

Identification, characterisation, and correction of artefacts in electroencephalographic data in study of stationary and mobile electroencephalograph

Monika Kaczorowska^{1,*}

¹Lublin University of Technology, Electrical Engineering and Computer Science Faculty, Institute of Computer Science, Nadbystrzycka 36B, 20-618 Lublin, Poland

Abstract. The present paper covers the review of artefacts of EEG recording, as well as methods to eliminate them, i.e. PCA and ICA. The reviewed artefacts were of various origins: biological (muscle, eye blinking, eyeball moving), as well as technological artefacts caused mainly by the power network with the frequency of 50 Hz. The research has been conducted on a group consisting of 8 students. The research was carried out with the use of two defined models of electroencephalographs: stationary and mobile, then results have been compared. The comparison of two amplifiers combined with various artefact elimination methods has been presented. Presented research has shown that the record performed by the mobile amplifier contains the artefacts of significantly lower amplitude.

1 Introduction

Electroencephalography (EEG) is a technique applied for registering and monitoring an electrical activity of a brain. It is a non-invasive method of measurement of specific waveforms emitted by a human brain. The special type of amplifier is dedicated to record EEG signals with a small amplitude. The signal is gathered from the human scalp by electrodes which are placed on a special cap [1]. The EEG signals are applied in medicine and in the construction of brain-computer interfaces (BCI). They may be applied in rehabilitation as well [2].

EEG signal is a resultant of all the potential electromagnetic fields that are present nearby the examined person. Most often it is disturbed by many undesirable signals called artefacts [1, 3]. They can originate from the bioelectrical activity of an examined person, as well as from the external environment [1]. The aim of the paper is to characterise and compare the process of artefact correction in EEG signals using two defined models of amplifiers, correspondingly: stationary and mobile. The comparison was made by using dedicated software (WinEEG). The task of artefact elimination is still a challenging one. This paper presents two methods applied to remove artefacts: PCA (Principal Component Analysis) and ICA (Independent Components Analysis).

The signals have been gathered from the representatives of two types of amplifiers: stationary and mobile. The mobile electroencephalograph is a comparatively new type of electroencephalograph convenient for implementing BCIs [4, 5]. Mobile EEG does not hamper the examined person due to the

absence of cables connecting it with a computer. It also meets the needs of creating BCIs not only because of its mobility but also due to the fact, that this type of EEG is equipped with several auxiliary devices, such as accelerometer, which helps to correct the artefacts from lower levels of recording [6]. The goal of the research was to compare an old-type stationary electroencephalograph with a modern type mobile EEG, in terms of noise level and artefacts.

2 Related works

2.1 Artefacts

Most artefacts in an EEG recording, such as eye-blinking and eye-movement artefacts, can be seen as a sudden peak. The other type of artefacts, such as muscle ones (especially jaw muscle-related artefacts) may be observed as a local increase of the high-frequency component. What is more, the technical artefacts, related to the power network with the frequency of 50 Hz, usually appear as a permanent presence of the 50 Hz frequency component in the EEG recording. The artefacts usually have a higher amplitude than the amplitude of the brain potentials. Most often artefacts can be registered by many or even all the electrodes attached to the scalp of the examined person. Along with that, the EEG components occurring in a typical EEG recording are usually quite local [6, 7]. Artefact correction is an essential part of pre-processing data to obtain a clean EEG recording [1].

Artefacts may be divided into two groups. The first group includes technical artefacts related to signal

* Corresponding author: m.kaczorowska@pollub.pl

registration. This group includes: inappropriate electrode attachment localisation, insufficient contact of an electrode with the skin, influence of the power network's electromagnetic field, electromagnetic field of outer electronic devices, unsuitable conditions of an examination, e.g. the result of the EEG examination can be affected by walking near an examined person [8].

The second one includes the biological artefacts concerning the motions of the examined person's body. The signal picked up by the electrodes is amplified by a specialised EEG amplifier. This amplifier is a source of noise. A very important parameter of the amplifier is the value of the amplitude of the output noise related to its input [8]. This group includes participants sweating, blinking, head muscle artefacts, muscle shivering and heart's electrical activity. The most popular biological artefacts which can be observed in most EEG recordings are eye-blinking artefacts, eye-movement artefacts and muscle activity artefacts. The disturbances may be observed in one or many electrodes, in different proportions [8, 9].

