works on ICA based artifact removal, various methods have been proposed for the identification of ICs related to the artifacts from multi-channel physiological signals such as EEG and MEG. The measurement used by most of these methods focuses on a characteristic of artifact components related to topographic patterns, statistical features, scalp map, or spectral properties [13]–[17]. However, there were no studies that take into account the combination of linear and nonlinear features. Available features and the automation of ICA based artifact removal would be helpful a wider application of MEG studies.

In this paper, we proposed several linear and nonlinear features for automatic artifact removal including kurtosis (K), probability density (PD), central moment of frequency (CMoF), spectral entropy (SpecEn), and fractal dimension (FD). These proposed features are employed to identify the artifactual components such as cardiac (EKG), ocular (EOG), muscular (EMG), and sudden high-amplitude changes (HAM). ICA method is on the basis of FastICA algorithm to decompose the underlying sources in MEG data. The individual features of artifacts in terms of linear and nonlinear methods are discussed. We also present the simple idea of reliable method to remove artifact automatically. The block diagram of the proposed method to identify and to remove artifacts is depicted in Fig. 1.

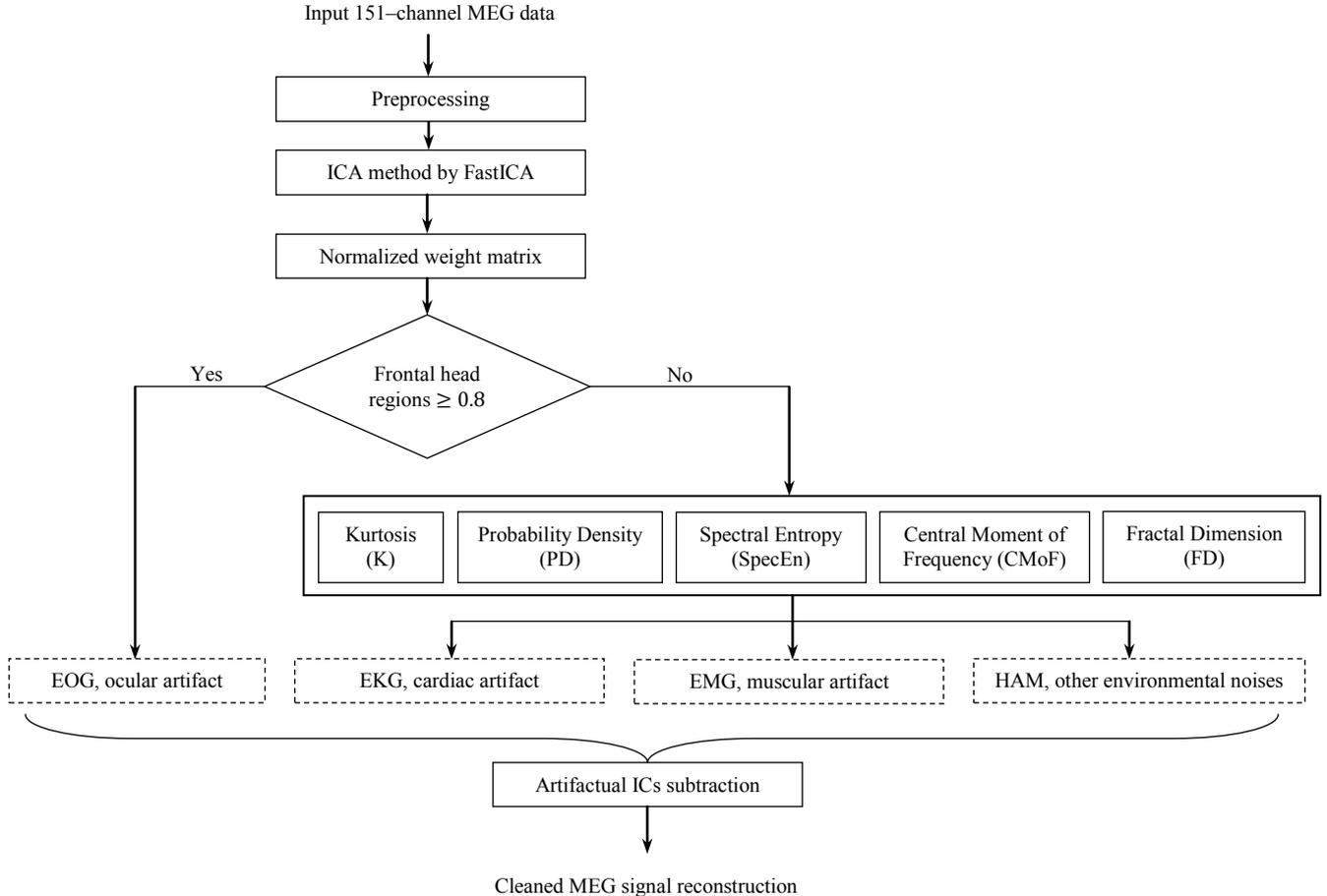


Fig. 1. Block diagram of the proposed method for artifactual ICs identification and removal.

