APPLICATION OF WRAPPER METHODS TO NON-INVASIVE BRAIN-STATE DETECTION: AN OPTO-ELECTRIC APPROACH

A THESIS IN

Electrical Engineering

Presented to the Faculty of the University
of Missouri-Kansas City in partial fulfillment of
the requirements for the degree

MASTERS OF SCIENCE

by

VIKAS GOTTEMUKKULA

B. Tech, Scient Institute of Technology, 2008

Kansas City, Missouri

2010

APPLICATION OF WRAPPER METHODS TO NON-INVASIVE BRAIN-STATE DETECTION: AN OPTO-ELECTRIC APPROACH

Vikas Gottemukkula, Candidate for the Master of Science

University of Missouri-Kansas City, 2010

ABSTRACT

Using a classification guided feature selection (wrapper method) in conjunction with a new performance metric, I present a solution to multi-class subject invariant Brain Computer Interface (BCIs) using electroencephalography (EEG) and near infrared spectroscopy (NIRS) signals, a complex problem known to be prone to trivial classification. In a data-driven multi-class BCI, evaluation of the one versus rest (OVR) classifier is a major challenge using error rate. The hence derived multiclass OVRs using wrapper methods with error rate as the classifier feedback can show degeneracy in terms of imbalance in sensitivity and specificity, leading to trivial classification. This imbalance can be removed by the usage of a scalar quality factor as the performance metric. The error rate is replaced by a simple scalar quality factor that adjusts the simple correct rate with the ratio of sensitivity and specificity. A 4-class subject invariant EEG-based BCI using signals from 10 untrained subjects is presented here to prove the efficacy of the quality metric. Left hand, right hand, left leg, and right leg movements are classified using Naïve Bayesian, Gaussian SVM, Polynomial SVM, and k-Nearest neighbor classifiers. Extending the same method to optical signals, here I present an NIRS-based BCI using signals from two subjects to classify left hand and right hand movements. The same quality-metric based wrapper methods could

identify the salient time samples of oxy-hemoglobin (HbO) and deoxy-hemoglobin (Hb) channels from NIRS signals to achieve 100% classification rate, sensitivity, and specificity.

APPROVAL PAGE

The faculty listed below, appointed by the Dean of the School of Computing and Engineering, have examined a thesis titled " Application of Wrapper Methods to Non-Invasive Brain-State Detection: An Opto-Electric Approach," presented by Vikas Gottemukkula, candidate for the Masters of Science degree, and certify that in their opinion it is worthy of acceptance.

Supervisory Committee

Reza Derakhshani, Ph.D, Committee Chair
Department of Computer Science Electrical Engineering

Gregory W. King, Ph.D
Department of Mechanical Engineering

Ghulam M. Chaudhry, Ph.D
Department of Computer Science Electrical Engineering

Walter D. Leon-Salas, Ph.D
Department of Computer Science Electrical Engineering

CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Brain Computer Interface is an evolving research topic since early 1970's (Ortiz, 2007). BCI is a thought translation device that can be used to operate an external device based on the brain activity with many potential applications (Birbaumer et al., 2000). The progress of the BCIs was enhanced as many studies hypothesized the correlation between EEG signals (or brain activity) and imagination (Anderson et al., 1998 and McFarland et.al 2000). The aim of the BCI is to translate human thoughts into a useful control signal. BCIs can act as a communication channel for paralyzed, locked-in, spinal cord injury, amyotrophic lateral sclerosis and brainstem stroke patients (Wolpaw et al., 2002). Applications of BCI are computer speller (Felton et al., 2007 and Millan et al., 2003), wheel chair control (Millan et al., 2003 and Valsan et al., 2009), smart home technology (Chan et al., 2008), games (Martinez et al., 2007) and many more.

The activity of the brain can be measured using various methods like Electroencephalography (EEG), Magneto-encephalography (MEG), Positron emission tomography (PET), Functional magnetic resonance imaging (fMRI) and Near Infrared Spectroscopy (NIRS). Observing the disadvantages like massive size of MEG and fMRI, and the exposure to radioactive traces using PET; EEG drew the attention of many researchers with advantages like portability, ease of usage, and price (Wolpaw et al., 2002 and Coyle et al., 2007). NIRS signals provide a clear picture of brain hemodynamic activity besides some disadvantages like limited light penetration depth (Huppert et al., 2006).

Electroencephalogram is a device to record the electrical activity of the brain that can be observed as a result of the excited nerve cell, Neuron. Neurons are the electrically active cells that carry information through electrochemical signaling method through a complex network system (Gonon, 2003). Every thought or action involves several thousand's of neurons firing across the brain (Wolpaw et al., 2002). EEG is composed of small scalp-induced potentials from the electrical activity of the outer layers of the brain. However, as the electrical activity of the neurons is in terms of millivolts the real world application of these signals require several noise cancellation and filtering techniques. These signals are recorded using electrodes placed on the scalp (EEG caps) or by inducing in the scalp Electrocorticogram (ECoG). The increase in number of electrodes in an EEG cap improves spatial resolution and signal-to-noise ratio of the collected EEG signals. A standard 10-20 placement EEG cap with 21 electrodes is used in this study.

As several thousands of neurons fire across the brain for each thought and action, they require more oxygen from blood to metabolize the consumed glucose (Heeger et al., 2002). The required oxygen for the neurons is supplied by oxy-hemoglobin (chromophore) in the blood. After the required oxygen is acquired at capillary beds the oxy-hemoglobin turns into deoxy-hemoglobin. . NIRS is an optical technique that uses the absorption of near infrared lights to measure concentration changes in deoxy-hemoglobin (Hb) and oxy-hemoglobin (HbO) (Coyle et al., 2007 and Schmitz et al., 2009). The changes in the tissue oxidation levels due to brain activity, regulates the scattering and absorption of the NIRS photons (Coyle et al., 2007). The relative changes in hemoglobin levels are determined by NIRS signals using Beer-Lamber law. The light emitted by the sources like Light emitting diode or Lasers with wavelength of 700 to 900 nm penetrates into the skin (here - cortex) up to a depth of few centimeters (0.5 to 2 cm in adult human (Fukui et al., 2003)), is collected back using detectors around the source. The spatial resolution of the collected signals is modulated by the geometry of the probe (Sitaram et al., 2007).

The objective of this study is to develop an automated wrapper method applicable to BCIs that choose features and select classifier parameters with balanced sensitivity and specificity. This method is applicable to the BCIs that operate on electric (EEG) and optic (NIRS) signals. A Wrapper method is an integrated data driven methodology to choose the features and select classification model based on a quality metric. In a multi class BCI, the evaluation of the performance of a two class classifier can be done using one vs. one or one vs. rest classifiers (OVR). In a N-class problem evaluation using one vs. one classifier we need to evaluate all possible pairs of classes, which lead for the evaluation of N(N-1)/2 classifiers. The count of the classifiers which need to be evaluated grows rapidly with the increase in the number of classes. OVRs on the other hand solve this problem with N OVRs. The performance measure of OVR classifier using a single scalar metric like correct rate might lead to degeneracy. In an multi-class BCI, the performance of an OVR need to be determined considering multiple metrics like correct rate, sensitivity, and specificity, rather than a single correct rate. For instance in a 5-class BCI with equal number of samples in each class, an OVR can generate correct rate of 80% without classifying none of the in-class samples (i.e., sensitivity = 0 and specificity = 1), trivial classifier. This problem is more prominent with the increase in number of classes in a BCI. By using a Quality factor (please refer section 2.5) as wrapper method feedback for the OVR we can achieve a non-degenerative solution regardless of the number of class. The significance of this BCI is its subject invariance, as it is developed on 10 untrained subjects EEG signals.

NIRS signals for the motor movements do have distinct patterns in the hemodynamic responses from the motor cortex (Sitaram et al., 2007). In the approach to classify motor movements using NIRS signals, Sitaram et al., 2007 achieved an average 5-fold cross validation recognition rates of 87.5% for linear SVM and 93.4% for hidden Markov model, and Niide et al., 2009 achieved 93.4% for linear SVM and 93.9% for Gaussian SVM. They used same dataset with 4 subjects, in which each subject's movements are classified separately and their recognition rates were averaged for each movement. Niide et al., 2009 presented a maximum 5-fold cross validation rates of 97.1% for a subject. Here we present an automated wrapper method guided by quality factor to choose features and select classifier parameters for a two class subject invariant BCI (using wrapper methods) that generated 100% 5-fold cross validation recognition rates using time samples of the highly active NIRS channels as features.

Science team for EEG experiment: Data collection and feature extraction was done by Jesse Sherwood. The feature ranking and selection, and classification were my own work under the guidance of Dr. Reza Derakhshani. The idea of Quality metric is brought up by RD.