Artefacts may also appear because of muscle tension caused by movement. Movement creates an electrical potential (EMG – electromyography), which is many times higher than the level of the EEG signal. This kind of artefacts is caused by movements of the limbs, neck muscles, swallowing, tongue movement, or clamping of the jaw. Blinking of the examined person may cause a potential up to an amplitude of 100 μ V [8, 10].

2.2 Parameters for artefact detection

There are several possibilities to classify channels as artefactual. One of the parameter is the mean of the channel correlation coefficients with the other channels. The channel with technical artefacts related to high impedance has a low correlation with other channels. The other parameter, which describes appearance of artefacts, is the variance of the channel. Despite of quite correlation, there is a chance that due to the artefacts, the variance is high. The next parameter is the Hurst exponent. The Hurst exponent is a measure of long-range dependence within a signal. To detect artefacts, it is possible to compute the amplitude range of the epoch. The deviation from each channel's average value is the following parameter. In case if the deviation's value is large, the probability of artefact appearance is high. It is important to choose the appropriate type of measure to reveal the artefacts in the signal [11].

2.3 Methods of artefact correction

There are several methods of artefacts reduction. However, they can never be eliminated completely. The first way is to reduce the movements and muscle tension of the examined person [3].

In the past, the fragment of the signal which contained artefacts was rejected [3]. Nowadays, there are several methods to detect artefacts automatically. One of them is the method called "frequency method". The main idea of artefact identification is to determine

what frequencies dominate in the recorded signal. If the signal frequencies deviate significantly from the frequency of the EEG signal, the classifier selects this fragment as the artefact [6, 12, 13].

The second type of methods of artefact correction are methods based on subtracting the interfering signal from the EEG signal. These methods may be applied in eye movement artefacts. Besides EEG recordings gathered from the scalp, electrooculography (EOG) electrodes are applied to record the electrooculography signal. There is a possibility to subtract the signal including the artefacts originating from the electrooculography activity. It may reduce artefacts in a significant way [9, 14, 15].

The most popular methods of removing artefacts are PCA and ICA [6,16]. This paper contains a comparison between removing artefacts from the signal gathered from a stationary and a mobile amplifier.

3 Applied methods

3.1 Principal Component Analysis

Two methods of correcting artefacts are reviewed in the paper. The first method is PCA, i.e. Principal Component Analysis. It is one of the Factor Analysis Methods. It is typically used to reduce the number of dimensions, eliminating low variety dimensions as they do not contain significant information [16].

The idea of the method may be described in the following way: the signal collected from a certain number of channels n within the time interval of k periods may be understood as a cloud of k points in the n -dimensional space or as a matrix with n rows and k columns (1).

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \quad (1)$$

Clearly, the basis of PCA is matrix calculus. In accordance with it, the following steps should be taken, to obtain a principal components matrix:

1. the mean values for the columns should be calculated;

$$X_{mean} = \frac{1}{n} [\sum_{i=1}^n x_{i1} \quad \sum_{i=1}^n x_{i2} \quad \dots \quad \sum_{i=1}^n x_{ik}] \quad (2)$$

2. the deviation matrix should be computed in accordance with (3):

$$\Delta X = X - \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_n \cdot X_{mean} \quad (3);$$

3. based on the Singular Value Decomposition (SVD) method three matrixes: U , Σ and V^* should be calculated, where:

- U is the left singular matrix with the size of $n \times n$, containing the eigenvectors of the matrix product $\Delta X \Delta X^T$;

- V^* is the conjugate transpose of the matrix containing the eigenvectors of the matrix product $\Delta X^T \Delta X$, its size is $k \times k$;
- Σ is the diagonal matrix containing the singular values of the deviation matrix.

Singular values are arranged in the descending order and represent principal components of the signal. The first principal component corresponds to the most common information contained in the signal, i.e. the part that is present in all the channels. The successive principal components correspond to the progressively more individual details of the signal [17].

