Using Combination of μ,β and γ Bands in Classification of EEG Signals

Mina Mirnaziri [1], Masoomeh Rahimi [1], Sepidehsadat Alavikakhaki [2], Reza Ebrahimpour [1 *]

1. Brain and Intelligent Systems Research Laboratory (BISLab), Department of Electrical and Computer Engineering, ShahidRajaee Teacher Training University, Tehran, Iran.
2. Department of Computer Science, School of Mathematics, Statistics and Computer Science, University of Tehran, Iran.

Article info:
Received: 12 October 2012
First Revision: 2 December 2012
Accepted: 8 January 2013

Key Words:
Brain – Computer Interface (BCI),
Electroencephalogram (EEG),
Common Spatial Pattern (CSP),
Multi – Layer Perceptron (MLP),
Linear Discriminant Analysis (LDA),
Radial Basis Function (RBF).

ABSTRACT

Introduction: In most BCI articles which aim to separate movement imaginations, μ and β frequency bands have been used. In this paper, the effect of presence and absence of γ band on performance improvement is discussed since movement imaginations affect γ frequency band as well.

Methods: In this study we used data set 2a from BCI Competition IV. In this data set, 9 healthy subjects have performed left hand, right hand, foot and tongue movement imaginations. Time and frequency intervals are computed for each subject and then are classified using Common Spatial Pattern (CSP) as a feature extractor. Finally, data is classified by LDA[1], RBF[2] MLP[3], SVM[4]and KNN [5] methods. In all experiments, accuracy rate of classification is computed using 4 fold validation method.

Results: It is seen that most of the time, combination of μ,β and γ bands would have better performance than just using combination of μ and β bands or γ band alone. In general, the improvement rate of the average classification accuracy is computed 2.91%.

Discussion: In this study, it is shown that using combination of μ, β and γ frequency bands provides more information than only using combination of μ and β in movement imagination separations.

1. Introduction

Many diseases such as Amyotrophic Lateral Sclerosis (ALS) can destroy neuromuscular channels which help the brain to communicate with the body. A simple solution for that problem is to provide new communicational way for brain not relying on muscles. Namely, a brain-computer interface that transfers brain messages to the external world.(Wolpow, Birbaumer, McFarland, Pfurtscheller& Vaughan, 2002).

In most previous studies that used Electroencephalogram (EEG) signals for motor imagery in frequency domain, extracted features were composed of μ and β bands. (Ramoser, Muller-Gerking & Pfurtscheller, 2000; Muller-Gerking, Pfurtscheller & Flyvbjerg, 1999; Kwang-Eun & Kwee-Bo, 2011; Schloegl, Neuper & Pfurtscheller, 1997; Heung-II & Seong-Whan, 2011; Acar, 2011). Among brain frequency bands, α overlaps μ band in 8-12 Hz. Although μ band is dedicated to movement imaginations, α band is produced in resting times or when the eyes are shut and it is not related to movement imaginations or activities. Among all fre-

1. Linear Discriminant Analysis (LDA)
2. Radial Basis Function (RBF)
3. Multi Layer Perceptron (MLP)
4. Support Vector Machine (SVM)
5. K-Nearest Neighbor (KNN)

**** Corresponding Author:***
Reza Ebrahimpour, Ph.D.,
Brain and Intelligent Systems Research Laboratory (BISLab), Department of Electrical and Computer Engineering, ShahidRajaee Teacher Training University, Tehran, Iran.
E-mail: ebrahimpour@ipm.ir

quency bands, only μ, β and γ are related to movement imaginations and activities (Niedermeyer & DA Silva, 2005).

Since γ frequency band has a wider frequency range, it is mostly used in ElectroCorticoGraphic (ECoG) Recording method. It has shown that γ band provides more information than μ and β bands in 35-50 Hz and 75-100 Hz using ECoG Recording method in movement imaginations. (Crone, Miglioretti, Gordon & Lesser, 1998).

It is seen that during movement imaginations using ECoG Recording method, μ and β amplitudes are decreased whereas γ band is increased simultaneously. Furthermore, γ band can reflect small and tiny movements while μ and β bands do not contain important information for these types of movements. (Aoki, Fetz, Shupe, Lettich & Ojemann, 1999). It is also shown that subjects can control a cursor movement in both up and down directions up to 70-100% accuracy rate using μ, β and γ bands in ECoG Recording method in less training time than EEG method. (Leuthardt, Schalk, Wolpow, Ojemann & Moran, 2004). High γ band power is increased during arbitrary motions without any visual cues in invasive recordings in epilepsy patients. (Ball, Nawrot, Schulze-Bonhage, Aertsen&Mehring, 2004; Miller et al., 2007; Oharaet al., 2000;Pfurtscheller, Graimann, Huggins, Levine&Schuh, 2003).

In healthy people, EEG Recording method could detect motor activities in γ bands over 30Hz as invasive recording methods, in epilepsy patients. (Ball et al., 2008).

This paper aims to show using combination of μ, β and γ bands would have better performance than using μ and β. To do so, 8-30Hz, 8-40Hz, 8-50Hz and 30-50Hz frequency ranges are separated using Fourier Transformation and Band Pass Filtering. Next, Common Spatial Pattern (CSP) is used as feature extractor and Multi Layer Perceptron (MLP), Radial Basis Function (RBF), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) are used as classifiers.

In this study, dataset 2a from BCI competition IV was used in which 9 healthy subjects performed left hand, right hand, foot and tongue motor imagery.

2. Methods

2.1. Preprocessing and Feature Extraction

Database

We used dataset 2a from BCI competition IV which was provided by Graz University in 2008. It consists of EEG signals from 9 healthy subjects who have 4 tasks: left hand (class 1), right hand (class 2), foot (class 3) and tongue (class 4). The signals were recorded in two sessions on different days with 6 runs separated by short breaks in each day. Each run contains 48 trials (12 trials for each class), making a total of 288 trials in each session. To record the EEG singles, 22 silver/silver chloride (Ag/AgCl) electrodes where used which were placed in a 3.5 cm distance from each other. The signals were sampled with 250 Hz. Next, they were filtered between 0.5 Hz and 100 Hz using Band Pass filter. For each subject, 288 train samples and 288 test samples were considered. Experimental models are displayed in Figures 1 and 2.

