A REAL-TIME DATA CLASSIFICATION SYSTEM FOR ACCURATE ANALYSIS OF NEURO-BIOLOGICAL SIGNALS

ZHANFENG JI1,2, TAKENAO SUGI1, SATORU GOTO1, XINGYU WANG2, and MASATOSHI NAKAMURA3

1 Department of Advanced Systems Control Engineering
Saga University
1 Honjomachi, Saga 840-8502, Japan
jzf1228@gmail.com

2 Department of Automation
East China University of Science and Technology
130 Road Meilong, Shanghai 200237, China

3 Research Institute of Systems Control Institute for Advanced Research and Education
Saga University
Yoka-machi, Saga 840-0047, Japan

Received February 2010; revised July 2010

ABSTRACT. Neuro-biological signals have much importance in clinical diagnosis. However, those signals are inevitably contaminated by various artifacts. The analysis without considering the artifacts would lead to mis-interpretation results. Therefore, an available signal selection technique is necessary to be developed for obtaining accurate and reliable results. In this paper, a model of real-time data classification system for accurate analysis of neuro-biological signals was proposed. The proposed model can be utilized for various research purposes. As an example, a physical system to analyze the relationship between characteristics of electroencephalogram (EEG) and electrocardiogram (EKG) was developed. The results indicated that the real-time system was effective for available signals selection. With the developed real-time system, the quality of signals can be monitored during the recording process.

Keywords: Real-time data classification, Accurate analysis, Artifacts detection, EEG, EKG

1. Introduction. The neuro-biological signals containing electroencephalogram (EEG), electrocardiogram (EKG), electrooculogram (EOG) and electromyogram (EMG) have close relationship with human mental/physical functions. EEG is the reflection of brain activities. EKG reflects the muscle movement of heart and the variability of autonomic nervous system. EOG comes from the movement of eyeball. EMG records the muscle activities at a certain part of the body.

In clinical diagnosis, EEG and EKG are more important than EOG and EMG. EEG analysis is widely used in monitoring and diagnosis, such as infants monitoring and epilepsy diagnosis. EKG analysis is useful to predict cardiac diseases and to investigate the autonomic nervous system since heart rate variability (HRV) analysis has become a popular tool [1]. The relationship analysis between EEG and EKG are important to understand the mechanisms of human brain and heart in future. However, there are no correlative researches being reported.

Artifact contamination problem is inevitable during the real clinical recording process. EEG and EKG are easy to be contaminated by various artifacts, which may be caused by loosing electrode, eye movement, blink, muscle activity, sweating, deep breath, or
movements of head or body. Such artifacts would lead to mis-interpretation results. In order to obtain reliable results, only available signals should be used for either visual inspection or automatic analysis.

For short-term neuro-biological signals analysis, available signals can be selected by visual inspection [2]. For long-term EEG recording, such as sleep data analysis, intensive care unit (ICU) monitoring, and cardiac disease monitoring, it is hard and laborious to check the artifacts by visual inspection. On the other hand, visual inspection is difficult for unskilled clinicians. In order to obtain reliable results and reduce the heavy burden of visual inspection, automatic artifacts detection and data classification for EEG and EKG are necessary to be developed.

Many papers reported artifacts detection/removal method for EEG and EKG analysis [3-13]. Various advanced methods have been applied to detect and remove artifacts in EEG signals, such as independent component analysis (ICA) [3-6], support vector machine (SVM) [4], wavelet analysis [7] and autoregressive (AR) model [8]. These methods were appropriate for offline analysis. For online application, the artifact detection algorithms need to be simplified. Agarwal et al. reported an approach of automatic analysis of segmented-EEG [9]. The authors applied a multi-level artifact rejection method during long-term recording. Durka et al. presented an automatic artifact detection system for polysomnographic (PSG) recordings [10]. However, these artifacts detection/removal methods seldom discriminated the artifacts in details. The importance of artifact detection for HRV analysis was emphasized by Berntson et al. [11]. Xu et al. proposed an automatic detection method for HRV analysis [12]. Sapoanikov et al. also presented artifact removal method for EKG [13]. In those methods, only heart beat interval series were considered. Until now, few research results have been reported regarding to the data classification method concerning both EEG and EKG for real-time application.

