Expert Systems with Applications 38 (2011) 10499–10505

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Classification of electroencephalogram signals with combined time and frequency features

Zafer Iscan [a,*], Zümray Dokur [a], Tamer Demiralp [b]

[a] Department of Electronics and Communication Engineering, Istanbul Technical University, 34469 Maslak, Istanbul, Turkey
[b] Department of Physiology, Istanbul Medical Faculty, Istanbul University, 34390 Capa, Istanbul, Turkey

ARTICLE INFO

Keywords:
Electroencephalogram
Power spectral density
Cross correlation
Least Squares Support Vector Machine
Classification
Divergence analysis

ABSTRACT

Epilepsy is a neurological disorder that causes people to have seizures and the main application field of electroencephalography. In this study, combined time and frequency features approach for the classification of healthy and epileptic electroencephalogram (EEG) signals is proposed. Features in the time domain are extracted using the cross correlation (CC) method. Features related to the frequency domain are extracted by calculating the power spectral density (PSD). In the study, these individual time and frequency features are considered to carry complementary information about the nature of the EEG itself. By using divergence analysis, distributions of the feature vectors in the feature space are quantitatively measured. As a result, using the combination rather than individual feature vectors is suggested for classification. In order to show the efficiency of this approach, first of all, the classification performances of the time and frequency based feature vectors in terms of overall accuracy are analyzed individually. Afterwards, the feature vectors obtained by the combination of the individual feature vectors are used in classification. The results achieved by different classifier structures are given. Obtained performances in the study are comparatively evaluated by the help of the other studies for the same dataset in advance. Results show that the combination of the features derived from cross correlation and PSD is very promising in discriminating between epileptic and healthy EEG segments.

1. Introduction

The studies on the analysis and classification of the electroencephalogram (EEG) signals have become popular in the recent years. In most of the analysis problems, statistical methods and artificial neural networks (ANN) are used. Besides, by the integration of different modalities (e.g. magnetoencephalogram (MEG), electromyogram (EMG), functional magnetic resonance imaging (fMRI), diffusion optical tomography (DOT), etc.) with EEG, the quantity and the quality of the obtained information from the brain are increased (Ales, 2007; Cao, 2007; Mourad, 2007).

Epilepsy is the main application area of EEG. In the diagnosis of epileptic seizures, EEG and MEG are the two main techniques. Even the fMRI, computerized tomography (CT), magnetic resonance spectroscopy (MRS) and positron emission tomography (PET) are used in the diagnosis, thanks to its high time resolution, only EEG or MEG can directly measure the electrical activity in brain. On the other hand, spatial resolutions of EEG and MEG are not sufficient to determine the cortical areas related to seizure directly. Intracranial EEG has very good time resolution. However, this technique is used only in exceptional situations due to the high risk potential for the subject (Calderon, 2007).

EEG is used in brain computer interface (BCI) designs as the control signal. BCI technology has great potential to provide functional increase and independence to individuals with severely disabled people. The studies in BCI research are increasing in the recent years (Felton, 2007; Müller-Putz & Pfurtscheller, 2008).

EEG classification is used also for sleep analysis (Kassebaum, 2008), understanding the depth of anesthesia (Linares-Perdomo, 2007), detecting drowsiness (Nodine, 2008), etc. There is ongoing research for the application potential of EEG to the diseases like bipolar disorder and hyperactivity (Walshaw, 2007).

All these studies in the literature show the importance and the potential of EEG classification.

Feature extraction is one of the most important steps in signal classification. When the features are not appropriate for the given classification problem, obtained performances are unsatisfactory. In this case, even the classification algorithm is optimally determined for the problem, because of the improper features, the algorithm cannot generate high performance. Therefore, it is mandatory to find and extract suitable features from the raw signals to be able to obtain good classification results.

* Corresponding author. Tel.: +90 212 2856635; fax: +90 212 2853679.
E-mail addresses: iscanz@itu.edu.tr, zaferiscan@yahoo.com (Z. Iscan), dokur@itu.edu.tr (Z. Dokur), demiralp@istanbul.edu.tr (T. Demiralp).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.02.110

Table 1
Recording conditions for each of the sets.

Set	Recording conditions
A	(Healthy) Relaxed in an awake state with eyes open
B	(Healthy) Relaxed in an awake state with eyes closed
C	Recorded from the hippocampal formation of the opposite hemisphere of the brain (seizure-free)
D	Recorded from within the epileptogenic zone (seizure free)
E	During seizure activity

Table 2
Recording settings.

