

Quality Assessment of EEG Signals based on Statistics of Signal Fluctuations

Huijuan Yang, Cuntai Guan, Kai Keng Ang, Kok Soon Phua and Chuanchu Wang

Institute for Infocomm Research,

Agency for Science, Technology and Research (A*STAR), Singapore 138632.

Email: {hjyang, ctguan, kkang, ksphua and ccwang}@i2r.a-star.edu.sg

Abstract—The quality of the non-invasive EEG signals was always affected by the changes in the contact impedances and the artifacts from eye blinking, eye movements and body movements. An effective quality assessment method is needed to assess the qualities of the EEG signals. This paper proposed a novel method to assess the signal quality of EEG signals based on block-based measurements of the fluctuations of the second-order power amplitudes of the EEG signals. The initial signal quality scores were generated by fusion of the mean power amplitudes and the signal fluctuations of the motor imagery state with reference to background idling state. These scores were subsequently mapped to different quality levels by using fuzzy-c means clustering. Experimental results were conducted on the basis of 3 data sets of 15 healthy subjects performing motor imagery of hand movements and idle, for both gel-based and gel-less electrodes. The results obtained demonstrated that the proposed method was capable of evaluating the quality of the EEG signals, as supported by the clear separation of the assigned quality levels between gel-based and gel-less electrodes. This further validated the assumption that generally the quality of the EEG signals acquired based on the gel-based electrodes was better than that of the gel-less electrodes.

I. INTRODUCTION

Electroencephalography (EEG) is a non-invasive technique to monitor the electrical activity of the brain. Successfully decoding of the EEG signals has established an effective way to translate the intention of the paralyzed patients to drive external assistive devices. However, many factors affect the quality of the EEG signals. Firstly, the contact impedances between the skin and electrodes should be sufficiently low to obtain EEG signals of high qualities, which can be achieved with the use of conductive gel. Nevertheless, the signal would get noisier with the gel getting dried up after several hours of use. How the quality of the EEG signals was affected by the electrodes of low and high impedances under different recording environments was investigated in [1], where the EEG signal was recorded during an oddball task. The results revealed that the low-frequency noise was increased at the high-impedance electrode site when P3 amplitude was measured, especially at the warm and humid recording conditions. Recently, some dry electrodes were used which did not require gel or skin preparation. This makes both a home-based and ambulatory brain computer interface (BCI) application possible since a fast setup can be expected [2]. However, the presence of motion artifacts formed a major obstacle for such systems, which were mainly caused by the contact changes between the skin and electrodes. Monitoring the changes of the electrode-tissue contact impedance can thus facilitate detection and reduction of the motion artifacts, e.g., to improve the prediction of the

motion artifacts with the use of the in-phase and quadrature components of the contact impedance [2].

Eye blinking and eye movements, and other body movements related artifacts complicated the interpretation of EEG signals [3], [4]. A quantitative investigation on the signal quality (e.g., in terms of artifacts contamination) of simultaneously recorded invasive and non-invasive EEG signals was carried out [4]. Comparisons of the blink-related artifacts in prefrontal and motor cortical regions indicated that significant blink-related artifacts were found in invasive recordings in the prefrontal region. A simple way to remove the ocular artifacts was to visually inspect the EEG data and set a threshold to reject the contaminated trials or channels [3], which would result in the loss of data. Regression-based method generally require a reference channel not containing any brain activity, which was not the case in real applications. Independent component analysis-based eye movements and eye blinks artifacts removal firstly decomposed the EEG signals into independent components [3]. The topographies and power spectral density features of those components were then extracted. Eventually a peak detection algorithm was applied to identify the components of eye blinking and a SVM classifier was trained to identify the components of eye movements.

Improving the signal-to-noise ratio by soft thresholding for better interpretation of the event-related potential was investigated in [5]. The denoising-based methods were generally sensitive to the robustness of the threshold. Furthermore, it was difficult to differentiate the instrumental noise from that of the background brain activity and movement noise. Denoising the nonlinear time series by adaptive filtering based on weighted polynomial fitting demonstrated the superior performance compared with that of the wavelet shrinkage-based denoising [6]. The degree of variability detected by the second-order difference plot of the EEG time series, between eye closing and eye opening, was used as a switching mechanism to control the assistive devices [7].