In this paper, a model of real-time data classification for EEG and EKG analysis is proposed. Kinds of advanced algorithms, such as ICA [3-6], SVM [4], fuzzy cluster analysis [14], self organizing maps (SOM) [15] and group sequential analysis [16], have been used for data classification. But they are not proper for online detection. In order to be integrated into the real-time application, the artifact detection algorithm should be significantly simplified comparing with offline detection methods. As an example, the model is applied on a physical system for analyzing the relationship between characteristics of EEG and EKG signals. At first, the artifacts in EEG and EKG were automatically detected. The artifacts detection method was newly constructed based on our previous work [17]. The quality of the signals was then evaluated based on detected artifacts by adjusting weight coefficients according to clinical interpretation purposes. Finally, the evaluated signals were classified by thresholds which were adaptable to clinical interpretation purposes. With the developed real-time system, the quality of signals could be monitored during the recording process, which would be helpful to obtain satisfactory recording and better understanding of artifacts. The proposed data classification method can also be extended for various research purposes by selecting artifacts for removal, defining weight coefficients, and adjusting thresholds.


2.1. Outline of data classification algorithm. The experimental paradigm is shown in Figure 1 (a). The task environment should be designed according to the purpose of neuro-biological signals interpretations. The measurement equipment records neuro-biological signals when subjects perform the tasks. The recorded data are sent to the
real-time data classification system. In this paper, task environment, subjects, and measurement equipment are explained in section 2.2.

Figure 1 (b) illustrates the flowchart of the real-time data classification system. First of all, neuro-biological signals are received from the measurement equipment, filtered for parameter calculation, and segmented for real-time detection. After that, all the parameters are calculated for artifact detection. Artifacts are then detected according to the neuro-biological signals interpretation purpose. Finally, the signals are scored with given weight coefficients and classified by using adaptive thresholds. The weight coefficients and adaptive thresholds are determined by the interpretation purpose of data analysis.

The model of the real-time data classification system can be applied to EEG and EKG analysis. In the following sections, the model will be explained with a physical system and an experiment. The purpose is to analyze the relationship between characteristics of EEG and EKG during long-time mental calculation and rest.

![Flowchart of experiment paradigm and real-time data classification system](image)

**Figure 1.** Flowchart of experiment paradigm and real-time data classification system

### 2.2. Data acquisition.

#### 2.2.1. EEG and EKG data. The EEG data investigated in this research were recorded from eight healthy male adults, aged 22-25 years old, from System Control Laboratory, Saga University of Japan. The subjects were informed to abstain from alcohol and caffeine 24 hours before experiment. They were requested to have a good sleep at the last night. During the data recording, the subjects were settled in a quiet electrical shield room with the temperature of 24~26 °C. All the experiments started at 10:00 AM. The subjects gave their informed consent prior to the experiments.

EEG electrodes were placed on the scalp according to the international 10-20 system at the following areas: Fpz, Fp2, F3, F4, O1, O2 against ipsilateral earlobe electrode (A1 or A2), and Fz, Cz against the average of A1 and A2 (Aav). A pair of EKG electrodes was placed according to standard limb lead II, in order to obtain the largest amplitude of R peak. The electrical impedance was kept under 10 kOhms for all electrodes. EEG and EKG signals were recorded on a digital electroencephalograph (Nihon-Koden EEG 2110) with the sampling frequency of 200 Hz, the upper cut-off frequency at 60 Hz, and the lower cut-off frequency at 0.016 Hz.
2.2.2. **Experiment tasks.** Experiment tasks consisted of four sequential sections: (1) normal rest for 5 minutes; (2) mental calculation for 120 minutes; (3) music relaxation for 15 minutes; and (4) mental calculation for 30 minutes. In mental calculation tasks (section 2 and 4), the subjects had to complete math sums (for example, $63 + 74$) and input answers with a numeric key pad as correctly as possible. The two 2-digit random numbers appeared on a 17 inches computer monitor. In the third section (music relaxation), Mozart Eine Kleine Nachtmusik was used. Subjects were free to open or close their eyes when they were listening to music with an earphone.

