JBEMi JOURNAL OF BIOMEDICAL ENGINEERING AND MEDICAL IMAGING

SOCIETY FOR SCIENCE AND EDUCATION UNITED KINGDOM

VOLUME 2 ISSUE 5
ISSN: 2055-1266

Classification of EEG Signals Produced by RGB Colour Stimuli

[1]Saim Rasheed and [2]Daniele Marini
[1]Department of Information Technology, Faculty of Computing and IT
King Abdulaziz University, KSA
[2]Dipartimento di Informatica, Università degli Studi di Milano, Italia
srahmed@kau.edu.sa

ABSTRACT

In this paper we have presented results for classification of electroencephalograph (EEG) signals produced by the random visual exposure of primary colours i.e. red, green and blue to the subject while sitting in a dark room. Event-related spectral perturbations (ERSP) are used as features for support vector machine (SVM). Our objective was to classify the EEG signals as Red, Green and Blue classes and we have successfully classified the three visual conditions having accuracy of 84%, 89% and 98% with linear, polynomial and radial basis function kernels respectively with in all the groups of data among all the subjects.

1 Introduction

In virtual and physical environment, a person can control the devices such as movement of cursor on monitor screen and controlling wheel chair using EEG-based non-invasive Brain-Computer interface (BCI) applications (Pires et al. 2008; Leeb et al. 2006). The electrical activity produced by the brain is recorded using electrodes placed over the scalp. The experiments are usually designed to record EEG signals as P300, or mu rhythm based event-related potentials or ERP that can more precisely describe the dynamics of brain related activities. For instance, BCI users concentrate on imagination of hand or foot movement (Pfurtscheller & Neuper 2001) or fixate their gaze on monitor screen to interact with Speller applications (Farwell & Donchin 1988). These mental activities are selected in such a way to activate the related brain area and may vary widely across different experiments, subjects and the related application. Such BCI systems performances may be improved with the help of feedback. In (Wolpaw et al. 2002), Wolpaw et al has presented an overall generic review of brain-computer interface systems that provide comprehensive detail on fundamental studies. The most meaningful information exist in the frequency bands, delta (0.1 – 4 Hz), theta (4 – 8 Hz), alpha (8 – 12 Hz) and beta (12 – 30 Hz). These are most commonly known frequency bands and depend on researcher's purpose for the analysis of a particular band to be utilized in a certain case. Usually, EEG signals recorded from scalp are very weak and may vary in the range from 5 to 100 microvolts during recording and amplified later for further signal processing. During recording of EEG signals, noise such as main's electrical impulse and other artifacts like EOG/EMG signals may interfere. Although, the subjects are instructed to avoid eye blinking and/or muscle movement, even then the signals are processed to eliminate such information using mathematical algorithms e.g. Independent component analysis (ICA) (Hallez et al. 2009; Jung et al. 2000) and principal component analysis (PCA) (Foresta et al. 2009) before performing classification task.

DOI: 10.14738/jbemi.25.1566
Publication Date: 29th October 2015
URL: http://dx.doi.org/10.14738/jbemi. 25.1566

The ultimate task of any BCI system is the classification of EEG signals (Chiappa & Barber 2006; Guan & Chen 2008; Lotte et al. 2007; Benjamin et al. 2002; Krusienski et al. 2008) which reflects the growing interest of researchers in EEG-based BCI and present published results doing analysis, evaluation and assessment of classification algorithms. For BCI applications, it is necessary to implement online classification in order to see real time execution which is a challenging task for signal processing and machine learning experts. However, an earlier offline analysis of EEG signals helps us in improving our classification accuracies. Once the brain signal is classified, it is fed to the outer world application to perform the desired operation. In this paper, we have performed an experiment and presented the results for offline classification of EEG signals recorded from the scalp, produced by primary colours stimuli, red, green and blue, presented at random. The purpose of the experiment is to verify, if either the observation of different real colours or their corresponding imagination of colours can be detected in the selected EEG frequency combination, and to select best frequency combination to maximize differences through colour signals in order to find a Way-In to further establish our argument and to provide a baseline to be compared with more complex visual stimuli to evaluate the effects of colours to navigate in an immersive VR environment for developing future BCI applications (Leeb et al. 2006). For example, we can develop a virtual reality (VR) application to navigate into VR environment. In a first and preliminary version, developed application will exhibit the traffic light signals. A moving vehicle may be stopped when turning ON the Red light, upon recognition of red colour through EEG signals. Similarly, upon recognition of the green light, a stopped vehicle will start moving. We suggest this procedure may be a faster and more effective communication channel than those based on motor imagery described in [2]. This way, we may use colours as controllable parameter in VR environments. Another application of such colour recognition system could be for colour blind and/or blind people (Dobelle 2000). These applications are quite novel in their fields and needs extensive collaborative research work in different domains.