## II. METHODOLOGY

### A. MEG Data Acquisition

MEG data were recorded with a 151-channel superconducting quantum interference device (SQUID) whole-head coaxial gradiometer MEG system (PQ 1151R; Yokogawa/KIT) in magnetically shielded room (Diode Steel). MEG data were digitized by the sampling rate of 1 kHz 16-bit resolution. Ten healthy children (5 boys, 5 girls) participated in this experiment. The participants had a mean chronological age of 65.4 months (39–81)<sup>1</sup>. All subjects were given to perform the resting state for three minutes. During MEG record-

ing, children lay on the bed and viewed a cartoon animation programs with stories through a TV monitor especially attractive to young children. The position of the head to be in the center of the MEG helmet was confirmed by measuring the magnetic fields after passing currents through coils that were attached at three positions as fiduciary points with respect to the landmark (bilateral mastoid processes and nasion). Off-line analysis of MEG data was performed with Yokogawa MEG Reader Toolbox Rev1.4 and Matlab (MathWorks).

### B. Preprocessing

Raw MEG data were pre-processed by an infinite impulse response (IIR) Butterworth type with a third order bandpass filter 1.5–80 Hz for initial environmental noise removal, power line noise at 50Hz has been removed by the notch filter (Q-factor = 35).

<sup>1</sup> All subjects had normal brain function and data collection was approved by the Ethics Committee of Kanazawa University Hospital, all which were in accordance with the Declaration of Helsinki.

### C. Independent Component Analysis (ICA) Model

ICA method has recently become an important tool for modeling and understanding empirical datasets. This method is a separating out independent sources from linearly mixed data, and belongs to the class of general linear models. Subsequently, ICA can be able to separate the underlying sources mixed in the raw MEG data. Due to the fact that magnetic fields of different bioelectric current sources superimpose linearly, the measured MEG data through of the SQUID array sensor can be modeled by mixing of a blind source [14]. The independent components (ICs) assumed to be mutually independent,  $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_m(t)]^T$ , and the observations of the mixed MEG data,  $\mathbf{x}(t)$ , can be modeled as

$$\mathbf{x}(t) = \sum_{k=1}^m \mathbf{m}_k \mathbf{s}(t) = \mathbf{M} \mathbf{s}(t) \quad (1)$$

where  $\mathbf{M}$  is a mixing matrix dimension  $m \times n$ . All observations are assumed to be mutually independent. The number of underlying signals is at most equal to the number of ICs  $n \geq m$ . The task of ICA is to determine the independent components (ICs),  $\hat{\mathbf{s}}(t)$ , from the observations by computing a separating matrix (or unmixing matrix),  $\mathbf{W}$ , which can be defined as

$$\hat{\mathbf{s}}(t) = \mathbf{W} \mathbf{x}(t) \quad (2)$$

The estimated ICs, in case of MEG data, are supposed to be independent from each other, i.e., EOG, EKG, EMG, HAM, and so on. This definition is based on the concept of mutual information that non-Gaussianity of the ICs is indispensable for identifying the model. In this study, we selected the ICA based on the fixed-point algorithm proposed by Hyvärinen and Oja [18], which is called FastICA<sup>2</sup>. This algorithm is for the estimation of ICs with Newton iteration to optimize an objective that measures independence from the observed data also shows a short computing time and to be less sensitive to white noise less than other ICA algorithms [26].

To determine the weight vectors in the separating matrix, we set the following parameters: maximum iteration 1,000 runs, contrast function is a cubic function,  $g = u^3$ , because time, cost, and features show this function to be more suitable than any other functions [19], and error rate for reaching an optimal condition,  $\varepsilon < 0.0001$ . In addition, there were two approaches for estimating the ICs, the deflation approach and symmetric approach. In this paper, we selected the deflation approach because it is more desirable in cases with specific ICs. After the process of ICA, the weight values in the separating matrix can be presented as shown in Fig. 2. Also, we

use the magnetic coil positions to construct the 3D ellipsoidal 151-channel topology based MEG for source localization.

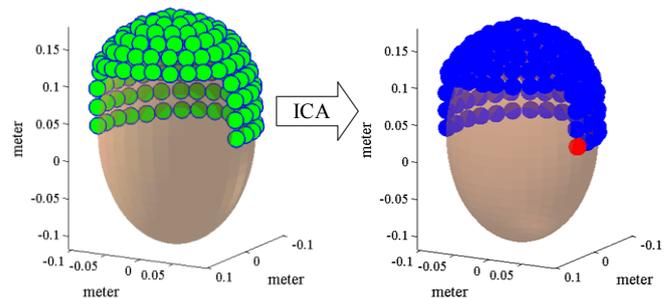


Fig. 2. The 3D ellipsoidal 151-channel topology and its ICA process.

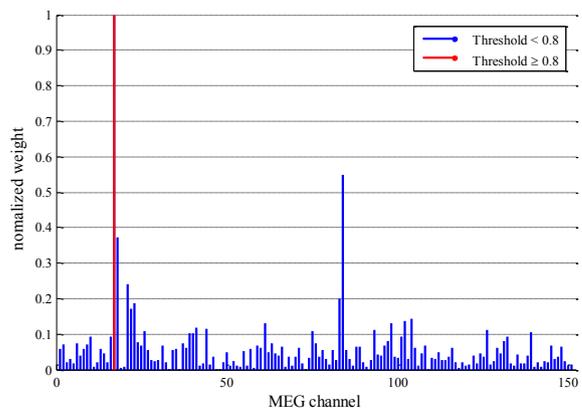


Fig. 3. Thresholding of the normalized weight value at 17th channel (IC-017).