Science team for NIRS experiment: Using the dataset provided by *NIRx GmbH* the feature ranking and selection, and classification were done by me, under the guidance of Dr. Reza Derakhshani. Artifacts like Mayer waves, respiration and hair movements, were removed by the filtering techniques provided by *NIRx GmbH*.

CHAPTER 2

METHODS

In this study, a four class BCI system was developed using EEG signals collected from 10 untrained volunteers performing four body movements, namely left hand, right hand, left leg, and right leg (designated as Movement 1, Movement 2, Movement 3, and Movement 4, respectively). Based on the literature and previous study by Dr. Reza Derakhshani and Jesse Sherwood, the selection of the four motor movements is done (Curran et al., 2003, Doynov et al., 2008, Mason et al., 2007 and Doyle et al., 2006).

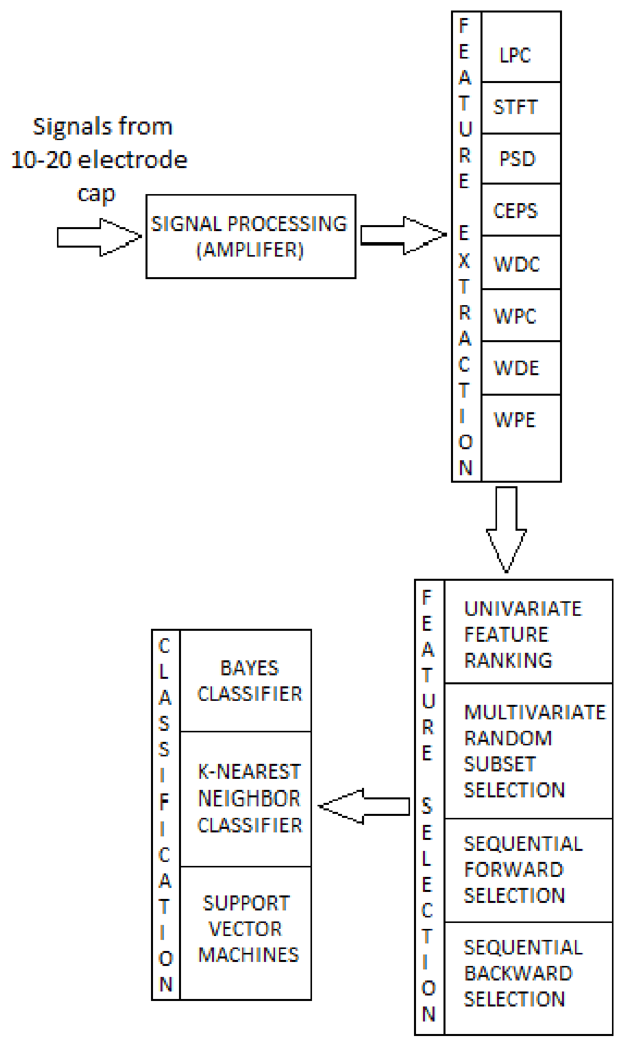

Figure 1: A block diagram of the overall EEG method.

Further preprocessing, feature extraction, feature selection and classification were done on these signals (Figure 1). A block diagram of the entire EEG analysis is shown in figure 1. To increase the statistical power of the results and eliminate over-fitting problem, all the classification results for the 4-class BCI did undergone a 5-fold cross validation with 10 Monte Carlo repetitions.

A two class BCI was developed using NIRS signal collected from two different volunteers performing two body movements, namely left hand and right hand movements. NIRS signals with the wavelength of 760 nm and 850 nm are used to measure HbO and Hb respectively. The recorded signals thereafter undergo preprocessing (artifact removal), feature selection and ranking, and classification (Figure 2). A 5-fold cross validation is done by creating each fold manually so as to divide the classes and subjects in equal proportions for all the folds. All the forty samples were divided into 5 eight sample folds, which have 2 left hand movements of subject 1, 2 left hand movements of subject 2, 2 right hand movements of subject 1, and 2 right hand movements of subject 2.

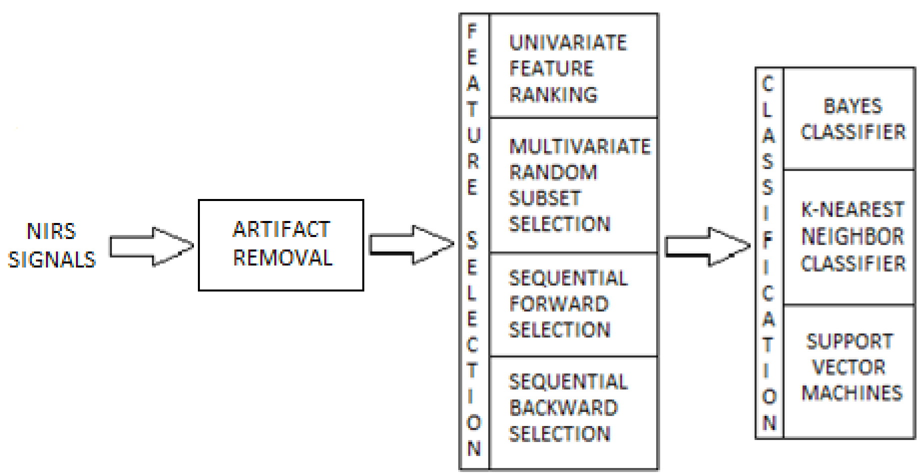

Figure 2: A block diagram of the overall NIRS method.

2.1. Data Collection

EEG signals are recorded under the UMKC IRB approval #090218, from 10 untrained volunteers for 4 movements using standard 10-20 electrode placement with two ear references (Fig 3). The EEG signals are amplified and filtered to reject power line interference using a NeuroPulse Systems MS-24R bioamplifier with 1.5 – 34 Hz bandpass filter at a sampling rate of 256 Hz. The subjects were asked to perform four body movements for 12 times in each of the two sessions without any preliminary training. Each sessions conducted is 4-6 weeks apart. Each of the four intended movements was repeated for 24 times in both the sessions, producing 960 eight-second epochs. So, each movement has 240 samples from 10 different subjects. During each session they were asked to perform 4 intended movements 12 times each over the course of 2 hours. All the instructions to the subjects were given using a computer screen prompt. All the subjects were instructed not to blink during the task performance. They positioned their both feet on a foam pad, loosely gripped rubber balls with each hand, and applied slight pressure based on the displayed instruction (pseudo movements). A short audio tone followed by 8 second command prompt on a screen is used to specify the movement. This is followed by a 10 second break before the next pseudo movement was displayed on the screen. The first 2 second recordings of all the 8 second recordings are taken for the analysis. The motor cortex signals are captured from C3 and C4 electrodes (Fig. 3). Large Laplacian filters were applied for better spatial resolution, where the reference potentials were derived from the difference of C3 and the average of F3, T3, P3, and Cz (left hemisphere). Similarly, C4 was referenced to the average of F4, T4, P4, and Cz (right hemisphere). Left and right hemisphere signals were filtered separately.

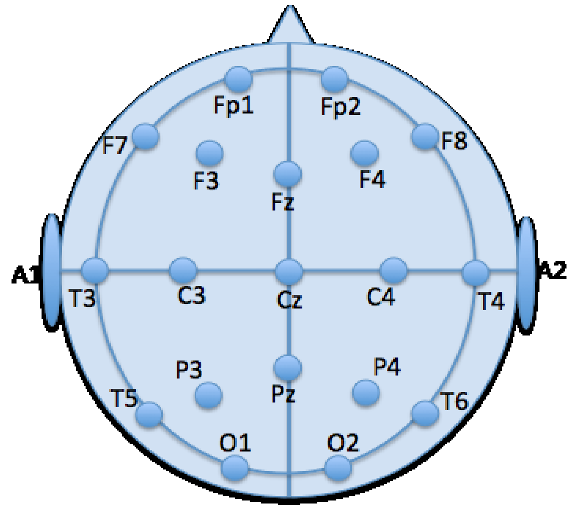

Figure 3: 10-20 EEG electrode placement system.