2 Experimental Design

Concerning the experimental settings, seven male subjects have undertaken the experiment, age ranging from 20 to 36 years. All subjects were free of neurological and psychiatric disorders and have normal colour vision. Experiment was performed in a dark room in front of a large curved screen on which the colour was presented as a square with size of 10 degree angle on a much wider gray background. The subjects were seated on a comfortable chair. Distance between the subjects and the screen was 3.5 meters and 1.5 meters high from the ground. The luminosity of each colour was kept constant at 4.5 cd/m2 and measured using the device Minolta SPOTMETER F. Screw-able gold EEG electrodes were used on the subject's scalp at P3, P4, O1 and O2 sites and referenced to right ear lobe and grounded at site AFz as shown in figure 1.

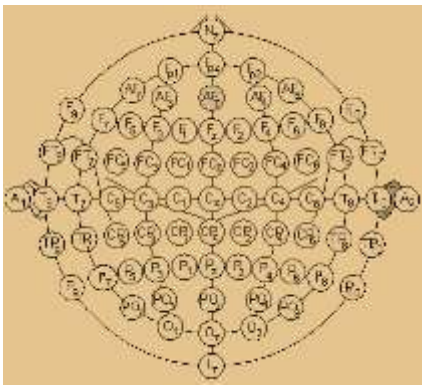

Figure 1. Electrode placement layout.

Experiment protocol is shown in figure 2 that depicts the duration of one sequence. In this protocol, each colour was presented for three seconds, twice in one sequence and the subject was instructed to imagine the same colour for three seconds in the same sequence twice. There are eight events in one sequence. Only one colour is presented in one sequence. After all the events are occurred in a sequence then the next sequence is started and another colour is presented. Uniform gray background colour was displayed before and after every target colour to balance possible after-effects of the RGB stimuli (Yoto et al. 2007). All impedances were kept below 5kΩ. EEG signals were recorded using BCI2000 (Schalk et al. 2004) with g.tec's g.MOBIlab+ portable device sampled at 256 Hz, processed and analyzed offline using EEGLAB (Delorme & Makeig 2004) that runs with MATLAB. Signals were band pass filtered from 0.1 Hz to 30 Hz. Each colour was presented 60 times randomly, resulting in 60 trials of each colour from each subject. Each trial contains 768 data points after the onset of stimulus. Once, all the EEG signals were recorded they were brought into EEGLAB for offline processing and analysis. EEGLAB has a vast compatibility to import EEG signals into MATLAB's workspace which are recorded with different devices using BCI2000. Epochs were extracted from continuous data for each colour. Each epoch lasts for three seconds i.e. one second before the event occurred and two seconds afterwards. To reduce the effects of abnormal values, those trials for which EEG signals crossed ± 60 µV were rejected and some segments were also dropped by visual inspection that did not cross ± 60 µV but were unreliable in further computations. Some trials were dropped due to artifacts like EOG/EMG signals, from each colour for each subject.

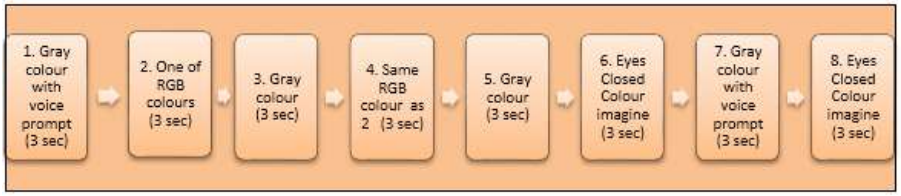

Figure 2. Experiment protocol

2.1 Artifact Rejection

ICA is a method that provides us the capability to solve the problem of Blind Source Separation. It can identify N independent source signals, $S=\{s_1(x),s_2(x),\cdots s_N(x)\}$ (e.g. different voices, music or other sound sources) from N linear mixtures matrix $M=\{m_1(x),m_2(x),\cdots m_N(x)\}$, that can be modeled by multiplying the source matrix S by an unknown square matrix T such that

$$M = TS \tag{1}$$

In order to discriminate the source signals from mixtures, it is assumed that sources s_i are independent and their mixtures are not. Having no prior knowledge of source signals or about the process, how the signals were mixed, the objective is to approximate another matrix $U = \{u_1(x), u_2(x), \cdots u_N(x)\}$, identical to source signals by specifying the filter to linearly invert the mixing process such that

$$U = VM \tag{2}$$

where V is a square matrix.

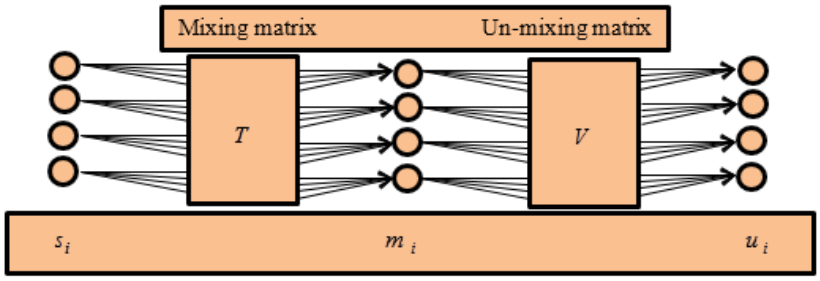

Figure 3. The general process of ICA.