As mentioned above, artefacts have higher amplitude compared to EEG signal and therefore, affect all or most of the channels, while brain signals are usually local and can be detected by one electrode or small group of electrodes. Hence, the highest principal component corresponds to the artefacts contained in the signal. That is why one or two first components are usually eliminated i.e. set to zero. The corrected signal is obtained by adding means to the matrix product $U\Sigma'V^*$, where Σ' is the principal component matrix with one or several principal components set to zero [17].

3.2 Independent Component Analysis

The other method that is considered in the paper is ICA, i.e. Independent Component Analysis. The main idea of the method is to divide the signal into a series of components, statistically independent of each other, and having frequency distribution maximally different from normal distribution [16]. Kurtosis is commonly used as the measure of a distribution's normality. Kurtosis of a signal \mathbf{x} may be defined as:

$$Kurt(\mathbf{x}) = E(\mathbf{x}^4) - 3(E(\mathbf{x}^2))^2 \quad (4)$$

Kurtosis may be described as a measure of probability density flatness. The value of kurtosis of a normal distribution equals zero, while for non-Gaussian distribution it may possess the following values:

1. kurtosis is positive, the probability density is more concentrated than in normal distribution,
2. kurtosis is negative, the probability density is less concentrated than in normal distribution [18].

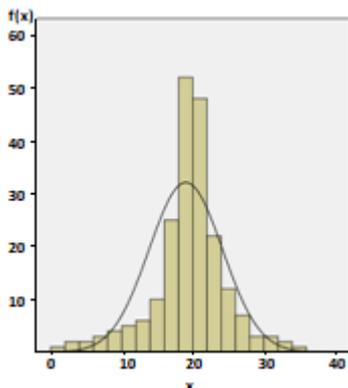


Fig 1a. Close to normal probability distribution [19].

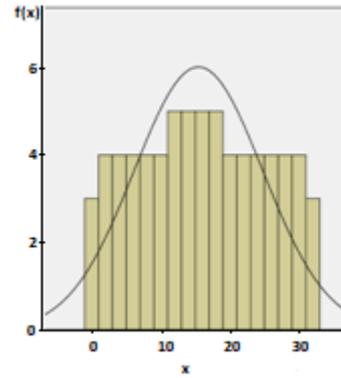


Fig 1b. Probability distribution differing significantly from normal [19].

Figure 1 presents a comparison of the probability density of two variables. Finding an independent component with a maximally non-Gaussian probability density, distribution is based on the maximisation or minimisation of the kurtosis value [19].

ICA is used in many fields, such as face and voice recognition, communication, and even stock price research [20]. Artefacts correction may be also performed by means of ICA. The main feature of the method is that the number of the components, that the signal can be divided into, is not greater than the channel number [16, 18].

3.3 Frequency based method

Another method used, was the frequency based method. The idea of this method is based on the use of various filters, such as lowpass, high-pass and band-stop (notch). The Fig.2 presents the amplitude response of band-stop filter. In the current application, the value of frequency requiring attenuation was equal to 50 Hz.

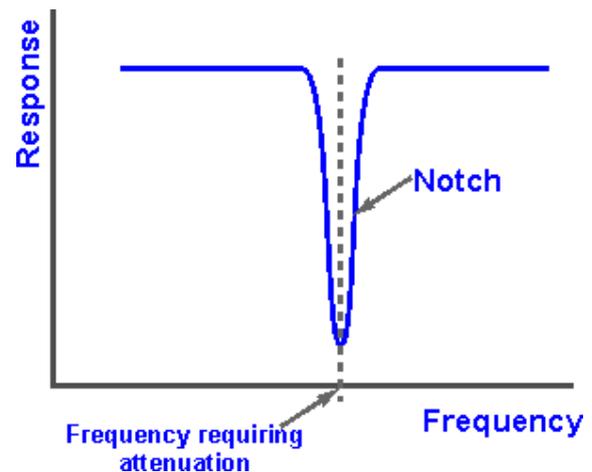


Fig 2. Amplitude response of a notch filter.

4 The case study

Eight participants took part in an experiment, in a quiet room of a dedicated laboratory. The room was illuminated with standard fluorescent light and outside

light was blocked. The participants were of similar age. During the test, the person examined was left alone.