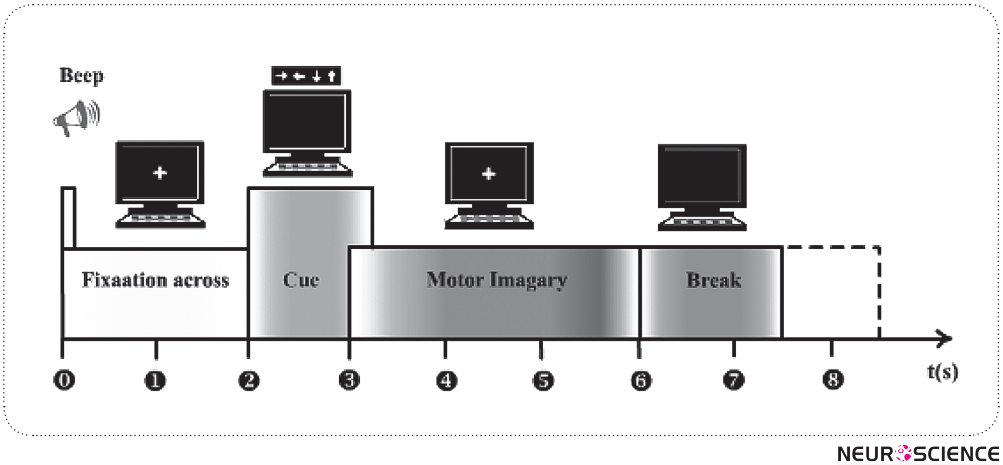


Figure 1. Timing scheme of the paradigm

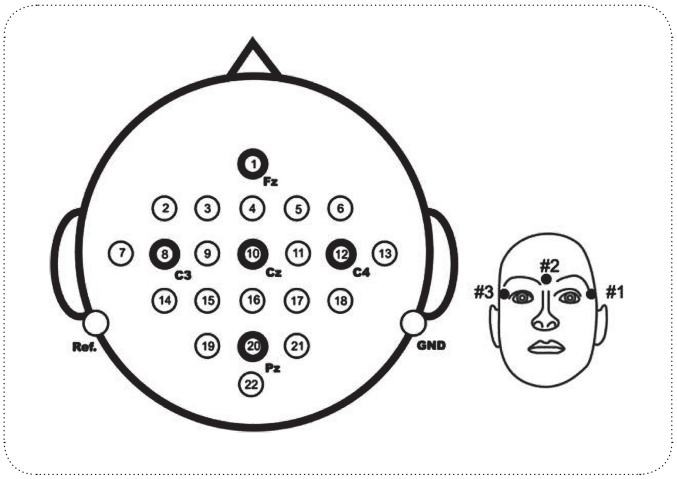


Figure 2. Electrode positions in Dataset 2a from BCI Competition IV.

In this stage, the suitable frequency intervals are separated using Discrete Fourier Transform (DFT) and Band Pass Fourier filter.

In DFT approach, data is transformed into the frequency domain by the following formula. (Dym & McKean, 1985; Yosida, 1968)

$$Z[k] = \sum_{n=0}^{N-1} z[n] e^{\frac{2\pi i}{N} nk} \quad (1)$$

k=0,…,N-1

Where z[n] is the input signal and N is data length.

2.2. Common Spatial Pattern (CSP)

CSP is a binary method in which variance of one class is increased whereas the variance of the other class is decreased in order to classify data more accurately. There are several approaches for multi class CSP. We used One Versus Rest (OVR) method. In this method, spatial patterns are found for one class versus the others. (Wu, Gao, X., & Gao, S. ,2006).

CSP is based on simultaneous diagonalization of two covariance matrixes. In this paper, we kept the first and last rows of spatial filters as our first pair, the second and the (n-1)th rows as our second pair and the third and (n-2)th rows as our last pair and omitted the rest rows. (Ramoser, Muller-Gerking & Pfurtscheller, 2000).

2.3. Classification

2.3.1. K-Nearest Neighbor (KNN)

K-Nearest Neighbor classifier is one of the simplest classifiers. To classify a new data sample, we compute its distance to all of its k nearest neighbors. This sample is assigned to the class in which most of these k neighbors are located. By distance we refer to Euclidean distance. (Denoeux, 2008)

2.3.2. Linear Discriminant Analysis (LDA)

In this method, the main idea is to find a hyper plane which separates data of two classes and maximize between class distances and minimize inter class distances. Thus, the optimized separation is guaranteed.

In LDA, data distribution is assumed to be Gaussian with equal covariance matrices for both classes. (Balakrishnama & Ganapathiraju, 1998). Middle and inner class scattering are obtained from the following formulas.

$$S_w = \sum_j p_j \times (cov_j) \quad (2)$$

$$S_b = \sum_j (\mu_j \times \mu_3) \times (\mu_j \times \mu_3)^T \quad (3)$$

Where S_w is the within-class scatter and S_b is the between-class scatter, p_i is prior probability, cov_j is the covariance of class j.

μ is the total average of classes and μ_j is the average of class μ_j. The aim is to find a measure which maximize the between class scatter and minimize the within class scatter which is defined as Equation 4.

$$criterion=inv^{6}(S_W)\times S_b \quad (4)$$

Thus, LDA tries to maximize this criterion. (Balakrishnama & Ganapathiraju, 1998).

2.3.3. Support Vector Machine (SVM)

Supporting Vector Machine is a classifying method proposed by V.Vapnik. (Cortes, & Vapnik, 1995). This method is a powerful tool in solving different problems. The main idea of this method is mapping the nonlinear input vectors to a high dimensional feature space in which a hyper plane exists that separates features linearly.