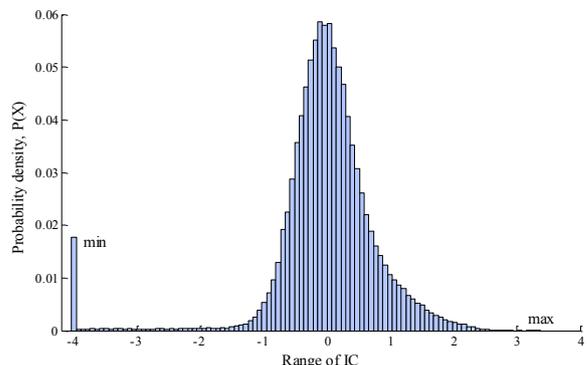


Fig. 4. Typical example of histogram space to identify cardiac artifact.

<sup>2</sup> FastICA is freely available at <http://www.cis.hut.fi/projects/ica/fastica>

#### D. Frontal Head Region (FHR) Thresholding

An ocular artifact (EOG) cause changes to the electric field around the eyes, e.g., eye blinking and eye movements. In measuring the MEG data, we are typically interested in the underlying neural potentials, which the recording is of the combination of neural and ocular artifacts. To remove such ocular artifact, FHR thresholding based on subject's nearest distance to the eye was proposed. In (2), we selected the twenty MEG channels to identify ocular artifact by the following threshold

$$T_{FHR} = \begin{cases} 1, & \text{if } \hat{w}_{ch} \geq 0.8 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

where  $\hat{w}_{ch}$  is the normalized weight vector of ICs at  $ch$ th channel: 17,21,25,49,130,51,137,84,88,93,114,117,118,147,148,149,129,138,145 and 146. At least one normalized weight value in the specific  $ch$ th channel that was greater than or equal to 0.8; it will considerably be artifact. For example, the normalized weight of ICs is detected at 17th channel (in this case  $T_{FHR} = 1$ ) shown in Fig. 3.

#### E. Kurtosis (K)

Kurtosis indicates the horizontal distribution of the observation comparing with the Gaussian distribution [25]. The kurtosis has been successfully used to detect ocular and cardiac [15],[17]. In this paper, we used the kurtosis to detect HAM artifact. The normalized kurtosis of a distribution is defined as

$$K = \frac{1}{3} \frac{E\{(x - \mu)^4\}}{\sigma^4} \quad (4)$$

where  $\mu$  is the mean of  $x$ ,  $\sigma$  is the standard deviation of  $x$ , and  $E\{\cdot\}$  represents the expected value. From (4), IC with kurtosis represents  $K = 1$  for the Gaussian distribution. Distributions that are more outlier-prone than the Gaussian distribution have  $K > 1$  whereas the distributions that are less outlier-prone have  $K < 1$ .

#### F. Probability Density (PD)

A cardiac artifact or EKG is associated with the sharp peak R component at minimum and maximum points, ICA method can extract the IC that related to this artifact. In this study, PD was proposed to measure the sharp peaks. In (5), probability densities are computed on the basis of histogram space. In this study, the optimal bin size of histogram that was chosen actually a theoretical measure from Scott's formula [21]:  $h = 3.5\sigma/\sqrt[3]{N}$  where  $h$  is the bin width,  $\sigma$  is the stand-

ard deviation and  $N$  is the number of data point length from the observed MEG dataset which can be defined by

$$PD = \frac{P(\mathbf{x} = x_{min}) + P(\mathbf{x} = x_{max})}{\max P(\mathbf{x})} \quad (5)$$

where  $P(\cdot)$  is a probability density at interval input vector  $\mathbf{x}$ ,  $x_{min}$  and  $x_{max}$  are the minimum and maximum ranges of IC, respectively. Fig. 4 shows the typical example of PD feature related to the cardiac artifactual IC.

#### G. Central Moment of Frequency (CMoF)

To measure the frequency response, fast Fourier transform (FFT) with a spectral resolution of 0.1Hz was applied. We propose the first order CMoF that can be determined by

$$CMoF = \frac{1}{M} \sum_{f=1.5Hz}^{40Hz} f P_n(f) \quad (6)$$

where  $P_n$  is a normalized power spectral density (PSD) which can be computed by  $\sum P_n(f) = 1$  and  $M$  is the number of frequency values. This feature is used to measure the most influent frequency response from the ICs such as EMG artifacts.

#### H. Spectral Entropy (SpecEn)

In order to quantify the flatness of the frequency spectrum, several studies have already applied SpecEn to analyze EEG and MEG [1],[13]. Therefore, SpecEn is computed to determine the spectral changes in time domain. SpecEn based on Shannon's entropy can be defined as

$$SpecEn = \frac{-1}{\log M} \sum_{f=0.5Hz}^{40Hz} P_n(f) \log [P_n(f)] \quad (7)$$

In (6) and (7),  $CMoF$  and  $SpecEn$  are then scaled from 0 to 1.

#### I. Fractal Dimension (FD)

This method is nonlinearly determined by the Hurst exponent,  $H$ , for dealing with complex systems. A method of estimating dimension has been widely used to describe complexity of a pattern or the quantity of information especially embodied in biological process as well as since it has been found useful for the analysis of physiological data especially EEG and MEG [27]–[29]. The basic idea of calculation is based on the power law relationship between the variance of the amplitude increments of the observation, which was produced by a dynamic process over time. In this study, we selected the variance fractal dimension (VFD) for estimating the