It is noted that a proper method to assess the qualities of the EEG signals is needed. In this paper, we proposed a novel quality assessment measure to assess the qualities of the EEG signals. The idea of the proposed method was to measure the irregularity that arises from the short time noisy fluctuations of the EEG signals. Existing quality measures for images can be classified into three categories [8], [9]: *full-reference*, *no-reference* and *reduced-reference*, whereby a complete reference, no reference, and partial reference were available for the assessment. In our problem, the ground-truth values of EEG quality were not available, hence, the proposed

method was a *no-reference* or *reduced reference* approach. To evaluate the quality of the EEG signals, a data set consisting of three groups of EEG data collected using the gel-based (two groups) and gel-less (one group) EEG electrodes were used. Furthermore, we assume that the signal quality of the EEG data acquired by the gel-based electrodes was better than that acquired by the gel-less electrodes, as confirmed by the impedances of the data of these two types of electrodes. This assumption would be validated by the signal quality levels assigned to the EEG signals acquired by the two types of electrodes by our proposed signal quality assessment method. The idea was motivated by the facts that the EEG signals tend to fluctuate a lot due to the noisiness caused by the high impedances between the electrodes and scalp. Hence, we proposed to measure the signal quality of the EEG signals by employing the statistics of the higher-order power amplitude fluctuations at a particular frequency of the motor imagery state with reference to background idling state.

II. PROPOSED METHOD

We now describe our proposed scheme for the quality assessment of EEG signals based on statistics of signal fluctuations (QS-EEG-SSF), which was illustrated in Fig. 1. The proposed EEG signal quality assessment scheme consisted of several components, as can be seen from Fig. 1: *EEG signal pre-processing*, *score generation* and finally the *score mapping*. These three components would be described in further details in the followings.

1. *EEG signal preprocessing* (shown in the blue dashed box in Fig. 1). Preprocessing the acquired EEG signals was firstly performed before the score generation and mapping process. Different frequency bands were studied to filter the EEG signal, which were: *mu* (8-13Hz), *low-beta* (13-16Hz), *middle-beta* (16-20Hz), *high-beta* (20-28Hz), *wide-beta* (16-28Hz) and *whole band* (4-40Hz). The signal was firstly filtered at the selected frequency band based on the butterworth filter of n th order, where $n=5$ was selected in the experiments. Secondly, removing the baseline (i.e., signals before onset of the cue) was carried out to cancel out the baseline activities, the so-obtained EEG signal was denoted as E_{fb} . The second-order power amplitude of the EEG signal for channel c and class y (denoted as $P_a(c, y)$) was calculated by

$$P_a(c, y) = \sqrt{d^2(c, y, t) + d^2(c, y, t + 1)} \quad (1)$$

where $d(c, y, t)$ was given by

$$d(c, y, t) = E_{fb}(c, y, s(t + 1)) - E_{fb}(c, y, s(t)) \quad (2)$$

where $y \in \{1, 2\}$ denoted class 1 and 2. The sample index $s(t)$ was given by $s(t) = t \cdot l_s - (n_s - t)$, where n_s and t were the studied time lag and t th lag ($t \leq n_s$), respectively, $n_s = 3$ was selected in the implementation; l_s was the signal length for channel c and class y . In this way, $s(1) = 1 : l_s - 2$, $s(2) = 2 : l_s - 1$ and $s(3) = 3 : l_s$.

2. *Score generation* (shown in the green dashed box in Fig. 1). It was noted that our proposed method was block-based. After obtaining the second-order power amplitude of the EEG signal, i.e., $P_a(c, y)$, it was then divided into blocks. The total number

of blocks was given by $n_k = \lfloor (l_s - (n_s - 1)) / l_b \rfloor$, where l_b was the block length. The signal at i th block (denoted as $P_a(c, y, i, s)$) was given by

$$P_a(c, y, i, s) = P_a(c, y)(s(i)) \quad (3)$$

where the index of the block $s(i)$ was given by $s(i) = (i - 1) \times l_b + 1 : i \times l_b$. The score generation consisted of the following steps.

- Calculated the maximum fluctuation of second-order power amplitude of the EEG signal (denoted as F_u) for the i th block of channel c and class y , which were given by

$$F_u(c, y, i) = |\max(P_a(c, y, i)) - \min(P_a(c, y, i))| \quad (4)$$

where $\max(x)$ and $\min(x)$ gave the maximum and minimum of elements in x , respectively; $|x|$ gave the absolute value of x .