Let $\{(\chi_i, y_i) \mid i=1,\ldots,l\}$ be the training data. x_i is a train sample and $y_i \in \{-1,1\}$ is the label of sample i.

A hyper plane can be defined as follows.

$$w \cdot x - b = 0 \quad (5)$$

Where w is a vector perpendicular to the hyper plane, x is one point on the hyper plane and b is the bias. If $w.\chi_i+b\geq+1$, x_i is assigned to class +1 and if $w.\chi_i+b\leq-1$, xi is assigned to class -1. The optimized hyper plane is the one which maximize the margin value. The optimized margin is the one with this property that all samples of class -1 are located on one side of it and all samples in class 1 are located on the other side.

As we have $m=2/\|w\|$ (where m stands for margin), so $\|w\|$ should be decreased to produce an optimized hyper plane (Lodder, 2009).

In this paper we used, LIBSVM toolbox or linear kernel function (Chang and Lin, 2001).

2.3.4. Multi-Layer Perceptron (MLP)

Multi Layer Perceptron (MLP) consists of several layers of neurons or computational nodes: an input layer, one or more hidden layers and an output layer. Learning rule in this network is based on Back Propagation (BP) algorithm. In this approach, input value is multiplied by neuron weights in order to calculate the output value. Next, the output value of network is subtracted from the desired output value and an error value is computed. Finally, the weights of the network are updated. (Bishop, 1995; Haykin, 1999).

Activation function of neurons is a nonlinear function known as sigmoid:

$$y_j = \frac{1}{1+\exp(-v_j)} \quad (6)$$

Where v_j indicates weighted summation of all synaptic inputs added to bias of neuron j and y_j is the output of the neuron. To update weights of network, MLP uses the following formulas.

$$w = w + \Delta w \quad (7)$$

$$\Delta w = \eta \sum (t-o_h)\, o_h (1-o_h)\, x \quad (8)$$

Where η is learning rate, t is the desired output (target) and o_h is the output of hidden layer neurons and x is the input of the network.

2.3.5. Radial Basis Functions (RBF)

This neural network uses a nonlinear function for mapping input data into a higher dimensional space. This process is done based on the Cover Theory. According to this theory, data would be linear separable in a higher dimensional space. This network consists of one input layer, one hidden layer, and an output layer. Mapping from input space into output space is done using the following formula.

$$F(\vec{x}) = \sum_{}^{N} w_i \varphi(\|\vec{x} - \vec{x}_i\|) \quad (9)$$

Where $\varphi(\|\vec{x} - \vec{x}_i\|) = e^{-\|\vec{x}-\vec{x}_i\|^2}$ and w_i is the synaptic weight connecting hidden layer neuron to output layer. (Haykin, 1999).

3. Results

To find the best classifier among MLP, RBF, LDA, SVM and KNN, each classifier is performed on 3-6 seconds time interval and 8-30 Hz frequency interval and the results are shown in Table 1. All results are obtained by 4-fold cross validation process.

4. fold Cross Validation

To obtain the accuracy rate of a classifier, first it is trained using training data and then it is tested using validation data. In this approach, data is divided into *n* different parts and (n-1) part of it is used as training data and the remaining part is used as validation data. This process is repeated on all n different parts of data so that each part is used as validation data once and (n-1) times as training data. Finally, the average of all computed accuracies is considered as the final accuracy rate.

6. Inverse

In this study, we used 4-fold cross validation process. In the other words, we divided data into 4 parts and each classifier was trained 4 times. Finally the computed accuracies are averaged.

The classifiers are being compared according to their ranking results. So ranks 1,2,3,4 and 5 are given to the best, the second, the third, the forth and the fifth classifier, respectively. Finally, the ranks of all classifiers are cumulated. The lowest rank belongs the best classifier. As it can be seen in Table 1, the best rank is assigned to MLP and SVM.

Since MLP and RBF classifiers have several parameters to be set in their learning process, some experiments are done to find out the optimum parameters.

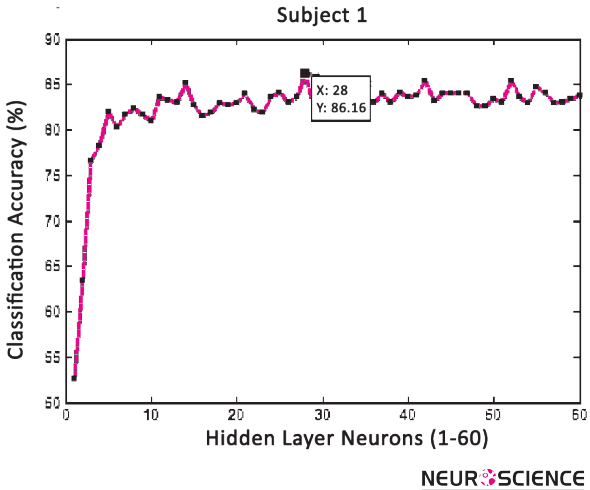
Table 1. Determination of the best classifier among MLP, RBF, LDA, SVM and KNN with time interval of 3-6 sec and frequency range of 8-30 Hz. In this experiment, classification accuracies were computed by 4-fold cross validation process.

Subjects	MLP(%)	RBF(%)	LDA(%)	SVM(%)	KNN(%)
1	72.82	70.57	75.34	72.22	67.36
2	41.92	46.70	59.72	61.45	61.11
3	80.20	75.60	76.38	74.65	67.36
4	40.36	44.70	52.77	58.33	49.65
5	48.26	33.85	39.58	40.27	48.95
6	44.44	34.37	45.83	45.48	37.84
7	76.73	64.84	71.53	68.75	57.63
8	80.20	62.41	75.34	77.08	61.80
9	70.83	63.54	66.66	68.40	67.01
Rank	**2.33**	4.22	2.44	**2.33**	3.66

NEUROSCIENCE

To find the appropriate number of nodes in the hidden layer of the MLP classifier, we performed the classifier with 1 to 60 nodes in its hidden layer on the validation data and calculated its accuracy rate each time. The best accuracy rate belongs to the MLP with the optimized number of neurons in its hidden layer.

