Evaluation of Time-Domain Features for Motor Imagery Movements using FCM and SVM

Conference Paper · July 2012

DOI: 10.1109/JCSSE.2012.6261918

Aida Khorshidtalab

Momoh Salami

International Islamic University Malaysia

Mahyar Hamedi

Evaluation of Time-Domain Features for Motor Imagery Movements using FCM and SVM

A. Khorshidtalab#1, M. J. E. Salami#2, M. Hamedi*3

#Department of Mechatronics Engineering
International Islamic University Malaysia, Gombak, Malaysia
1aida.khorshid@student.iium.edu.my
2momoh@iium.edu.my

* Faculty of Biomedical and Health Science Engineering
Universiti Teknologi Malaysia, Johor Bahru, Malaysia
3hamedi.mahyar@gmail.com

Abstract— **Brain–Machine Interface is a direct communication pathway between brain and an external electronic device. BMIs aim to translate brain activities into control commands. To design a system that translates brain waves and its activities to desired commands, motor imagery tasks classification is the core part. Classification accuracy not only depends on how capable the classifier is but also it is about the input data. Feature extraction is to highlight the properties of signal that make it distinct from the signal of the other mental tasks. Performance of BMIs directly depends on the effectiveness of the feature extraction and classification algorithms. If a feature provides large interclass difference for different classes, the applied classifier exhibits a better performance.**

In order to attain less computational complexity, five time-domain procedure, namely: Mean Absolute Value, Maximum peak value, Simple Square Integral, Willison Amplitude, and Waveform Length are used for feature extraction of EEG signals. Two classifiers are applied to assess the performance of each feature-subject. SVM with polynomial kernel is one of the applied nonlinear classifier and supervised FCM is the other one. The performance of each feature for input data are evaluated with both classifiers and classification accuracy is the considered common comparison parameter.

Keywords-Electroencephalogram; Feature extraction; Motor imagery; Brain-Machine Interface; Support Vector Machine;Fuzzy C-Means.

I. INTRODUCTION

Brain machine interfaces (BMI) aim to translate brain activities into control commands. Not only BMIs can remarkably improve quality of life of those with neuromuscular difficulties, but also it opens a wide variety of new ideas of communication and control in a new dimension for people with normal abilities. A BMI replaces the usage of muscles and nerves produced through ordinary movements with electrophysiological signal produced by brain, through brain activities, by merely thinking about the similar movement. Brain activities for motor imagery movements are generally measured in the form of Electroencephalography (EEG) signals. BMIs are mostly designed based on pattern recognition approaches, such as feature extraction and classification, to identify and differentiate different mental tasks [1-2].

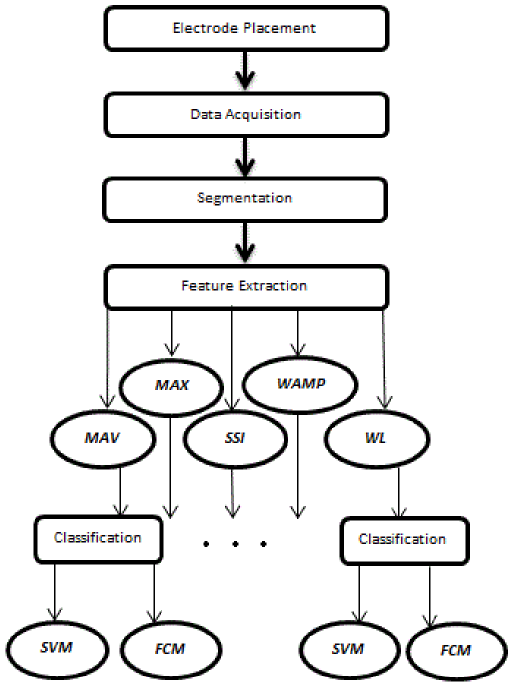

Figure 1.The processing block diagram

To convert the raw signal recorded from one's scalp to desired commands EEG data should be processed. EEG signals are one of the most challenging bio-signals in term of processing due to their nature. EEG signals are known by their poor signal to noise ratio and their high dimensionality. Moreover, their non-stationary characteristics and rapid variation over time and over sessions of recording pose some difficulties.

