An Eye Blinking Artifact Rejection Method Using Single-Channel Electroencephalographic Signals and 2 Step Non-Negative Matrix Factorization

Suguru Kanoga¹ and Yasue Mitsukura²

Abstract— Many biological signal analyses have been proceeded in various research fields. In particular, electroencephalographic (EEG) signal analyses have attracted attention because the EEG signal includes a mixture of endogenous brain activities. However, the EEG signal is often mixed physiological artifacts such as the eye blinking, ocular motion, heart beat, or muscle activity. Specifically, humans are unable to keep gazing at something without the eye blinking. In other words, the eye blinking artifacts absolutely invade the EEG signals while a subject wears an EEG device. Therefore, it is important to remove the artifacts from the EEG signals when researchers attempt to analyze the brain activities accurately. An independent component analysis (ICA) has been used for removing the eye blinking artifacts effectively with over 90% of the reconstruction for original EEG signals. However, a drawback of the ICA is that this method can only manage overdetermined mixtures, which entails many EEG electrodes. There is no numerical approach of eye blinking artifact rejection method for single-channel EEG signals. Therefore, in this paper, we propose an eye blinking artifact rejection method using single-channel EEG signals and 2 step non-negative matrix factorization (NMF). We acquired 14 EEG and 1 vertical electrooculographic signals from 10 subjects who blink every 5 seconds. Furthermore, we performed the proposed method to reject the eye blinking artifacts using single-channel EEG (Fp1) signals and the ICA using multi-channel EEG signals to prepare the target for comparison. High SNR between reconstructed signals by the ICA and the proposed method was represented. Moreover, over 99% of the reconstruction for original EEG signals was showed by the proposed method. Therefore, we confirmed the validity of the proposed method for the eye blinking artifact rejection method using only single-channel EEG signals.

I. INTRODUCTION

Many biological signal analyses have been proceeded in various research fields. For example, the rehabilitation [1][2], neuromarketing [3], and behavioral analysis [4]. In particular, electroencephalographic (EEG) signal analyses have attracted attention because the EEG signal includes a mixture of endogenous brain activities.

The EEG signal measurement devices are divided roughly into the following two types. One is a cap type device having multi-channel electrodes. This device measures brain activities extensively, therefore EEG signal researchers use this device to investigate entire brain activities. The other is a head band type device having a single-channel electrode. Though this device measures limited brain activities, it can be worn easily and give a less oppression to the wearer than the cap type device.

Physiological artifacts often mix in the EEG signals when the EEG signal researchers measure the EEG signals with either EEG device. For example, the eye blinking, ocular motion, heart beat, and muscle activity. Such artifacts invade the EEG signal and make the EEG analysis difficult because the EEG energy is generally lower than the artifact energy [5]. Therefore, these artifacts rejection from the EEG signal is important when the EEG signal researchers attempt to analyze the brain activities accurately. Specifically, humans are unable to keep gazing at something without eye blinking. In other words, the eye blinking artifacts absolutely invade the EEG signals while a subject wears the EEG device.

An independent component analysis (ICA) is the most popular method for the eye blinking artifact rejection [6][7]. The ICA has been confirmed that it can remove the eye blinking artifact effectively, and its reconstruction is over 90% of original EEG signal. The major constraint of the ICA is that it can only manage overdetermined mixtures. In other words, this method entails preparing electrodes at least as many as the sources of artifacts plus one for getting the meaningful information. Therefore, the ICA is not available for single-channel EEG signal analyses.

There is no numerical approach of eye blinking artifact rejection method for single-channel EEG signals. It did not paid much attention because most of EEG signal analyses used multi-channel EEG device until last decade. However, the single-channel EEG device has come to draw attention on account of its usability for measurement and portability. Moreover, the single-channel EEG signal is received focus as real-time processing for smartphone and tablet computer nowadays. With this change, an eye blinking artifact rejection method using only single-channel EEG signals is required.

Research concerning an eye blinking artifact removal method using single-channel EEG signals and non-negative matrix factorization (NMF) had been reported [8]. However, the method was not formulated the accuracy.

