Removal of artifacts and fluctuations from MEG data by clustering methods

Toru Kumagai*, Akio Utsugi

Research Institute of Human Science and Biomedical Engineering, National Institute of Advanced Industrial Science and Technology (AIST), AIST Tsukuba Central 6-11, 1-1 Higashi 1-Chome, Tsukuba, Ibaraki 305-8566, Japan

Received 1 September 2001; received in revised form 1 July 2002; accepted 1 August 2002

Abstract

We proposed to cluster a set of magnetoencephalogram (MEG) records using mixture of factor analyzers to remove those outlying records that were contaminated with artifacts. We showed the effectiveness of the proposed clustering approach by applying it to visual and auditory evoked MEG data.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Magnetoencephalography; Electroencephalography; Artifact rejection; Clustering; Mixture of factor analyzers; k-means; Visual evoked field; Auditory evoked field

1. Introduction

Artifact removal is an important step in analysis of sensory evoked magnetoencephalograms (MEGs) and electroencephalograms (EEGs) since it requires triggered averaging to eliminate spontaneous waves from raw signal data. Many studies have specifically addressed artifact removal [1,3].

The rejection technique is a common and intuitive one. In this method, visual inspection or simple automatic detection criteria identify trials in which artifacts occur. This method’s power is restricted by the difficulty of selecting detection criteria [12]. For instance, simple threshold criteria cannot exclude artifacts that are smaller than basic rhythm activity.

* Corresponding author. Tel.: +81-298-61-6674; fax: +81-298-61-6634.
E-mail address: kumagai.toru@aist.go.jp (T. Kumagai).

0925-2312/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
The correction technique, which eliminates the detection criteria problem, has been studied widely [4,11,13,15,16]. This technique estimates artifacts on each channel. An uncontaminated EEG or MEG is approximated by subtraction. Various estimation techniques have been proposed mainly for ocular artifacts and cardiac artifacts. The capability of the correction technique is limited by the accuracy of artifact estimation. Naturally, it cannot remove unknown artifacts.

Recently, the separation technique, based on independent component analysis (ICA), has received attention [6,7,9,10,14]. ICA is a method for solving the blind source separation problem. After statistical separation of independent components by ICA, an artifact-free MEG or EEG is obtained by projecting artifact-free components. The separation technique also has limitations. It may be difficult to know which component contains artifacts. Some separated components may contain both artifactitious components and encephalographic signals. It requires sufficiently numerous and smooth time-series data.

This study examines a different approach to remove artifacts. We propose to cluster a set of MEG or EEG records to remove artifact-contaminated records. We compared two clustering methods: mixture of factor analyzers (MFA) [5,8] and \( k \)-means [2]. After clustering, we estimated brain activity for a gravity center of each class by an equicurrent dipole method; subsequently, we evaluated the clustering result by the estimation error. This criterion is proper because all contamination and fluctuations increase the estimation error. The estimation error is minimized within the limitation of the equicurrent dipole method if all artifact contamination and fluctuations are removed.

Intuitively, the clustering approach provides some advantages over other techniques. It has a wide range of the application and can be applied easily because it does not require detection criteria, as does the rejection technique; in addition, it does not require a model of artifacts, as does the correction technique. Furthermore, it requires time point data, not the necessary time-series data of the separation technique. On the other hand, the rejection technique and the correction technique are applicable online, but the clustering approach is not. The separation technique uses time series information, but the clustering approach does not. The clustering approach and the other techniques are complementary.

2. Clustering method

MFA provides a clustering method that groups records into classes that are described by the factor analysis model [4]. The data generation model of factor analysis is written as

\[
x = \Lambda z + u,
\]

where \( x \) is \( D \times 1 \) observed data, \( \Lambda \) is a \( D \times K \) factor loading matrix, and \( u \) is Gaussian noise according to \( N(0, \Psi) \), which is called the unique factor, where \( \Psi \) is a diagonal matrix. In addition, \( z \) is a \( K \times 1 \) matrix called the common factor. It is assumed to be
We can write the data generation model of mixture of factor analyzers as

\[ P(x) = \sum_{j=1}^{M} \left\{ P(w_j) \int P(x | z, w_j) P(z | w_j) dz \right\}, \]

where \( \mu_j \) is the average value of \( x \) for sub-model \( j \). Also, \( w_j \) is an event in which data is generated by sub-model \( j \). The learning algorithm of a mixture of factor analyzers, which yields, iteratively, maximum likelihood estimators of \( \pi, \mu, A \). Finally, \( \Psi \) is derived using an EM algorithm. It is known that MFA is superior to conventional methods in its ability to discriminate in a noisy environment [5,8]. It should be ideal for clustering of evoked responses upon which large spontaneous waves are superimposed.