Noise reduction and segmentation are the two main stages of preprocessing and processing mostly consists of feature extraction and then either classification or regression. However, in many researches dimension reduction has been mentioned a necessity of an effective processing [19-20/3-4].

Feature extraction is to highlight the properties of signal that make it distinct from others. Numbers of different algorithms using varied dimensions, complexity, and efficiency have been suggested for motor imagery EEG signals [5-6-7].

After feature extraction, there could be other stages before classification named feature selection or dimension reduction. In case many extracted features are redundant or not enough discriminant or if numbers of features are too many to be processed, dimension reduction and feature selection are recommended.

There are plenty of researchers who focus their efforts on selecting the most discriminant data which would enable classifiers to produce more accurate results. For motor imagery data, classifiers mostly identify different patterns of brain activities [8]. Hence, a BMI system can be considered partly as a pattern recognition system [9-10-11]. Performance of pattern recognition systems directly depends on the effectiveness of feature extraction and classification algorithms. The general idea of this paper is depicted in figure 1.

The organization of this paper is as follows. In the next section, EEG signal acquisition is explained. Section 3, is about feature extraction methods. Section 4, is for the applied classifiers. Results, discussion and conclusion are presented in section 5, 6 respectively.

II. EEG Acquisition

EEG signals are recorded from multiple electrodes placed on the subject's scalp, resulting in multichannel time series. Multichannel EEG signals typically have low signal to noise ratio giving a rather blurred image of brain activities [12]. In our experiment, three electrodes, known as C_3, C_4, C_z were located on the subject's scalp based on the International 10-20 electrode placement system. EEG signals were recorded via g.tec equipment, which is known to be one of the most accurate with high resolution devices available for recording bio-signals [13].Signals were recorded at the rate of 512Hz and a combination of high pass and low pass filters, from the g.tec software, was applied to the recorded data. Besides applied mentioned filters, no more filters were applied to recorded signal. Also, there was no eliminating or ignoring signal partially.

Data was collected from ten mentally healthy people of which six of them are male. Subjects were in the range of 20 to 36 years old and all of them were university students.

During recording, there were no actual movements. In fact, the subjects were asked to replace the desired movement with imagination of related movements. Subjects were asked to think about the movement of their right hand, left hand, movement of their tongue in the right side and in their left side of their mouth.

In the recording environment, only intense sound disturbances were avoided. There also has been no use of bio feedbacks to help subjects to perform these thinking tasks better. As it has been mentioned in several papers that biofeedback could provide a remarkable aid for subjects to have better performances [14-15-16-17-18].

Each record took one minute and imagination of the same movement is recorded three times. Each trial is divided to hundred segments, each segment's length is two hundred fifty five milliseconds, and segments do not have overlap.

A feature value is extracted from each data segment. Hence, one hundred feature values would be obtained from the whole signal. Each class has three related signals which are acquired from three afore mentioned channels. Therefore, for each class three signals and from each signal hundred values as feature, which means for each class three hundred values, known as feature, were obtained for further processing, figure 3.

III. Feature Extraction

Feature extraction from EEG signals is carried out using the five afore mentioned techniques which are subsequently discussed.

A. Mean Absolute Value

Mean absolute value (MAV) estimates the mean absolute value of each segment by adding the absolute value of all the values x_i –ith point, the current point, of signal x- and dividing it by the length of the segment ,figure 2, so that:

$$MAV_K = \frac{1}{N}\sum_{i=1}^{N}|x_i| \tag{1}$$

The extracted feature, MAV, from one channel of left hand signal is represented in figure 3. Figure 4 depicts the classification accuracy of SVM with black bars and FCM with gray bars where MAV is the applied feature.