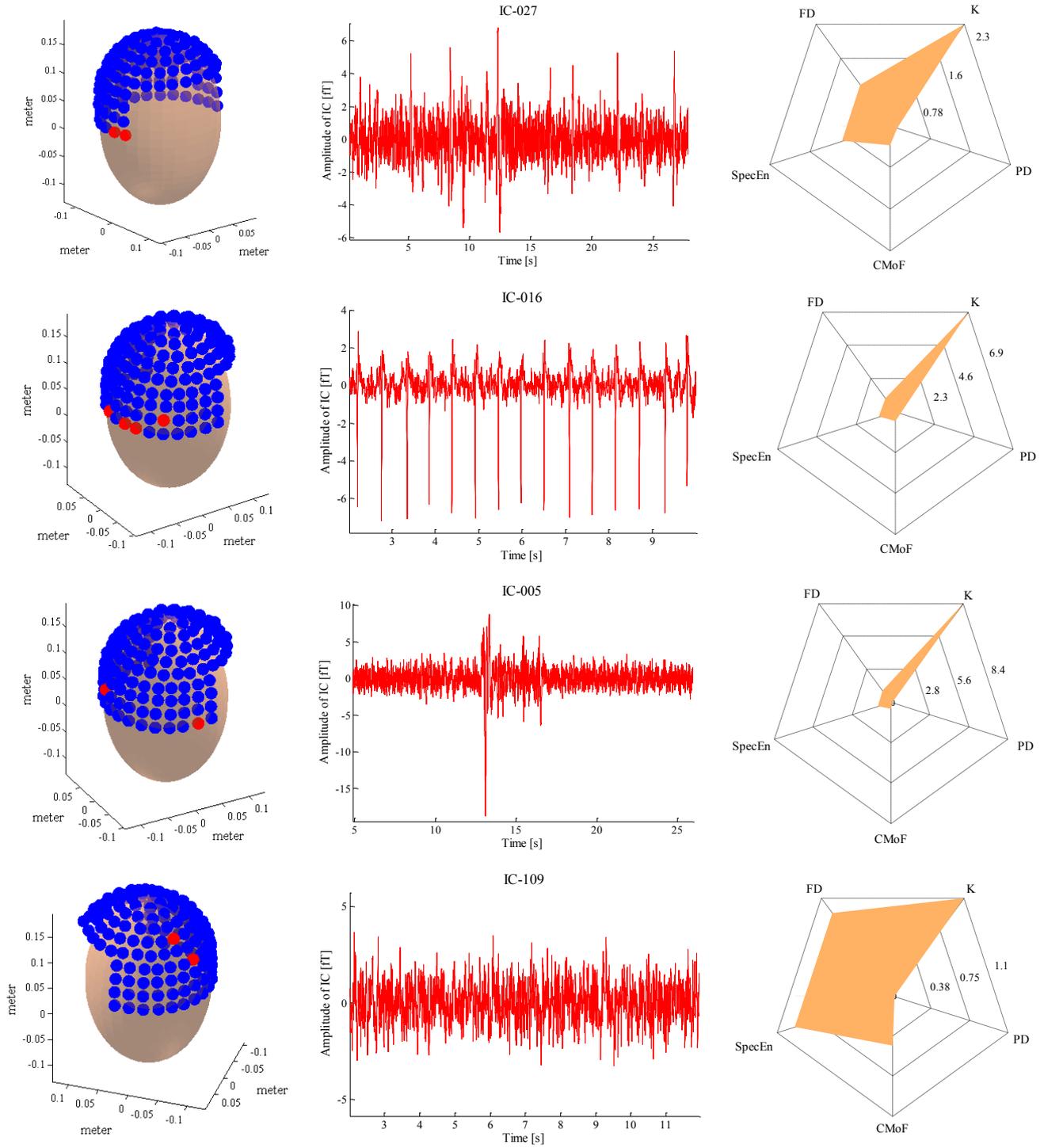


Fig. 5. Example of the processed MEG artifact identification based on the proposed features. (Left column) MEG source location by ICA method and FHR thresholding; (Middle column) IC-027, IC-016, IC-005 and IC-109 in time domain are identified as ocular artifact (EOG), cardiac artifact (EKG), sudden high-amplitude noise (HAM) and detected MEG component, respectively; (Right column) Radar plots of the normalized features corresponding to the ICs.

FD of MEG data since the main advantage of VFD was its support of the real-time computation [20] and has been used successfully in EEG classification [19]. The amplitude increments of a datum over a time interval  $\Delta t$  adhere to the following power law relationship  $\text{Var}[x(t_2 - t_1)] \propto |t_2 - t_1|^{2H}$ , the Hurst exponent can be calculated by using a log-log plot then given by

$$H = \lim_{\Delta t \rightarrow 0} \left( \frac{1}{2} \frac{\log_2(\text{Var}[(\Delta x)_{\Delta t}])}{\log_2(\Delta t)} \right) \quad (8)$$

The variance in each window per stage  $k$  can be calculated as follows

$$\text{Var}[\Delta x_{\Delta t}] = \frac{1}{(N_k - 1)} \left[ \sum_{j=1}^{N_k} (\Delta x)_{jk}^2 - \frac{1}{N_k} \left( \sum_{j=1}^{N_k} (\Delta x)_{jk} \right)^2 \right] \quad (9)$$

$$(\Delta x)_{jk} = x(jn_k) - x((j-1)n_k), \quad \text{for } j = 1, 2, \dots, N_k \quad (10)$$

The least-square linear fitted line corresponds to the slope of the plot  $\log_2(n_k)$  and  $\log_2(\text{Var}[\Delta x]_k)$ ,  $H = s/2$  where  $s$  is the obtained slope. Finally, FD can be estimated as

$$D = 2 - H \quad (11)$$

The process of calculating the FD essentially involves segmenting the entire input time series data into numerous subsequence (or window). The values  $k$  represents the integer range chosen such that each window of size  $N_T$  contains a number  $n_k = 2^k$  of smaller windows of size  $N_k = \lfloor N_T/n_k \rfloor$ . In the experiment, we set  $k = 1, \dots, \lfloor \log_2 N \rfloor - 2$ . Then, the normalized FD is given by  $FD = D/2$ .