Settings	Value
Channels	128
Reference	Average common
ADC	12 bit
Sampling frequency	173.61 Hz
Band-pass filter	0.53–40 Hz
Electrode placement	International 10–20 system

Time and frequency domain based feature extraction methods are commonly used in the feature extraction processes of biological signals. In this paper, combined time and frequency features for the classification of electroencephalogram signals are considered. As time domain features, cross-correlations of the EEG time series are selected (Chandaka, Chatterjee, & Munshi, 2009) due to their high performance in classification. As frequency domain features, spectral power ratios of the main EEG bands (theta, delta, alpha, beta and gamma) were calculated after application of discrete Fourier transform (DFT) to the EEG series. Calculating power spectrum for different EEG frequency bands is a popular method in literature (Nguyen-Ky, Wen, & Li, 2009).

Up to now, many classifier structures have been used in the classification of EEG signals. Lotte, Congedo, Lécuyer, Lamarche, and Arnaldi (2007) analyzed the classification algorithms in the EEG based BCI designs. They compared the performances of different algorithms in the literature and mentioned the important points to take into consideration in determining the appropriate classification algorithm for a specific BCI.

Müller, Anderson, and Birch (2003) gave some examples for the application of EEG data to linear and non-linear methods in their study. In the end, they summarized the advantages and disadvantages of these methods. As a result, simplicity usually generated the best result. Therefore, they preferred linear methods. On the other hand, they showed that non-linear methods generated better results (especially in complex or large datasets) in some applications. According to this study, it can be concluded that the linear and the non-linear methods can be appropriate in different classification problems. In the EEG based BCI systems, it is considered that the kernel based methods like Support Vector Machines (SVMs) are the most advanced methods (Zhong, Lotte, Girolami, & Lécuyer, 2008).

In Section 2, EEG dataset used in the study is described. In Section 3, methods used for feature extraction and feature selection will be explained. Section 4 shortly mentions about the classifiers which are utilized in the study. Obtained classification results by using these classifiers are given in Section 5. Finally, conclusions derived from the study and future work are described in Section 6.

2. EEG dataset

In this study, the dataset used in Andrzejak et al. (2001) paper is considered. In the dataset, there are 100 single channel EEG segments of 23.6 s acquired from different subjects for each of 5 different sets. In Tables 1 and 2, recording conditions and settings for each of the sets are given, respectively.

In this study, only 2 sets (classes) are considered. Set *A* was formed with the segments taken from healthy surface EEG recordings that were carried out on five healthy volunteers (eyes open) using a standardized electrode placement scheme. Set *E* was formed with the segments taken from five patients and contains seizure activity. More information about the dataset can be found in the given reference. In Fig. 1, healthy and epileptic EEG examples from the database are given.

3. Methods

3.1. Feature extraction

Feature extraction is a very crucial step in signal classification problems, because it directly influences the classification performance. When the features are not distinctive, it is less likely to achieve good classification performances. In this study, features based on time and frequency domain are extracted and then combined for classification of two different (healthy or epileptic) EEG states.

3.1.1. Cross-correlation

Cross-correlation (CC) features are derived in the same way as Chandaka et al. (2009) paper. The cross-correlation (R_{xy}) of two discrete signals (e.g. $x(n)$ and $y(n)$) is given by the following formula:

$$R_{xy}(m) = \begin{cases} \sum_{n=0}^{N-m-1} x_{n+m} y_n & m \geqslant 0 \\ R_{yx}(-m) & m < 0 \end{cases} \tag{1}$$

In Eq. (1), m is the time shift parameter. N is the length of $x(n)$ for R_{xy} or length of $y(n)$ for R_{yx}. If $x(n)$ and $y(n)$ have the same length (e.g. M), length of R_{xy} is $2 \times M - 1$. In Fig. 2a and b, cross correlo-

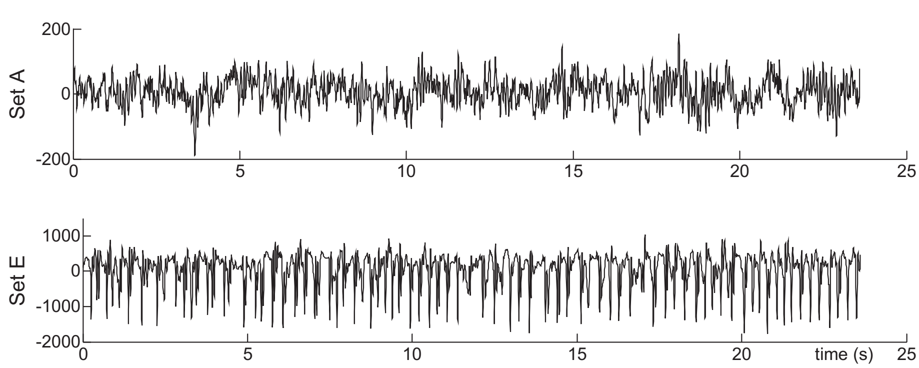

Fig. 1. EEG examples for two different sets: set *A* (healthy) and set *E* (epileptic).