2.3. **Artifacts contamination.** Since subjects were not requested to keep still, several kinds of physical activities occurred during recording process. The observed artifacts in EEG and EKG data were explained respectively as below.

2.3.1. **EEG data.** During the long-time task, subjects were too tied to keep still all the time. They preferred to change postures for comfort. Since body movement affected the potential of all the EEG electrodes, base-line drift and electrode artifact were considered.

The most frequent EOG artifacts were blink and eye movement, which affected the potential of frontal electrodes. During the mental calculation task, subjects kept eyes open. Therefore, blink and eye movement artifacts happened frequently. During rest, subjects closed eyes. There were seldom blink and eye movement artifacts.

EMG artifacts were serious for EEG analysis. When subjects felt uncomfortable and moved their head or neck, EMG artifacts would appear at occipital electrodes. If large EMG artifacts occurred, all the EEG electrodes would be contaminated.

2.3.2. **EKG data.** The body movement is the major reason of artifacts in EKG. EMG artifacts were common in EKG studies, which could be caused by sudden body movements. RR intervals (beat to beat interval series) were the most important information for HRV analysis. The distortion of RR interval including base-line drift and R peak missing was considered as one kind of artifact. The base-line drift was induced by slow body movements. R peak missing was caused by large variability of electrode impedance, which was also caused by body movements.

2.4. **Parameters calculation.**

2.4.1. **EEG data.** Four parameters are employed for artifacts detection in EEG: (a) amplitude ($\mu V$): $A_z(x) = 6\sqrt{S_z(x)}$; (b) symmetry (%): $P_z(x, y) = 6\sqrt{S_z(x)}/6\sqrt{S_z(x + y)} \times 100$; (c) extension (%): $E_z(x, y) = 6\sqrt{S_z(y)}/6\sqrt{S_z(x)} \times 100$; and (d) correlation: $R(x, y)$, where $x$ and $y$ represent specific electrodes of F$_{P1}$, F$_3$, O$_1$, F$_{P2}$, F$_4$, O$_2$, F$_Z$, and C$_Z$; $z$ denotes respective EEG component: $L$ (0-0.5 Hz), $\delta$ (0.5-4 Hz), $\theta$ (4-8 Hz), $\alpha$ (8-13 Hz), $\beta$ (13-25 Hz), and $H$ (35-50 Hz). The following items are employed in the definition:

- $S_z(x)$ is the summation of periodogram with the frequency band of $z$ in channel $x$;
- $S_z(x - y)$ is the summation of periodogram with the frequency band of $z$ in channel $x$ and $y$, in which the EEG time series of channel $y$ is subtracted from that of channel $x$;
- $S_z(x + y)$ is the summation of periodogram with the frequency band of $z$ in channel $x$ and $y$, in which the EEG time series of channel $y$ adds that of channel $x$.

2.4.2. **EKG data.** Three parameters are employed for artifacts detection in EKG: (a) RR interval series: $y(n)$, (b) estimation of base-line drift: $E_{BLD}$, and (c) amplitude: $A_{H2}$. $E_{BLD}$ and $A_{H2}$ are calculated by Equations (1) and (2), where $x(n)$ represent the base-line estimated by a discrete Fourier transform (DFT) low pass filter with a upper cut-off
frequency of 1 Hz; H2 denotes high frequency band of (45-59 Hz) and (61-100 Hz); and
S_H2 is the summation of frequency component of H2 band.

\[ E_{BLD} = \max\{x(n)\} - \min\{x(n)\} \]  
\[ A_{H2} = 6\sqrt{S_{H2}} \]  

2.5. Artifacts detection.

2.5.1. EEG data. The criteria for artifacts in EEG are shown in Table 1. The left column
is the type of artifacts in EEG, and the right column is the judgment condition respectively.
The combinations of four parameters defined in section 2.4.1 among the EEG
channels are utilized for artifact detection in EEG data.