After setting the nodes and their optimized corresponding weights in the hidden layer, MLP is performed on test data. In this classifier, the learning rate is set 0.2 and weights are initialized randomly. Also, the algorithm used for updating weights is based on Back Propagation. To make this clear, accuracy rates gained by MLP classifier for the first subject are plotted in order to find the optimized number of neurons in the hidden layer. The desired case is determined in the Figure 3.



NEUROSCIENCE

Figure 3. The classification accuracy rates(%) of MLP(vertical axis) and the Optimized number of hidden layer neurons(horizontal axis) computed for the first subject.

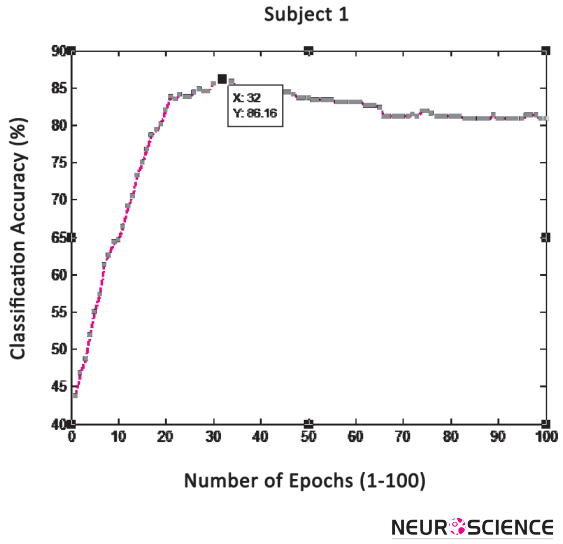


Figure 4. Determination of the optimum number of epochs for MLP classifier based on BP algorithm for Subject 1. Horizontal axis: number of epochs, vertical axis: classification accuracy (%).

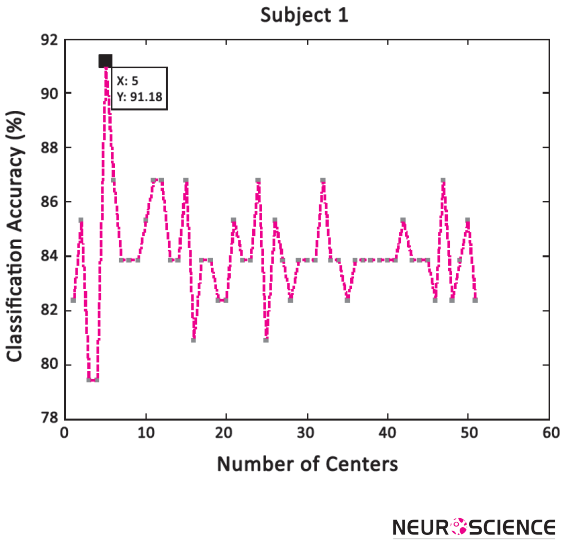


Figure 5. Determination of optimized number of centers in RBF for the 1st subject. Horizontal axis: number of centers, vertical axis: classification accuracy (%).

Then, for the detected optimum number of hidden layer neurons, the experiment was repeated 1 to 100 epochs as it is shown in Figure 4.

To find the appropriate number of centers in the RBF classifier, we varied them from 1 to 60 centers. The RBF was performed on validation data with different number of centers and the accuracy rate was computed each time. The number of centers corresponding to the best accuracy rate was set as the number of centers of the RBF classifier. Next, RBF was performed on the test data.

In this classifier, we used K-means algorithm to find out the location of each center as follows. First, we initialized centers using data randomly. Then, each data was classified according to its distance from all centers. When all data were classified, the mean of each class was computed and was set as its new center.

The classification process was performed several times till no significant changes to were found in centers locations. For example, the accuracy rates of the RBF classifier performed on first subject's data is plotted in Figure 5.

Then, for detected optimum number of centers, the experiment was repeated 1 to 100 epochs as it is shown in Figure 6.

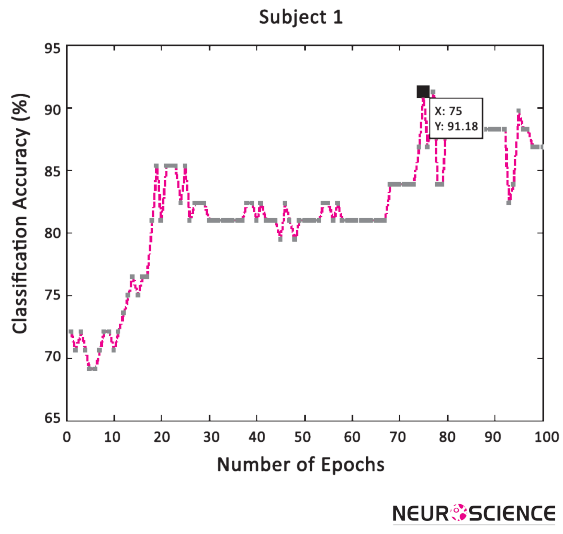


Figure 6. Determination of the optimum number of epochs for RBF classifier using k-means algorithm for Subject 1. Horizontal axis: number of epochs, vertical axis: classification accuracy (%).

Since each person's brain functionality is unique, his response to stimulus shown on the screen might be different from others. For the same reason, each subject responses in specific time and frequency intervals which might be different from the others (Wolpow, Birbaumer, McFarland, Pfurtscheller & Vaughan, 2002).

Thus, in this paper, we examined different time and frequency intervals for each subject and determined his

proper time and frequency intervals. This procedure is called Calibration in BCI system. (Wolpow, Birbaumer, McFarland, Pfurtscheller & Vaughan, 2002)

To find out the best time intervals each subject who performed motor imagery, different time intervals were examined with CSP feature extraction method and MLP classifier. Performances obtained in different time intervals are displayed in Table 2.