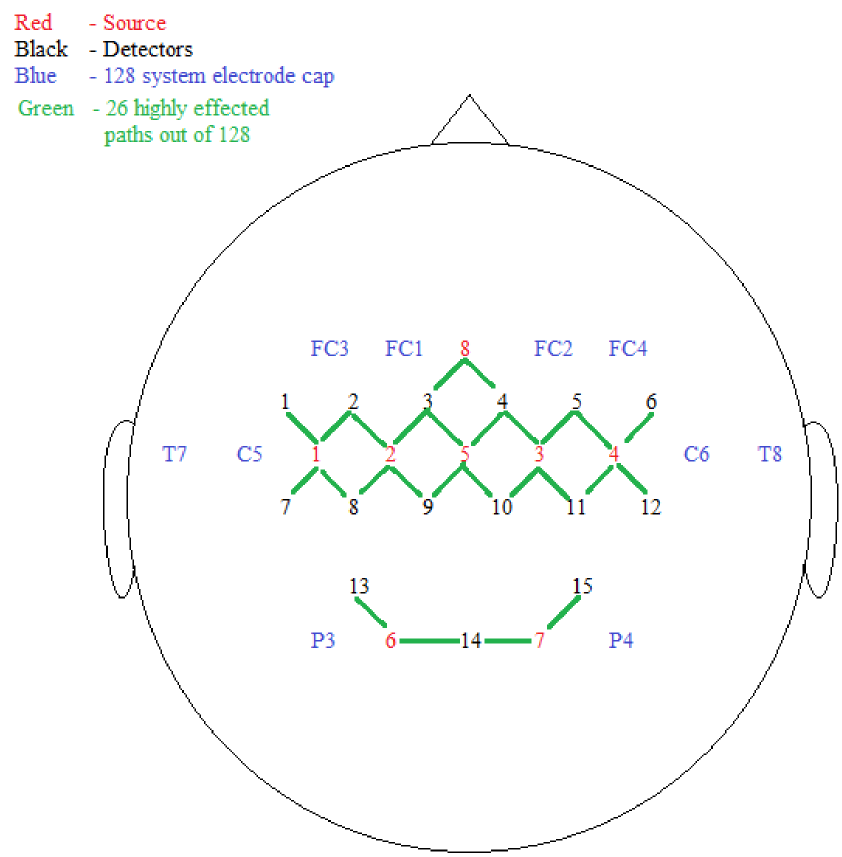

Figure 4: Source and detector placement for NIRS signal recording.

NIRS signals are recorded at a sampling frequency of 6.25 Hz from two subjects during left hand and right hand movements while clutching a ball in each hand. The NIRS device used for the data collection has 8 sources and 16 detectors placed as shown in the figure 2. Source-detector placements resemble those of a 128 electrode EEG montage. Here, the 16 detectors for each of the 8 sources measured 128 different optical signals. In order to evaluate the major effective paths, the signals between each source and its immediate detector are considered, that resulted to 26 effectively active paths (shown in the green color in figure 4). The 26 selected signals, in terms of their source and detector paths, are in order: Signal 1: 1-1, Signal 2: 1-2, Signal 3: 1-7, Signal 4: 1-8, Signal 5: 2-2, Signal 6: 2-3, Signal 7: 2-8, Signal 8: 2-9, Signal 9: 3-4, Signal 10: 3-5, Signal 11: 3-10, Signal 12: 3-11, Signal 13: 4-5, Signal 14: 4-6, Signal 15: 4-11, Signal 16: 4-12, Signal 17: 5-3, Signal 18: 5-4, Signal 19: 5-9, Signal 20: 5-10, Signal 21: 6-13, Signal 22: 6-14, Signal 23: 7-14, Signal 24: 7-15, Signal 25: 8-3, and Signal 26: 8-4 (All the red colored numbers represent the sources).

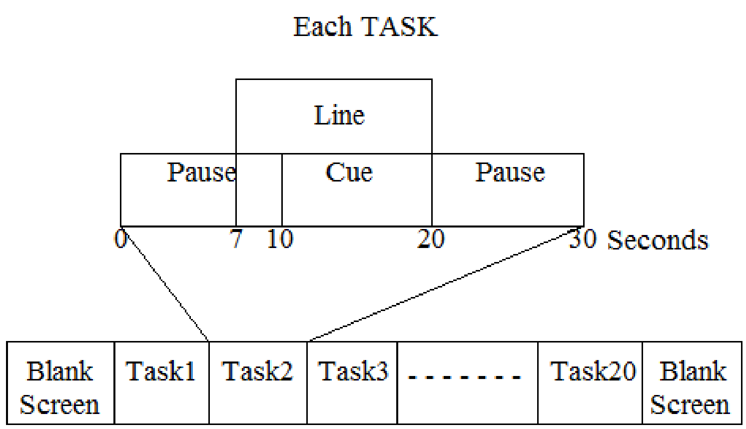

Figure 5: Timing of experiment.

Subjects were prompted for 10 left and 10 right hand movements, in a random order using a computer screen. Each recording session was initiated with 30 seconds of pre-waiting period with a blank screen, followed by twenty prompted hand pseudo movements trails. Each trial lasts for 30 seconds, where the first 10 second is pause period without movements, followed by a 10 second task execution (left or right hand), and finished with another 10 second pause. A vertical green line was appeared on the screen three seconds prior to each command to make the subject ready for task (Figure 5).

2.2. Feature Extraction

All the following feature extraction techniques were applied to the EEG signals by Jesse Sherwood. The features for all the modalities are extracted separately for left and right hemisphere EEG signals. All the features vector derived from all the modalities were checked for the constant rows (Matlab command – removeconstantrows), that is all the features with their samples under a variance of 10^{-8} were removed considering that they are too hard to classify along with other high variance features.

2.2.1. Linear Predictive Coefficients

In BCI systems, the p^{th} order forward linear predictive coding (LPC) coefficients have been successfully used as time domain features (Anderson et al., 1995). The 256 Hz signal was down sampled by 3 and then based on the classification rates obtained using the SVM classifiers, 18^{th} order LPC features were used in the study. The increase in the order made the penalized error decrease and then relatively flat and the increased thereafter up to the order of 20 (Roberts et al., 2006). Increasing the order beyond 20 reduced the training quality and generalization ability of the trained classifiers (Jain et al., 2000). LPC had a total of 36 features, in which 18 are from left and the rest from right hemisphere.

2.2.2. Short Time Fourier Transforms

As EEG signals can be represented by the class of non-stationary signals, the time-frequency features can be extracted using the short time Fourier transforms (STFT) (Mitra et al., 2006). A window length of 1 second with a 255 sample overlap and a range of 1-48 Hz aggregated into four 12 Hz frequency bins were used in this study. All this was done on the

EEG signals obtained from left and right hemisphere, separately. STFT had a total of 32 features, in which 12 are extracted from the left hemisphere EEG signals and 12 from right.

2.2.3. Power Spectral Density

EEG signals representing imagined motor tasks are believed to be attributed to event-related desynchronization of neuron signals. This can be reflected in changes in spectral energy levels within discrete frequency bands of the EEG signals. Based on this characteristic the selections of spectral features were chosen in this study (Herman et al., 2008). The Welch periodogram spectral estimation method with the Hamming window length of 33 and 97% overlap is used for the calculation of the string of numbers representing frequency bin energies of power spectral density. A 1-48 Hz PSD span divided was into twelve 4 Hz frequency bins and then the energy was calculated under each non-overlapping bin. The above setting was derived by trial and error method. This contributed the feature vector with a total of 24 features for PSD (12 from left and 12 from right hemisphere EEG signals), that are used in the study.

2.2.4. Cepstrum

Cepstrum is a non-linear signal transform derived from the Fourier transform (Oppenheim et al., 1975). It was included as it provided a non-linear frequency-based feature set. A total of 1022 features (511 features from left and 511 features from right hemisphere EEG signals) are derived, and used in this study.

2.2.5. Wavelet

The drawbacks of the STFT such as time-frequency resolution in the non-stationary signals can be removed using a multi resolution time-frequency signal transform, known as Wavelet decomposition (Mallat, 2001). For the signal reconstruction, regular wavelet decomposition and wavelet packet provide similar information. The features from the Wavelet packet are more classifiable, as they are produced as a result of a subdivision of high frequency details into sub-bands which provides advantages in extracting high frequency features (Vetterli et al., 1992; Ting et al., 2008).

All the wavelet families available in the Matlab toolbox were tested using SVM classifiers and then the best families based on the minimum classification error were selected. Reverse biorthogonal and symlet families are used for this modality (Sherwood et al., 2009). Two kinds of wavelet features were constructed: as the filter banks outputs or aggregation of sub-band energies, where the latter foregoes temporal information in the interest of shift-invariance by marginalizing over the shift parameter (Mallat, 2001). The wavelet decomposition filter bank output coefficients (WDC) were calculated by reverse biorthogonal 3.7 wavelets and wavelet packet filter banks output coefficients (WPC) were calculated with symlet 15 wavelets. The energy of wavelet decompositions were marginalized over time shift and their energies across different frequency scales were calculated as WDE feature sets using reverse biorthogonal 3.1 wavelets. Similarly wavelet packets energy features (WPE), were calculated from wavelet packets using symlet 15 wavelets.

A total of 1048-WDC features, 1032-WPC features, 48-WDE features, and 96-WPE features were derived from their respective feature extraction methods. All these feature vectors were checked for constant rows and removed.