Therefore, in this paper, we propose an eye blinking artifact rejection method using single-channel EEG signals and 2 step NMF. We acquire 14 EEG and 1 vertical electrooculographic (EOG) signals from 10 subjects who blink every 5 seconds. Furthermore, we perform the proposed method to reject the eye blinking artifacts using single-
channel EEG (Fp1) signals and the ICA using multi-channel EEG signals to prepare the target for comparison. The results of our proposed method are compared with the results of ICA to detect appropriate basis number of NMF for the artifact rejection. Moreover, the reconstruction signals of the proposed method are compared with the original Fp1 signals.

II. EXPERIMENTS

A. Electroencephalographic and Vertical Electrooculographic Signal Measurements

In this paper, we used a cap type device, g.tec having multi-channel electrodes because we need to compare the proposed method with the ICA. The cap type device measured EEG and vertical EOG signals with the sampling rate of 256Hz.

The EEG signals were measured from Fp1, Fp2, F3, Fz, F4, T3, C3, C4, T4, P3, Pz, P4, O1, and O2 positioned referring to the international 10-20 system. The vertical EOG signals were measured as finite difference from upper and lower the right eye by using two disposable electrodes. The reference and ground electrodes were put at A1 and Fpz, respectively.

In the ICA, these 14 EEG and 1 vertical EOG signals were used for the eye blinking artifact rejection. In the proposed method, only Fp1 signals were used for the eye blinking artifact rejection because the head band type EEG device can measure only this position.

B. Experimental Condition

8 males and 2 females aged 21-27 years old were participated in the experiments. Each subject was asked to sit on a chair and blink every 5 seconds. The task was continued until the blink number approaches 30 times. The subjects received an explanation of informed consent and permitted it prior to their participations.

III. ARTIFACT REJECTION METHODS FOR EYE BLINKING

A. Independent Component Analysis

An independent component analysis (ICA) is a major method for the eye blinking artifact rejection using the multi-channel EEG and vertical EOG signals [6][7]. This method is based on spatial filtering and does not rely on having a clean reference channel. The multi-channel EEG signals are decomposed temporally independent components effectively. The independent component which has the highest correlation with the vertical EOG signal in all the independent components, will be removed as the eye blinking artifact. In this paper, we dealt with Fp1 signals after applying the ICA as target signals.

B. Non-Negative Matrix Factorization

A non-negative matrix factorization (NMF) is a multivariable analysis method. It can additively factorize non-negative matrix such as amplitude spectrum and power spectrum to two non-negative matrices [9]-[11]. The NMF have been used in the various research fields. For example, the automatic transcription [12], sound emphasis or separation [13], and band spreading [14]. About EEG signal analysis, this method has been used for EEG feature extraction to separate classification [15]. However, it is not used as artifact rejection method until recent years.

A multivariate $M$-dimensional non-negative data vector $y_n$ is placed in the columns of $M \times N$ matrix $Y$ where $N$ is a number of data vectors in the dataset and is called an observation vector. The matrix $Y$ is approximately factorized into $M \times K$ matrix $H$ and $K \times N$ matrix $W$ where $K$ is a number of “basis” that is optimized for linear approximation of the data in the matrix $Y$. It can be rewritten by (1),

$$y_n = \sum_{k=1}^{K} h_k w_{k,n} \quad (n = 1, ..., N) ,$$

where $h_k$ and $w_{k,n}$ mean an entry of $H$ and $W$, respectively. In other words, respective Fp1 signal vector $y_n$ is approximated by a linear combination of the basis $h_k$, weighted by the components of $w_{k,n}$. Therefore, it can be rewritten by following equation,

$$Y \simeq HW .$$

For finding an approximate factorization, we first need to define iterative algorithms that quantify the quality of approximation. Such an iterative algorithm can be constructed using some measure of distance between two non-negative matrices $Y$ and $X$. The algorithm is not called “distance”, because it is not symmetric in $Y$ and $X$, therefore the algorithm is referred to as the ‘divergence’ of $Y$ from $X$ [10]. Currently, there are various kinds of distance and divergence, for example, the Euclidean (EU) distance, Kullback-Leibler (KL) divergence, and Itakura-Saito (IS) divergence. These algorithms can be written as

$$D_{\text{EU}}(H, W) = \sum_{m,n} D_{\text{EU}}(y_{m,n} | h_{m,k} w_{k,n}) ,$$

where $(,)$ indicates respective algorithms, EU, KL, and IS. In this paper, we defined the IS divergence for the iterative algorithm because it minimizes the objective function $D$ with the generating process of signals. The IS divergence repeats the following multiplicative update rules.