A conventional clustering method that is intended to minimize the sum-of-squares of distances within classes is \( k \)-means [2]. This study denotes the class number of \( k \)-means as \( K_m \).

3. Clustering of visual evoked fields

In this section, we apply clustering methods to visually evoked field (VEF) records contaminated with eye movement artifacts. We presumed that VEF records would be classifiable into two major groups: those contaminated with ocular artifacts, and others.

3.1. Measurements of VEF

We measured VEF elicited by luminance onset stimuli (120 trials) using a 64 channel MEG system (CTF Systems, Inc.). Each signal channel was filtered with a low-pass filter (40 Hz). Initial noise reduction was performed using software provided by CTF Systems, Inc. The simple average of VEF had the first remarkable peak at a latency of 75 ms. Fig. 1 shows its topography (\( \mu_a \)). Magnetic sources were estimated

![Fig. 1. An isocontour map of the averaged data. White areas indicate a source of magnetic flux. Gray areas indicate a sink of magnetic flux. One contour line equals 10 fT.](image-url)
in the bilateral occipital lobe. The estimation error was 27%. This component was presumed to be evoked by visual stimuli. However, the large error suggested that some records were contaminated. In fact, we confirmed that some records were contaminated with extraneous eye movement artifacts.

3.2. Clustering through MFA and k-means

We clustered the VEF records at latencies of 75 and 78 ms into four classes using MFA with $K=3$ and $M=4$. We then reanalyzed these VEF records using $k$-means with $K_m=12$. Parameters of MFA and $k$-means were determined to obtain the best clustering result by preliminary analyses. We repeated the analysis ten times using different initial values. After MFA and $k$-means, we estimated the magnetic sources for a gravity center of each class using the equicurrent dipole method for evaluating the analysis results. In many cases, the class with the smallest estimation error among classes of each analysis was presumed to model visual evoked responses. Their magnetic sources were located in the bilateral occipital lobe. Fig. 2 shows the estimation error of the simple average and that of the class with the smallest estimation error among classes of each MFA and $k$-means for the five best results. The MFA and $k$-means methods are markedly superior to the simple average in terms of estimation error: MFA also appears to be superior to the $k$-means method.

3.3. Meaning of MFA parameters

In this section, we examine MFA parameters and confirm validity of the clustering results. Fig. 3 shows acquired MFA parameters. The average of class 4 ($\mu_4$) had similar topography to the simple average ($\mu_s$). Its magnetic sources were located in the occipital lobe. The estimation error was only 11%, which was much less than that of the simple average. The averages of class 1 and 2 ($\mu_1$ and $\mu_2$) showed different topography from the simple average. These waves presumably reflect spontaneous waves because their topography is similar to that seen in a single trial record for this subject, without regard to latency. The average of class 3 ($\mu_3$) showed large
activity in the anterior area, which seems to result from eye movements. These results suggested that MFA indicated clusters of VEF records that were: artifact-contaminated records, records which contained mainly spontaneous waves, and records which contained mainly evoked responses.

Other MFA model parameters also supported that MFA worked effectively. The variance of the common factor had a peak value in the anterior area. This value seemed to reflect artifacts such as eye movements. Factor loading was only shown for the fourth class. Here, we denote the factor loading of class 4 as $A_4 = [A_{41} A_{42} A_{43}]$. $A_{41}$ and $A_{42}$ show symmetrical topography. This result is very plausible because many brain activities and eye movement artifacts show a symmetrical structure. $A_{43}$ was presumed to reflect spontaneous waves as well as $A_1$ and $A_2$.

4. Clustering of auditory evoked MEG data

In the previous section, we clustered VEF records which were highly variable because of eye movements. In this section, we attempted to apply this method to fine auditory evoked field (AEF) records that were not contaminated with distinct artifacts.