Then, 8-30 Hz (μ,β), 8-40 Hz (μ,β and 10 Hz of γ), 8-50 Hz (μ,β and 20 Hz of γ) and 30-50 Hz (20 Hz of γ) were separated by Discrete Fourier Transform (DFT) and Band Pass filter on the best obtained time interval.

To demonstrate the meaningful differences between using combination of μ and β and combination of μ, β and γ frequency bands we used t-test which is a statistical test. The results are shown in Table 4.

To do so, first, the best time interval was selected for each subject separately. Then, we separated different frequency intervals (i.e. ranges 8-30 Hz, 8-40 Hz, 8-50 Hz and 30-50 Hz) .The MLP was performed on each of those frequency intervals with optimized number of hidden layer and epochs, iterated 10 times. Finally, we compared them two by two and the t-values were found. The degree of freedom was computed 18.

As it can be seen, in all cases there exist meaningful differences among different bands. For the first subject, there is no meaningful difference between 8-30 Hz and 8-40 Hz frequency bands. However, t-values for other bands are much higher than the t-value with the p-value equal to 0.0001 which shows that there are meaningful differences among bands with the accuracy rate 99.9% gained by the used classifier.

For numbers 2, 5 and 7, the computed t-values are higher than the t-value with the p-value equal to 0.0001.

For the third subject, a meaningful difference is seen between 8-30 Hz and 8-40 Hz bands (i.e. p-value<0.05) and the accuracy rate is 95%. The rests have 99.9% accuracy rate.

There are no meaningful differences between 8-30 Hz and 8-50 Hz bands for the fourth subject. But the computed values for t in 8-30 Hz and 8-40 Hz are higher than the t-value with the p-value equal to 0.05 as well as t-values for 8-40 Hz and 8-50 Hz bands. For the remaining bands, meaningful difference can be seen with 99.9% accuracy rate.

For subject 6, the accuracy rate for 8-40 Hz and 8-50 Hz (p-value <0.01) is 99% as well as 8-50 Hz and 30-50 Hz frequency bands. For the other frequency bands it is calculated (p-value<0.001) 99.9%.

Table 2. Determination of the best time intervals (CSP feature extractor and MLP classifier with optimum number of hidden layer neurons and epochs and learning rate=0.2) (S1 to S9 stand for Subject1 to Subject9). Classification accuracies of all time intervals from 2 to 7 seconds were computed by 4-fold cross validation process. The best accuracy rates were bolded.

Time Intervals (Sec)	S1 (%)	S2 (%)	S3 (%)	S4 (%)	S5 (%)	S6 (%)	S7 (%)	S8 (%)	S9 (%)
2-7	**74.47**	48.52	72.30	3038	38.19	46.87	63.19	62.84	64.93
2-6	73.95	48.61	**80.38**	50.00	53.81	47.22	70.13	73.95	77.08
2-5	74.30	51.38	71.52	**62.84**	53.12	**47.56**	79.51	73.95	**78.81**
2-4	63.88	42.27	65.27	61.02	48.61	45.13	78.81	65.27	78.47
2-3	44.96	40.01	42.70	54.07	**57.63**	43.05	69.79	66.31	56.59
3-7	63.02	**58.15**	77.86	29.51	46.52	46.18	58.33	68.75	51.38
3-6	72.82	41.92	80.20	40.36	48.26	44.44	76.73	**80.20**	70.83
3-5	59.89	51.21	73.95	50.00	53.81	40.27	**79.86**	76.38	72.22
3-4	58.42	40.71	68.66	46.52	50.69	38.54	65.97	63.88	68.75
4-7	59.20	50.95	68.92	27.34	42.01	44.44	50.69	63.19	46.17
4-6	62.58	53.55	68.57	30.29	39.93	42.70	65.97	65.62	57.98
4-5	57.63	46.87	61.63	42.70	40.97	39.23	67.36	63.88	64.58
5-7	52.77	38.62	58.50	26.90	36.11	38.19	46.52	44.79	31.59
5-6	58.50	47.56	59.72	26.04	40.62	35.06	60.41	48.95	36.80
6-7	39.49	46.61	43.22	26.47	34.02	32.63	47.91	44.09	29.51

There is no meaningful difference between 8-30 Hz and 8-50 Hz frequency bands for subject 8. However, t-values for other frequency bands are higher than the t-value with the p-value equal to 0.001.

The accuracy rate of 8-30 Hz and 8-40 Hz frequency bands is 95% (p-value <0.05) and for the other frequency bands it is (p-value<0.001) 99.9% for subject 9.

Table 3. t-values and p-values from t-table with 18 degrees of freedom.

p_value	t_value
0.05	2.10
0.01	2.88
0.001	3.92

NEUROSCIENCE

Table 4. Calculated t-values in t-test. (S1 to S9 stand for Subject1 to Subject9).

Frequency Intervals (Hz)	S1	S2	S3	S4	S5	S6	S7	S8	S9
(8-30) and (8-40)	1.81	9.19	2.04	2.06	13.81	8.75	4.52	8.76	2.15
(8-30) and (8-50)	7.06	5.66	17.65	0.56	9.28	7.89	8.65	0.60	35.67
(8-30) and (30-50)	87.54	31.27	191.26	84.61	44.80	28.17	77.03	52.34	32.55
(8-40) and (8-50)	4.34	36.98	24.33	2.43	21.77	3.44	12.87	8.89	46.28
(8-40) and (30-50)	82.34	50.40	271.68	74.98	53.26	13.92	89.31	35.47	38.04
(8-50) and (30-50)	82.66	81.33	256.78	77.73	36.69	3.02	55.37	50.51	5.55

NEUROSCIENCE

Then features are obtained by CSP feature extraction method and classification accuracy are obtained by MLP classifier. Results of this experiment are displayed in Table 5.