Two 21-channel amplifiers were applied during the experiments. The first of them was the stationary – Mitsar EEG 201 amplifier, and the other – mobile EEG Smart BCI. The EEG amplifiers transmitted data to a computer with native software (WinEEG). The stationary amplifier used USB protocol to communicate with computer, and the mobile amplifier used Bluetooth. Every participant was examined with stationary amplifier first, and then with the mobile one.

The samples were recorded at a frequency of 500 Hz. The signals were gathered from 19 electrodes placed in a special EEG cap, which kept them in the right position. The cap was compatible with international 10-20 EEG standard and located over an examined person's scalp.

The ground electrode was placed in the centre of the frontal lobe and two reference electrodes (A1, A2) were placed on the ears.

The participants were asked to perform the following activities:

- sitting with eyes opened and closed,
- blinking the eyes,
- moving the jaw,
- writing words on a sheet of paper,
- adding up even numbers – cognitive activity data,
- squeezing the hands,
- moving the eyeballs.

Frequency filtration was applied to remove the noisy part of the signal containing low level information. It is the signal which has frequency lower than 1 Hz or higher than 45 Hz, including signals originating from external electronic devices. EEG signals have low amplitude, so quality and high signal-to-noise ratio is very important [8].

EEG signal pre-processing covered signal filtration and artefact elimination. In the study, Principal Component Analysis and Independent Component Analysis were applied to remove the artefacts.

5 Data processing results

Tables 1 and 2 show the amplitude for the raw EEG signal and the signal after elimination of the first component, the first two components and the first three components for PCA and the selected components of ICA algorithms. The tables show also the amplitude of the EEG signal of relaxation state. This value is used as the baseline for other data in the table. It was observed that the increase in the number of the principal components eliminated caused a decrease of the artefact amplitudes. This effect may be explained by the fact that the principal components with the highest values contain the most common part of the signals recorded by the channels, so the common part, i.e. the artefacts, was removed because of removing the highest principal components. The individual information from the parts of the head closest to the corresponding electrodes was left unaffected [9, 17, 21].

Data in Tables 1 and 2 shows that the artefact amplitude is higher after using ICA, than the artefact amplitude of the signal obtained because of using PCA. The values of the amplitudes were measured with dedicated software (WinEEG).

The results show that the amplitude of artefacts decreased after correction. The obtained results, i.e. the artefact amplitude values, deviated from mean amplitude values insignificantly, so mean values are sufficient to demonstrate the results.

Table 1. Mobile amplifier. Amp - amplitude of EEG signal with artefacts, Amp. rel. - amplitude of EEG signal of relaxation state.

Type of artefacts	Amp. [μ V]	PCA, amplitude [μ V]			ICA, amp. [μ V]	Amp. rel. [μ V]
		1.	2.	3.		
Eyes blinking	150	15	4	4	30	8.5
Jaw moving	220	26	22	18	50	
Eyeballs moving	140	16	9	7.5	90	
Hands squeezing	20	14	8	4.5	17	
Writing words on the sheet of paper	14	12	6	5	9	

Table 2. Stationary amplifier. Amp - amplitude of EEG signal with artefacts, Amp. rel. - amplitude of EEG signal of relaxation state.

Type of artefacts	Amp. [μ V]	PCA, amplitude [μ V]			ICA, amp. [μ V]	Amp. rel. [μ V]
		1.	2.	3.		
Eyes blinking	170	15	13	13	90	11
Jaw moving	380	80	60	50	130	
Eyeballs moving	100	60	24	13	40	
Hands squeezing	40	20	20	9	25	
Writing words on the sheet of paper	22	18	18	8	18	

As it is known, the ICA method operates in a different way. It does not separate the signal into parts describing progressively less common information of the signal, but rather divides the signal (being a mix of independent signals) into independent components. On the basis of that feature, it is theoretically not enough to eliminate several components with the highest values. The components corresponding to particular artefacts should be found and removed [15]. The research confirmed this statement: the average number of rejected independent components was significantly higher than in the case of PCA [12, 17].