The electrode artifact indicates the ear-lobe reference electrode artifact. Since it is
mainly caused by the impedance variety of the reference electrode, electrode artifact
would only affect the channels at the same site. If the waveforms are similar in all the
adjacent channels on the same site but different with the other site, the related electrode
artifact will be discriminated [17]. Left and right sites are discriminated respectively. \( \delta \) band is employed to detect the oscillation at each channel, and correlation coefficients
are employed to judge the similarity between waveforms of different channels. In order to
detect the electrode artifact, at least three pairs of electrodes are required. Here, \((F_{P1}, F_{P2}), (F_3, F_4), \) and \((O_1, O_2)\) are employed.

The base-line drift artifact is a type of electrode artifact. The main reason is the variety
of electrode impedance, which could be brought by body movement or sweating. Since
the impedance variety caused by body movement is slow, \( L \) band is employed to judge the
oscillation at each channel. Correlation coefficient is used to judge the similarity between
the channel pairs on both site of the scalp. If all the conditions are satisfied, base-line drift
will be discriminated. For detecting base-line drift artifacts, the electrodes for detecting
electrode artifacts are utilized.

Blinks and eye movements are frequent while eyes open. They affect frontal electrodes
\((F_{P1} \text{ and } F_{P2})\) seriously. The power of low frequency components \((A)\) is employed to detect
the oscillation at \( F_{P1} \) and \( F_{P2} \). Since left and right eyes blink together, the symmetry \((P)\)
is employed to judge the similarity between \( F_{P1} \) and \( F_{P2} \). Because the central region is
also affected at the same time, the extension to central region \((E)\) is employed to judge the
influence on \( F_3 \) and \( F_4 \). Based on visual inspection of the waveforms contaminated by
blink and eye movement, \( \delta \) band is selected to discriminate blink, and \( L \) band is selected
to discriminate eye movement. \( F_{P1}, F_{P2}, F_3, \) and \( F_4 \) are essential for detection of blink
and eye movement.

EMG artifacts are detected respectively for each electrode. The power of \( H \) band and
\( \beta \) band are used for judgment [17].

2.5.2. EKG data. The criteria for artifacts in EKG are shown in Table 2. The left column
is the type of artifacts in EKG and the right column is respective judgment conditions.
The three parameters defined in 2.4.2 are utilized for artifact detection in EKG data.

RR intervals (beat to beat interval series) are between 0.4 and 1.2 seconds for normal
people in quiet status. Abnormal RR intervals should be rejected for HRV analysis. This
kind of distortion of interval is large. The large base-line drift reflects body movements
or deep breathing, which indicates that EKG signals contain the information of physical
activities. If the large base-line drift is detected, it is not proper to analyze the relationship
between EKG signals and mental activity. However, HRV could be correctly calculated.
This kind of distortion of interval is small.
The power of high frequency components is employed to identify EMG artifacts. If the power is larger than 200 μV, which may affect R peak detection, large EMG artifacts will be determined. If the power is larger than 58 μV, which indicates that EKG signals include the information of physical activities, small EMG artifact will be determined.