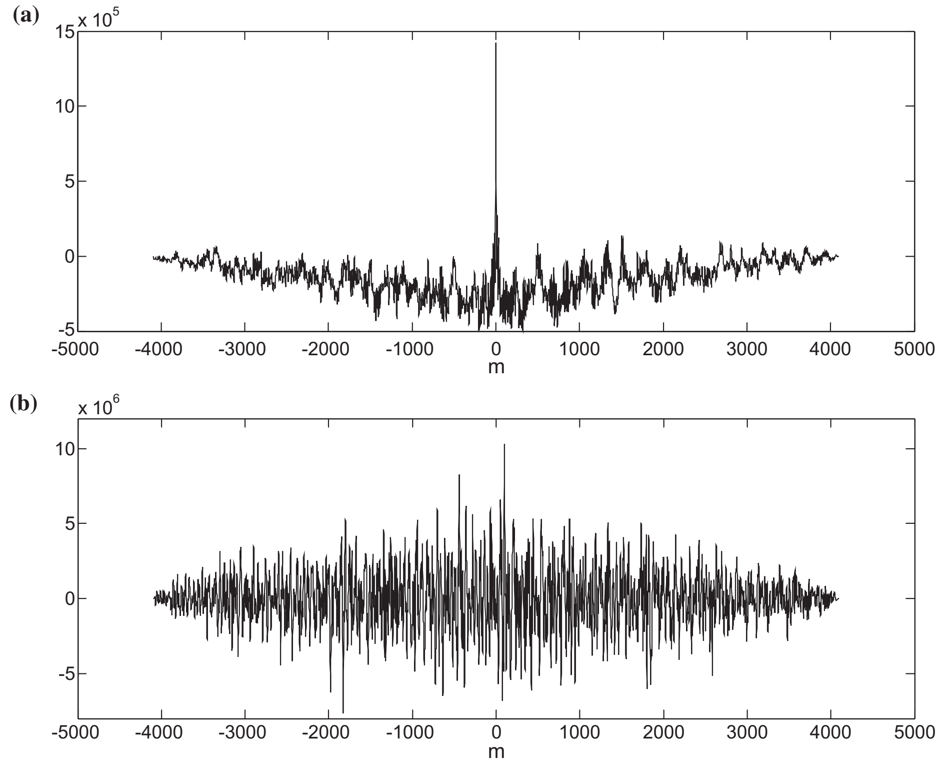

Fig. 2. (a) Cross correlogram of healthy EEG segments and (b) cross correlogram of a healthy and an epileptic EEG segment.

gram of two healthy EEG segments and cross correlogram of a healthy and an epileptic segment are given, respectively.

The five features derived from R_{xy} are listed below:

(1) *pv*: peak value of the cross correlogram.
(2) *ins*: instant value corresponding to the peak value.
(3) $centroid = \frac{\sum_{m=-M}^{M} m \cdot R(m)}{\sum_{m=-M}^{M} R(m)}$ (2)
(4) $eqwidth$: equivalent width $= \frac{\sum_{m=-M}^{M} R(m)}{pv}$ (3)
(5) msa : mean square abscissa $= \frac{\sum_{m=-M}^{M} m^2 \cdot R(m)}{\sum_{m=-M}^{M} R(m)}$ (4)

In feature extraction step of the cross-correlation method, first of all, one of the EEG segments in the database is chosen as reference. Afterwards, cross-correlations of all other EEG segments with the reference one are calculated. Therefore, the number of feature vectors is decreased by 1. In this paper, the reference EEG segment is chosen as the first healthy EEG segment in the database.

3.1.2. Spectral band powers

In EEG literature, EEG frequency spectrum is generally analyzed according to some specific frequency bands. These sub-bands are called delta (δ), theta (θ), alpha (α), beta (β) and gamma (γ) bands. However, there are no strict frequency ranges for these different bands. In Table 3, frequency band ranges in several references are given.

In this paper, ranges are selected as: delta (0.5–4), theta (4–8), alpha (8–13), beta (13–25) and gamma (25–40). Dastidar (2007) states that the individual frequency sub-bands may be more representative of brain dynamics than the entire EEG. At the end of his thesis, he observed that the sub-bands gave more information about the underlying neuronal activities and some kind of changes in EEG could be noticed only when all sub-bands were analyzed separately. Using this observation, in this paper, total powers of the EEG sub-bands over the whole band are extracted as features. First of all, discrete Fourier transforms (DFT) of EEG segments in the database are calculated (Eq. (5))

Table 3
EEG frequency band ranges (Hz) in several references.