### J. Artifact-free MEG Data Reconstruction

The output results of all features based on linear and nonlinear approaches were used to identify the actual ICs regarded as MEG data and the artifactual ICs regarded as noises. From (1) and (2), the identification was used to create the  $i$ -th component of the diagonal matrix  $\mathbf{A}$  as

$$\mathbf{x}_c(t) = \mathbf{x}(t) - \mathbf{W}^{-1} \mathbf{A} \hat{\mathbf{s}}(t) \quad (12)$$

Equation (12) is the artifact-free reconstruction in which the artifacts were clearly subtracted from MEG recordings,  $\mathbf{x}_c(t)$ , with the weight values of inverse separating matrix  $\mathbf{W}$ . The selection of each element of  $\mathbf{A}$  will be set equal to 1 if the component is identified as noise, and equal to 0 component is identified as MEG signal.

## III. EXPERIMENTS AND RESULTS

Raw 151-channel MEG data were processed by 50-trial ICA computations. As a priori knowledge, the identification at the first process was done by manual inspection. The results showed that this approach revealed different feature between MEG and other artifacts, which led a future useful identification method. All features present the different patterns individually with artifact contaminations as shown in Fig. 5. In this experiment, we tested all MEG data from ten subjects to show the performance and effectiveness of the proposed features. The obtained results are shown in Table I. Figs. 6 to 10 show the distinction of actual MEG signal and artifact, we determined the statistically significant results by using a one-way analysis of variance (ANOVA).

TABLE I. COMPARISON RESULTS OF THE AVERAGE VALUE AND STANDARD DEVIATION IN BRACKET FOR ALL FEATURES

| Component | K                  | PD                 | CMoF               | SpecEn             | FD                 |
|-----------|--------------------|--------------------|--------------------|--------------------|--------------------|
| EKG       | 5.8627<br>(3.0493) | 0.2516<br>(0.0465) | 0.4585<br>(0.0376) | 0.9568<br>(0.0129) | 0.9582<br>(0.0132) |
| EOG       | 1.3830<br>(0.4913) | 0.0205<br>(0.0279) | 0.4522<br>(0.0354) | 0.9468<br>(0.0146) | 0.9651<br>(0.0161) |
| HAM       | 6.3029<br>(4.5864) | 0.0484<br>(0.0389) | 0.4301<br>(0.0325) | 0.9382<br>(0.0115) | 0.9598<br>(0.0152) |
| MEG       | 1.2189<br>(0.1384) | 0.0148<br>(0.0053) | 0.4590<br>(0.0315) | 0.9494<br>(0.0126) | 0.9684<br>(0.0135) |

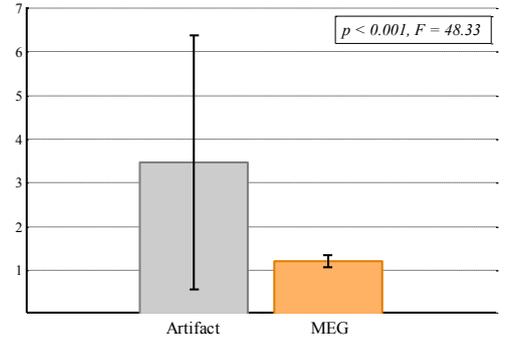


Fig. 6. Result of the normalized kurtosis, K.

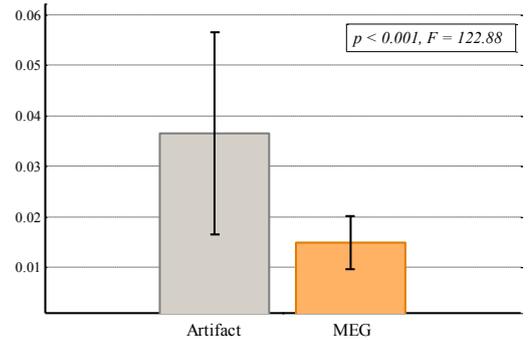


Fig. 7. Result of the probability density, PD.

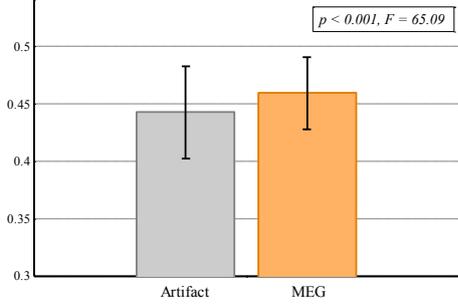


Fig. 8. Result of the central moment of frequency, CMoF

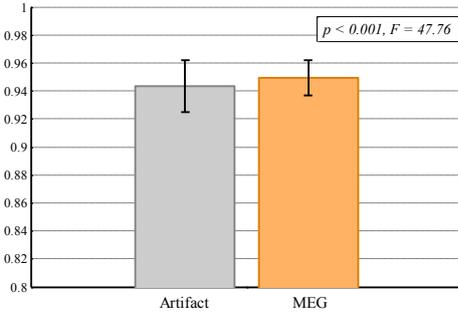


Fig. 9. Result of the spectral entropy, SpecEn.