Further, calculated the mean and variance of the second-order power amplitude of the EEG signal for the i th block of channel c and class y (denoted as $P_m(c, y, i)$ and $v_b(c, y, i)$), which were given by

$$P_m(c, y, i) = \frac{1}{l_b} \sum_{k=1}^{l_b} P_a(c, y, \tilde{s}(i, k)) \quad (5)$$

$$v_b(c, y, i) = \frac{1}{l_b} \sum_{k=1}^{l_b} (P_a(c, y, \tilde{s}(i, k)) - P_m(c, y, i))^2 \quad (6)$$

where $\tilde{s}(i, k)$ denoted k th index of sample index of block i .

- Normalization. F_u and P_m were then normalized by the variance of the second-order power amplitude of the EEG signal, which were given by

$$\tilde{F}_u(c, y, i) = F_u(c, y, i) / v_b(c, y, i) \quad (7)$$

$$\tilde{P}_m(c, y, i) = P_m(c, y, i) / v_b(c, y, i) \quad (8)$$

- Generated the quality assessment score (S_c). The score was finally generated by fusion of the changes of the mean and fluctuation of the second-order power amplitude of the EEG signal for the motor imagery state with relative to the background idling state, which was given by

$$S_c(c) = \frac{1}{n_k} \sum_{k=1}^{n_k} (w_1 * D_m(c, k) + w_2 * D_u(c, k)) \quad (9)$$

where

$$D_m(c, k) = |\tilde{P}_m(c, 1, k) - \tilde{P}_m(c, 2, k)| \quad (10)$$

$$D_u(c, k) = |\tilde{F}_u(c, 1, k) - \tilde{F}_u(c, 2, k)| \quad (11)$$

where w_1 and w_2 ($w_1, w_2 \in \{0, 1\}$) were the weights provided by the user for the fusion. In the experiments, $w_1 = w_2 = 0.5$ was used.

3. *Score mapping* (shown in the red dashed box of Fig. 1). The generated quality assessment scores were mapped to different quality levels for the convenience of assessment. For

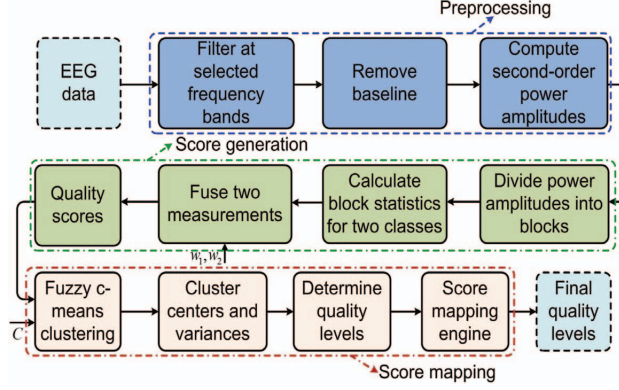


Fig. 1. A schematic illustration of the proposed quality assessment of EEG signals based on statistics of signal fluctuations (QS-EEG-SSF).

for this purpose, the scores were firstly normalized based on the minimum and maximum scores for all the training data, e.g., $\tilde{Q}_s = \frac{(Q_s - Q_m)}{Q_x - Q_m}$, where Q_m and Q_x denoted the minimum and maximum of the quality scores, respectively. It should be noted that the training data should contain the data that were both of good quality (e.g., gel-based EEG data) and bad quality (e.g., gel-less EEG data). The normalized quality scores were subsequently mapped to different quality levels based on the predefined quality levels (Q_l) to obtain the so-called “initially mapped quality levels”, which was given by $Q_{ml} = \tilde{Q}_s Q_l$. The initially mapped quality levels were post-clustered based on Fuzzy-C means clustering [10], [11] to again group the quality levels into two dominate clusters for the EEG signals of “good” and “bad” qualities, based on the minimization of the following objective function, which was given by

$$J_m = \sum_{i=1}^{N_s} \sum_{j=1}^C u_{ij}^m \|q_i - C_j\|^2 \quad (12)$$

where the fuzzy parameter m was a real number and $m > 1$. The default value of $m=2$ was used. C denoted the number of clusters and N_s denoted the number of data instances, i.e., the initially mapped quality levels Q_{ml} ; u_{ij} denoted the degree of membership of q_i belonged to cluster j ; q_i and c_j denoted the i th data of Q_{ml} and the j th cluster center; and $\|x\|^2$ denoted the L_2 norm. An iterative clustering process was carried out with the membership u_{ij} being updated by

$$u_{ij} = \frac{1}{\sum_{v=1}^C \left(\frac{\|q_i - c_j\|}{\|q_i - c_v\|} \right)^{\frac{2}{m-1}}} \quad (13)$$

The cluster center C_j was given by

$$C_j = \frac{\sum_{i=1}^{N_s} u_{ij}^m \cdot q_i}{\sum_{i=1}^{N_s} u_{ij}^m} \quad (14)$$

The iteration stopped when $\max_{ij} |u_{ij}^{(k+1)} - u_{ij}^{(k)}| \leq \epsilon$, where ϵ was a termination criterion and $0 < \epsilon < 1$, k was the iteration step.