Reference	delta (δ)	theta (θ)	alpha (α)	beta (β)	gamma (γ)
Garcia Molina (2004)	2–4	4–8	8–13	13–30	>30
Allison (2003)	0.5–4	4–8	8–13	13–40	–
Dastidar (2007)	0–4	4–8	8–13	13–30	30–60
Krepki (2004)	0.5–3.5	3.5–8	8–13	13–25	25–40
Calderon (2007)	0.5–4	4–8	8–13	13–25	25–70

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-\frac{j2\pi k \cdot n}{N}} \quad k = 0, 1, \ldots, N-1 \tag{5}$$

In Eq. (5), $x(n)$ represents the discrete samples of EEG data. N is the length of the EEG data.

After calculation of DFT ($X(k)$) of EEG samples, square of the absolute value of $X(k)$ is computed to obtain the power spectrum of EEG (Eq. (6))

$$\text{Power spectral density} = P(k) = |X(k)|^2 \tag{6}$$

In Fig. 3(a) and (b), power spectrums of a healthy and an epileptic EEG segment are presented, respectively.

Total powers are calculated for each sub-band ($\delta, \theta, \alpha, \beta, \gamma$) using Eq. (6). Features are defined as the ratios (prdelta, prtheta, pralpha, prbeta, prgamma) of each sub-band power (pdelta, ptheta, palpha, pbeta, pgamma) to the whole band (0.5–40 Hz) power (ptotal). Thus, there are totally five features derived from the EEG time series.

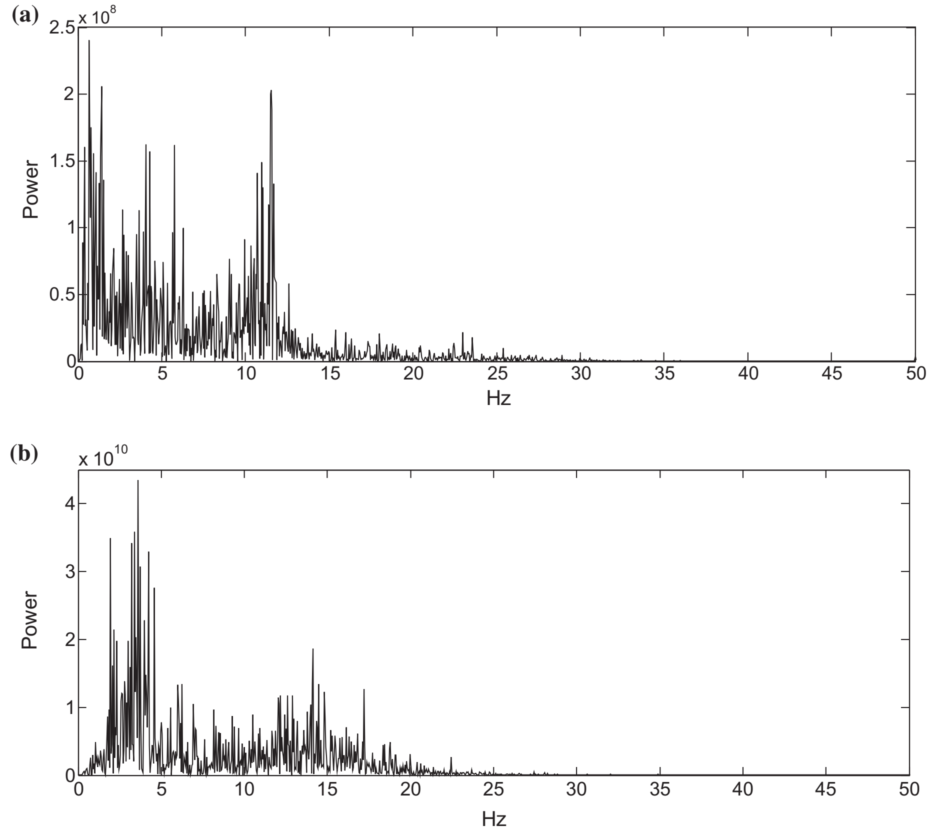

Fig. 3. Power spectrums of a healthy (a) and an epileptic (b) EEG segment.

Table 4
The features used in the study.

Feature no.	1	2	3	4	5	6	7	8	9	10
Features	*pv*	*ins*	*centroid*	*eqwidth*	*msa*	prdelta	prtheta	pralpha	prbeta	prgamma

The values of the features in the feature vectors were normalized between 0 and 1 before classification step.

In Table 4, the features used in the study are given.