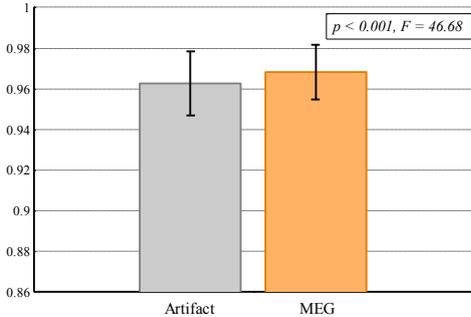


Fig. 10. Result of the normalized fractal dimension, FD.

#### IV. DISCUSSIONS

By using ICA method based on FastICA, the residual of ICs that out of the optimal condition is supposed to be white Gaussian noise (WGN) since FastICA requires pre-whitening and it is sensitive to WGN. In this experiment, approximately 10–20% of selected ICs based on the selection of setting parameters become WGN. These ICs were excluded automatically by the process of ICA method. By using ANOVA to determine statistical significant results between artifact and non-artifact (or supposed to be MEG), the results showed the all features present  $p < 0.001$ . This means we can effectively utilize them in identifying artifactual ICs from MEG record-

ing. Since the artifact-free MEG signal is not known, a universal filtering method for determining the success of automatic artifact removal in MEG data is needed. In our experimental results, the optimal performance of our proposed method depended on the choice of all features we need to use an appropriate thresholding for identifying MEG data and noise are as:

*FHR Thresholding*: in (3), we selected the twenty MEG channels to identify EOG artifact on the basis of the most EOG occurrence that intuitively comes from the MEG nearest positions to the eyes. In our experiment, the optimal threshold value to be 0.8 the reason for this was that in (2) normalized weight values in separating matrix presented at least 80% of amplitude which relatively correlated to the EOG artifactual component in time domain. This value was also used to identify the topography based MEG for source localization.

*K*: this parameter was used in FastICA, order and sorting  $K$ , the optimal condition we used the Bayesian model, i.e., an empirical Bayesian thresholding posterior densities of decision thresholds estimated if the normalized value in (4) is to be judged by WGN, then the  $K$  is one. This happens because in practice the HAM artifact is often occurred electrically by sensors. Empirically, we found that the optimal value of choosing  $K$  between  $1.2 < K_o < 1.8$  for the identification of MEG component since the actual MEG component in resting-state relatively closes to WGN. However, some ICs in the optimal range can be detected as artifact, i.e., EOG artifact, by FHR thresholding. In Fig.11, the increase of  $K$  presents the HAM artifacts that there was some  $K > 10$  while the IC within optimal range presents actual MEG component.

*PD*: we proposed this parameter to detect cardiac artifact (EKG). In the experiment, we found that the maximum PD was the optimal value to present EKG component. For the combination of all features, we found that the optimal condition is  $PD_o < 0.015$ . Figure 12 shows the example of PD based EKG signal detection from 10-channel MEG data and the artifact removal.

*CMoF*: this proposed parameter was used to measure the most influent frequency component due to muscular (EMG) artifacts or other environmental noise. If the normalized value in (6) was to be judged by perfect flatness of spectrum in frequency domain, then the  $CMoF$  is 0.5, we found that the optimal condition is  $0.45 < CMoF_o < 0.48$ .

*SpecEn*: this parameter was on the basis of frequency approach for quantifying the flatness of spectral response from MEG artifact because the EMG has typical spectra with the entire frequency range of interest [15]. We found that the optimal condition is  $0.94 < SpecEn_o < 0.96$ .

*FD*: the usefulness of this parameter is to measure irregularity of data especially for the MEG resting state. We found that the optimal condition is  $0.96 < FD_o < 0.97$ .

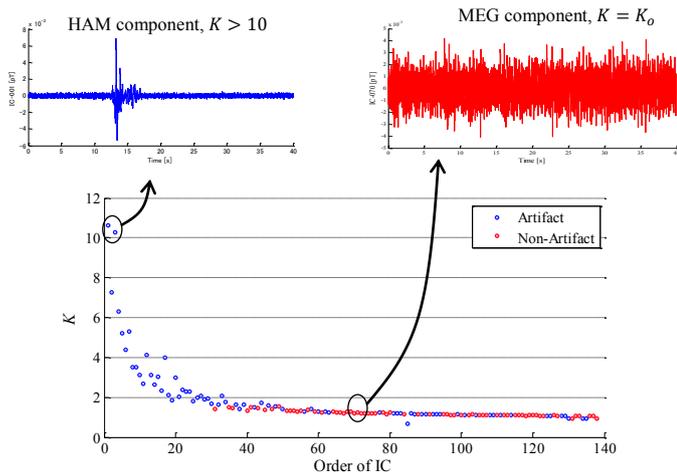


Fig. 11 Distribution of  $K$  values between artifact and non-artifact.

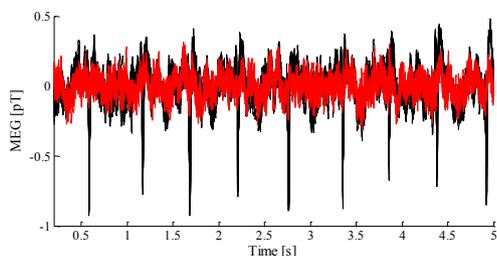


Fig. 12. The EKG artifact removal by using optimal PD. (Black line) Detected artifactual MEG from 10-channel; (Red line) Cleaned MEG.