The outputs of the iterative clustering processing were the two cluster centers ($c_s(i)$), and variances of the two cluster

centers for all the channels ($v_c(i)$), where $i \in \{1, 2\}$ denoted the indexes of the clusters 1 and 2. Subsequently, the initially mapped quality levels that fall into the two clusters were further mapped to the final quality levels (q_l). For example, assume the number of quality levels, i.e., $N_q=5$, one way to determine the final quality levels was: $q_l = \{c_s(1) - v_c(1), c_s(1), c_s(2), c_s(2) + v_c(2), c_s(2) + 2v_c(2)\}$.

Once the final quality levels ($q_l(k)$, $k=1, 2, \dots, N_q$) were determined, the initially mapped quality levels (q_i , $q_i \in Q_{ml}$) were mapped to the nearest quality levels by

$$\hat{k} = \arg \min_{k=\{1, 2, \dots, N_q\}} |q_i - q_l(k)| \quad (15)$$

where $|x|$ denotes the absolute value of x . The pseudo codes for mapping the initially mapped quality levels to the final determined quality levels were detailed in Table I.

TABLE I. PSEUDO CODES FOR MAPPING THE INITIALLY MAPPED QUALITY LEVELS TO THE FINAL DETERMINED QUALITY LEVELS

Inputs: initially mapped quality levels (q_i , $q_i \in Q_{ml}, i=1, 2, \dots, N_s$) and the determined quality levels ($q_l(k), k=1, 2, \dots, N_q$)
Outputs: the final quality levels (Q_{fl})
Select quality level $q_l(k)$ according to Eq. (15).
if $q_i < q_l(k) \& (k \neq 1)$
$Q_{fl}(i) = q_l(k-1)$;
elseif $q_i < q_l(k) \& (k=1)$
$Q_{fl}(i) = q_l(1)$;
elseif $q_i > q_l(k) \& (k \neq n)$
$Q_{fl}(i) = q_l(k+1)$;
elseif $q_i > q_l(k) \& (k=N_q)$
$Q_{fl}(i) = q_l(N_q)$;
endif

III. EXPERIMENTAL RESULTS

The data set used in the experiments to validate the proposed quality measure consisted of three data sets of 15 healthy subjects performing the motor imagery of hand movements (MI-HM) versus idle, namely DS1, DS2 and DS3. The EEG data for a total of 6, 5 and 4 healthy subjects were collected for the first, second and third experiments. It should be noted that

the EEG data for the first and third data sets (DS1 and DS3) were based on gel electrodes, and the second data sets (DS2) were based on gel-less electrodes. The experimental protocol consisted of: 1 s of preparation, which was shown as a fixation at the center of the computer screen; 2 s of visual cue which was shown as a virtual hand holding a peg (for MI-HM) or a stop sign (for idle), 4 s of action time to perform motor imagery of hand movements for MI-HM or do nothing for idle; and 6 s of rest. In the experiments, the subjects were instructed to perform motor imagery of hand movements by holding the peg towards different directions following the peg cue shown at the center of the computer screen; whereas a still hand was shown for idle state to instruct the subject not to do anything. One session consisted of two runs with each run consisted of 40 trials of MI-HM and 40 trials of idle. The EEG data were collected using Shanghai NCC model B (for DS1 and DS2), and model C (for DS3) EEG system, with the EEG cap consisting of 20 (and 32) channels. The model was equipped with a bluetooth wireless amplifier and the sampling rate was 256 Hz. The 16 EEG channels used in the experiments were: F3, Fz, F4, FC3, FCz, FC4, T3, C3, Cz, C4, T4, TP7, CP3, CPz, CP4 and TP8 for 20 channels settings. While the 22 channels used in the experiments were: Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Sp1, Sp2, Fz, Cz, Pz, Oz for 32 channels settings. The locations of electrodes were consistent with the international 10-20 system for electrodes placements. In the experiments, the impedances for the gel-data were between 5 K Ω to 30K Ω . However, the impedances for the gel-less data at the frontal electrodes (e.g., with little hair) were from 50 K Ω to 100 K Ω , whereas the impedances for those electrodes lie in the middle of brain (e.g., with thick hairs) were from 100 K Ω to 200K Ω . Hence, there is a clear separation between the qualities of gel and gel-less data. This served as the groundtruth or partial information to validate the assessment scores and quality levels generated using our proposed method.