$$h_{m,k} \leftarrow h_{m,k} \left( \frac{\sum_{n} y_{m,n} w_{k,n} x_{m,n}^2}{\sum_{n} w_{k,n} x_{m,n}} \right)^{1/2}$$

$$w_{k,n} \leftarrow w_{k,n} \left( \frac{\sum_{m} y_{m,n} h_{m,k} x_{m,n}^2}{\sum_{m} h_{m,k} x_{m,n}} \right)^{1/2} ,$$

where

$$x_{m,n} = \sum_{k} h_{m,k} w_{k,n} .$$

If the basis vectors $H$ find structure that is latent in the data, the dimension of $H$ is smaller than the dimension of $Y$, and it can be said that good approximation was achieved. It is known that the basis number $K$ should be determined
less than half of $M$ when we use the NMF. However, the appropriate basis number is unknown in the case of EEG signal analysis, therefore we attempt to detect it using the proposed method.

C. Preparing Datasets

We acquired 14 EEG and 1 vertical EOG signals on a subject from the experiments. Each length of EEG signal is 155 seconds (multiply 30 trials by 5 seconds plus margin).

Firstly, we performed the ICA using acquired all signals. We acquired reconstructed Fp1 and removed Fp1 signals by the ICA. The reconstructed Fp1 signal indicates that the signal was reconstructed from original Fp1 signal without the component which has the highest correlation with the vertical EOG signal. The removed Fp1 signal indicates that the signal was removed from original Fp1 signal as the eye blinking artifact.

Secondly, the original, reconstructed, removed Fp1 signals, and vertical EOG signals were separated into 30 signals respectively. Each signal has the peak of amplitude at 2.25 second (the 576th sampling point). The amplitude variation was raised by the eye blinking. These 4 types of signals are shown in Fig. 1. After this processing, we acquired 120 signals (multiply 30 signals by 4 types) whose length of signal is 5 seconds (1280 sampling points). These signals are used to compare the accuracy of time series.

Thirdly, we performed a short-time Fourier transform (STFT) to the measured entire original Fp1 signal with 256 sampling point hamming window and 128 sampling point shifting. Furthermore, amplitude spectra were calculated. Since we acquired 5 seconds of original Fp1 signals, 330 amplitude spectrum data (multiply 30 signals by 11 windows) were calculated. The data included folding noise, therefore, they were modified to 129-dimensional amplitude spectrum data. The 129-dimentional data were required for reconstruction of the original Fp1 signal because the first and 129th components in the data are unique components.

In this paper, we defined that the data was mixed the eye blinking artifact if the value of amplitude spectrum under 10Hz was over 2000. For this definition, we noticed that each fifth, sixth, and seventh original amplitude spectrum data are mixed the eye blinking artifact in the low frequency region (see Fig. 2). Such amplitude spectrum data were separated from original amplitude spectrum data.

After these processing, we acquired 2 types of amplitude spectra. In other words, the original amplitude spectrum data was separated 90 and 240 129-dimensional data.

D. Proposed Method

In this paper, we propose 2 step NMF for the eye blinking artifact rejection using the single-channel EEG signals. The outline of the proposed method is shown in Fig. 3.

Original amplitude spectrum data not including the eye blinking artifacts (blue windows on Fig. 2) are used to decompose into $H$ and $W$ in the second step as $Y_2$. We defined these matrices as $H_{2nd}$ and $W_{2nd}$. The elements of matrix $H_{1st}$ has no relation to the elements of matrix $H_{2nd}$ by the multiplicative update rules.