**Table 1. Criteria for artifacts in EEG**

<table>
<thead>
<tr>
<th>Artifacts</th>
<th>Judgment conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Electrode</strong></td>
<td></td>
</tr>
</tbody>
</table>
| Left:           | \( \min \{ A_5(F_{P1}), A_5(F_3), A_5(O_1) \} \geq 25 \mu V, \)  
|                 | \( \min \{ R(F_{P1}, F_3), R(F_3, O_1), R(O_1, F_{P1}) \} > 0.9, \)  
|                 | \( \min \{ R(F_{P1}, F_{P2}), R(F_3, F_4), R(O_1, O_2) \} < 0.8. \)  
| Right:          | \( \min \{ A_5(F_{P2}), A_5(F_4), A_5(O_2) \} \geq 25 \mu V, \)  
|                 | \( \min \{ R(F_{P2}, F_4), R(F_4, O_2), R(O_2, F_{P2}) \} > 0.9, \)  
|                 | \( \min \{ R(F_{P1}, F_{P2}), R(F_3, F_4), R(O_1, O_2) \} < 0.8. \)  
| **Base-line drift** | \( \min \{ A_L(x) \} \geq 60 \mu V, \)  
|                 | \( \min \{ R(F_{P1}, F_{P2}), R(F_3, F_4), R(O_1, O_2) \} > 0.8 \)  
| **Blink**       | \( \min \{ A_6(F_{P1}), A_6(F_{P2}), A_6(F_{P1} + F_{P2})/2 \} > 40 \mu V \)  
|                 | \( P_6(F_{P1}, F_{P2}) < 55\% \)  
|                 | \( \max \{ E_6(F_{P1}, F_3), E_6(F_{P2}, F_4) \} \leq 85\% \)  
| **Eye movement**| \( \min \{ A_L(F_{P1}), A_L(F_{P2}), A_L(F_{P1} + F_{P2})/2 \} > 60 \mu V \)  
|                 | \( P_L(F_{P1}, F_{P2}) < 55\% \)  
|                 | \( \max \{ E_L(F_{P1}, F_3), E_L(F_{P2}, F_4) \} \leq 85\% \)  
| **EMG**         | \( A_H(x) \geq 10 \mu V, \)  
|                 | \( A_H(x) \geq A_5(x) \)  

*The thresholds are established for the recording filtered between 0.53 Hz and 60 Hz.

**Table 2. Criteria for artifacts in EKG**

<table>
<thead>
<tr>
<th>Artifacts</th>
<th>Level</th>
<th>Judgment conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distortion of interval</td>
<td>large</td>
<td>( y(i) &gt; 1.2 ) s, or ( y(i) &lt; 0.4 ) s ( E_{BLD} &gt; 900 \mu V )</td>
</tr>
</tbody>
</table>
| **EMG artifacts** | small | \( A_{H2}(x) > 200 \mu V \)  
|                 | large | \( A_{H2}(x) > 58 \mu V \) |

*The thresholds are established for the recording filtered between 0.08 Hz and 60 Hz.

2.6. **Data classification.** According to the purpose of data analysis, weight coefficients are given to the artifacts for quality evaluation of each segment. The segments are then scored on the basis of detected artifacts.

The score of EEG or EKG is equal to the summation of the weight of detected artifacts. The score will be revised as 1, if larger than 1. 0 stands for the best signals, and 1 for the worst. The weight coefficients for all the artifacts are shown in Table 3.

A weight of 0.1 is given to the artifacts in EEG. Since EMG artifact is detected respectively for each channel, the total weight of 8 channels is 0.8. A weight of 1.0 is given to large distortion of interval and large EMG artifacts. A weight of 0.5 is given to small distortion of interval. A weight of 0.3 is given to small EMG artifacts. At last, the segments are scored by the higher score of EEG and EKG, and are classified as three groups (best, available, and worst) by two given thresholds for the best and the worst. In this paper, the segments will be classified as the worst for a score larger than 0.7, and the best for a score less than 0.2.
Table 3. Weight coefficients for quality evaluation

<table>
<thead>
<tr>
<th>Artifacts in EEG</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrode</td>
<td>0.1</td>
</tr>
<tr>
<td>Base-line drift</td>
<td>0.1</td>
</tr>
<tr>
<td>Blink</td>
<td>0.1</td>
</tr>
<tr>
<td>Eye movement</td>
<td>0.1</td>
</tr>
<tr>
<td>EMG artifacts</td>
<td>0.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Artifacts in EKG</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distortion</td>
<td></td>
</tr>
<tr>
<td>of interval</td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>small</td>
</tr>
<tr>
<td>EMG artifacts</td>
<td></td>
</tr>
<tr>
<td></td>
<td>large</td>
</tr>
<tr>
<td></td>
<td>small</td>
</tr>
</tbody>
</table>

2.7. **Real-time realization by DA-AD converter.** The whole algorithm was realized in real-time. The data classification system received the signals from the digital electroencephalograph through a DA-AD converter. EEG signals (channel 0-7) were transferred with the lower cut-off frequency at 0.53 Hz, and EKG signals were transferred with the lower cut-off frequency at 0.08 Hz (channel 8). The recording was divided into 5.12 s segments. The algorithm was repeated for each segment.