A common increase of  $CMoF$ ,  $SpecEn$ , and  $FD$  were close to the WGN when recording the normal resting state of MEG data. Moreover, the combination of using all features made the automatic artifact removal become available. To confirm this, the result of artifact-free MEG data reconstruction by applying Equation (12) with the above suggested parameters is shown in Fig. 13. Based on the obtained results, these feature information could be used in categorizing artifactual ICs. In this paper, the EMG artifacts from motor movements could not be identified because the MEG recording was carried out for only resting state. However, in this case, the specific selection of ICs that regarded as muscular artifacts over motor area, i.e., premotor and sensorimotor cortexes, can be used as feature.

## V. CONCLUSIONS

We proposed the simple method for automatic artifact removal from MEG based on linear and nonlinear features including kurtosis, probability density, central moment of frequency, spectral entropy, and fractal dimension to identify artifactual ICs. The results showed that the proposed features could provide different patterns for reconstructing MEG data from noisy environment and could remove artifact effectively for potential MEG based applications. For the future work, we

plan to investigate MEG responses of children with their cognitive-behavioral skills and optimizing these features with other classification algorithms in detecting artifacts while preserving actual MEG data.

## ACKNOWLEDGMENT

This research was supported by the Japan Society for the Promotion of Science (JSPS) and the Hokuriku Innovation Cluster for Health Science (MEXT Program for Fostering Regional Innovation).

## REFERENCES

- [1] R. Hornero, J. Escudero, A. Fernández, J. Poza, and C. Gómez, "Spectral and nonlinear analysis of MEG background activity in patients with Alzheimer's disease," *IEEE Trans. Biomed. Eng.*, vol.55, no.6, pp.1658–1665, Jun 2008.
- [2] R.M. Yulmetyev and et al, "Dynamic effects and information quantifiers of statistical memory of MEG signals at photosensitive epilepsy," *Math. Biosci. Eng.*, vol.6, no.1, pp.189–206, 2009.
- [3] A. Kondo, H. Tsubomi, and K. Watanabe, "Inter-trial effect in luminance processing revealed by MEG," *The 33rd European Conf. Visual Perception*, Lausanne, Switzerland 2010.
- [4] R.M. Yulmetyev and et al, "Statistical memory of MEG signals at photosensitive epilepsy," *Int. J. Bifurcat. Chaos*, vol.18, no.9, pp.2799–2805, 2008.
- [5] M. Kikuchi and et al., "Lateralized theta wave connectivity and language performance in 2- and 5-year-old children," *J. Neurosci.*, vol.31, no.42, pp.14984–14988, Oct 2011.
- [6] A. Knoblauch and G. Palm, "What is signal and what is noise in the brain?," *BioSystems*, vol.79, pp.83–90, 2005.
- [7] R. Vigário and E. Oja, "BSS and ICA in neuroinformatics: From current practices to open challenges," *IEEE Rev. Biomed. Eng.*, vol.1, pp.50–61, 2008.
- [8] J.F. Gao, Y. Yang, P. Lin, P. Wang, and C.X. Zheng, "Automatic removal of eye-movement and blink artifacts from EEG signals," *Brain Topogr.*, vol.23, pp.105–114, 2010.
- [9] F. Bießmann, S. Plis, F.C. Meinecke, T. Eichele, and K.-R. Müller, "Analysis of multimodal neuroimaging data," *IEEE Rev. Biomed. Eng.*, vol.4, pp.26–58, 2011.
- [10] G. Bartels, L.C. Shi, and B.L. Lu, "Automatic artifact removal from EEG - A mixed approach based on double blind source separation and support vector machine," *The 32nd Int. Conf. IEEE EMBS*, pp.5383–5386, Sep 2010.
- [11] R. Vigário, J. Sarela, V. Jousmaki, M. Hamalainen, and E. Oja, "Independent component approach to the analysis of EEG and MEG recordings," *IEEE Trans. Biomed. Eng.*, vol.47, no.5, pp.587–593, 2000.
- [12] J.W. Kelly, D.P. Siewiorek, A. Smailagic, J.L. Collinger, D.J. Weber, and W. Wang, "Fully Automated Reduction of Ocular Artifacts in High-Dimensional Neural Data," *IEEE Trans. Biomed. Eng.*, vol.58, no.3, pp.598–606, Mar 2011.
- [13] D. Mantini, R. Franciotti, G.L. Romani, and V. Pizzella, "Improving MEG source localizations: An automated method for complete artifact removal based on independent component analysis," *NeuroImage*, vol.40, pp.160–173, 2008.
- [14] A. Ziehe, G. Nolte, T. Sander, K.R. Müller, and G. Curio, "A comparison of ICA-based artifact reduction methods for MEG," *The 12th Int. Conf. Recent Adv. Biomagnetism*, Helsinki University of Technology, 2001.
- [15] F. Rong and J.L. Contreras-Vidal, "Magnetoencephalographic artifact identification and automatic removal based on independent component analysis and categorization approaches," *J. Neurosci. Methods*, vol.157, pp.337–354, 2006.