In this paper, the matrix $H_{1st}$ was used as fixed value in the second step. For this constraint, the matrix $H_{2nd}$ attempts to express the matrix $Y_2$ using the rest of bases $(K_2)$. Therefore, the artifacts mixing with the original Fp1 signals are stored in the rest of bases. In the both steps, we performed the NMF with the basis number $K_1$ and $K_2$ whose length is ranged between 2 and 64 respectively, because the sampling rate was set as 256Hz.

Furthermore, we used the following equation for recon-
Fig. 3. The outline of the proposed method for the eye blinking artifact rejection using the single-channel EEG signals and the 2 step NMF.

struction of amplitude spectrum data vector.

Reconstructed spectra

\[ Y = \frac{\sum_{k=1}^{K_1} H_{1st} k W_{2nd} k}{\sum_{k'=1}^{K_1+K_2} H_{2nd} k' W_{2nd} k'} \]  

(7)

Estimated eye blinking artifact spectra

\[ Y = Y_2 - \text{Reconstructed spectra} \]  

(8)

By using (7) and (8), we acquire reconstructed and estimated eye blinking artifact amplitude spectra about fifth, sixth, and seventh windows. The original fifth, sixth, and seventh window’s amplitude spectra are replaced the reconstructed amplitude spectra, and they are transformed into time series signal by using inverse Fourier transformation. The estimated eye blinking artifact amplitude spectra are also transformed into time series signal. However, the windows except fifth, sixth, and seventh are substituted by zero. The estimated eye blinking artifacts by proposed method are compared with the removed data by the ICA to detect the appropriate basis number, \( K_1 \) and \( K_2 \) for the eye blinking artifact rejection.

The reconstructed signal by the ICA has been complemented by signals obtained from other electrodes. However, the matrix \( H_{2nd} \) and \( W_{2nd} \) were only based on the original Fp1 amplitude spectra, therefore the reconstructed signal by the proposed method will be similar to the original Fp1 signal. In other words, the phase of reconstructed signal by the ICA and the phase of reconstructed signal by the proposed method are a difference absolutely. Thereby, we only compared the removed data.

The comparative method is signal-to-noise ratio (SNR).

\[ SNR = 10 \log_{10} \frac{S}{N} \]  

(9)

where \( S \) is the variance of the removed signal by the ICA and \( N \) is the variance of the estimated eye blinking artifact signal by the proposed method. We calculated 4096 SNRs (multiply 64 bases by 64 bases) for each signal. Therefore, we calculated 1228800 SNRs (multiply 4096 patterns by 30 signals by 10 subjects) to detect the appropriate basis number for the eye blinking artifact rejection.

IV. RESULTS AND DISCUSSIONS

A. Comparison of SNRs

The result of average SNRs is shown in Fig. 4. From this figure, we have to note that the SNR is high when a low first basis number \( K_1 \) and a large second basis number \( K_2 \) were selected.

In the NMF, good approximation can only be achieved when the basis vectors found structure that is latent in the EEG data. The matrix \( H_{1st} \) has to get the information of pure EEG data in the first step. The good approximation is easy to be found when the basis number is less than half of the effective frequency range of the target data. The effective frequency range of EEG data is usually less than 60Hz. Therefore, we assumed that the SNR will be high when \( K_1 \) is ranged between 2 and 30 in the first step.

The assumption was generally correct about the first step from the Fig. 4. We also assumed that the SNR will be high when \( K_2 \) is ranged between 2 and 30 in the second step. The results of average SNRs were supposed to draw symmetrical pattern in a diagonal line traversing the Fig. 4 from the lower left corner to the upper right corner. However, this assumption was markedly different because the SNR is high when a large second basis number \( K_2 \) is selected. Thereupon, we investigate several waveforms to reveal the reason.

B. Comparison of Spectra in Frequency Domain

Respective amplitude spectra are shown in Fig. 5. About Fig.5, the top shows the original Fp1 amplitude spectrum, removed amplitude spectrum by the ICA, and 3 types of estimated eye blinking artifact amplitude spectra by the proposed method. The red line is the results of ICA, therefore if the other line is similar to the red line, the line can be said good performance. From the Fig. 5, we could know that the NMF found good approximation with each basis number. The lower basis number of \( K_1 \), the better accuracy.