The interface contains the following items: 1) subject information, 2) original waveforms, 3) detected artifacts, 4) history trends of selected artifact, and 5) history trends of total score. By these items, the quality of recording can be properly estimated. In order to obtain satisfactory recording, when unexpected artifacts were detected, the user should find out the origin as soon as possible. If the artifacts were induced by physical activities, the user should remind subjects to keep still. If the artifacts came from equipment such as loose electrode, the experiment might have to be paused and performed again.

3. **Results.**

3.1. **Real-time data classification.** The interface of real-time data classification system is shown in Figure 2. Part A is subject information. Part B shows time series of three continuous segments. Detected artifacts are listed in Part C. The notation “YES” denotes detected, while blank denotes not detected. In the line of EMG artifact (EEG), the number below “YES” is the amount of detected channels. Part D illustrates the history trends of selected artifact events, which could be changed with the menu button of History. Part E displays the history trends of total weights. The upper dashed is the threshold of 0.7 for the worst signals, and the lower is the threshold of 0.2 for the best.

3.2. **Effect of data classification.** The power of $\alpha$ band ($A_\alpha$) and the power of $\beta$ band ($A_\beta$) are common characteristics in EEG analysis. Heart rate (HR), power of low frequency components (LF) and high frequency components (HF) are common for EKG analysis. As an example, $A_\beta(F4)$ and HR are used to analyze the relationship between EEG and EKG.

As shown in Figure 3, the relationships between $A_\beta(F4)$ and HR are analyzed by three different data selection method. Although all the recordings form eight subjects were analyzed by real-time data classification system, only the results of two subjects are shown in Figure 3. The horizontal axis indicates EEG parameters of $A_\beta(F4)$. The vertical axis indicates the EKG parameters of HR. As shown in Figure 3 (a), 263 segments from subject A and 152 segments from subject B, whose score is less than 0.2, are selected as the best segments by proposed selection method. As shown in Figure 3 (b), 124 segments
Figure 2. The interface of real-time data classification system

from subject A and 124 segments from subject B, whose segment number can be exactly divided by 20, are selected as random segment. As shown in Figure 3 (c), 202 segments from subject A and 65 segments from subject B, whose score is larger than or equal to 0.7, are selected as the worst segments by proposed selection method. With proposed selection method, the best segments are clustered, while the worst segments are scattered. Additionally, a few randomly selected segments are scattered from the others.

4. Discussion.

4.1. Weight coefficients.

4.1.1. Weight setting. The weight coefficients for each artifact should be given properly according to the purpose of data analysis, which determines the importance level of contaminated information. During the EEG data analysis, frequency components of δ, θ, α, and β are mostly studied. In addition, the information of α band and β band is more important than that of δ band and θ band. During the EKG data analysis, continuous RR intervals are the most important information for HRV interpretation.

Ear-lobe reference electrode artifact, base-line drift, blink, and eye movement affect low frequency components (usually δ band, sometimes θ band, seldom α band). Although blink and eye movement can be frequently encountered when eyes open, but both of them only affect the frontal electrodes. Therefore, all these four artifacts are given a low weight of 0.1.

EMG artifacts affect both EEG and EKG signals. For EEG analysis, EMG artifacts increase the power of β band seriously and sometimes also increase the power of α band.
For EKG analysis, small EMG artifacts indicate possible physical activities, so that the signals are not proper for mental activity analysis; large EMG artifacts may induce wrong detection of R peak, which are not available for HRV analysis. Therefore, a total weight of 0.8 was given to EMG artifacts in EEG, a weight of 0.3 was given to small EMG artifacts in EKG, and a rejection weight of 1.0 was given to large EMG artifacts in EKG.