The scatter plot of the initially generated quality assessment scores using a wider beta frequency band for all the channels and subjects were shown in Fig. 2. It can be observed from Fig. 2(a)(b) that the initially generated quality assessment scores had a good tendency to separate the “good” (DS1 or DS3) versus the “poor” quality data (DS2), except for some noisy channels. However, the scores for the two groups of gel-based data appeared to be of similar quality, as can be seen from Fig. 2(c). This was true considering the similar impedances of the groups of gel-based data, and the much higher impedances of the gel-less data compared with that of gel-based data.

It is noted that the initially generated quality assessment scores varied a lot for the EEG signals of different qualities, or in other words, the ranges of the scores were too large. In order to generate quality assessment levels based on the predefined levels, these initially generated quality assessment scores were normalized and mapped to obtain “initially mapped quality levels”, which were plotted in Fig. 3. The results showed in the figure further demonstrated that there was a clear separation between the initially mapped quality levels between gel-based data (DS1 and DS3) versus that of the gel-less data (DS2). The variations were constrained at the predefined quality levels after mapping.

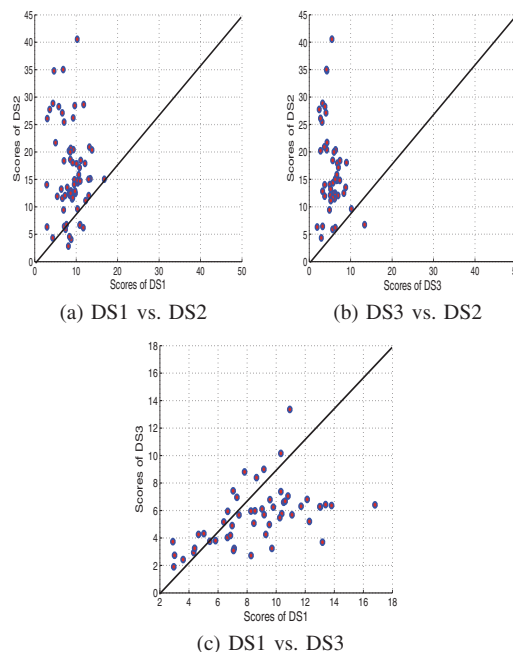


Fig. 2. Scatter plot of the initially generated quality assessment scores for data set 1 (DS1) versus data set 2 (DS2) (a), data set 3 (DS3) versus data set 2 (DS2), and data set 1 (DS1) versus data set 3 (DS3).

Finally, one quality score was generated for each subject by mapping the median score of selected channels that contributed mostly to the mental tasks. The final quality levels generated for each subject using DS1, DS3 and DS2 were shown in Table II. Note that the wider beta (16Hz to 28Hz) frequency band was employed in the experiments. The results shown in the table revealed that employing a wider beta frequency band was effective in generating the quality levels based on proposed method for the assessment of the qualities of EEG signals, as confirmed by the quality levels which correlated well with the overall impedances of the subjects. By closely examining the quality levels for each individual channel, it was found that overall 19.79%, 1.25% and 26.92% of 96 channels for the 6 subjects of DS1, 52 channels for 4 subjects of DS3 and 80 channels of 5 subjects of DS2 produced a quality level which was not consistent the quality level assigned for the subject. This was not surprising since some channels located in the frontal, occipital and temporal locations in the brain can be quite noisy due to body movements and eye blinking/movements etc.

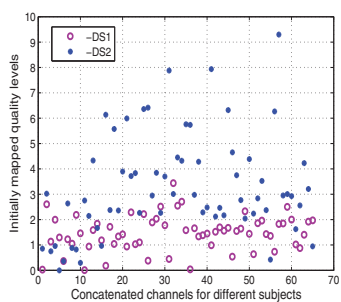
IV. CONCLUSION

In this paper, we investigated the novel problem of how to assess the qualities of the EEG signals on the basis of the facts that the quality of the non-invasive EEG signals was always affected by the changes in contact impedances and the artifacts from eye blinking, eye movements and other body movements. To achieve this goal, we proposed a novel method to evaluate

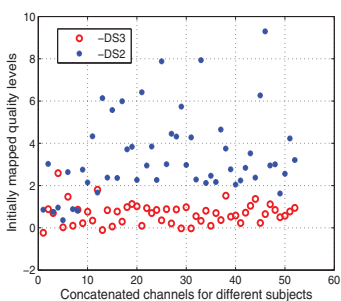
TABLE II. THE FINAL QUALITY SCORES/LEVELS ASSIGNED TO EACH SUBJECT FOR THE THREE DATA SETS.