2.3. NIRS Feature Generation

A band stop 6^{th} order Butterworth filter with a stop band of 0.09- 0.45 Hz is used in order to minimize the artifacts like Mayer waves, respiration and hair movements. The baseline changes of the signals are removed using a high pass filter (0.01 Hz, Butterworth order 5). Baseline shifts are further removed by subtracting the mean of each 10-second preceding pause period signal from its following 10-second task execution signal. All the above filtration techniques are suggested by *NIRx GmbH*. Finally, these signals were decimated by 4. Considering the original sampling rate of 6.25 Hz, and the fact that the majority of the spectral contents of the fNIRS signals seemed to be under 0.75 Hz, and we deemed this new sampling rate of 1.5 Hz of the decimated signal to be sufficient. As the signals were recorded at a sampling rate of 6.25 Hz, each ten second task execution period have 62.5 samples which were round to 64 samples, so a 16 sample signal was derived after decimating each 10 second signal of each pseudo movement from all the 26 signals. Based on our pilot studies and related reports in the literature, we chose all these 16 points from 26 channels as features in this study (Sitaram et al., 2007; Niide et al., 2007). In addition mean and variance of each signal are also taken as features. Thus, for each 10 second task-execution span 936 feature elements were produced, with 416 HbO features (16 samples * 26 channels), 416 Hb features, 26 features as mean of HbO, 26 features as mean of Hb, 26 features as variance of HbO, and 26 features as variance of Hb. Given the 10 left hand and 10 right hand movements from two subjects, a total of 40 feature vectors of length 936 was produced in the manner. Figure 6 and 7 shows the average of the entire 10 records for each subject and for each movement across all the 26 channel highly effective channels in terms of HbO and Hb signals. Each signal resembles 18 samples in which the middle 16 samples represents the second 10 seconds of each trial, the task execution period as shown figure 4, that is the leading pause followed by the task execution and end pause. It can be clearly seen that the temporal evolution of each signal is different. Thus the selection of salient time samples is done for a robust classification of the target pseudo movements from single trial fNIRS signals.

In figures 6 and 7, red and blue lines indicate the oxy-hemoglobin and deoxy-hemoglobin signals. Left and right hand movements are depicted with solid and dashed lines, respectively.

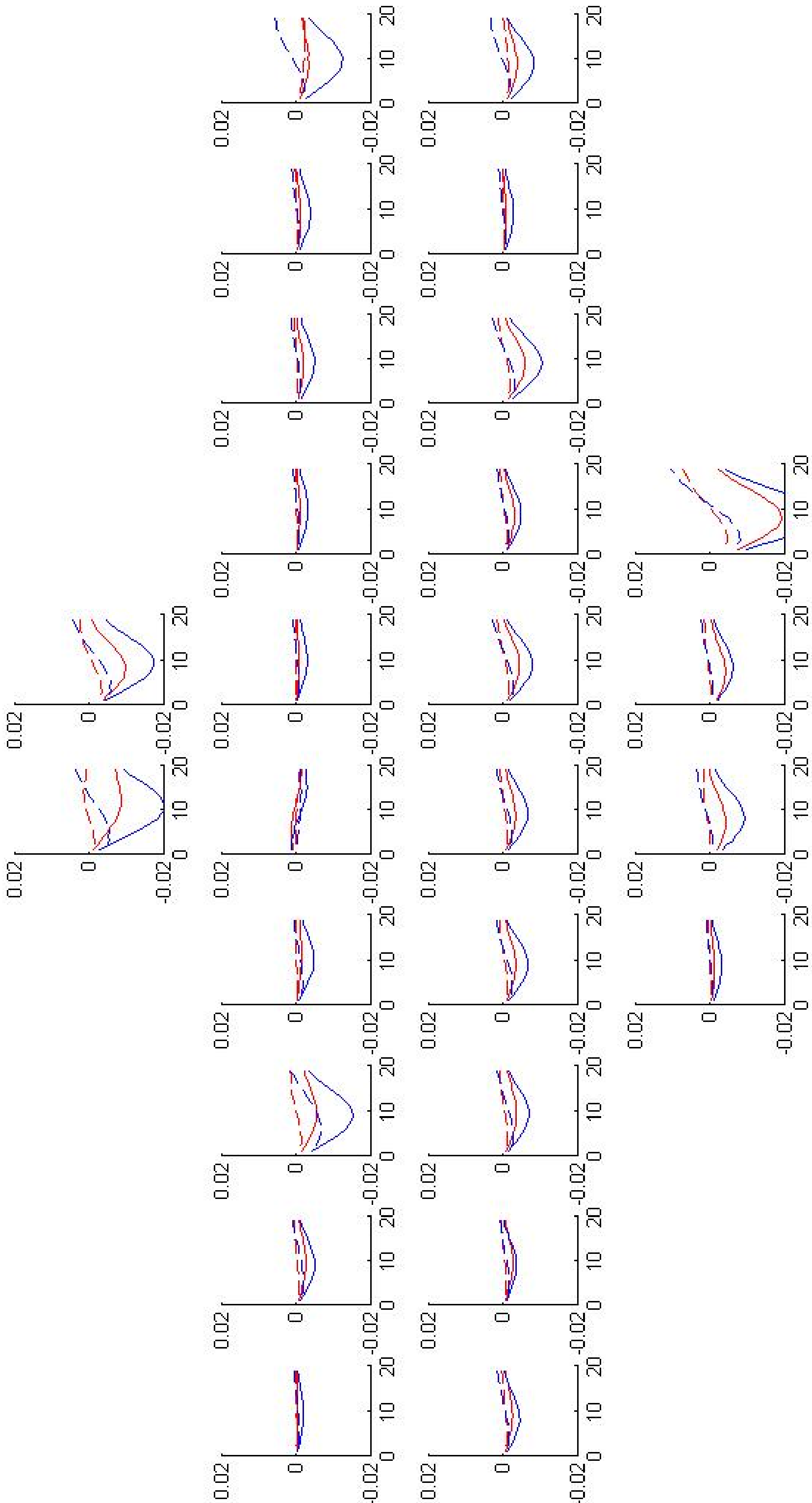

Figure 6: Average fNIRS signals for subject 1.

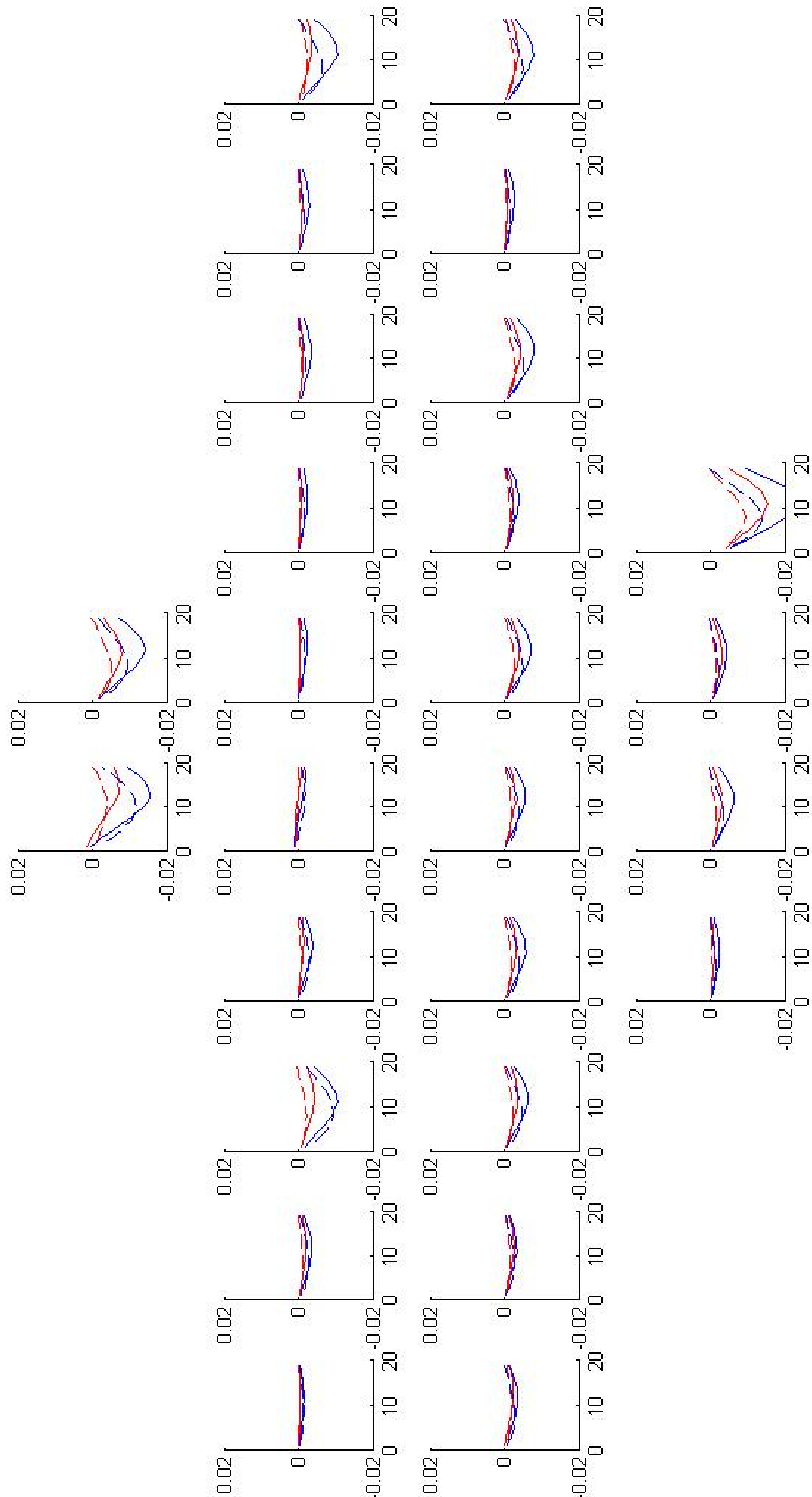

Figure 7: Average fNIRS signals for subject 2.