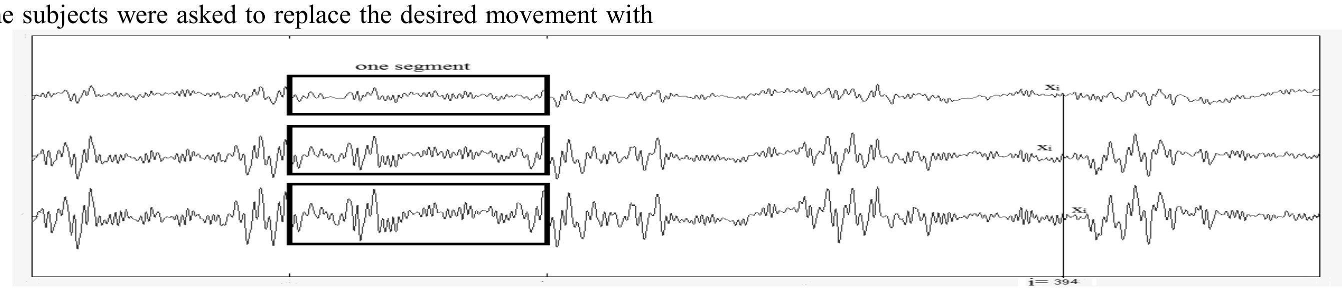

Figure 2.Channel C_z (first top), C_3 (second top), and C_4 (down), X_i where i=394 and one segment of each channel of signal

B. Maximum Value

Maximum peak value (MAX) refers to the maximum absolute value of each considered segment, that is:

$$x_k = \max|x_i| \tag{2}$$

Figure 5 depicts the classification accuracy of SVM with black bars and FCM with gray bars, where MAX is the applied feature.

C. Simple Square Integral

Simple Square Integral (SSI) calculates the energy of EEG signal according to:

$$SSI_K = \sum_{i=0}^{N}(|x_i^2|) \tag{3}$$

Figure 6 depicts the classification accuracy of SVM with black bars and FCM with gray bars for SSI.

D. Willison Amplitude

Willison Amplitud (WAMP) counts the number of times that absolute value of difference between EEG signal amplitude of two consecutive samples exceeds a predetermined threshold value.

$$WAMP_K = \sum_{i=1}^{N-1} f(|x_i - x_{i+1}|) \tag{4}$$

$$f(x) = \begin{cases} 1 & x > \varepsilon \\ 0 & otherwise \end{cases}$$

Figure 7 depicts the classification accuracy of SVM with black bars and FCM with gray bars, where WAMP is the applied feature.

E. Waveform length

Waveform Length (WL) is the cumulative length of the waveform over the segment. It indicates a measure of waveform amplitude, frequency and duration all within a single parameter.

$$WL_k = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{5}$$

Figure 8 depicts the classification accuracy of SVM with black bars and FCM with gray bars, where WL is the applied feature.

N: is the length of segment, k: is the current segment, X_i: is the current point of signal, and i: is the index of current point in all mentioned equations.

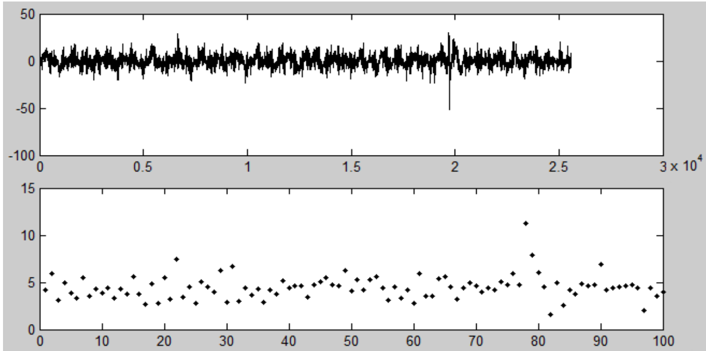

Figure 3.Channel C_3 signal of left hand movement (up) and MAV extracted feature (down)

IV. CLASSIFICATION

Support Vector Machine (SVM) and Fuzzy C-Means (FCM) are applied to data to evaluate the efficiency of mentioned features so as to provide a clear vision of how capable these features are at separating these four mental tasks. For making both classifiers each subject contributes and the evaluation also is based on the data from all subjects.