3.2. Divergence analysis

In the study, by using divergence analysis (Cohen, 1986); performances of different feature extraction methods are measured quantitatively and the best features among all the extracted ones within a feature set are searched. The divergence is a class separability measure and gives information about the distribution of class vectors in the feature space. Divergence value decreases for indiscriminately scattered class vectors. Formulations can be found in the given reference. In the study, obtained features are ordered from the most discriminative to the least according to the divergence values.

4. Classifiers

To be able to classify healthy and epileptic EEG data, in addition to the *K*-Nearest Neighbor (*K*-NN) classifier, MATLAB-based two different free toolboxes are used: LS-SVMlab1.5 (Pelckmans et al., 2003) and PRTools 4.0. LS-SVMlab 1.5 Toolbox is specialized for LS-SVM method. Classifiers chosen from PRTools 4.0 are: Support Vector Machines (SVMs), Parzen, Fisher Discriminant Analysis, Binary Decision Tree, Naive Bayes, Nearest Mean and Quadratic classifiers.

Table 5
Distribution of feature vectors in the training and the test sets.

Feature vectors	Healthy		Epileptic	
	Training	Test	Training	Test
CC	50	49	50	50
PSD	50	50	50	50
Combination	50	49	50	50

The *K* Nearest Neighbor (*K*-NN) is a widely used, instance based learning algorithm. In this algorithm a feature vector in the test set is classified by assigning it to the most frequently represented vector class in the *K* (in this study $K = 3$) nearest training vectors (Duda, Hart, & Stork, 2001). Deciding upon the *K* value is critical since it directly affects the classification performance.

Support Vector Machines rely on preprocessing the data to represent patterns in dimensions higher than the original feature space dimension. It accomplishes this task using an appropriate nonlinear mapping to higher dimension. By this way, data samples from two different classes become separable by a hyper-plane (Duda et al., 2001). Least Squares Support Vector Machines (LS-SVM) are reformulations to standard SVMs which lead to solving linear Karush–Kuhn–Tucker (KKT) systems (Suykens, Van Gestel,

Table 6
Performances for three different feature conditions and three different partitions.

Partition	Features	Classifiers and total accuracy (%)								
		K-NN	SVM	LS-SVM	Parzen	LDA	Quad	Decision Tree	Naive Bayes	Nearest Mean
1	CC	94.94	87.87	91.91	93.93	87.87	87.87	95.95	93.93	89.89
	PSD	90	80	97	85	89	96	90	85	86
	Combination	97.97	88.88	100	91.91	89.89	100	98.98	94.94	91.91
2	CC	92.92	86.86	95.95	93.93	85.85	92.92	95.95	94.94	85.85
	PSD	87	86	84	85	92	98	91	80	89
	Combination	91.91	87.87	97.97	89.89	90.90	98.98	95.95	96.96	86.86
3	CC	95.95	92.92	98.98	95.95	89.89	94.94	95.95	95.95	90.90
	PSD	90	92	88	87	93	94	86	79	86
	Combination	94.94	93.93	97.97	94.94	93.93	97.97	98.98	94.94	95.95

Table 7
Distribution of the extended feature vectors in the training and the test sets.

Feature vectors	Healthy		Epileptic	
	Training	Test	Training	Test
CC	800	799	800	800
PSD	800	800	800	800
Combination	800	799	800	800

De Brabanter, De Moor, & Vandewalle, 2002). In the study, LS-SVM classifier of LS-SVMlab1.5 is used. Tenfold cross validation and Bayesian initialization are done for determining parameters optimally.

In Parzen-window classification, the densities for each category are estimated and a test point is classified by the label corresponding to the maximum posterior. The decision regions depend on the choice of the window function (Duda et al., 2001).

In Fisher's linear discriminant analysis (LDA), a linear function is obtained which provides the maximum ratio of between-class scatter to within-class scatter (Duda et al., 2001). By adding additional terms involving the products of components of feature vectors, quadratic discriminant function is obtained (Duda et al., 2001). Thus, quadratic discriminant function has additional coefficients and generates more complicated separating surfaces. The separating surface is a second-degree or hyperquadric surface.

Decision Trees are formed by a sequence of queries which generate nodes along a path from root to leaf (Duda et al., 2001).

When the dependencies among the features are unknown, it is very common to make an assumption such that, given the category, the features are conditionally independent. This is the assumption behind the Naive Bayes classifier (Duda et al., 2001).

When the prior probabilities are the same for all classes, a feature vector is classified by measuring the Euclidean distances between each feature vector and each of class mean vectors (Duda et al., 2001). Afterwards, the feature vector's class is determined by the nearest mean vector (minimum distance classifier).

Table 9
Divergence values for the CC non-segmented case.