In our case, we have four electrodes that measure the four mixtures of signals. Applying ICA decomposes the mixtures into four independent source signals. The general overview of process of ICA is shown in figure 3. Electrode measurements m_i are assumed to be composed of a linear mixture of independent sources s_i . Un-mixing matrix V is produced using ICA, which decompose m_i to estimate the independent sources u_i . After identification of components, the one that contains EOG/EMG artifacts is subtracted and artifact free EEG is obtained for further feature extraction and classification. Software routines are available within EEGLAB for detection of artifacts based on ICA, however marking and rejection of artifacts is done manually before feature extraction and classification tasks. A detailed review on theory of ICA and their practical use is discussed in (Jung et al. 2000; James & Hesse 2005; Jung et al. 2000; Stone 2004).

2.2 Event-Related Spectral Perturbation as Features

ERSP is a measure to study the event-related brain dynamics. It reflects the information about variation in power at different frequencies at a certain time point. Makeig in (Makeig 1993) reported that according to several studies, event-related potentials (ERPs) are not capable of capturing maximum brain's response to events due to their instability and not being fully independent of EEG. Due to this limitation of ERPs being used as features may not classify satisfactorily, which we also tried to use in our classifier but attained less accurate results. On the other hand, we have used event-related spectral perturbation (ERSP) (Makeig et al. 2004; Makeig 1993; Huang et al. 2005) values in terms of time-frequency measurements as features for the classifier and the results are dramatically improved. "*To compute an ERSP, baseline spectra are calculated from the EEG immediately preceding each event. The epoch is divided into brief, overlapping data windows, and a moving average of the amplitude spectra of these is created. Each of these spectral transforms of individual response epochs are then normalized by dividing by their respective*

mean baseline spectra. Normalized response transforms for many trials are then averaged to produce an average ERSP" (Makeig 1993) .

2.3 Support Vector Machine as Classifier

We have used SVM as classifier because it has been proved to be very good in classification tasks (Kaper et al. 2004; Rakotomamonjy et al. 2005; Garrett et al. 2003; Benjamin et al. 2002; Peterson et al. 2005). SVM constructs a hyperplane in the simplest case that is used for classification to separate the data points belonging to two different classes. This hyperplane is chosen in such a way to maximize the distance margin δ , as shown in figure 4 and 5, to the closest training data points of any class. In general, the larger the margin is, the higher is the classification accuracy. SVM may broadly be categorized as hard margin and soft margin (Abe 2005). However, the SVM theory was developed by Vapnik (Vapnik 2000) and a good review on classification algorithms for BCI research is presented in (Lotte et al. 2007). To classify M, $m-$ dimensional training data points x_i ($i=1,\cdots,M$) which either belong to class 1 or class 2 and the associated labels be $y_i=1$ for class 1 or $y_i=-1$ for class 2. A decision function is defined as $F(x)=w^T x+b$, where w is an $m-$ dimensional vector and b is the bias term. If the data is linearly separable then no data points will satisfy

$$w^T x+b=0 \tag{3}$$

So, for $i=1,\cdots,M$

$$w^T x_i+b\begin{cases}\geq 1 & y_i=1\\ \leq -1 & y_i=-1\end{cases} \text{ for} \tag{4}$$

Equation (4) may be written as

$$y_i(w^T x_i+b)\geq 1 \text{ for } i=1,\cdots,M\ . \tag{5}$$

The function,

$$F(x)=w^T x+b=c \text{ for } -1<c<1 \tag{6}$$

defines the hyperplane that separates the data points x_i ($i=1,\cdots,M$). For $c=0$, the separating hyperplane lies in the middle of two hyperplanes with $c=1$ and -1. Pictorial representation of SVM in case of two class problem for hard margin is shown in figure 4. The optimal separating hyperplane can be obtained by solving the optimization problem,

Minimize,

$$Z(w)=\frac{1}{2}\|w\|^2 \tag{7}$$

subject to the constraints,

$$y_i(w^T x_i+b)\geq 1 \text{ for } i=1,\cdots,M \tag{8}$$

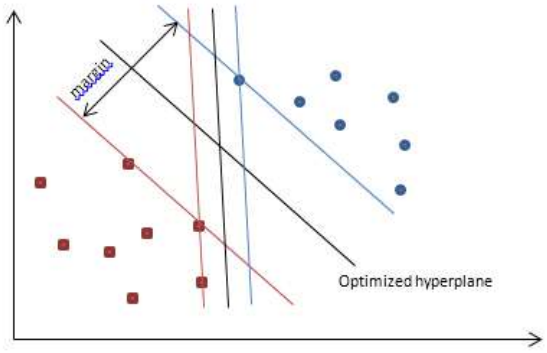

Figure 4. A separating hyperplane in two class problem for hard margin case.