Table 5. Determination of suitable frequency ranges using CSP feature extraction method and MLP classifier with optimum number of hidden layer neurons and epochs. The best accuracy rates were bolded. (S1 to S9 stand for Subject1 to Subject9). 8-30 Hz is μ and β frequency ranges and 30< Hz is γ frequency range.

Frequency Intervals (Hz)	S1 (%)	S2 (%)	S3 (%)	S4 (%)	S5 (%)	S6 (%)	S7 (%)	S8 (%)	S9 (%)
8-30	74.47	58.15	80.38	62.84	57.63	**47.56**	**79.86**	**80.20**	**78.81**
8-40	75.00	**70.48**	**83.33**	70.83	**63.54**	46.52	79.86	77.77	71.18
8-50	**78.12**	64.93	81.94	**73.61**	58.68	43.75	75.00	79.86	62.84
30-50	25.95	39.75	25.26	22.22	38.54	30.20	42.70	39.58	43.15

NEUROSCIENCE

As you can see in Table 3, for 8-50 Hz for subjects 1 and 4, 8-40 Hz for subjects 2 & 3 and 8-30 Hz for subjects 6, 8 and 9, MLP had the best performance. 8-30 Hz and 8-40 Hz for subject 7 showed the same performance. However, 8-40 Hz for subject 6 and 8-50 Hz for subject 8 had about 1% difference in performance.

We have done our experiment using other methods such as CSP-LDA (Lotte, 2011), CSP-SVM (Wang, Miao & Blohm, 2012) and CSP-KNN (Liyanage, Xu, Guan, Ang & Lee, 2010) on the obtained optimum time interval and different frequency ranges (8-30 Hz, 8-40 Hz, 8-50 Hz and 30-50 Hz), too. In KNN classifier, k was assumed 5.

Table 6 displays LDA accuracy rates determining the suitable frequency ranges for each subject.

As it can be seen, for subjects 1 and 3, 8-50 Hz frequency range, for subjects 2, 6,8 and 9, 8-40 Hz frequency range and for subjects 2, 5 and 7, 8-30 Hz frequency range have had the best accuracies. But for Subject 2 the obtained accuracy rate in 8-30 Hz frequency range had

Table 6. Determination of suitable frequency ranges using CSP feature extraction method and LDA classifier. The best accuracy rates were bolded. (S1 to S9 stand for Subject1 to Subject9). 8-30 Hz is μ and β frequency ranges and 30< Hz is γ frequency range.

Frequency Intervals (Hz)	S1 (%)	S2 (%)	S3 (%)	S4 (%)	S5 (%)	S6 (%)	S7 (%)	S8 (%)	S9 (%)
8-30	75.34	**59.72**	76.38	52.77	**39.58**	45.83	**71.53**	5.34	66.66
8-40	76.04	59.02	76.73	**58.68**	38.19	**47.22**	71.18	**76.74**	**74.31**
8-50	**77.77**	56.94	**78.82**	55.56	29.17	43.06	66.67	73.61	70.83
30-50	32.29	34.38	30.21	25.00	25.00	24.31	42.01	32.29	50.00

NEUROSCIENCE

0.7% difference in performance compared to 8-40 Hz. Subject 5 had 1% and subject 7 had 0.4% difference in performance compared to 8-40 Hz. Therefore, existence of the γ band have had a positive effect. Table 7 displays SVM accuracy rates determining the suitable frequency ranges for each subject.

Table 7. Determination of suitable frequency ranges using CSP feature extraction method and SVM classifier using linear kernel function. The best accuracy rates were bolded. (S1-S9 stand for Subject1- Subject9). 8-30 Hz is μ and β frequency ranges and 30< Hz is γ frequency range.

Frequency Intervals (Hz)	S1 (%)	S2 (%)	S3 (%)	S4 (%)	S5 (%)	S6 (%)	S7 (%)	S8 (%)	S9 (%)
8-30	72.22	61.45	74.65	**58.33**	40.27	45.86	68.75	**77.83**	68.40
8-40	**74.65**	**64.93**	74.30	54.81	**40.97**	**46.18**	**69.44**	77.08	**74.30**
8-50	71.87	61.11	**75.34**	52.77	33.33	44.44	67.36	75.34	64.23
30-50	35.06	43.05	29.51	25.34	25.00	27.43	37.50	45.13	31.59

NEUROSCIENCE

As can be seen in Table 7, for subjects 1, 2, 5, 6,7 and 9, 8-40 Hz frequency range have had the best accuracies. For subject 3, 8-50 Hz and for subjects 4 and 8 8-30 Hz has had the best performance. But for subject 8 the obtained accuracy for 8-30 Hz had 0.8% difference in performance compared to 8-40 Hz frequency range. Table 8 displays KNN accuracy rates determining the suitable frequency ranges for each subject.

Table 8. Determination of suitable frequency ranges using CSP feature extraction method and KNN classifier using Euclidean distance criterion. The best accuracy rates were bolded. (S1 to S9 stand for Subject1 to Subject9). 8-30 Hz is μ and β frequency ranges and 30< Hz is γ frequency range.

Frequency Intervals (Hz)	S1 (%)	S2 (%)	S3 (%)	S4 (%)	S5 (%)	S6 (%)	S7 (%)	S8 (%)	S9 (%)
8-30	67.36	**61.11**	67.36	49.65	**46.87**	37.84	**57.63**	61.80	67.01
8-40	61.80	59.02	**68.05**	54.51	43.05	**43.75**	56.59	**64.93**	**69.09**
8-50	**69.44**	57.63	**68.05**	**55.55**	25.69	39.93	49.30	**64.93**	64.93
30-50	32.29	40.62	28.47	28.47	25.34	31.25	37.50	39.58	47.56

NEUROSCIENCE

As it can be seen, KNN classifier (k=5) had the best accuracy rate for subjects 3, 6, 8 and 9 on 8-40 Hz frequency range and for subjects 1 and 4 on 8-50 Hz frequency range. Also, this classifier had the best performance on 8-30 Hz frequency range for subjects 2, 5 and 7. But for subject 2 and 3 the obtained accuracy rate for 8-30 Hz frequency range had 2% and 1% differences in performance compared to 8-40 Hz, respectively.