Feature no.	1	2	3	4	5
	pv	*pv* *eqwidth*	*pv* *eqwidth* *msa*	*pv* *eqwidth* *msa* *ins*	*pv* *eqwidth* *msa* *ins* *centroid*
Divergence	2.2705	2.3326	2.3507	2.3545	2.3581

Table 10
Divergence values for the CC segmented case.

Feature no.	1	2	3	4	5
	pv	*pv* *ins*	*pv* *ins* *eqwidth*	*pv* *ins* *eqwidth* *centroid*	*pv* *ins* *eqwidth* *centroid* *msa*
Divergence	7.2469	7.4545	7.6507	7.6866	7.8220

In conclusion, in the paper, a total of nine classifiers are utilized for the classification of the EEG data. Since the structures of the classifiers are published in the literature in advance, details of the classifiers are not given in this paper.

5. Results

All the classifications are performed in MATLAB software. Database is divided into training and test sets after feature extraction steps as shown in Table 5.

Table 8
Performances for extended feature vectors and three different partitions.

Partition	Features	Classifiers and total accuracy (%)								
		K-NN	SVM	LS-SVM	Parzen	LDA	Quad	Decision Tree	Naive Bayes	Nearest Mean
1	CC	100	100	100	100	100	100	100	100	96.87
	PSD	100	75	100	100	75	96.87	100	93.75	65.62
	Combination	100	100	100	100	100	100	100	100	96.87
2	CC	100	100	100	100	100	96.49	100	100	96.68
	PSD	100	50	100	100	75.12	96.87	100	93.06	65.87
	Combination	100	100	100	100	100	100	100	100	96.99
3	CC	100	100	100	100	100	100	100	100	96.74
	PSD	100	50	100	100	78.87	96.62	100	93.87	66.62
	Combination	100	100	100	100	100	100	100	100	96.81

Table 11
Divergence values for the PSD non-segmented case.

Feature no.	1	2	3	4	5
	prtheta	prtheta prbeta	prtheta prbeta prgamma	prtheta prbeta prgamma prdelta	prtheta prbeta prgamma prdelta pralpha
Divergence	0.3463	0.4798	0.7362	0.7978	0.8985

Table 12
Divergence values for the PSD segmented case.

Feature no.	1	2	3	4	5
	prbeta	prbeta prgamma	prbeta prgamma prtheta	prbeta prgamma prtheta pralpha	prbeta prgamma prtheta pralpha prdelta
Divergence	0.1627	0.3219	0.4210	0.4457	0.4621

Three different classification conditions are considered using:

- CC features,
- PSD features,
- combination of CC and PSD features.

Besides, three different partitions of training and test features are taken into consideration in order to show the dependency of the classification performances on the distribution of feature vectors in the training and test sets. In partition 1, training and test feature vectors are taken in the same order as given in the original dataset whereas in partitions 2 and 3, the feature vectors are selected in a random manner. Table 6 shows the performances for three different feature conditions and three different partitions of feature vectors.

In Table 6, performance measure is given as the total classification accuracy. Since the original EEG records have 23.6 s length (4096 samples), they cannot be regarded as stationary signals. Therefore, the records are segmented into 16 segments of 256 samples in another classification approach. Distribution of the extended feature vectors for this approach is given Table 7.

Table 8 shows the performances for three different feature conditions and three different partitions of extended feature vectors.

In order to show the appropriateness of the combination and segmentation approach, divergence analysis is performed for both the individual CC, PSD features and the combination of them in the segmented and non-segmented cases. Tables 9–14 show the total divergence values for the given feature combinations.

6. Conclusions and future work

Looking at the classification performances given in Table 6, it could be inferred that LS-SVM, Binary Decision Tree and Quadratic classifier generated higher classification accuracies as compared with the other classifiers. Moreover, combination of features generally yields the highest accuracy as compared to the individual CC and PSD features. It should be noted that there is no misclassification (100% accuracy) in the test set in the first partition for LS-SVM and Quadratic classifiers. In Table 8, it is clear that the segmentation of long EEG records enhances stationary assumption and generates better performances. There is no misclassification for K-NN, LS-SVM, Parzen and Decision Tree classifiers in any of the three different partitions. CC features generate better performances as compared to PSD features when they are utilized individually. On the other hand, the combination of the features reached the highest classification accuracies again.

Ordered features according to the divergence value in Tables 9 and 10 show that the usage of extended (segmented) records gen-

Table 13
Divergence values for the combined non-segmented case.