Equation (7) forms a quadratic optimization problem with inequality constraints expressed by (8). This constrained optimization problem can be solved by solving its equivalent unconstrained problem using the saddle points of the Lagrangian functional,

$$Z(w,b,\alpha)=\frac{1}{2}w^T w-\sum_{i=1}^{M}\alpha_i\left[y_i(w^T x_i+b)-1\right] \tag{9}$$

Where $\alpha=(\alpha_1,\cdots,\alpha_M)^T$ and α_i are the non-negative Lagrange multipliers. Lagrangian must be minimized with respect to w and b and needs to be maximized w.r.t $\alpha_i \geq 0$. This problem can be solved into its primal space of parameters w and b and also in dual space of lagrange multipliers α_i. The approach of dual space gives us insightful results as it has the number of variable as the number of training data. Equation (9) is manipulated with Karush-Kuhn-Tucker (KKT) conditions (Abe 2005) to get the dual pro

Maximize,

$$Z(\alpha)=\sum_{i=1}^{M}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{M}\alpha_i\alpha_j y_i y_j x_i^T x_j \tag{10}$$

with respect to α_i subject to the constraints

$$\sum_{i=1}^{M} y_i\alpha_i=0 \text{ and } \alpha_i \geq o \qquad \text{for } i=1,\cdots,M \tag{11}$$

which is a hard margin SVM. Using lagrangian functional in (9) with KKT conditions we obtain

$$w=\sum_{i=1}^{M}\alpha_i y_i x_i \tag{12}$$

and

$$\sum_{i=1}^{M}\alpha_i y_i=0 \tag{13}$$

to finally get the decision function as

$$F(x)=\sum_{i\in S}\alpha_i y_i x_i^T x+b \tag{14}$$

where S is the set of support vector indices along with

$$b=\frac{1}{|S|}\sum_{i\in S}\left(y_i-w^T x_i\right) \tag{15}$$

Hence the unknown datum x belongs to either class 1 if $F(x) > 0$, or class 2 if $F(x) < 0$ and remains unclassifiable if $F(x) = 0$.

Now SVM introduces slack variables $\xi_i \geq 0$ in case of soft margins and the optimization problem turned into minimizing

$$Z(w,b,\xi) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{M}\xi_i^l \tag{16}$$

subject to the constraints

$$y_i(w^T x_i + b) \geq 1 - \xi_i \quad \text{for } i = 1,\cdots,M\,. \tag{17}$$

Where $\xi = (\xi_1,\cdots,\xi_M)^T$ and C is the margin parameter that determines the trade-off between the maximization of the margin and minimization of the classification error. The values for l are either 1 or 2 calling SVM either L1 soft margin SVM (L1SVM) or L2 soft margin SVM (L2SVM) respectively. Pictorial representation of SVM in case of two class problem for soft margin is shown in figure 5. Considering L1SVM, Lagrangian multipliers α_i and β_i are introduced likewise linearly separable case and we get,

$$Z(w,b,\xi,\alpha,\beta) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{M}\xi_i - \sum_{i=1}^{M}\alpha_i(y_i(w^T x_i + b) - 1 + \xi_i) - \sum_{i=1}^{M}\beta_i\xi_i \tag{18}$$

where $\alpha = (\alpha_1,\cdots,\alpha_M)^T$ and $\beta = (\beta_1,\cdots,\beta_M)^T$ with $\alpha_i + \beta_i = C$.

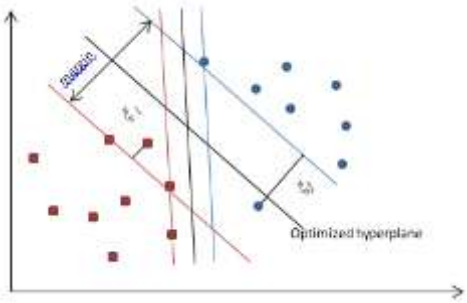

Figure 5. A separating hyperplane in two class problem for soft margin case.

The Dual problem for L1SVm is the same as dual problem for hard margin SVM described in (10) and (11). The only difference is that α_i cannot exceed C i.e. causing an extension in one of the constraints in (11) is $C \geq \alpha_i \geq 0$ for $i = 1,\cdots,M$. For three class or multiclass problem, readers may see (Steinwart & Christmann 2008; Abe 2005) for details.

If the data is not linearly separable then we can define a function that map the original input lower dimensional data into higher dimensional feature space also called dot-product space in order to enhance the linear separability. Such functions that perform mapping from lower to higher dimensional feature space are called Kernels which give us the advantage of not treating the higher dimensional feature space explicitly. This approach is also known as *kernel trick*. Given that the non-linear vector function

$$q(x) = (q_1(x),\cdots,q_n(x))^T \tag{19}$$

maps the $m-$dimensional input vector x into the $n-$dimensional feature space and the linear decision function in the feature space is defined as

$$F(x) = w^T q(x) + b \tag{20}$$

where w is an $n-$dimensional vector and b is a bias term. Hilbert-Schmidt theory says that if a function $K(x, x')$ is symmetric and satisfies

$$\sum_{i,j=1}^{M} k_i k_j K(x_i, x_j) \geq 0 \tag{21}$$

then there exists a function $q(x)$ that maps x into the dot product feature space and satisfies

$$K(x, x') = q^T(x) q(x') \tag{22}$$

This $K(x, x')$ is used in training and classification instead of $q(x)$. The dual problem in the feature space using kernel is as follows