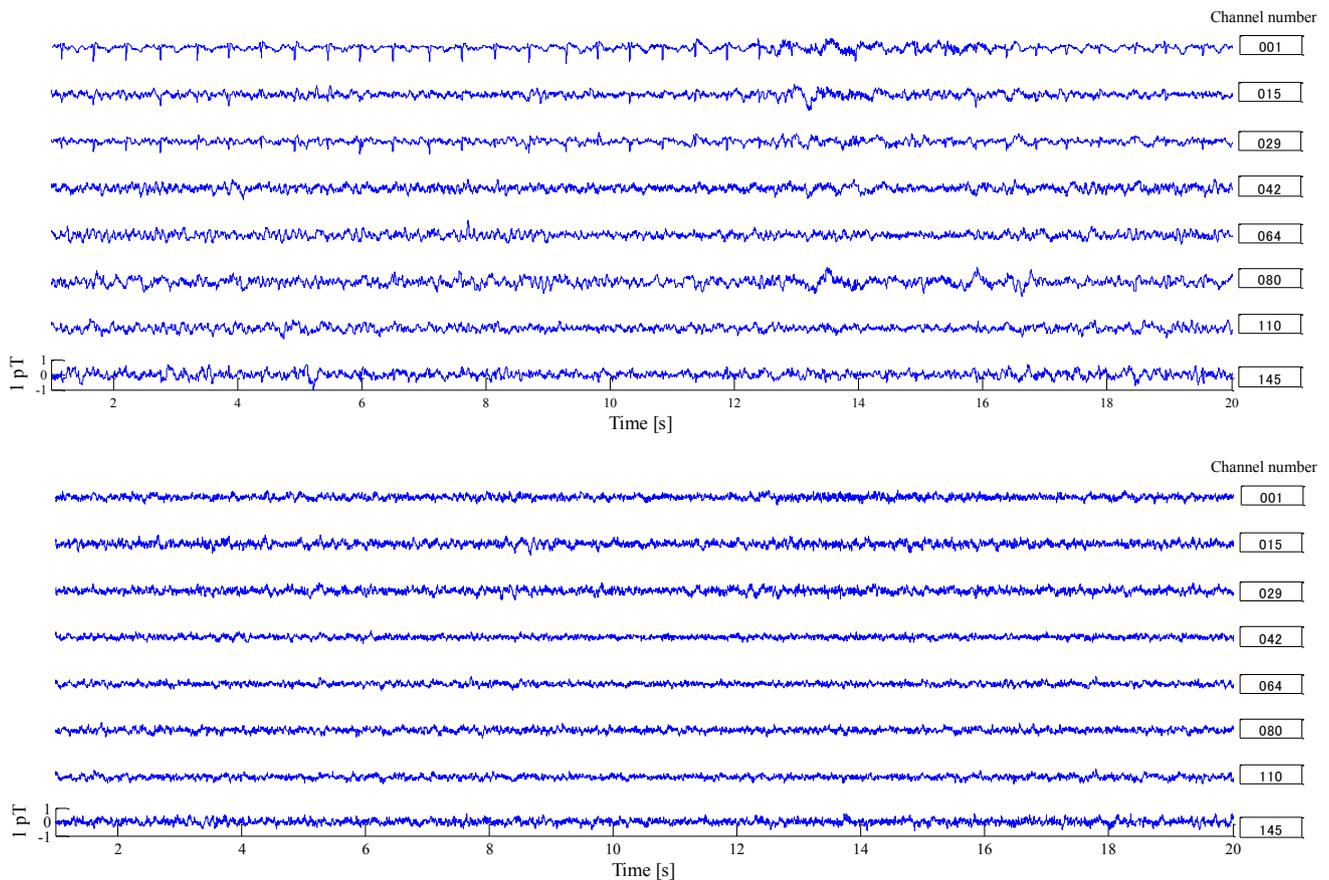


Fig. 13. Result of before and after automatic artifact removal by the proposed method. (Top) Raw MEG data with artifacts; (Bottom) Cleaned MEG data.

- [16] J. Escudero, R. Hornero, D. Abásolo, A. Fernández, and M. López-Coronado, "Artifact removal in magnetoencephalogram background activity with independent component analysis," *IEEE Trans. Biomed. Eng.*, vol.54, no.11, pp.1965–1973, Nov 2007.
- [17] Y.-P. Liao, and et al., "The automatic detection of artifacts in magnetoencephalographic signals based on Fast ICA," *The 4th Int. Conf. Biomed. Eng. & Informatics*, pp.862–866, 2011.
- [18] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Comput.*, vol.9, no.7, pp.1483–1492, 1997.
- [19] M. Phothisonothai and M. Nakagawa, "EEG-based classification of motor imagery tasks using fractal dimension and neural network for brain-computer interface," *IEICE Trans. Info. & Syst.*, vol.E91–D, no.1, pp.44–53, 2008.
- [20] W. Kinsner, "Batch and real-time computation of a fractal dimension based on variance of a time series," Univ. Manitoba Canada, *Tech. Report*, del94–6, 1994.
- [21] D.W. Scott, "On optimal and data-based histograms," *Biometrika*, vol.66, no.3, pp.605–610, 1979.
- [22] S. Ikeda and K. Toyama, "Independent component analysis for noisy data-MEG data analysis," *Neural Network*, vol.13, pp.1063–1074, 2000.
- [25] L.L. DeCarlo, "On the meaning and use of kurtosis," *Psy. Methods*, vol.2, no.3, pp.292–307, 1997.
- [26] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Netw.*, vol.10, no.3, pp.626–634, May 1999.
- [27] C. Gómez, A. Mediavilla, R. Hornero, D. Abásolo, and A. Fernández, "Use of the Higuchi's fractal dimension for the analysis of MEG recordings from Alzheimer's disease patients," *Med. Eng. Phys.*, vol.31, no.3, pp.306–13, Apr. 2009.
- [28] C.J. Stam, "Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field," *Clinical Neurophysiology*, vol.116, pp.2266–2301, 2005.
- [29] M. Phothisonothai and M. Nakagawa, "A classification method of different motor imagery tasks based on fractal features for brain-machine interface," *J. Integr. Neurosci.*, vol.8, no.1, pp.95–122, 2009.