Feature no.	1	2	3	4	5	6	7	8	9	10
	pv	*pv* prtheta	*pv* prtheta pralpha	*pv* prtheta pralpha prgamma	*pv* prtheta pralpha prgamma *msa*	*pv* prtheta pralpha prgamma *msa* *centroid*	*pv* prtheta pralpha prgamma *msa* *centroid* *ins*	*pv* prtheta pralpha prgamma *msa* *centroid* *ins* prdelta	*pv* prtheta pralpha prgamma *msa* *centroid* *ins* prdelta prbeta	*pv* prtheta pralpha prgamma *msa* *centroid* *ins* prdelta prbeta *eqwidth*
Divergence	1.5133	2.1197	2.3014	2.3492	2.3660	2.4373	2.4504	2.4531	2.6479	2.6551

Table 14
Divergence values for the combined segmented case.

Feature no.	1	2	3	4	5	6	7	8	9	10
	pv	*pv* prbeta	*pv* prbeta prgamma	*pv* prbeta prgamma prdelta	*pv* prbeta prgamma prdelta pralpha	*pv* prbeta prgamma prdelta pralpha prtheta	*pv* prbeta prgamma prdelta pralpha prtheta *ins*	*pv* prbeta prgamma prdelta pralpha prtheta *ins* *eqwidth*	*pv* prbeta prgamma prdelta pralpha prtheta *ins* *eqwidth* *centroid*	*pv* prbeta prgamma prdelta pralpha prtheta *ins* *eqwidth* *centroid* *msa*
Divergence	7.2469	7.8257	8.5197	8.7449	8.9293	9.2718	9.5787	9.5906	9.5910	9.6034

Table 15
Classification accuracies in the literature for the same dataset.

Reference	Features	Classifier	Total accuracy (%)
Subasi (2007)	Wavelet	ME	93.2
	Wavelet	MLP	94.5
Chandaka et al. (2009)	CC	LS-SVM	95.95
Polat and Günes (2007)	FFT	Decision Tree	98.72
Polat and Gunes (2008a)	DFT	Decision Tree	99.02
	AR	Decision Tree	99.32
	DWT	Decision Tree	92
	DFT, DBDR	Decision Tree	99.12
	AR, DBTR	Decision Tree	98.94
	DWT, DBDR	Decision Tree	89.50
Polat and Gunes (2008b)	PCA FFT	AIRS	100

erates higher divergence values. Besides, *pv* feature is the most discriminative feature among all other CC features. In Tables 11 and 12, it is clear that PSD based features have lower divergence values than the CC features. On the other hand, there is a reduction in the divergence value for the segmented records. In Tables 13 and 14, the combined features for the segmented records generate highest divergence values. These results support the appropriateness of the combined features and the segmented records.

Obtained results in Tables 6 and 8 give the highest classification accuracies as compared to the other studies (Table 15) encountered in the literature which use the same dataset.

In the evaluation of the classification results of different studies, partition methods of training and test sets should be known. In Subasi (2007), each of the EEG time segments was divided into eight parts before feature extraction step. Thus, totally 1600 feature vectors are generated. Seventy percent of the training vectors are used for training and the remaining 30% for test step. In a recent study using the same dataset (Chandaka et al., 2009), two different sizes of training and test set were used. In the first study, partitions of the feature vectors were performed in the same way as described in Table 5. In the second case, EEG time segments were windowed by a rectangular window composed of 512 discrete data. Thus, a total of 1600 EEG segments were generated. The overall accuracies were 95.96% and 95.5%, respectively, for the two different partitions. In Polat and Günes (2007), again, EEG time series of 4096 samples were windowed by a rectangular window composed of 256 discrete data. Thus, a total of 3200 EEG segments were generated. Five and tenfold cross-validations were performed in classification. 98.68% and 98.72% classification accuracies were achieved, respectively. In Polat and Gunes (2008a), each of the EEG time series were divided into 16 parts. Similarly, totally 3200 feature vectors are generated. However, distance based data reduction (DBDR) method was used to reduce from 3200 vectors to 1600 vectors. Results were given for the partitions with and without DBDR. Tenfold cross validation method was used in classification. In Polat and Gunes (2008b), each of the EEG time series were divided into 16 parts. Different training and test set ratios were considered: 50–50%, 70–30%, and 80–20%. The obtained test classification accuracies for these partitions were 99.81%, 100% and 100%, respectively. Hundred percent accuracy was obtained for tenfold cross validation.

In this study, individual performances of the classifiers are taken into consideration. However voting of classifiers may be capable of generating more robust classification performances which can decrease the sensitivity of the results to the different partitions of the feature vectors.

References

Ales, J. M. (2007). *Disambiguating electrical activity in visual cortex by combining multifocal evoked potentials and FMRI*. PhD thesis, Vision Science, University of California, Berkeley.