2.4. Feature Ranking and Selection

2.4.1. Univariate Feature Ranking (ROC ranking)

The features within each modality are ranked using the area under Receiver Operating Characteristic (ROC AUC) as classifiability criterion (Fawcett, 2006). ROC is a plot of true positives rate (y-axis) versus false positives rates across different classification thresholds for univariate class likelihood. A higher area under the curve is indicative of better overall sensitivity and specificity. For the EEG data set ranking for the 8 feature modalities has been done individually and also for all the eight modalities so as to obtain the ranked list of the features in a specific modality and among all the modalities to pick the best modality. Using this ranked list, features were aggregated from the top till the classification quality metric was maximized. Since the ranking has been done to discriminate the 4 movements across all the subjects, the results provide features with higher movement-discrimination power irrespective of subjects. For the NIRS, all the 936 features were ranked using ROC AUC criterion and the ranked list is used in the same way as in EEG.

2.4.2. Classification-Guided Subset Selection and Ranking

First, within each modality, a number of fixed-length feature subsets that can attain a certain classification rate (Threshold) are selected and merged (Waske et al., 2006). The feature with highest frequency is ranked first, and so on. Here in the EEG analysis, the classification rates of the randomly sampled subsets were calculated using a four class k-NN classifier with a configuration of k varied from 5 to 20 (5, 10, 15, & 20), a Euclidean distance metric and majority voting method. In NIRS analysis k is varied from 4 to 10. For a set of configuration (Pool size, subset size and k) the threshold was started with 60% and ended up with the maximum achieved by my machines, in this order for NIRS analysis three ranked list obtained using 3 different configurations are used in this study; 1. Subset size – 10, Pool Size – 150, Threshold – 0.905; 2.Subset size – 15, Pool Size – 80, Threshold – 0.87; 3.Subset size – 20, Pool Size – 80, Threshold – 0.88. For EEG analysis a subset size – 12, Pool Size – 100, Threshold – 0.69 was used.

2.4.3. Sequential Selection

The sequential algorithm selects the features from a set of candidates so that an ensuing classification metric is optimized (Bishop, 2006). Sequential forward (SF) and sequential backward (SB) selection methods are used in this analysis. SF starts from an initial feature element adding more elements one at a time, as long as the classification metric is improving. Sequential backward selection works similarly but in the opposite direction: all the features are initially selected into the best set, and then elimination of the features occurs with one element at a time based on the feedback from the selected classification metric.

2.5. Quality Factor

As mentioned during the introduction, classification rates might not be a successful performance measure of the OVR model. This can be observed when the sensitivity and specificity are lopsided, leading to degeneracy and trivial classification despite of high classification rates. For instance in a 4-class OVR classifier with sensitivity and specificity of 10 and 90 respectively will come up with a classification rate of 70%, despite of poor sensitivity. To be more specific, in an N class BCI design with equal number of samples in each class we need N individual OVR's, wherein each compare two groups. Each OVR need to classify two groups where all the samples of one class are in group 1 and rest of all the samples of N-1 class are in group 2. In such a case with unbalance between each group, the measure of performance using classification rates can mislead.

In the approach to solve this problem a confusion matrix can be used to measure performance. Sensitivity and specificity can be used to reduce the number of scalar metrics, which need to be considered in the confusion matrix (True positives, true negatives, false positives, and false negatives). But for the classification of the signals using wrapper methods multiple scalar metrics cannot be used, as it uses single scalar metrics as classifier feedback to its feature search and selection routine.

Most of the studies do avoid this issue just by presenting the classification rates (instead of sensitivity and specificity), or by using other classifier arrangements. A method to solve this issue is presented in BCI2005 competition through Cohen's Kappa coefficient (Schlögl et al. 2005, and Schlögl et al. 2007). A scalar value that falls in chance-corrected agreement statistics which measures the correlation between predicted and an actual class is used. This scalar, derived from the confusion matrix, was proved to have one deficiency by Gwet (Gwet, 2002). Here I present an approach to solve the problem using a single scalar metric which uses confusion matrix, which improved the classification quality metric thus called as Quality factor, Q. This reflects the recognition rate and also addresses the N-1/N classification rate problem with the use of sensitivity and specificity. Q is defined as the ratio of classification rate to the sensitivity vs. specificity or its inverse, whichever greater.

$$Quality\ factor = \frac{1 - \text{Overall error}}{\max[\left(\frac{Sensitivity}{specificity}\right) or \left(\frac{Specificity}{sensitivity}\right)]}$$

2.6. Classification

2.6.1. k-NN Classifier

The k-NN classifier classifies an unknown sample based on majority vote of k nearest samples of the classified training samples (Cover et al., 1967). Here in this study, a Euclidean metric is used to measure the distance between test sample and its k nearest training samples.

A better classifier is selected, not only using classification rate, but with an equal importance given to k value and classification rate. For instance in an unbalanced two class classifier with a ratio of 1:4 samples in class1 vs. class 2 of the training and test sets, if the k value is greater than 2/5th of training dataset, no matter what ever the test sample it would be classified into class 2. For EEG data, k values are varied from 1 to 100 and for NIRS it's 1 to 10.

2.6.2. Naïve Bayesian Classifier

Bayes classifiers assume normal distribution of classes, and thus operate on the basis of quadratic or linear discriminants that result from intersection of normal log-likelihoods (Duda et al., 2001). Two Bayesian classifier configurations were chosen. One used a linear discriminant function with a diagonal covariance estimate pooled across the classes (DL). The second method used a quadratic discriminant function with the diagonal covariance estimates stratified by classes (DQ).

2.6.3. Support Vector Machines

Support Vector Machines (SVM) separates the data in a feature space using a maximum margin hyperplane (Sebald et al., 2000). The data points near the hyper-plane are called support vectors and are used to determine SVM discriminant. The SVM determines its parameters based on the solution of a convex optimization problem, and as such the local solution can be a global optimum (Bishop C.M., 2006). The Gaussian and Polynomial estimates for the calculation of the hyper-plane are used in this analysis. In a Gaussian SVM, the training data is projected into a feature space spanned by Gaussian kernels, whose spread represented by sigma (σ), where in a polynomial SVM the classes are separated by polynomial functions. The tested range on both datasets for the spread of the Gaussian kernel in the Gaussian SVM is 0.1 to 25 and the order of the polynomial in the polynomial SVM is 2 to 5. The box constraints (or C parameter) values were changed from 0.1 to 100 to control sensitivity of SVM boundaries to outliers using “soft” or “hard” margins, allowing trade-off between the slack variables, misclassification penalty, and the discriminant rigidity (Bishop C.M., 2006). Increasing the C value from 0.1 to 100 refers to a change of margin type from soft to hard, affecting outlier tolerance and training error.

CHAPTER 3

RESULTS

3.1.EEG Analysis

Using the ranked list from univariate feature ranking, and classification guided subset selection and ranking methods for each feature modality, the feature vector to the classifier are given starting from the top ranked and then adding the next best until the Q factor is maximized. All these feature vectors are tested by the classifiers with all the different configurations possible (within the test range mentioned). The set of features obtained for the sequential forward and sequential backward selection are tested with all the classifier configurations and the best classifier is selected based on the Q factor. The tested range for all the classifiers are Naïve Bayesian (Diagonal linear-DL, Diagonal Quadratic-DQ), kNN (k = 1-100), and SVMs (Polynomial with n=2-5; Gaussian with σ=1-25; C=0.1-100 in both cases). For all the movement classifications, the top 5 best classifier used the univariate feature ranked list of WPC and WDC modality. All results presented here are 5-fold cross-validation results with 10 Monte Carlo repetitions to reflect upon predictive power of the results (generalization). The automated wrapper method guided by error rate is initially applied to each class and most of the obtained top ranked classifiers had imbalance in the sensitivity and specificity, signifying the importance of Q factor based wrapper methods. The error factor guided wrapper methods were tested with k = 1, 5, and 10; C = 0.1, 1, 5, and 10; σ = 1 and 25; n = 2 and 3. Although it was a small test range, most of them ended with Q factor <50%, that was due to low sensitivity and high specificity. Thus, Q factor based wrapper methods eliminates the imbalance in sensitivity and specificity.