---

**Fig. 3.** An isocontour map of the averaged data. White areas indicate a source of magnetic flux. Gray areas indicate a sink of magnetic flux. One contour line equals 20 fT.
4.1. Measurements of AEF

We measured AEF elicited by 1 kHz tone burst stimuli to the right ear (128 trials) using a 64 channel MEG system (CTF Systems, Inc.). Each signal channel was filtered with a low-pass filter of 40 Hz, and noise reduction software (CTF Systems, Inc.) was applied. The simple average of VEF showed the primary peak at latency of 116.8 ms. Its magnetic sources were estimated to be in the bilateral primary auditory areas. The estimation error was 11%. This component was presumably evoked by auditory stimuli.

4.2. Clustering by MFA and k-means

We clustered the AEF records into four groups at latencies of 116 and 116.8 ms by MFA with $K = 1$ and $M = 4$. We also analyzed the data using $k$-means with $K_m = 4$. The parameters of MFA and $k$-means were determined to obtain the best clustering result by preliminary experiments. We repeated the analysis ten times using different initial values of MFA and $k$-means. Most analysis results indicated that the class with the smallest estimation error was presumed to model auditory evoked responses. Those magnetic sources were located in around the primary auditory area. Fig. 4 shows the estimation error of the simple average and that of the class with the smallest estimation error in each MFA and $k$-means analysis for the best five results. MFA and $k$-means are markedly superior to the simple average in the estimation error. No significant difference was shown between results of the MFA and $k$-means methods.

Fig. 5 shows average values of classes acquired using MFA. Their topographies were similar to one another. The average of class 2 ($\mu_2$) had the smallest estimation error of the four classes. One likely cause of the diversity of response patterns was the phase of spontaneous waves: the records were clustered according to the phase of spontaneous waves. Fluctuations of subject responses and contamination of artifacts could have also explained some of the variability.
5. Discussion

The clustering approach worked well not only for the VEF data that was heavily contaminated, but also for the fine AEF data. This means that the clustering approach works for fluctuations and minor artifacts, which are not treatable by the rejection and correction technique. The clustering approach may also work to cluster different responses of subjects in event-related experiments such as an oddball paradigm. Clustering of responses during long latency may also be a good application of the clustering approach, but it does not utilize time series information.

For clustering of the VEF records that were contaminated with distinct artifacts, MFA was superior to $k$-means in its estimation error. MFA allows the maximum likelihood estimation of $\exp(2)$, where $k$-means partitions data into clusters to thereby minimize the sum of the squared distances to the cluster centers. That is, the clusters obtained by MFA have a structure provided by factor analysis, but those of $k$-means have no structure. Presumably, this character of MFA enhances its performance for VEF data that were a combination of VEF, ocular artifacts, and spontaneous waves. This superior performance is plausible because it is known that MFA is superior to $k$-NN and PCA for recognition tasks of noisy samples [8]. In contrast, no significant difference in the clustering of the fine AEF records was shown between MFA and $k$-means. This equivalence suggests that MFA offers no advantage over $k$-means for data without a distinct structure. However, MFA still offers the advantage of revealing the data structure.

This study demonstrates that the clustering approach works well and that MFA is better than $k$-means for clustering of artifact-contaminated records. The clustering approach could serve as a preprocessing technique for another technique. Moreover, ICA could be a good preprocessing technique for the clustering approach.

References


Toru Kumagai joined Industrial Products Research Institute in 1991. He had been a researcher of National Institute of Bioscience and Human-technology since 1993. He was a planning officer of National Institute of Advanced Industrial Science and Technology (AIST), planning office in 2001 and has been a senior researcher of Institute for Human Science and Biomedical Engineering, AIST since 2002. His primary areas of interest are neural networks and nonlinear information processing. He received his BE and ME in Mechanical Control Engineering from University of Electronic Communications in 1989 and 1991. He received his Ph.D. in Systems and Information Engineering from Hokkaido University in 1997.

Akio Utsugi received the M.E. degree in biophysical engineering from Osaka University, Japan in 1986. Then he joined Industrial Products Research Institute (IPRI), Ministry of International Trade and Industry (MITI), Japan. He is currently working at National Institute of Advanced Industrial Science and Technology (AIST). His research interests include statistical modeling of cognition and its engineering applications.