A. Support Vectore Machine

SVM is one of the widely used and capable classifiers with low complexity compare to neural network or fuzzy classifiers. SVM uses discriminant hyperplane to identify different classes of data. Selected hyperplane is the one that maximizes the margins. Maximizing the margins is known to increase the generalization capabilities. SVM uses a regularization parameter C that enables accommodation to outliers and allows error on the training set.

With small increase of the classifier's complexity, linear SVM can make nonlinear decision boundaries by using "kernel trick". Generally, it is done by mapping the data to another space, mostly of much higher dimensionality, with help of kernel function.

SVM has several advantages due to its margin maximization and regularization term. SVM is known to have good generalization properties. It is also insensitive to overtraining and dimensionality. It has a few hyperparameters that need to be defined manually such as, the regularization parameter, C [19]. In our work, polynomial kernel function has been applied; as in many cases it could obtain similar results with Gaussian SVM but for some feature-subject it carried out better results. To have a fair assessment for all subjects and methods, polynomial kernel is applied and all other parameters of classifier remained unchanged for each time of processing.

B. Fuzzy C-Means

Fuzzy C-means is a method of clustering that recently has been applied for classification of various kinds of biomedical signals in the fields of Human-Machine Interface (HMI) and BMI [20]. This method tends to be more applicable for detecting patterns of data which the other methods face difficulties detecting them. Fuzzy clustering methods like FCM are more precise than crisp ones. Since a total commitment of a vector to a given class is not essential for each iteration. The flexibility of FCM is defined by the membership value of each data point in feature space in the interval of (0, 1) and it allows data to belong to two or more clusters. The advantage of a fuzzy model is that all the variables are continuous and differentiable. Therefore, the formulated problem is easier to be solved and the cost of computation is less. In the current work, the supervised version of FCM is applied to create a suitable direction during training procedure.

As in this work four different imagery movements need to be classified, the numbers of clusters is given as four. By applying the extracted features as input data and adjusting all initial parameters, FCM determines the position of each cluster center and the membership value of each feature to each cluster. This procedure is repeated until the optimized value of each cluster center and the optimized membership value of each point is achieved.

V. RESULTS AND DISCUSION

In this section, result of the assessment performed to portray the recognition ability of each combination of mentioned feature and classifier is presented. Performances of five time-domain features are evaluated with two different classifiers, SVM and FCM.

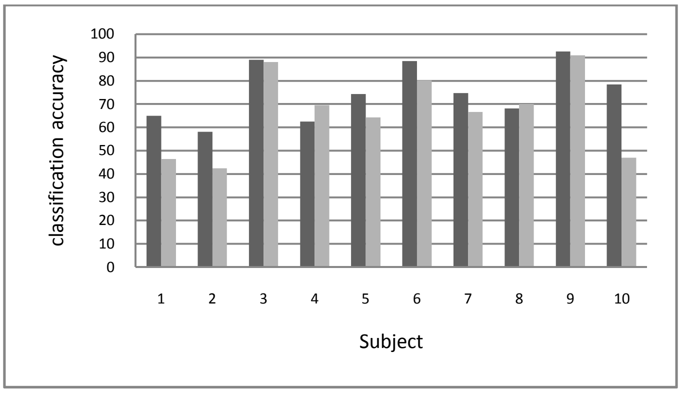

Figure 4.Evaluation of classification accuracy with SVM and FCM respectively for Mean Absolute Value as feature