3.1.1. Movement 1

Table 1: Best 5 classifiers for movement 1.

Classifier	*Configuration*	*Selected Features*	*Feature type*	*Sensitivity*	*Specificity*	*Correct rate*	*Q factor*
Naïve Bayes	Diagonal Linear	ROC 1-79	WPC	0.7	0.801	0.832	0.727
Naïve Bayes	Diagonal Linear	ROC 1-83	WPC	0.699	0.803	0.833	0.725
Naïve Bayes	Diagonal Linear	ROC 1-81	WPC	0.695	0.803	0.838	0.725
Naïve Bayes	Diagonal Linear	ROC 1-66	WPC	0.697	0.804	0.836	0.725
SVM	σ=15, C=2	ROC 1-17	WPC	0.689	0.734	0.749	0.703

Using the ranked list obtained from univariate feature ranking method, the Naïve Bayesian OVRs outperformed with a best Q factor of approximately 72%, followed by Gaussian SVMs. All the top 5 classifiers used the univariate ranked list of the WPC feature modality (Table 1). For the Gaussian SVMs, box constraints (C) and σ values were tested with all the possible combination (within the test range), and the values of the Gaussian SVM shown in the table 1 is chosen based on the best-attained Q factor. Here in general, for the movement 1 classification Q factor decreased with C while keeping σ constant. This indicates a preference for soft margins that are less sensitive towards outliers, which is expected given the low signal to noise ratio EEG-based BCI dataset. On the other hand, while maintaining a constant C, Q factor initially increased with σ, followed by an abrupt decline for the values >15. Here k-NN classifiers with all different feature ranking and selection methods performed poorly, with less than 50% Q factor. Polynomial SVM with order 2 and 3 performed poorly, and the higher orders struggled to converge with few feature modalities. Best 10 classifiers used univariate ranked list of WPC features followed by WDC. Although sequential forward and sequential backward feature sets yielded good error rates their Q factor were poor.

3.1.2. Movement 2

Table 2: Best 5 classifiers for movement 2.

Classifier	*Configuration*	*Selected Features*	*Feature type*	*Sensitivity*	*Specificity*	*Correct rate*	*Q factor*
SVM	σ=15, C=0.2	ROC 1-70	WPC	0.72	0.804	0.827	0.741
SVM	σ=15, C=0.1	ROC 1-65	WPC	0.715	0.792	0.813	0.734
Naïve Bayes	Diagonal Linear	ROC 1-80	WPC	0.701	0.808	0.839	0.728
Naïve Bayes	Diagonal Linear	ROC 1-78	WPC	0.701	0.803	0.833	0.727
Naïve Bayes	Diagonal Linear	ROC 1-68	WPC	0.700	0.801	0.834	0.727

For the classification of movement 2, the Gaussian SVM OVR's outperformed other classifiers with the best Q factor of 74.1%, followed by naïve Bayesians using WPC features ranked by univariate feature ranking method (Table 2). k-NN OVRs were on the bottom of the list, with the best Q factor of 47% using WPC features at k=18. Most of the SVMs with high Q factor had smaller C values (0.1-0.2), indicating that more slack was needed to allow for softer margins. This states that outliers in the training datasets needed to be misclassified to yield a more sensible maximum margin boundary. Similar to the movement 1, WPC features ranked by univariate feature ranking method produced better results followed by the univariate ranked list of WDC features. As seen, the correct rate of fifth best classifier is better than the correct rate of first best Q factor based classifier due to the imbalance in sensitivity vs. specificity.

3.1.3. Movement 3

Table 3: Best 5 classifiers for movement 3.

Classifier	*Configuration*	*Selected Features*	*Feature type*	*Sensitivity*	*Specificity*	*Correct rate*	*Q factor*
Naïve Bayes	Diagonal Linear	ROC 1-90	WDC	0.702	0.757	0.772	0.716
Naïve Bayes	Diagonal Linear	ROC 1-88	WDC	0.703	0.755	0.769	0.716
Naïve Bayes	Diagonal Linear	ROC 1-89	WDC	0.698	0.751	0.769	0.715
Naïve Bayes	Diagonal Linear	ROC 1-86	WDC	0.700	0.752	0.768	0.715
SVM	σ=15, C=0.1	ROC 1-80	WPC	0.700	0.755	0.770	0.714

Naïve Bayes OVRs were the best performed classifiers for movement 3 using univariate feature ranked list of WDC feature modality, followed by the Gaussian SVMs using univariate feature ranked list of WPC feature modality (Table 3). kNNs and polynomial SVMs again failed to match these numbers with their best result. kNNs yielded a maximum Q factor of 52% with WDC features, and k value of 31. Polynomial SVMs with higher order and high box constraints struggled to converge with few modalities. In terms of feature selection and ranking, univariate feature ranked list performed better followed by classification guided subset selection and ranking, keeping sequential forward and sequential backward selection to the end of the list. Although sequential backward selection feature sets from PSD and CEPS modalities showed low error rates with naive bayes and SVM classifier respectively, they faced imbalance problem with sensitivity and specificity and landed-up with low Q factor.

3.1.4. Movement 4

Classifying this movement was the most challenging case compared to rest. Gaussian SVMs with σ=15 and C=0.1 yielded the best results using WPC features ranked using univariate feature ranking. The better performances of Gaussian SVM over the rest of the classifiers indicate a more nonlinear decision boundary. kNNs and polynomial SVMs performed poorly with <50% Q factors, although the top three classifier based on error rates were kNN classifiers they had high imbalance in sensitivity and specificity. All the features for the top 5 ranked classifiers were derived for the ROC ranked list (Univariate feature ranking), using WPC features followed by WDC.

Table 4: Best 5 classifiers for movement 4.

Classifier	*Configuration*	*Selected Features*	*Feature type*	*Sensitivity*	*Specificity*	*Correct rate*	*Quality factor*
SVM	σ=15, C=0.1	ROC 1-51	WPC	0.677	0.719	0.732	0.689
SVM	σ=15, C=0.2	ROC 1-50	WPC	0.679	0.718	0.726	0.687
SVM	σ=15, C=0.1	ROC 1-50	WPC	0.679	0.722	0.730	0.687
SVM	σ=15, C=0.1	ROC 1-40	WPC	0.674	0.707	0.718	0.685
Naïve Bayes	Diagonal Linear	ROC 1-24	WDC	0.617	0.648	0.659	0.628

3.2.NIRS Analysis

Similar to the EEG analysis all combination of the feature selection and ranking methods, and classifiers were analyzed. The features for the classifier from univariate ranking method and classification guided subset selection and ranking methods are taken in the order of the top ranked feature list similar to EEG analysis. For the sequential forward (SF) and backward (SB) selections the number under the feature column represent the total number of elements selected in the feature set out of 936 total features. The tested range for all the classifiers are Naïve Bayesian (Diagonal linear-DL, Diagonal Quadratic-DQ), kNN (k = 1-10), and SVMs (Polynomial with n=2-5, and Gaussian with σ=1-25; C=0.1-10 in both cases). Here in the two-class BCI with equal number of samples in both classes, usage of the Q factor does not make much significance but the same method used for electric signals is applicable to the optical signals. All the results presented here in the NIRS analysis are the 5-fold cross validation results. Here I present the classifiers (Table: 5, 6, 7, 8) which are best of each different feature selection and ranking methods with each different classifier. Ranked list from the classification guided subset selection and ranking dominated the rest of the feature selection methods.

3.2.1. kNN Classifier

Although the test range of the k is 1 to 10, the classifiers with k value above 3 were selected for consistency in the results. The best performed kNNs produced a Q factor of 90.03% using the ranked feature list obtained from classification guided subset selection and ranking with subset size of 10, pool size of 150 and threshold of 0.905. The best quality factors for Univariate ranking methods, Classification guided subset selection and ranking, Sequential forward selection, and Sequential backward selection seen in the table 5.

Table 5: k-NN.

	Movement 1					Movement 2				
Ranking method	*K*	*Features*	*Sens*	*Spec*	*Q*	*K*	*Features*	*Sens*	*Spec*	*Q*
ROC	7	1 – 96	0.8	0.8	0.8	7	1 - 96	0.8	0.8	0.8
Random	4	1-26	0.95	0.9	0.9003	4	1 - 22	0.95	0.9	0.9003
SF	8	5	0.7	0.9	0.7055	4	3	0.80	0.75	0.7504
SB	7	985	0.7	0.7	0.7	7	984	0.7	0.7	0.7

3.2.2. SVM Classifiers

Table 6: Polynomial SVM.