Allison, B. (2003). *P3 or not P3: Toward a better P300 BCI*. PhD thesis in Cognitive Science, University Of California, San Diego.

Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., & Elger, C. E. (2001). Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E, 64*, 061907.

Calderon, R. (2007). *Brain computer interface and neuroprosthetics*. MSc thesis, The College of Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL.

Cao, N. (2007). *Estimating extended brain sources using EEG and diffuse optical tomography*. PhD thesis, Department of Electrical and Systems Engineering, Washington University, Saint Louis, Missouri.

Chandaka, S., Chatterjee, A., & Munshi, S. (2009). Cross-correlation aided support vector machine classifier for classification of EEG signals. *Expert Systems with Applications, 36*(2), 1329–1336.

Cohen, A. (1986). *Biomedical signal processing* (Vol. 2). Boca Raton, FL: CRC Press (pp. 75–79).

Dastidar, S. G. (2007). *Models of EEG data mining and classification in temporal lobe epilepsy: Wavelet-chaos-neural network methodology and spiking neural networks*. PhD thesis in Biomedical Engineering, Ohio State University.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification* (2nd ed.). New York: John Wiley & Sons.

Felton, E. A. (2007). *Human factor studies of brain computer interfaces: Performance and mental effort for able and physically disabled*. PhD thesis in Biomedical Engineering, University of Wisconsin, Madison.

Garcia Molina, G. N. (2004). *Direct brain–computer communication through scalp recorded EEG signals*. PhD thesis, Institut de Traitement des Signaux, Swiss Federal Institute of Technology, Lausanne.

Kassebaum, J. A. (2008). *Application of frequency domain state-space analysis to sleep*. PhD thesis, Purdue University, West Lafayette, Indiana.

Krepki, R. (2004). *Brain–computer interfaces: Design and implementation of an online BCI-system for the control in gaming applications and virtual limbs*. PhD thesis, Technische Universität, Berlin, Germany.

Linares-Perdomo, O.-J. (2007). *Quantification of the impact of anesthesia by multivariate data analysis and nonlinear analysis of human brain electrical activity*. PhD thesis, Department of Bioengineering, The University of Utah.

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain–computer interfaces. *Journal of Neural Engineering, 4*.

Mourad, M. M. (2007). *An integrated functional mapping and source localisation platform for brain research*. PhD thesis, Electrical Engineering, Florida International University, Miami, FL.

Müller, K.-R., Anderson, C. W., & Birch, G. E. (2003). Linear and nonlinear methods for brain-computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 11*(2), 165–169.

Müller-Putz, G. R., & Pfurtscheller, G. (2008). Control of an electrical prosthesis with an SSVEP-based BCI. *IEEE Transactions on Biomedical Engineering*(1).

Nguyen-Ky, Tai, Wen, P., & Li, Y. (2009). Theoretical basis for identification of different anesthetic states based on routinely recorded EEG during operation. *Computers in Biology and Medicine, 39*(1), 40–45.

Nodine, E. (2008). *The detection of drowsy drivers through driver performance indicators*. MSc thesis in Human Factors Engineering, Tufts University.

Pelckmans, K., Suykens, J. A. K., Van Gestel, T., De Brabanter, J., Lukas, L., Hamers, B., et al. (2003). *LS-SVMlab toolbox user's guide version 1.5*. Belgium: Katholieke Universiteit Leuven.

Polat, K., & Günes, S. (2007). Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. *Applied Mathematics and Computation, 187*(2), 1017–1026.

Polat, K., & Gunes, S. (2008a). A novel data reduction method: Distance based data reduction and its application to classification of epileptiform EEG signals. *Applied Mathematics and Computation, 200*(1), 10–27.

Polat, K., & Gunes, S. (2008b). Artificial immune recognition system with fuzzy resource allocation mechanism classifier, principal component analysis and FFT method based new hybrid automated identification system for classification of EEG signals. *Expert Systems with Applications, 34*(3), 2039–2048.

PRTools 4.0. The Netherlands: The Pattern Recognition Research Group of the TU Delft.

Subasi, A. (2007). EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications, 32*(4), 1084–1093.

Suykens, J. A. K., Van Gestel, T., De Brabanter, J., De Moor, B., & Vandewalle, J. (2002). *Least squares support vector machines*. Singapore: World Scientific.

Walshaw, P. D. (2007). *Inhibition in bipolar disorder and attention-deficit hyperactivity disorder*. PhD thesis in Graduate School, Temple University.

Zhong, M., Lotte, F., Girolami, M., & Lécuyer, A. (2008). Classifying EEG for brain computer interfaces using Gaussian processes. *Pattern Recognition Letters, 29*, 354–359.