Maximize,

$$Z(\alpha) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{M} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \tag{23}$$

subject to the constraints

$$\sum_{i=1}^{M} y_i \alpha_i = 0, \quad C \geq \alpha_i \geq 0 \qquad \text{for } i = 1, \cdots, M \tag{24}$$

The decision function that contain kernel expression becomes

$$F(x) = \sum_{i \in S} \alpha_i y_i K(x_i, x) + b \tag{25}$$

where b is given by

$$b = \frac{1}{U} \sum_{j \in U} (y_j - \sum_{i \in S} \alpha_i y_i K(x_i, x_j)) \tag{26}$$

Here U is the set of unbounded support vector indices.

Generally used kernels in SVM are linear kernel, polynomial kernel and Radial Basis Function (RBF) kernel. However, three-Layer Neural Network Kernels and some other kernels are briefly discussed in (Abe 2005; Wang 2005). Most commonly used kernel in BCI community is RBF or Gaussian kernel. The mathematical expression for linear, polynomial of degree d and RBF kernels are given in (27), (28) and (29) respectively.

$$K(x, x') = x^T x' \tag{27}$$

$$K(x, x') = (x^T x' + 1)^d \tag{28}$$

$$K(x, x') = \exp(-\gamma \|x - x'\|^2) \tag{29}$$

Here γ is a parameter for controlling the radius.

SVM that uses RBF kernel is known as Gaussian SVM or RBF-SVM. RBF converts the data space to higher dimensional feature space making separation much more likely for nonlinear cases. An SVM with linear kernel, called linear SVM is also used widely. Linear SVM (Rakotomamonjy et al. 2005; Garrett et al. 2003;

Benjamin et al. 2002) and RBF-SVM (Kaper et al. 2004; Garrett et al. 2003), both have been applied to BCI research quite successfully.

3 Results and Discussion

Here we have presented results only for real colour exposure. We have EEG signals belonging to three classes in our experiment to be classified as Red, Green and Blue and used ERSP data as features that was computed using EEGLAB (Delorme & Makeig 2004) Toolbox. We have ERSP datasets for each colour i.e. red, green and blue, for each subject. To acquire these datasets, ERSP was calculated for each trial in terms of time-frequency framework, for each colour among all the subjects. An average ERSP was taken across the trials for each colour as shown in figure 6 which represents the ERSP response for electrode position O2 across all the subjects for the real colours presented to the subjects. It is clearly visible in all the colours that variation in power in a specific frequency band at a certain time point is quite discriminative and remains discriminative for individual subjects for each colour. After the onset of stimuli, the highest increase in power is seen in red and lowest in blue during an interval from 100 to 400ms within delta and theta band. However, the lowest decrease in power is found in blue and highest in green during an interval from 250 to 1000ms within the alpha frequency band. These features were fed into SVM classifier.

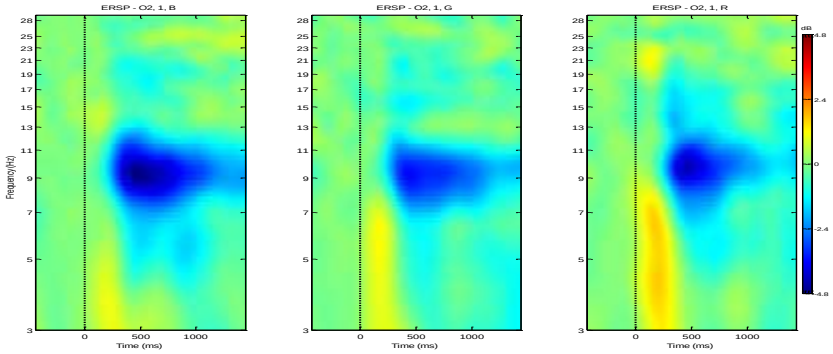

Figure 6. Averaged ERSP shown for electrode position O2. Right (Red), Middle (Green) and Left (Blue).

To classify EEG signals into red, green and blue classes, LIBSVM(Chang & Lin 2010)Toolbox was used which requires a specific format of data as follows in general,

<Label Class1> <index1> : <feature1> <index2> : <feature2> <index3> : <feature3> …

<Label Class2> <index1> : <feature1> <index2> : <feature2> <index3> : <feature3> …

<Label Class3> <index1> : <feature1> <index2> : <feature2> <index3> : <feature3> …