	Movement 1					Movement 2				
Ranking method	*Order & C*	*Features*	*Sens*	*Spec*	*Q*	*Order & C*	*Features*	*Sens*	*Spec*	*Q*
ROC	3 & 0.1	1-16	0.8	0.75	0.7504	3 & 0.1	1-16	0.8	0.75	0.7504
Random	3 & 0.1	1-15	1	1	1	3 & 0.1	1 - 24	1	1	1
SF	2 & 0.1	5	0.85	0.75	0.515	3 & 0.1	3	0.8	0.9	0.8015
SB	2 & 0.1	985	0.55	0.65	0.5519	2 & 0.1	984	0.55	0.65	0.5519

Sens –Sensitivity, Spec–Specificity.

Unlike the EEG analysis, polynomial SVM performed better than Gaussian SVM. Both polynomial and Gaussian SVMs yielded better results using classification guided random subset selection and ranking with subset size of 10, pool size of 150 and threshold of 0.905. High tolerance of the sensitivity and specificity in most of the SVM's for the changes in the C values reveal the information that they have very few outliers. The 100% 5-fold cross validation results of the polynomial SVM using the ranked list from classification guided subset selection and ranking method proves the significance of the NIRS signals and the wrapper methods in designing a BCI. Best classifiers based on quality factor using the features from the 4 different feature ranking and selection methods are tabulated below.

Table 7: Gaussian SVM.

	Movement 1					*Movement 2*				
Ranking method	*Sigma & C*	*Features*	*Sens*	*Spec*	*Q*	*Sigma & C*	*Features*	*Sens*	*Spec*	*Q*
ROC	1 & 0.1	1 - 21	0.75	0.75	0.75	1 & 0.1	1 - 21	0.75	0.75	0.75
Random	2 & 0.1	1 – 13	0.95	0.90	0.9003	1 & 0.1	1 - 17	0.95	0.95	0.95
SF	3 & 0.1	5	0.85	0.85	0.85	2 & 0.1	3	0.9	0.85	0.8503
SB	9 & 0.1	985	0.8	0.7	0.702	9 & 0.1	984	0.8	0.7	0.702

3.2.3. Navïe Bayesian Classifier

Navïe Bayes classifiers were at the end of list, with best Q factor of 80.04% and 80% for the two configurations Diagonal linear (DL) and Diagonal quadratic (DQ) respectively, using the features from sequential forward selection for left hand and right hand respectively. For most of the values DL and DQ yielded similar results.

Table 8: Navïe bayesian classifier

	Movement 1					*Movement 2*				
Ranking method	*Type*	*Features*	*Sens*	*Spec*	*Q*	*Type*	*Features*	*Sens*	*Spec*	*Q*
ROC	DL/DQ	1-8	0.7	0.7	0.7	DL/DQ	1-8	0.7	0.7	0.7
Random	DL	1-8	0.65	0.75	0.6517	DL/DQ	1-23	0.7	0.7	0.7
SF	DL/DQ	5	0.8	0.85	0.8004	DL/DQ	3	0.8	0.8	0.8
SB	DL	985	0.65	0.7	0.6504	DL	984	0.7	0.65	0.6504

CHAPTER 4

DISCUSSION

4.1. EEG

As the complexity of the BCI increases with the number of brain states, the usage of Q factor in OVR models will reduce the required minimum number of individual classifiers and the imbalance in the sensitivity and specificity. We here proved how the Q factor in conjunction with wrapper methods can be used for multi-class subject-invariant BCIs challenging the problem which lead to trivial OVR classification. Figure 8 shows a simple best-first wrapper method that uses ROC AUC ranked feature list for a navïe Bayesian classifier with Q as the OVR classifier feedback. This could successful realize the best 4-class subject invariant BCI by solving the earlier mentioned problems with not only classification rate but also sensitivity and specificity. More sophisticated wrapper methods such as sequential forward selection, sequential backward selection (Theodoridis et al., 2009), and classification guided subset selection (Li et al., 2004) without Q factor as a classifier feedback failed to achieve acceptable results using similar datasets as they degenerated to the trivial classifiers using the same OVR models.

Although we used 4 different feature selection and ranking methods, univariate feature ranking were proved to be the best of all. For the best classifier using univariate feature ranking, the input feature vector of the classifier contains first D elements of the ranked list whose Q factor is the maximum out of all possible D's from the list. Figure 8 shows one such wrapper feature selection by plotting Q factor vs. the number of included features. The best classifier for movement 1 classification has a Q factor of 0.727 whose input vector has top 76 ranked features from WPC modality using univariate feature ranking method. The presence of the correlated elements in the feature vector results the fluctuations in Q factor, also known as nesting problem (Theodoridis et al., 2009). By using wrapper process, it is possible to remove such features using a sequential float forward selection. To prove this we eliminated all the local minima's (features ranked with 4, 6, 34, 37 and 69; as shown in figure 8 with red dots), that resulted in an improved Q factor of 0.7305. Now we tried to add all different combinations of the local minima's and achieved 0.7378 Q factor by adding 14th ranked element. Although there is slight improvement, it clearly shows the nesting problem. More advanced sequential selection algorithms like sequential float forward selection can eliminate this problem.

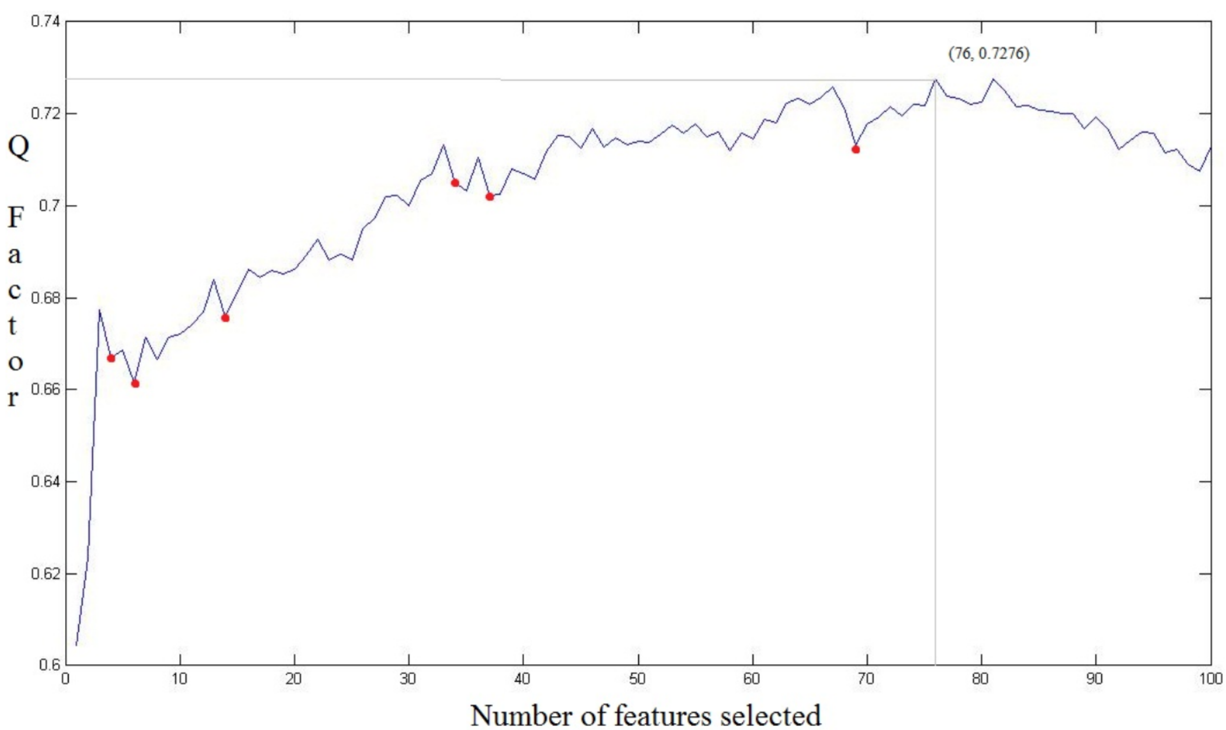

Figure 8: Quality factor verses number of features using a best-first wrapper with naïve Bayesian OVR. Q peaks at 0.728 using the first 76 features from ROC-AUC ranked WPC.

All the different classifiers were examined with all the modalities with error rate and Q factor as the classifier grading criterion. Talking about the overall performance of the classifiers, naïve Bayesian classifiers with diagonalized covariance matrices and normal distribution assumption, followed by Gaussian kernel SVMs, performed better than kNNs and polynomial kernel SVMs. The stability and robustness of the linear naïve Bayesian given the diagonalized covariance matrix proved to be the best for the EEG dataset. Wavelets features, both as packets (WPC) and regular decompositions (WDC), outperformed compared to the WPEs, WDEs, LPCs, PSDs, CEPS, and STFTs.