One of the measures presented in the section 2.2 was used for artefact detection. This measure was channel potential variance. The mean values of separate

electrode potential variance among all the participants for all types of biological artefacts mentioned in the paper were calculated. The tables 3, 4 and 5 present the selected mean values of separate electrode potential variances before and after PCA. On the basis of decreased value of variance and performed visual inspection, it was concluded that the artefacts were successfully eliminated to a high degree. The electrodes presented in the tables were chosen due to the fact that the variance value was the highest for these electrodes during presence of specific artefacts, i.e.: eyes blinking, eyeballs moving and jaw moving. As data in the tables show, artefacts affect less the record carried out by the mobile amplifier. That conclusion was made on the basis of comparison of variance values of signals recorded by stationary and mobile amplifiers in the Tables 3-5.

Table 3. Eyes blinking.

Electrodes	Before		After	
	Stationary	Mobile	Stationary	Mobile
Fp1	2041.32	715.03	7.92	4.86
Fp2	2041.32	932.99	23.79	9.94
F7	2152.57	399.01	33.03	6.59

Table 4. Eyeballs moving.

Electrodes	Before		After	
	Stationary	Mobile	Stationary	Mobile
Fp1	867.86	512.04	12.51	7.63
Fp2	813.06	528.49	10.14	5.88
F7	3402.70	2075.65	33.33	21.33
F8	2868.24	1720.94	18.21	10.01

Table 5. Jaw moving.

Electrodes	Before		After	
	Stationary	Mobile	Stationary	Mobile
F3	1939.47	1635.90	238.72	28.43
F4	3625.23	378.58	188.29	39.33
T3	3032.03	2393.03	206.31	59.89
C3	2041.64	1138.21	222.78	28.95
T4	3155.57	1770.70	271.70	62.89

The Figures 3-7 present the EEG recordings with artefacts of various types that were registered in the process of the research.

The Fig. 3 presents the example of frequency based method: the signal spectrum obtained by the means of Fourier Transform clearly indicates presence of undesired frequency component.

The muscle and eye-related artefacts have a very characteristic shape in the records of both amplifiers, but as Tables 1 and 2 show, the artefacts in the signal recorded by the mobile EEG demonstrate a considerably lower magnitude.

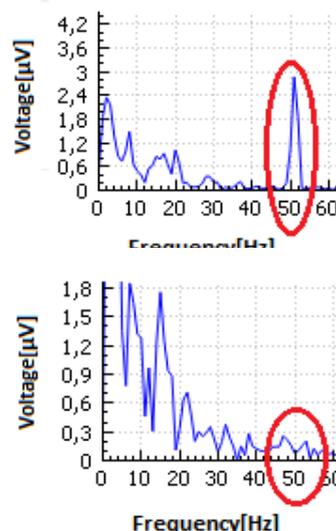


Fig 3. Technical artefacts related to a 50 Hz network (Fourier transform – frequency domain). Top – stationary, bottom – mobile.

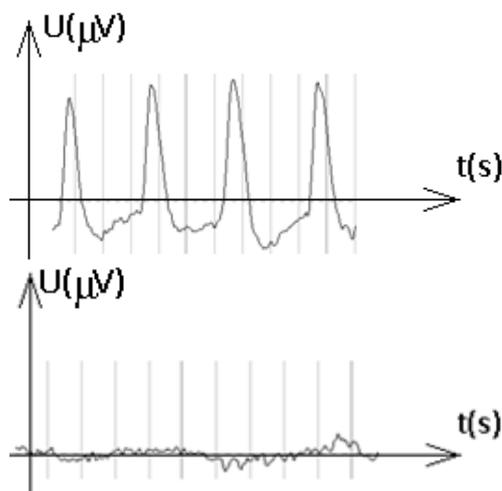


Fig 4. Eyes blinking. Top – original, bottom – corrected.

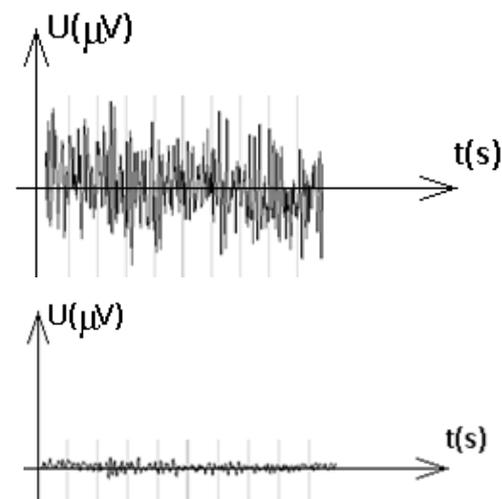


Fig 5. Jaw moving. Top – original, bottom – corrected.