⋮ ⋮ ⋮ ⋮

Each row represents an instance or observation and each column represents the feature after the label column. Data was labeled before feeding to the classifier i.e. red labeled as 1, green labeled as 2 and blue labeled as 3. Numerically, ERSP data is present in two dimensional dataset of 100*200 values which indicate that there are 100 frequency points along rows and 200 time points along columns. For instance, to convert the data of subject 1 for channel P3, all the subject 1's colours i.e. red, green and blue are taken into account and each time point along its frequency points is chosen randomly within a single colour and

placed according to above format into the target dataset which will be used for training and testing of classifier. Once a time point is chosen from red class then a time point is chosen from green and finally from blue. This sequence continues until all the time points are chosen randomly within a single colour. Since there are 100 frequency points indicating 100 feature values in the target dataset against each label. Having three classes in hand, there would be 600 instances for data coming from channel P3, in the target dataset as each class contains 200 instances of time points i.e. each time point becomes an instance along the row in the target dataset with the corresponding label. After conversion, target dataset is divided into two subsets, one for training of classifier and the other is used for testing the classifier. As we have used four channels, two from parietal lobe and two from occipital lobe i.e. P3, P4, O1 and O2, the data was fed into the classifier not only as an individual channel but also in combination of different channels i.e. P3, P4, O1, O2, P3P4 (parietal region), O1O2(occipital region), P4O2(right occi-parietal), P3O1(left occi-parietal) and P3P4O1O2 (All). We have used C-SVC with linear kernel, polynomial kernel and RBF kernel where $\gamma = 0.1$ and degree of polynomial kernel is 3 along with default parameters. Tables 1, 2 and 3 present classification accuracies for linear, polynomial and RBF kernels respectively, for all the groups of data within all the subjects.

Table 1 presents the classification accuracies with an average accuracy of 84% for seven subjects among different groups of data which were used with linear kernel.

	S1	S2	S3	S4	S5	S6	S7
P3	90.74	83.7	85.19	85.93	84.82	84.82	81.85
P4	81.11	83.7	84.44	84.07	87.78	84.07	84.82
O1	81.11	86.3	81.85	84.82	88.15	82.96	87.78
O2	82.22	84.07	82.22	87.04	84.44	87.04	84.44
P3P4	83.7	85.93	83.89	87.41	88.15	84.63	84.44
O1O2	85	84.26	83.89	84.63	83.15	85.56	82.59
P4O2	85	82.78	83.89	84.26	85.93	85.93	82.59
P3O1	85	85.37	85.74	84.63	84.82	86.67	82.59
All	84.63	84.17	85	84.17	84.85	83.98	80.83

Table 2 presents the classification accuracies with an average accuracy of 89% for seven subjects among different groups of data which were used with polynomial kernel of degree 3.

	S1	S2	S3	S4	S5	S6	S7
P3	90	85.93	84.07	88.52	85.56	93.7	90.37
P4	83.7	86.67	86.3	89.63	91.11	89.63	91.85
O1	85.56	87.41	86.67	90.37	87.78	91.11	92.22
O2	83.7	86.67	89.25	87.04	86.3	87.78	88.15
P3P4	88.7	87.22	92.96	91.85	96.3	93.15	89.82
O1O2	85.74	82.59	81.48	88.33	90.56	87.41	89.07
P4O2	90.19	87.96	87.41	88.7	86.48	82.96	87.96
P3O1	87.41	84.63	84.26	92.22	90.56	92.96	87.78
All	92.59	90.74	88.89	91.57	95.56	89.17	87.41

Table 3 presents the classification accuracies with an average accuracy of 98% for seven subjects among different groups of data which were used with RBF kernel.

	S1	S2	S3	S4	S5	S6	S7
P3	99.63	96.3	99.26	96.67	97.04	97.41	99.63
P4	96.3	95.56	97.78	100	97.78	97.04	98.89
O1	94.82	96.3	95.93	99.63	97.04	96.67	98.15
O2	95.56	97.04	99.26	100	93.7	92.22	98.15
P3P4	97.78	94.63	99.63	99.82	100	99.44	97.78
O1O2	97.41	95	99.96	99.44	97.59	90.37	97.04
P4O2	98.33	93.7	99.3	99.3	96.85	95.93	97.59
P3O1	98.52	97.04	98.7	98.33	97.41	98.7	98.89
All	99.26	95.37	98.8	99.72	98.43	96.11	97.87

Classification accuracies for the data from all the channels (P3P4O1O2) with in all the subjects is shown in figure 7 for the three kernels used. Linear kernel has come up with lowest accuracy and RBF with highest accuracy in all the subjects. Similar results for the rest of the groups of data were also seen with an ascending order from lower to higher accuracy for linear, polynomial and RBF kernels in all the subjects. However, in some cases linear kernel has proven to be slightly better than polynomial in occipital region. Results regarding linear, polynomial and RBF kernels performances are shown in figures 7, 8 and 9 for data groups 'All', 'parietal' and 'occipital' channels respectively. However, the results for individual channels and right and left occi-parietal channels are similar. Tools used in this study are BCI2000 (Schalk et al. 2004), EEGLAB (Delorme & Makeig 2004) and LIBSVM (Hsu et al. 2003; Chang & Lin 2010; Fan et al. 2005).

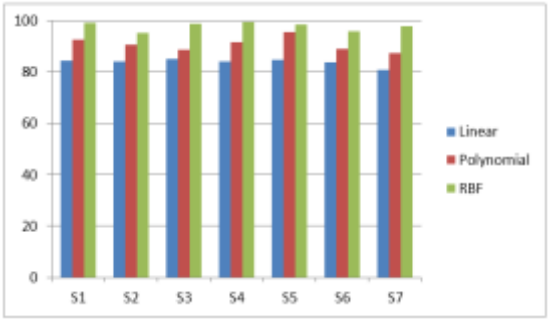

Figure 7. Classification accuracies in comparison among linear, polynomial and RBF kernels for data from all channels.