Large distortion of interval indicates interrupted RR intervals, which is caused by false or miss detection of R peak. Small distortion of interval indicates that RR intervals contain the information of not only mental activities, but also physical activities. Physical activities could be reflected by large base-line drift of EKG signals. Therefore, a rejection weight of 1.0 was given to large distortion of interval, and a weight of 0.5 was given to small distortion of interval.

4.1.2. Extended application. In this research, the purpose of experiment is to study the relationship between characteristics of EEG and EKG, i.e. both EEG and EKG signals are important, so that the final score is defined as the higher score of EEG and EKG.

Users should define all possible artifacts according to analysis purposes, and adjust the weight coefficients according to the importance of the information contaminated by corresponding artifacts. After that, users should define the score and adjust the thresholds for data classification for their own analysis purposes. For example, if the purpose is to analyze the characteristics of EEG, a weight of 0 should be given to the artifacts in EKG. If the purpose is to analyze the characteristics of EKG under certain physical activity, small EMG artifact and small distortion of interval, both of which contain information of physical activities, should not be considered as artifacts any more, and a weight of 0 should be given.
Moreover, the recording conditions are crucial for artifacts detection. The recording conditions presented in section 2.2 and 2.7 are common in clinical analysis. The criteria illustrated in section 2.5 are established under these conditions. The criteria should be adjusted to recording conditions in different experimental paradigm for the individual analysis purpose. The principles for judgment of each artifact are described in section 2.5. The criteria are allowed to be revised according to the principles. For instance, in order to identify electrode artifacts, at least three pairs of electrodes should be recorded. In this paper, \((F_{P1}, F_{P2}), (F_3, F_4), \) and \((O_1, O_2)\) are utilized. If \(O_1\) and \(O_2\) are not recorded, but \(T_3\) and \(T_4\) are recorded, then \((F_{P1}, F_{P2}), (F_3, F_4), \) and \((T_3, T_4)\) are also effective. On the other hand, some artifacts may not be able to be detected under certain recording conditions because of lacking information, e.g. blink will not be detected except all the necessary channels \((F_{P1}, F_{P2}, F_3 \) and \(F_4)\) are recorded.

4.2. Effect of data classification. The difference between the best and worst segments classified by proposed data classification method is obvious, as shown in Figure 3 (a) and (c). The worst segments include seriously contaminated signals. Most artifacts are caused by body movement, which could increase the power of \(\beta\) band in EEG, and affect the heart rate. That is the reason why the worst segments are scattered.

The quality of the recording for analysis is crucial for obtaining reliable characteristics. If the worst segments are used for data analysis, the accuracy of the results will be seriously affected, as shown in Figure 3. The best segments, which are selected by proposed data classification method, represent stable characteristics. The characteristic of randomly selected segments is disturbed by a few bad segments. The characteristic of the worst segments is vague. The presented data classification method is effective for meaningful analysis.

4.3. Real-time application. Artifacts detection is crucial for reliable analysis of EEG or EKG signals. However, in most research work, the artifacts are usually detected after recording. Therefore, it is very difficult to find out the actual reason. Moreover, for long-time experiment, it is necessary to monitor the quality of signals in real-time, in order to obtain satisfactory recording. If continuous serious contamination is detected after recording, long-time experiment has to be performed again. In this paper, a data classification method is developed for real-time application. The major part is artifacts detection algorithm. Most papers reported artifacts detection methods with offline mode [3-8]. They used complicated algorithms to detect or separate artifacts from original signals. But our research aims to obtain satisfactory recording, and provide available data for offline analysis. Therefore, our artifacts detection algorithm is significantly simplified comparing with others. With the model and artifacts detection algorithm, the user could monitor the recording quality in real-time, and perceive contaminated signals promptly. It is also possible to find out the reason of artifacts during recording process.

5. Conclusions. In this paper, a real-time data classification model for accurate analysis of neuro-biological signals was proposed. As an example, a physical system to analyze the relationship between characteristics of EEG and EKG was developed. The results suggested that the proposed data classification method was effective. Comparing the best and worst segments, it was observed that contaminated signals would lead to misinterpretation. With this system, it is convenient to monitor the quality of signals, acquire better understanding of artifacts, and obtain satisfactory recording.
REFERENCES