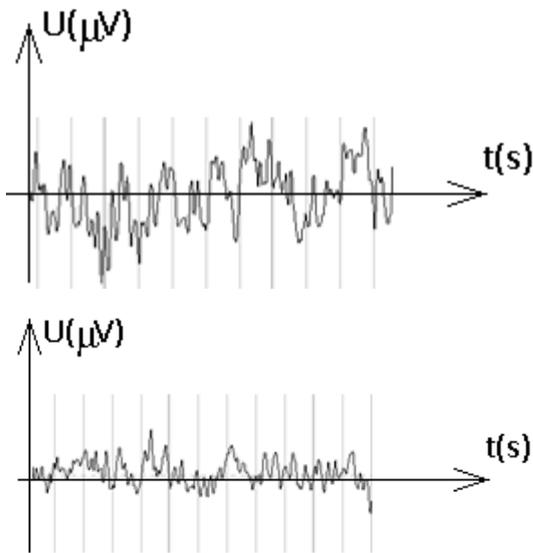


Fig 6. Hands squeezing. Top – original, bottom – corrected.

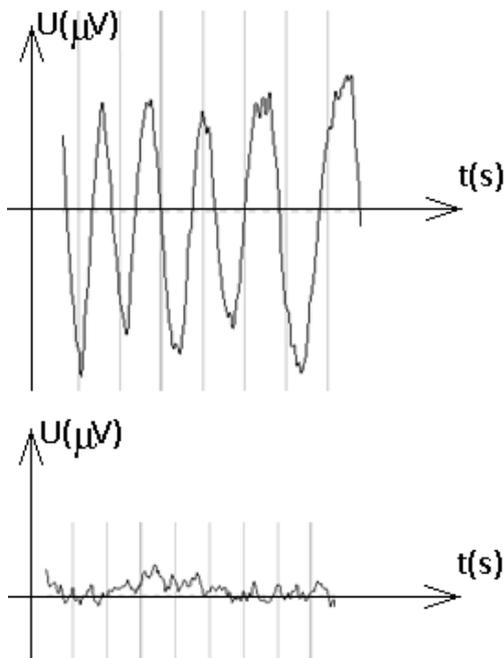


Fig 7. Eyeballs moving. Top – original, bottom – corrected.

Another type of artefacts, being worth mentioning and necessary to discuss, are technical artefacts. After comparing the records performed by the two types of electroencephalograph, it may clearly be seen that the mobile EEG is significantly less sensitive to the effect of the power network with the frequency of 50 Hz. As it is known, the technical artefacts of that kind are caused by the potentials induced in the long cables used to transmit the potential sensed by the electrodes placed on the scalp of the person examined to the EEG [8]. The length of cables used in combination with stationary EEG usually equals to 2 m, while the cables used for the mobile encephalograph are considerably shorter (20 cm), so the length of cables is not enough for the noisy 50 Hz potentials to be generated.

The artefacts in the signals gathered from standard and mobile amplifier were compared. Data are

presented in Tables 1 and 2. The research has shown that jaw mowing generated the artefacts with the highest amplitude. The artefacts of the lowest amplitudes in the research were generated during the part covering hand writing. The said effects have been observed on both amplifiers. When comparing the muscle artefacts in the signals from the two amplifiers, it was noticed that the signal recorded by the mobile amplifier contained the muscle artefacts of a significantly (up to 1/3) lower amplitude. The influence of 50 Hz power network has been proven to be considerably lower on the mobile electroencephalograph.