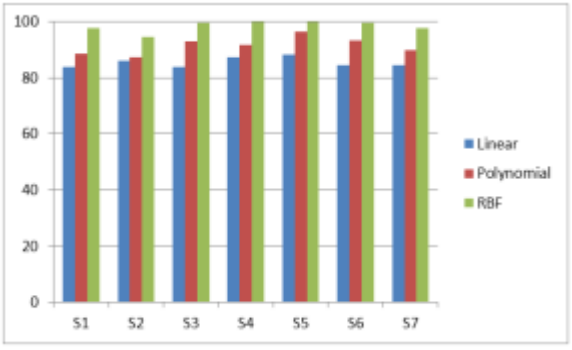

Figure 8. Classification accuracies in comparison among linear, polynomial and RBF kernels for data from parietal channels.

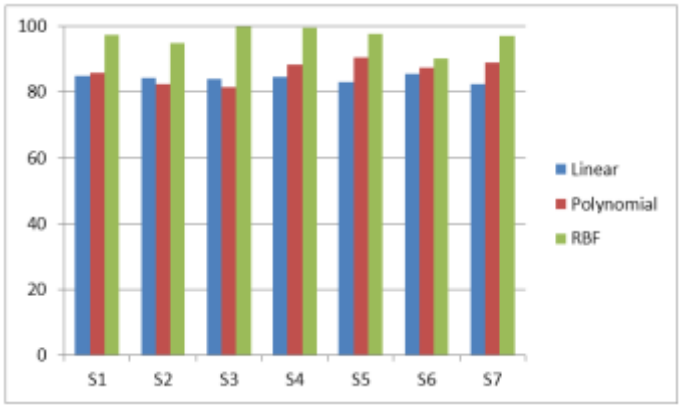

Figure 9. Classification accuracies in comparison among linear, polynomial and RBF kernels for data from occipital channels.

4 Conclusion

In this experiment we presented only primary colours to the subjects in random sequence in order to have discriminative EEG signals from parietal and occipital regions of brain. We have used ERSP data as features and C-SVM as classifier with three different kernels i.e. linear kernel, polynomial kernel and RBF kernel for the classification of EEG signals recorded as event-related potentials in response to RGB uniform colour stimuli. To interpret more accurately the differences among the exposure of primary colours, average event related spectral perturbation (ERSP) results in terms of time-frequency frame are presented in figure 5, that have shown significant power variations in the delta, theta and alpha bands. Red exposure has shown highest power increase starting from 100 ms to 400 ms in delta and theta bands than green and blue exposures, however the blue exposure has the lowest increase in power within the same time period and frequency bands. Moreover, a discriminative decrease in power among all the colours is seen in alpha band within 1000ms after the onset of stimulus. We have seen that these differences against the visual conditions in terms of single uniform colours are successfully classified using support vector machine. We have found very good accuracy results, on average 84%, 89% and 98% for linear, polynomial and RBF kernels respectively, with in all the groups of data among all the subjects. Highest accuracy was found in RBF kernel with nowhere less than other kernels in any group of data in any subject. As a next step, we propose to do a comparative study on different classifiers in conjunction with dimensionality reduction to select the most suitable data features in order to eliminate redundant information.

ACKNOWLEDGEMENTS

With special thanks to Prof. Daniele Marini, Prof. Raffaella Folgieri, Prof. Alessandro Rizzi, Prof. Davide Gadia and Adnan Abid whose suggestions continually kept improving the work. Further thanks to the subjects who participated in the experiment. This work has been partly funded by project VIRTHUALIS, VI PQ, contract n. 515831-2.

REFERENCES

[1] Abe, S., 2005. Support Vector Machines for Pattern Classification S. Singh, Japan: Springer.

[2] Benjamin, B., Gabriel, C. & Müller, K., 2002. Classifying Single Trial EEG : Towards Brain Computer Interfacing. In Advances in Neural Inf. Proc. Systems. MIT Press, pp. 157--164.

[3] Chang, C. & Lin, C., 2010. LIBSVM : a Library for Support Vector Machines. Department of Computer Science, National Taiwan University, Taiwan, 1-30.

[4] Chiappa, S. & Barber, D., 2006. EEG classification using generative independent component analysis. Neurocomputing, 69, 769-777.

[5] Delorme, A. & Makeig, S., 2004. EEGLAB : an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods, 134, 9-21.

[6] Dobelle, W.H., 2000. Artificial Vision for the Blind by Connecting a Television Camera to the Visual Cortex. ASAIO, 46, 3-9.

[7] Fan, R., Chen, P. & Lin, C., 2005. Working Set Selection Using Second Order Information for Training Support Vector Machines. Journal of Machine Learning Research, 6, 1889-1918.