In 4 above tables, 30-50 Hz frequency range (only γ band) have never had the best performance. Therefore, it can be concluded γ band merely has not sufficient information for movement imagination classification.

As a result, combination of μ,β and γ bands would have better performance than μ,β in motor imageries classification.

Kappa Coefficient

In order to compare the results of our experiments with previous studies, we had to find kappa coefficient. Kappa coefficient is a measure of classification interpretation and is computed using the following formula.

$$k = \frac{P_o - P_c}{1 - P_c} \quad (10)$$

Where P_o is the relative observed agreement between raters, and P_c is the hypothetical probability of chance agreement. (Sim & Wright, 2005)

The kappa values gained by first three winners' methods in the BCI Competition IV ,Multi- Segment Joint Approximate Diagonalization (MSJAD) (Gouy-Pailler, Congedo, Brunner, Jutten & Pfurtscheller, 2010), and Five-Stage Decoding of EEG (FSDE) (Wang, Miao & Blohm, 2012) methods and our proposed method are displayed in Table 9. As it can be seen, our proposed method achieved significant results and got better results in 5 subjects than first three winners' methods in the BCI competition IV. (Subjects 1,2,4,5 and 6). Furthermore, our method got better performance compared to MSJAD method for 6 subjects (Subjects 1,2,4,6,7 and 9), and in comparison with FSDE method for 5 subjects (Subjects 1,2,3,4 and 9).

In general, it was superior to other approaches in 3 subjects and as no artifact removal was done in our method, this approach seemed to be the winner among all these methods.

Table 9. kappa values obtained by first three winners' methods in the BCI Competition IV, MSJAD and FSDE methods and our proposed method. Best kappa values are bolded.

Subject	1stwinner's Method	2ndwinner's Method	3rd winner's Method	MSJAD Method	FSDE Method	Proposed Method
S1	0.68	0.69	0.38	0.66	0.56	**0.70**
S2	0.42	0.34	0.18	0.42	0.41	**0.53**
S3	0.75	0.71	0.48	**0.77**	0.43	0.72
S4	0.48	0.44	0.33	**0.51**	0.41	**0.52**
S5	0.4	0.16	0.07	**0.5**	**0.68**	0.43
S6	0.27	0.21	0.14	0.21	**0.48**	0.3
S7	0.77	0.66	0.29	0.3	**0.8**	0.62
S8	**0.75**	0.73	0.49	0.69	0.72	0.69
S9	0.61	**0.69**	0.44	0.46	0.63	0.66
Mean	0.57	0.51	0.31	0.5	0.57	0.57
Std.	0.18	0.23	0.15	0.18	0.15	0.14

NEUROSCIENCE

In order to show that γ band is effective in performance improvement, we counted the number of frequency bands in which any of 4 classifiers (MLP, LDA, SVM and KNN) had the best accuracy rates and placed them in Table 10. As it can be seen, 8-40 Hz frequency range has the first rank. Thus, it can be deduced that presence of γ band is effective in performance improvement. Also, we can see 8-40 Hz range ranks first, 8-30 Hz ranks second and 8-50 Hz ranks third. In other words, wider frequency range may not improve performance necessarily.

Table 10. The ranking of frequency intervals. (S1 to S9 stand for Subject1-Subject9)

Frequency Intervals (Hz)	S1	S2	S3	S4	S5	S6	S7	S8	S9	Total Wins
8-30	0	2	0	1	2	1	3	2	1	12
8-40	1	2	2	1	2	3	1	2	3	17
8-50	3	0	3	2	0	0	0	1	0	9
30-50	0	0	0	0	0	0	0	0	0	0

NEUROSCIENCE

4. Discussion

In this paper, it's been showed that using combination of μ, β and γ bands would have better performance than combination of μ,β in motor imageries classification. Different time and frequency intervals were experimented in order to find the best time and frequency intervals for each subject. CSP was used as feature extractor and MLP, RBF, LDA, SVM and KNN were used classifiers.

The average of classification accuracy using combination of μ, β and γ is 72.25% and it is 69.34% using combination of μ and β. Thus, the improvement rate for all subjects is 2.91%.

Future works are going to study different feature extraction and classifier combination methods.

5. Acknowledgement

The authors wish to thank Shahid Rajaee Teacher Training University for funding this project.

References

Acar, E. (2011). Classification of Motor Imagery Tasks in EEG Signal and its Application to a Brain-Computer Interface for Controlling Assistive Environmental Devices. Master Thesis. The Graduate School of Natural and Applied Sciences of Middle East Technical University, 59-65.

Ang KK, Chin ZY, Wang C, Guan C, Zhang H. (2012). Filter Bank Common Spatial Pattern Algorithm on BCI Competition IV Datasets 2a and 2b, Institute for Infocomm Research, Agency for Science, Technology and Research (ASTAR) Singapore.

Aoki, F., Fetz, E. E., Shupe, L., Lettich, E., &Ojemann, G. A. (1999).Increased gamma-range activity in human sensorimotor cortex duringperformance of visuomotor tasks.Clinical Neurophysiology, 110(3), 524-537.

Balakrishnama, S., &Ganapathiraju, A. (1998). Linear Discriminant Analysis - A brief tutorial, Institute for Signal and Information Processing [Adobe Digital Editions version].Retrieved from: www.music.mcgill.ca/~ich/classes/mumt611_07/classifiers/lda_theory.pdf.

Ball, T., Demandt, E., Mutschler, I., Neitzel, E., Mehring, C., Vogt, K., et al. (2008). Movement related activity in the high gamma range of thehuman EEG. NeuroImage, 41, 302-310. doi:10.1016/j.neuroimage.2008.02.032.

Ball, T., Nawrot, M.P., Pistohl, T., Aertsen, A., Schulze-Bonhage, A., &Mehring, C. (2004).Towards an implantable brain-machine interface based on epicortical field potentials.Biomed.BiomedizinischeTechnik, Biomedical Engineering,49, 756-759.