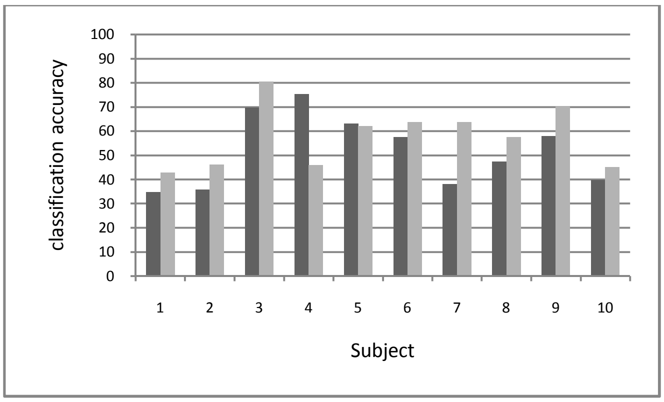

Figure 5.Evaluation of classification accuracy with SVM and FCM respectively for max value as feature

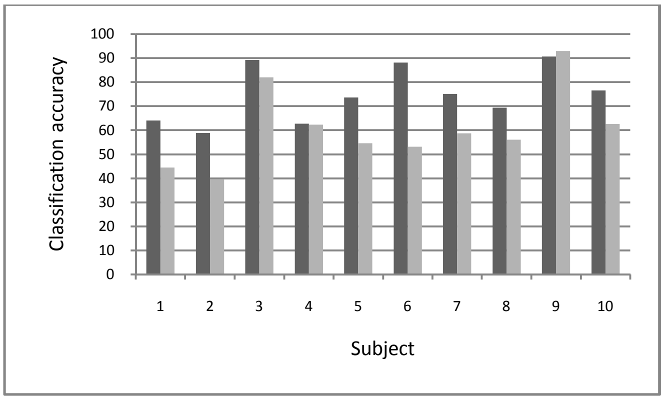

Figure 6.Evaluation of classification accuracy with SVM and FCM respectively for Simple Square Integral as feature

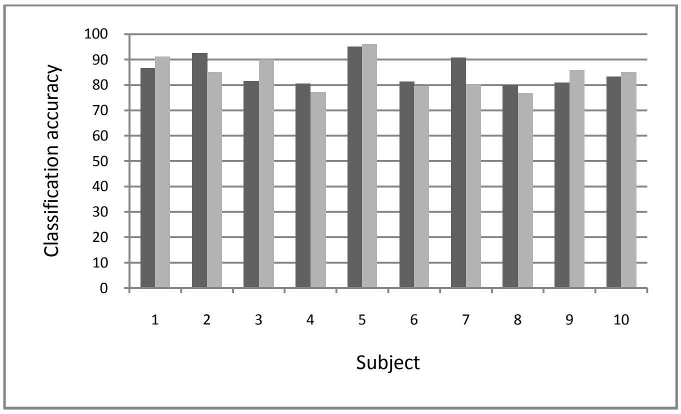

Figure 7.Evaluation of classification accuracy with SVM and FCM respectively for Willison Amplitude as feature

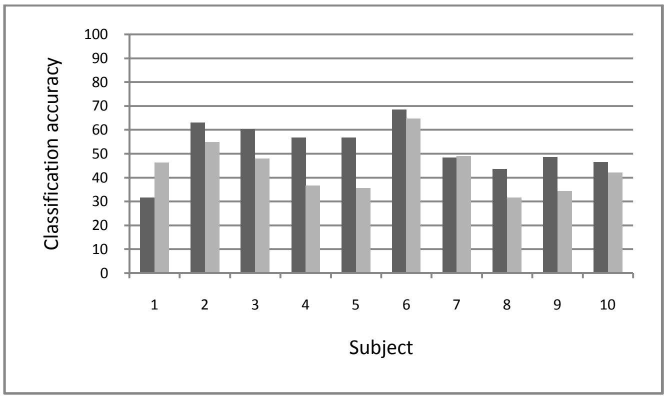

Figure 8.Evaluation of classification accuracy with SVM and FCM respectively for Waveform Length as feature