[8] Farwell, L. & Donchin, E., 1988. Talking off the top of your head - Toward a mental prosthesis utilizing event-related brain potentials. Electroencephalography and Clinical Neurophysiology, 70(6), 510-523.

[9] Foresta, F.L., Mammone, N. & Morabito, F.C., 2009. PCA – ICA for automatic identification of critical events in continuous coma-EEG monitoring. Biomedical Signal Processing And Control, 4(3), 229-235.

[10] Garrett, D. et al., 2003. Comparison of Linear, Nonlinear, and Feature Selection Methods for EEG Signal Classification. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 11(2), 141-144.

[11] Guan, J. & Chen, Y., 2008. Single-trial EEG classification using in-phase average for brain-computer interface. Frontiers of Electrical and Electronic Engineering in China, 3(2), 194-197.

[12] Hallez, H. et al., 2009. Removing muscle and eye artifacts using blind source separation techniques in ictal EEG source imaging. Clinical Neurophysiology, 120(7), 1262-1272.

[13] Hsu, C., Chang, C. & Lin, C., 2003. A Practical Guide to Support Vector Classification, National Taiwan University.

COPYRIGHT© SOCIETY FOR SCIENCE AND EDUCATION UNITED KINGDOM

[14] Huang, R., Jung, T. & Makeig, S., 2005. Analyzing Event-Related Brain Dynamics in Continuous Compensatory Tracking Tasks. In IEEE-EMBS 27th Annual International Conference of Engineering in Medicine and Biology. Shanghai, pp. 5750-5753.

[15] James, C.J. & Hesse, C.W., 2005. Independent component analysis for biomedical signals. Physiological Measurement, 26, R15-R39.

[16] Jung, T. et al., 2000. Independent Component Analysis of Biomedical Signals. In Proc. Int. Workshop on Independent Component Analysis and Signal Separation. pp. 633-644.

[17] Jung, T. et al., 2000. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clinical Neurophysiology, 111(10), 1745-1758.

[18] Kaper, M. et al., 2004. BCI Competition 2003 — Data Set IIb : Support Vector Machines for the P300 Speller Paradigm. IEEE Transactions on Biomedical Engineering, 51(6), 1073-1076.

[19] Krusienski, D.J. et al., 2008. Toward Enhanced P300 Speller Performance. Journal of Neuroscience Methods, 167(1), 15-21.

[20] Leeb, R. et al., 2006. Walking by Thinking : The Brainwaves Are Crucial , Not the Muscles ! PRESENCE, 15(5), 500 -514.

[21] Lotte, F. et al., 2007. A review of classification algorithms for EEG-based brain – computer interfaces. Journal of Neural Engineering, 4(2), R1-R13.

[22] Makeig, S. et al., 2004. Mining event-related brain dynamics. Trends in Cognitive Sciences, 8(5), 204-210.

[23] Makeig, S., 1993. Auditory Event-Related Dynamics of the EEG Spectrum and Effects of Exposure to Tones. Electroencephalography and Clinical Neurophysiology, 86(4), 283-293.

[24] Peterson, D.A. et al., 2005. Feature Selection and Blind Source Separation in an EEG-Based Brain-Computer Interface. EURASIP Journal on Applied Signal Processing, 19, 3128-3140.

[25] Pfurtscheller, G. & Neuper, C., 2001. Motor Imagery and Direct Brain – Computer Communication. In Proceedings of the IEEE. pp. 1123-1134.

[26] Pires, G., Castelo-branco, M. & Nunes, U., 2008. Visual P300-based BCI to steer a Wheelchair: a Bayesian Approach. In 30th Annual International IEEE EMBS conference. Vancouver, pp. 658-661.

[27] Rakotomamonjy, A. et al., 2005. Ensemble of SVMs for Improving Brain Computer Interface P300 Speller Performances. In Artificial Neural Networks: Biological Inspirations – ICANN. Berlin: Springer, pp. 45-50.

[28] Schalk, G. et al., 2004. BCI 2000 : A General-Purpose Brain-Computer Interface (BCI) System. IEEE transactions on Biomedical Engineering, 51(6), 1034-1043.

[29] Steinwart, I. & Christmann, A., 2008. Support Vector Machines M.Jordan, J. Kleinberg, & B. Scholkopf, New York: Springer.

[30] Stone, J.V., 2004. Independent Component Analysis: A Tutorial Introduction, MIT Press.

[31] Vapnik, V.N., 2000. The Nature of Statistical Learning Theory M. Jordan et al., Springer.

[32] Wang, L., 2005. Support Vector Machines: Theory and Applications J. Kacprzyk, Springer.

[33] Wolpaw, J.R. et al., 2002. Brain – computer interfaces for communication and control. Clinical Neurophysiology, 113, 767-791.

[34] Yoto, A. et al., 2007. Effects of Object Color Stimuli on Human Brain Activities in Perception and Attention Referred to EEG Alpha Band Response. Journal of Physiological Anthropology, 26(3), 373-379.

COPYRIGHT© SOCIETY FOR SCIENCE AND EDUCATION UNITED KINGDOM