Fig. 4. The result of average SNRs between the removed signal by ICA and the estimated eye blinking artifact signal by proposed method.
of estimating. In particular, the waveform which bases are 5 and 50, was similar to the removed spectrum by the ICA.

The bottom shows the original Fp1 amplitude spectrum, reconstructed amplitude spectrum by the ICA, and 3 types of reconstructed amplitude spectra by the proposed method. The lower basis number of $K_1$, the better accuracy of estimating. This result indicates that the NMF has to be applied with less bases to find good approximation in the first step.

The waveform caused by the eye blinking don’t necessarily shape the same waveform because the waveform was generated based on the movements of eyelid [16]. Equalizing the adjustment of the force or variation time is nearly impossible, therefore each waveform caused by the eye blinking is slightly different. Also, it is the same for the waveform caused by the brain activity.

When $K_1$ is small, the matrix $Y_1$ is decomposed into few bases. For the effective frequency range of EEG data, good approximation is achieved with few bases in the first step. If $K_1$ is over 30, a basis expresses a component such as 4Hz. Conversely, if $K_1$ is 1, a basis expresses the training data without changes. This condition is not said to be sparse and good approximation. Therefore, $K_1$ has to be set between 2 and 30.

On the other hand, the matrix $Y_2$ is decomposed into many bases according to the matrix $H_{1_{st}}$ in the second step. As for good approximation is achieved in the first step, the bases of the second step can represent the waveform caused by the eye blinking.

In the NMF, the waveform is decomposed in the range of the original waveform based on the property of NMF. If good approximation is achieved in the first step, $K_2$ had little to do with approximation. Therefore, the determining parameter on $K_1$ should be prudent. Empirically, we recommend deciding the basis number $K_1$ and $K_2$ as 5 to 30 and 30 to 50 for finding good approximation.

C. Comparison of Waveforms in Time Domain

Respective waveforms in time domain are shown in Fig. 6. The top shows the original Fp1 signal, removed signal
by the ICA, and 3 types of estimated eye blinking signals by the proposed method. From the top of Fig. 6, we clearly represented that the NMF could remove the eye blinking artifacts effectively when the first basis number $K_1$ is low and the second basis number $K_2$ is large.

The bottom shows the original Fp1 signal, reconstructed signal by the ICA, and 3 types of estimated eye blinking signals by the proposed method. The waveforms are similar to the original signal because the waveforms were calculated from only single-channel signals.

D. Reconstruction of Original Signals

Furthermore, we focus on the reconstruction of the proposed method. The original Fp1 data and sum of the reconstructed and removed data in frequency and time domain where the number of basis is 5 and 50 are shown in Fig. 7.

Each sum of reconstructed and removed waveform (orange line) has an overlap with the original waveform (black line). The average of SNR was 55.12dB and average of similarity ratio was over 99%. From this result, we could obtain relatively good result.

The 2 step NMF could decompose a non-negative matrix to two non-negative matrices with high accuracy. The reconstruction of the proposed method was higher than the ICA, therefore, we confirmed the validity of the 2 step NMF for the eye blinking artifact rejection method using only single-channel EEG signals.

V. CONCLUSIONS

In this paper, we proposed an eye blinking artifact rejection method using the single-channel EEG signals and the 2 step NMF. We acquired 14 EEG and 1 vertical EOG signals from 10 subjects who blink every 5 seconds. Furthermore, we performed the proposed method to reject the eye blinking artifacts using the single-channel EEG (Fp1) signals and the ICA using the multi-channel EEG signals to prepare the target for comparison.

As a result, we represented the SNRs between waveforms from the ICA and the proposed method in frequency and time domain, and the reconstruction of the proposed method. We could show that the SNR is high when a low basis number $K_1$ and a large basis number $K_2$ were selected, and the average of reconstruction of the proposed method was over 99%. Therefore, we confirmed the validity of the proposed method for the eye blinking artifact rejection method using only single-channel EEG signals.

REFERENCES