The obtained results show that in average, SVM and FCM have more or less similar capability in recognition and differentiation of these mental tasks. As this paper is not into comparing the results delivered by FCM and SVM, just some points are highlighted to provide a clearer vision of these applied features' capability. In term of complexity, FCM is more complicated as it consumed considerable more time to deliver the final results compare to SVM. For four out of five features-subjects, FCM had the minimum classification accuracy. However, it could gain three out of five maximum classification accuracies. It shows that variation of the delivered results by FCM is more comparing to SVM. Although they both had similar results, most of the maximum and minimum values belong to FCM.

In term of subjects, interesting results have been obtained by both classifiers. Subject two and subject nine are two subjects with quite different performances. Subject two has four of the minimum performances for features MAV and SSI with both FCM and SVM classifiers. What makes the result more notable is that subject nine has four of the highest performances exactly for the same combination of features-classifiers; which means, while MAV and SSI could be defined as two of the weakest features for subject two, they could prove themselves as two of the most powerful features for subject nine.

Subject one and eight have three of the minimum performances and subject three, four, five, and six have at least one of the highest classification accuracy per feature. Subject seven and ten have neither the highest nor the lowest accuracy and these two subjects performed just as average. Another point about subjects is that none of the subjects have the highest performance for one feature and at the same time lowest among all for another feature. All these points indicate that how subject specific EEG signals are.

All the subjects have at least one of their performances with more than 80% accuracy; plus, six out of ten subjects have at least one of their performances with more than 90% accuracy. These results show that for BMI applications that can tolerate a few errors, these feature-classifiers are suitable.

The best result was delivered by WAMP and SVM with average of 85.07% accuracy for ten subjects with standard deviation 5.7 percentage. While the worst one belongs to WL and SVM with average of 44.33% accuracy with standard deviation 10.31percentage, figure 9.

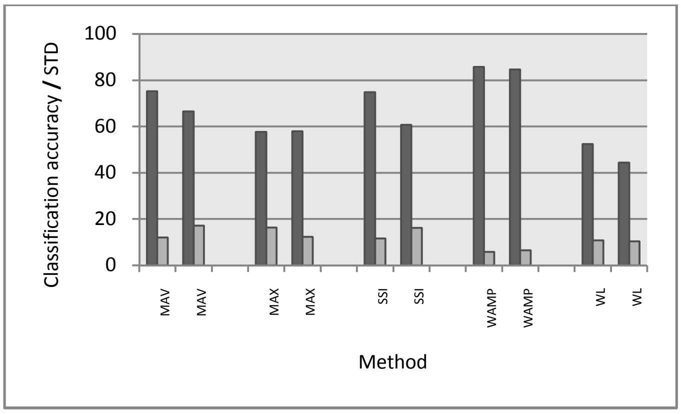

Figure 9.Average of classification accuracy and standard deviation of ten subjects with SVM and FCM respectively

Table 1 shows the classification accuracy for the classifiers by applying the mentioned features for each subject. In order to have more accurate results, five times each class of data were processed and each time, test data and training data were different sets. Therefore, we obtained five different classification accuracy values per each subject-method. The value presented in table 1 for each subject-method is actually the mean value of these five times evaluation. In the table 1, values with one star are minimum values and those values with two stars are maximum values. However maximum and minimum of each combination of feature-classifier has been mentioned in the last two rows of table 1, these stars would help to track the performance of subject easier.

Regarding table1, small standard deviation indicates that regardless of feature ability in distinction, for different people EEG signal, feature performs similarly. On the contrary, a large standard deviation exhibits that feature is not robust and it had difficulty dealing with chaotic behavior of EEG signal. Obviously, the best feature is the one with the highest mean value and lowest standard deviation value.