6 Conclusions

The goal of the present paper was to compare two methods of artefact correction and two electroencephalographs. The research has shown that the PCA method is more universal and simpler in use, since the artefacts are usually visible at all the electrodes, and originate from one or two sources, so it is usually enough to eliminate one or two principal components [8]. The ICA method divides the signal into a series of independent components, their number is equal to the number of the channels [16]. However, the identification of a component as the artefact is more complicated. The localisation of the component, as well as its character, should also be analysed [15]. Another inconvenience was the fact that it was necessary to delete a comparatively larger number of components (6 IC compared to 1-2 PC).

A comparison of the two electroencephalographs has shown that the mobile encephalograph deals better with the artefacts of both biological and technological origin. The results obtained prove that the mobile EEG better meets the needs of BCI applications, as the signal recorded by this type of EEG contains a considerably lower number of biological artefacts and a significantly decreased number of technological artefacts related to the power network with the frequency of 50 Hz [5]. These features might facilitate achieving the following results:

1. the performance requirements for the system based on the mobile EEG would be lower, since it enables to process the obtained signal consuming less performance resources (simpler artefact correction method, absence of notch filter);
2. better possibilities to process data in the frequency domain in the range of higher frequencies (40+ Hz), which would be achieved thanks to the absence of the notch filter. That would allow to monitor the gamma component of brain waves, the significant part of which (the band between 45 and 65 Hz) would be thrown away if the notch filter was used.

References

1. Lopes da Silva F. H., Storm van Leeuwen W., Remond A., *Handbook of electroencephalography and clinical neurophysiology*. Revised series 1986; vol. 2. Amsterdam: Elsevier; pp. 15-62, (1986)

2. M. Plechawska-Wojcik, M. Kaczorowska, *10th International Technology, Education and Development Conference (INTEDE)*, pp. 844-852, (2016)
3. C. Jutten, J. Herault, *Signal Processing*, 24: pp. 1-10,(1991)
4. M. Plechawska-Wojcik, P. Wolszczak, R. Cechowicz, *9th International Conference on Human System Interactions (HSI)*, pp. 348-354, (2016)
5. M. Plechawska-Wojcik, P. Wolszczak, *12th International Scientific Conference on Beyond Databases, Architectures and Structures (BDAS)*, pp. 485-496, (2016)
6. Z. Zhang, H. Li, D. Mandic, *2016 IEEE 13th International Conference*, pp. 177-182, (2016)
7. N. C. Rogasch, R. H. Thomson, F. Farzan, B. M. Fitzgibbon, Neil. W. Bailey, J. C. Hernandez - Pavon, Z. J. Daskalakis, Paul B. Fitzgerald, Elsevier, pp. 425-439, (2014)
8. J. A. Uriguen, B. Garcia-Zapirain, *Journal of Neural Engineering* Vol. **12**, pp. 342-348, (2015)
9. Y. Li, Z. Ma, W. Lu, Y. Li, *Physiological Measurement*, 27 (4), (2006)
10. B. Singh, H. Wagatsuma, *Computational and Mathematical Methods in Medicine*, pp. 612-620, (2017)
11. H.Nolan, R.Whelan, R.B. Reilly, Elsevier pp. 152-162, (2010)
12. A. Cichocki, S. Amari, *Learning Algorithms and Applications*, Wiley, pp. 106-111, (2003)
13. S. Motamedi-Fakhr, M. Moshrefi-Torbati, M. Hill, C. M. Hill, P. R. White, Elsevier pp. 21-33,(2014).
14. Y. Ding, I. W. Selesnick, Elsevier, pp. 236-248, (2016)
15. G. Barbati, C. Porcaro, F. Zappasodi, P. M. Rossini, F. Tecchio, *Clin. Neurophysiology*, pp. 1220-1232,(2004)
16. M. Kaczorowska, M. Plechawska-Wojcik, M.Tokovarov, R. Dmytruk, *International Symposium on Advanced Topics in Electrical Engineering (ATEE) 2017*, pp. 262-267, (2017)
17. L. I. Smith: *A tutorial of Principal Components Analysis*, pp 218-223, (2002)
18. A. J. Bell, T. Sejnowski, *Neural Computation*,: pp. 1129-2259, (1995).
19. M. Kołodziej, Warsaw University of Technology, pp. 56-60, (2011)
20. J. Koronacki, WNT, pp. 250-256, (2005)
21. W. J. Krzanowski, *Oxford University Press*, (2000)