Data sets	Electrodes type (gel/gel-less)	Subj.	Quality scores/levels generated (using all trials and beta (16-28 Hz) frequency band)			
			quality assessment scores (median)	initially mapped quality levels	finally assigned quality levels	impedance ranges (K Ω)
DS1	gel	S1	0.4922	0.5378	1	<30
		S2	1.7706	1.8824	1	<30
		S3	1.6423	1.7401	2	<30
		S4	1.9564	1.9765	2	<30
		S5	1.5233	1.6430	2	<30
		S6	1.6681	1.6770	2	<30
DS3	gel	F1	0.2243	0.1745	1	<30
		F2	0.8792	0.9204	1	<30
		F3	0.5244	0.5193	1	<30
		F4	0.8529	1.0555	1	<30
DS2	gel-less	G1	5.3857	4.6820	4	>50
		G2	2.7384	3.1098	4	>50
		G3	3.6807	3.3539	4	>50
		G4	2.3235	2.4308	4	>50
		G5	2.9392	3.1128	4	>50

Note: quality levels: 1-very good, 2-good, 3-moderate good, 4-moderate bad, 5-bad, and above 5-very bad.



(a) DS1 vs. DS2



(b) DS3 vs. DS2

Fig. 3. Plot of the initially mapped quality levels for (a) DS1 vs. DS2 and (b) DS3 vs. DS2. Clear separation between gel-less data (DS2) with gel data (DS1 and DS3) can be observed except some noisy channels.

the signal quality of EEG signals based on the measurements of the irregularity of the block-based, second-order fluctuation of the power amplitudes of the EEG signals. The obtained quality assessment scores were eventually mapped to different quality levels based on fuzzy c-means clustering and dynamic mapping. Experimental results were conducted on three data sets (one data set was based on gel-less electrodes and another two were based on gel electrodes) consisting of 15 subjects

performing motor imagery of hand movements and idle. The obtained quality assessment scores and the mapped quality levels demonstrated that the proposed method was capable of evaluating the quality of the EEG signals, based on the EEG signals filtered at a wider beta frequency bands. The separation of the quality scores and quality levels between the gel-based and that of gel-less electrodes correlated well with the impedances of data acquired by the two types of electrodes.

REFERENCES

- [1] Emily S. Kappenman and Steven J. Luck. The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology*, 47(5):888-904, 2010.
- [2] A. Bertrand, V. Mihajlovic, B. Grundelner, C. V. Hoofy, and M. Moonen. Motion artifact reduction in EEG recordings using multi-channel contact impedance measurements. *2013 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 258-261, 2013.
- [3] J. F. Gao, Y. Yang, P. Lin, P. Wang, C. X. Zheng. Automatic removal of eye-movement and blink artifacts from EEG signals. *Brain Topogr.*, 23(1):105-14, 2010.
- [4] T. Ball, M. Kern, I. Mutschler, A. Aertsen, and A. Schulze-Bonhage. Signal quality of simultaneously recorded invasive and non-invasive EEG. *NeuroImage*, 46:708716, 2009.
- [5] T. Falkner, A. Abedi, M. Hayes and J. Paul. Improving the signal to noise ratio of event-related EEG signals in high risk newborn. *25th Biennial Symposium on Communications (QBSC)*, 28-31, 2010.
- [6] J. Gao, H. Sultan, J. Hu, and W.-W. Tung. Denoising nonlinear time series by adaptive filtering and wavelet shrinkage: a comparison. *IEEE Signal Processing Letters*, 17(3):237-40.
- [7] R. A. Thuraisingham, Y. Tran, P. Boord and A. Craig. Analysis of eyes open, eye closed EEG signals using second-order difference plot. *Med Biol Eng Comput.*, 45(12):1243-9, 2007.
- [8] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13(4): 600-12, 2004.
- [9] C. Chen, A. C. Kot and H. Yang. A two-stage quality measure for mobile phone captured 2D barcode images. *Pattern Recognition*, 46(9): 2588-98, 2013.
- [10] J. C. Dunn. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3:32-57, 1973.
- [11] J. C. Bezdek. Pattern recognition with fuzzy objective function algorithms. *Plenum Press, New York, NY*, 1981.