	MAV		**MAX**		**SSI**		**WAMP**		**WL**			
	FCM	**SVM**	**FCM**	**SVM**	**FCM**	**SVM**	**FCM**	**SVM**	**FCM**	**SVM**	***Mean***	***STD***
Subject 1	46.46	64.99	42.86*	34.83*	44.44	63.99	91	86.62	46.24	31.67*	55.31	20.64
Subject 2	42.43*	58.15*	46.28	35.83	39.5*	58.82*	85	92.5	54.87	62.98	57.63	18.75
Subject 3	88	89	80.41**	69.83	82	89.15	90	81.4	48	60.32	77.81	14.14
Subject 4	69.5	62.49	46	75.16**	62.32	62.67	77	80.37	36.78	56.83	62.91	13.76
Subject 5	64.29	74.34	62	63	54.55	73.6	96**	95.1**	35.71	56.7	67.52	18.30
Subject 6	80.21	88.5	63.64	57.49	53.13	88.15	79.43	81.32	64.6**	68.52**	72.5	12.66
Subject 7	66.67	74.68	63.64	38.16	58.7	75.15	80	90.7	48.98	48.3	64.49	16.26
Subject 8	70	68.15	57.45	47.48	56.12	69.39	76.77*	79.82*	31.63*	43.62	60.04	15.56
Subject 9	91**	92.64**	70	57.84	92.93**	90.67**	85.86	80.86	34.38	48.49	74.47	20.96
Subject 10	46.97	78.47	45	39.7	62.63	76.6	85	83.2	42.11	46.5	60.62	18.53
Mean	*66.54*	*75.14*	*57.93*	*57.63*	*60.63*	*74.81*	*84.6*	*85.07*	*44.33*	*52.39*		
STD	*17.14*	*11.97*	*12.3*	*16.24*	*16.1*	*11.5*	*6.4*	*5.7*	*10.31*	*10.78*		
Minimum	*42.43*	*58.15*	*42.86*	*34.83*	*39.5*	*58.82*	*76.77*	*79.82*	*31.63*	*31.67*		
Maximum	*91*	*92.64*	*80.41*	*75.16*	*92.93*	*90.67*	*96*	*95.1*	*64.6*	*68.52*		

Table 1: Classification accuracy, mean value and standard deviation of five features using SVM and FCM

VI. CONCLUSION

In the transformation of raw signal to feature vector, feature extraction highlights important data and eliminates redundant data which causes dimensionality reduction. The best collection of subsets is the one that minimize the probability of misclassification. In this paper, SVM and FCM were used for classification in order to assess the ability of mentioned time-domain features. However, in details, there are variation based on the subject's and features' ability for classifiers, the total performance shows that, the ability of SVM in classification of EEG data is better than FCM. Besides, in term of consumed timed by both classifiers, SVM needed much less time compare to FCM.

Considering some of the characteristics of this investigation would help to understand the ability of these features-classifiers better. No use of any filters for noise reduction, No elimination or rejection of signal partly, no use of artifact removal or thresholding and last but not least, no use of any biofeedback show that how robust these features-classifiers could perform, as all the mentioned additional processing help for better performances. The tradeoff between computational complexity and speed, in addition to the importance of real-time BMI systems was the main motivation for eliminating all the mentioned preprocessing steps. Thus the computationally burden for a BMI system is reduced.

Combining these time-domain features or a selection of two or more of these features and trying feature selection methods can be considered as future work. Another possibility is to examine fusion classifiers with these time-domain features for the best possible result.

REFERENCES

[1] M. Grosse-Wentrup, C. Liefhold, K. Gramann, and M. Buss. "Beam forming in noninvasive brain computer interfaces." IEEE Trans.on Biomed. Eng., 56(4):1209 –1219, 2009.

[2] G. Pfurtscheller and C. Neuper. "Motor imagery and direct brain-computer communication." proc. of the IEEE, 89(7):1123–1134, 2001.

[3] L. J. Cao, K. S. Chua, W. K. Chong, H. P. Lee, and Q. M. Gu, "A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine," *Neurocomputing*, vol. 55, no. 1-2, pp. 321-336, Sep. 2003.