Bishop, C. M. (1995). The Multi-Layer Perceptron, Neural Networks for Pattern Recognotion (pp. 116-122). Birmingham, UK: CLARENDON PRESS• OXFORD.

Brunner, C., Leeb, R., Muller-Putz, G. R., Schlogl, A., &Pfurtscheller, G. (2008). BCI Competition 2008 – Graz data set A [Adobe Digital Editions Version]. Retrieved from: http://www.bbci.de/competition/iv/desc_2a.pdf .

Chang,C.C.,andLin,C.J.(2001).Libsvm:ALibraryforSupportVectorMachines.Availableat:http://www.csie.ntu.edu.tw/scjlin/libsvm

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.

Criag, E. T., Jurgen, K. (2005). Reference-free quantification of EEG spectra: Combining current source density (CSD) and frequency principal components analysis (fPCA). Clinical Neurophysiology, 116, 2826-2846. doi:10.1016/j.clinph.2005.08.007.

Crone, N.E., Miglioretti, D.L., Gordon, B., & Lesser, R.P. (1998). Functional Mapping of Human sensorimotor cortex with electrocorticographic spectral analysis II.Event-related synchronization in the gamma band.Brain, 121, 2301-2315.

Denoeux, T. (2008). A k-nearest neighbor classification rule based on Dempster-Shafer theory. Classic work of the Dempster-Shafer theory of belief functions, 737-760.

Dym, H., & McKean, H. (Eds). (1972). Fourier Series and Integrals. Academic Press, New York/ London.

Haykin, S. (1999). Multilayer Perceptrons,Neural Networks A Comprehensive Foundation (2nded.). (pp. 159-167). Canada: Prentice Hall International, Inc.

Haykin, S. (1999). Radial-Basis Function Networks,Neural Networks A Comprehensive Foundation (2nded.).(pp. 257-260). Canada: Prentice Hall International, Inc.

Heung-Il, S., &Seong-Whan, L. (2011). A probabilistic approach to spatio-spectral filters optimization in Brain-Computer Interface.Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on, 19-24. doi: 10.1109/ICSMC.2011.6083636.

Koles, Z. J. (1991). The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG. Electroenc.Clin.Neurophys, 79(6), 440-447.

Kwang-Eun, K., &Kwee-Bo, S. (2011). Optimized Adaptive Neuro-Fuzzy Inference Systemfor Motor Imagery EEG signals Classifications. 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2, 669-672. doi: 10.1109/FSKD.2011.6019759.

Leuthardt, E. C., Schalk, G., Wolpow, J. R., Ojemann, J. G., & Moran, D. W. (2004). A brain–computer interface using electrocorticographic signals in humans. Journal Of Neural Engineering, 1(2), 63-71. doi:10.1088/1741-2560/1/2/001.

Liyanage, S. R., Xu, J. X., Guan, C., Ang, K. K., & Lee, T. H. (2010, August). Multi-Class Motor Motion Imagery using Common Spatial Patterns based on Joint Approximate Diagonalization. In IASTED Technology Conferences. ACTA Press.

Lodder, S. (2009). Single-trial classification of an EEG-based brain computer interface using the wavelet packet decomposition and cepstral analysis (Doctoral dissertation, Stellenbosch: University of Stellenbosch).

Lotte, F. (2011). Generating Artificial EEG SignalsTo Reduce BCI Calibration Time. In 5th International Brain-Computer Interface Workshop (pp. 176-179).

Miller, K. J., Leuthardt, E. C., Schalk, G., Rao, R. P. N., Anderson, N. R., Moran, D. W., et al. (2007). Spectral changes in cortical surface potentials during motor movement. The Journal of Neuroscience, 27(9), 2424-2432.doi: 10.1523/JNEUROSCI.3886-06.2007.

Müller-Gerking, J., Pfurtscheller, G., &Flyvbjerg, H. (1999). Designing optimal spatial filters for single-trial EEG classification in a movement task.Electroenc.Clin.Neurophys, 110, 787-798.

Niedermeyer, E., & DA Silva, F.L. (2005). Electroencephalography: Basic Principles, Clinical Applications and Related Fields (5th ed.). Philadelphia: LIPPINCOTT WILLIAMS & WILKINS.

Ohara, S., Ikeda, A., Kunieda, T., Yazawa, S., Baba, K., Nagamine, T., et al. (2000). Movement-related change of electrocorticographic activity in human supplementary motor area proper. Brain, 123, 1203-1215.

Pfurtscheller, G., Graimann, B., Huggins, J. E., Levine, S. P, &Schuh, L. A. (2003). Spatiotemporal patterns of beta desynchronization and gamma synchronization in corticographic data during self-paced movement. Clin.Neurophysiol, 114, 1226-1236.doi:10.1016/S1388-2457(03)00067-1.

Ramoser, H., Muller-Gerking, J., &Pfurtscheller, G. (2000).Optimal Spatial Filtering of Single Trial EEG DuringImagined Hand Movement.IEEE TRANSACTIONS ON REHABILITATION ENGINEERING, 8(4), 441-446.

Schloegl, A., Neuper, C., &Pfurtscheller, G. (1997).Subject specific EEG patterns during motor imagery.Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE, 4, 1530-1532. doi: 10.1109/IEMBS.1997.757001.

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Physical therapy, 85(3), 257-268.

Wang, D., Miao, D., & Blohm, G. (2012) Multi-class motor imagery EEG decoding for brain-computer interfaces. Frontiers in Neuroscience, 6, 151.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. Clinical Neurophysiology, 113, 767-791.

Wu, W., Gao, X., & Gao, S. (2006, January). One-versus-the-rest (OVR) algorithm: An extension of common spatial patterns (CSP) algorithm to multi-class case. In Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference.

Yosida, K. (1968). Functional Analysis. (2nded.). Berlin: Springer-Verlag.