[4] A. Widodo and B.-S. Yang, "Application of nonlinear feature extraction and support vector machines for fault diagnosis of induction motors," *Expert Systems with Applications*, vol. 33, no. 1, pp. 241-250, Jul. 2007.

[5] Z. Iscan, Z. Dokur, and T Demiralp. "Classification of electroencephalogram signals with combined time and frequency featurs. Expert Systems with Applications," Elsevier, 2001.

[6] C. Vidaurre, N. Kramer, B. Blankertz, and A. Schlogl. "Time domain parameters as a feature for EEG-based brain-computer interfaces," Neural Networks, Elsevier, 22(9): 1313-1319, 2009.

[7] S.M. Zhou, J. Q. Gan, and F.Sepulveda. "Classifying mental tasks based on features of higher-order statistics from eeg signals in brain-computer interface," Information Sciences, Elsevier, 178(6):1629-1640, 2008.

[8] D. J. McFarland, C. W. Anderson, K.-R. Muller, A. Schlogl, and D. J. Krusienski. "BCI meeting 2005-workshop on bci signal processing:feature extraction and translation," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 14(2):135 – 138, 2006.

[9] A.Vallabhaneni, T Wang, and B. He. "Brain computer interface,". In He B (Ed): Neural Engineering, Kluwer Academic / Plenum Publishers, pp 85-122, 2005.

[10] D. J. McFarland and J. R. Wolpaw. "Sensorimotor rhythm-based brain-computer interface(bci):feature selection by regression improves performance," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 13(3):372-379, 2005.

[11] A. k. Jain, R. p. W. Duin, and J Mao. Statistical pattern Analysis and Machine Intelligence, 22(1):4-37, 2000.

[12] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Muller, "Optimizing spatial filters for robust EEG single-trial analysis," IEEE Signal Processing Mag., vol. 25, no. 1, pp. 41–56, Jan. 2008.

[13] http://www.eniac-csi.org/CSI/node/21

[14] A. Chatterjee, V. Aggarwal, A. Ramos, S. Acharya, and N.V. Thakor." A brain-computer interface with vibrotactile biofeedback for haptic information," Journal of Neuro Engineering and rehabilitation, 4(40),2007.

[15] F. Cincotti, L. Kauhanene, F. Aloise, T. Palomaki, N. Caporusso, P. Jylanki, D. Mattia, F. Babiloni, G. Vanacker, M. Nuttin, M.G. Marciani, and J. delR. Millan. "Vibrotactile feedback for brian-computer interface operation," Computational Intteligence and Neuroscience,2007,2007.

[16] T. Hinterberger, N. Neumann, M. Pham, A. Kubler, A. Grether, N. Hofmayer, B. Willhelm, H. Flor, and N. Birbaumer. "A multimodal brain-based feedback and communication system," Journal Experimental Brain Research, 154(4):521-526, 2004.

[17] L. Kauhanen, T Palomaki, P. Jylanki, F. Aloise, M. Nuttin, and J. del R. Millan. "Haptic feedback compared with visual feedback for BCI," In Processings of the 3rd International BCI Workshop and Traning Course 2006, Graz Austria, pp 66-67, 2006.

[18] J. R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, and T.M. Vaughan. "Brain-computer interfaces for communication and control," Clinical Neurophysiology, 113(6):767-791,2002.

[19] C. J. C. Burges. A tutorial on support vector machines for pattern recognition. Knowledge Discovery and Data Mining, 2, 1998.

[20] M. Hamedi, Sh-Hussain Salleh, T.S. Tan, K. Ismail, J. Ali, C. Dee-Uam, C. Pavaganun, and P.P. Yupapin, " Human Facial Neural Activities and Gesture Recognition for Machine-Interfacing applications. ", International Journal of Nanomedicine, 2011:6 3461-3472.