All the codes were designed in MATLAB using the built-in functions available for features extraction methods, feature selection and ranking methods, and classifiers. The built-in command for the sequential forward and sequential backward selections did struggled a lot and performed poorly. In most of the cases, the sequential backward selection ended up with majority of the input feature list notifying its struggle in the deletion of the features and consistent working. On the other hand for the classification guided subset selection and ranking, pool size and subset size need to be selected based on the number of input features and also subset size need to be varied to maximize the threshold. This can be clearly observed while we have a look at the results of the NIRS analysis.

4.2. NIRS

Applications of the BCI in the real world need to consider environmental noise, portability of the equipment, robustness, and processing time. For the applications like wheel chair control or any other equipment for the locked-in or paralyzed patients, the portable NIRS equipment need to consider the processing time. Noting that the difference in the temporal evolution of NIRS signals, I present a BCI eliminating the complex and time consuming feature extraction methods used in the EEG analysis. Resulting with the processing time is less than two seconds involving feature generation, selecting the preranked features and classifying them using pre-tested configuration using a 2.5 GHz processor (Intel Core 2 Duo - T9300).

Considering the non-linear patterns of the limited NIRS data, polynomial SVMs outperformed the navïe Bayesian, Gaussian SVMs, and kNNs. Although I used 3 different configurations for classification guided subset selection and ranking lists, the one with maximum threshold performed well. Once again feature set from sequential backward selection remained at the end of the list with the poor performance. Also, the mean and variance of the HbO and Hb signals of all the 26 channels did not showed up in the first 50 top ranked list of the classification guided subset selection and ranking.

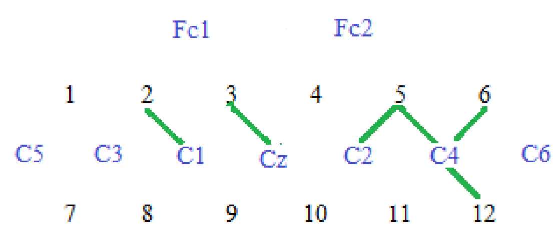

Figure 9: Selected paths for left hand classification.

The best classifier with 100% recognition rates for left hand and right hand classification is achieved by polynomial SVM using the features from classification guided subset selection and ranking.Table 9 shows the top 15 ranked time samples by the classification guided subset selection and ranking algorithm which were used by polynomial SVM to achieve 100% quality factor for left hand classification. The list mostly contains deoxy-hemoglobin time samples than that of oxy-hemoglobin. Figure 9 shows paths of the selected signals in the table 9 (5, 14, 17, 16, 10, 13), which come from the blood absorption around the electrodes C1, Cz, C2, and C4 of a 128 electrode placement style. Also, the majority of selected time samples for the best classifier are from the second half of the recorded 10 second signal. This shows that the fNIRS signals have intrinsic time delay.

Table 9: Best classified signals and their time samples.

Ranking method	*Signal*	*Time*
Hb	5	9.375
Hb	14	6.25
HbO	5	7.5
Hb	14	7.5
Hb	5	7.5
Hb	14	8.125
Hb	5	8.75
Hb	17	2.5
Hb	5	6.875
Hb	16	6.25
HbO	5	6.25
HbO	16	7.5
Hb	10	8,125
Hb	13	4.375
Hb	13	5.625

CHAPTER 5

CONCLUSION

Although the BCI performance evaluation have not become more significant till date, the future designs of more complex BCI (multiclass BCI) will prove the significance. So as to signify the importance of performance evaluation, I here present the 4-class subject invariant BCI based on the EEG data of 10 different untrained subjects given the fact that the spatio-temporal characteristics of the EEGs differ with respect to each subject and each trial with the subject (Christoforou et al., 2010). For the complex multiclass and subject invariant EEG data, the data driven OVR feature selection and classification using simple scalar metric like the classification rates lead to the degenerate multi-class BCI systems. This study signifies the usage of the Q factor in conjunction with wrapper methods for the classifier performance evaluation.

Naïve Bayesian classifiers with linear discriminant functions and Gaussian SVMs outperformed polynomial SVMs and kNN classifiers. WPC and WDC feature modalities outperformed the rest notifying that the features with proper time shift information are more salient. Also the Q-guided classification is needed for the feature extraction methods, classification guided subset selection and ranking, and sequential selection methods.

The best obtained OVRs for each movement were as follows (Tables 1 through 4, 5-fold cross-validation results with 10 monte-carlo repetitions): Movement 1: sensitivity 0.7, specificity 0.801, correct rate 0.832, and a Q factor of 0.727 (Naïve Bayesian classifier, top 79 WPC features ranked by univariate feature ranking method). Movement 2: sensitivity 0.72, specificity 0.804, correct rate 0.827, and a Q factor of 0.741 (SVM classifier, top 70 WPC features ranked by univariate feature ranking method). Movement 3: sensitivity 0.702, specificity 0.757, correct rate 0.772, and a Q factor of 0.716 (Naïve Bayesian classifier, using

90 WDC features ranked by univariate feature ranking method). Movement 4: sensitivity 0.677, specificity 0.719, correct rate 0.732, and a Q factor of 0.689 (SVM classifier, top 15 WPC features ranked by univariate feature ranking method). Without considering this complex subject invariant design over 10 different untrained subject using Q factor are evaluation performance, other studies state that 70% accuracy is adequate to control a two class BCI (Sebald et al., 2000).

The aforementioned wrapper methods used in the EEG analysis were also applied to the NIRS signals time sample to achieve perfect classification. In the NIRS analysis, I present a perfect classifier with 100% 5-fold cross validation recognition rates challenging the classification rates problem of the BCI's. The best 5-fold cross validation classification rate of 97.1% for single subject in the previous study (Niide et al., 2009) is being replaced by 100% 5-fold recognition rates using dataset of two subjects. Unlike the other studies who classified each subject separately and averaging the 5-fold cross validation results, here I combined two subject's data (subject invariant BCI) which is more complex. The best obtained classifiers for each movement were as follows (Tables 5 through 8, cross-validation results): Movement 1: sensitivity 1, specificity 1, correct rate 1, and a Q factor of 1 (Polynomial SVM classifier, top 15 features ranked by classification guided subset selection and ranking method). Movement 2: sensitivity 1, specificity 1, correct rate 1, and a Q factor of 1 (Polynomial SVM classifier, top 24 features ranked by classification guided subset selection and ranking method).

CHAPTER 6

FUTURE WORK

Future focus will be on the study of malfunctions in basal ganglia and corresponding regions. Given the location of the basal ganglia the signals produced by it are difficult to record on the scalp. Due to the presence of the dopamine pathways between the frontal lobe and motor cortex, and basal ganglia, the signals can be detected on the scalp. Study will focus on the patterns of the brain signals measured from EEG. As a case study I would like to analyze the EEG patterns of a dystonic subject with psychogenic dystonia subject.

Dystonia is a neurological movement disorder characterized by involuntary muscle movements involving twists, abnormal postures, and repeated movements in a part or whole body (Fahn, 1987). Dystonia can be considered as the inability to select the required muscles for a movement (Guehl et al., 2009). Although several theories evolved from past few decades on Dystonia, the cure is yet to be determined. Some studies defined the cause to be hereditary or Genetic disorder (Ozelius et al., 1997), birth-related (Ozelius et al., 1997), infection, reactions to pharmaceutical drugs, specific role of brain stem, basal ganglia nuclei (Bhatia et al., 1994 and Fross et al., 1987), premotor cortex (Feve et al., 1994), and brain disease (Marsden et al., 1976). Dystonia caused due to malfunction of basal ganglia can be observed in motor cortex using electroencephalography (EEG) through dopamine pathways connecting frontal lobe and motor cortex. Also Dystonia-specific early and late movement related Electroencephalographic (EEG) patterns have been observed in patient EEG signals during their dystonic episodes (Feve et al., 1994). EEG signals of some movements and thoughts have unique signatures observed across the brain and also their characteristics do differ with respect to subject and also repetition of single EEG experiment (Allison et al., 2007 and Christoforou et al., 2010). The subject participated in future study have psychogenic dystonia. I would like to compare his EEG signals to that of a Dystonic subject. However, all the efforts for controlling Dystonia or any other activity by humans are as result of the activity in brain and thus it benefits the study. The study and observation of the EEG patterns related to the psychogenic dystonia and dystonia is the primary concern. In the future study I would like to apply a variety of simple mathematical and complex machine learning algorithms related to Brain Computer Interfacing for recognizing the patterns of brain activity.

REFERENCES

Allison, B.Z., Wolpaw, E.W., and Wolpaw, J.R. 2007. Brain computer interface systems: progress and prospects. *Medical devices*, *4*, 463-474.

Anderson, C.W., Stolz, E.A., and Shamsunder, S. 1995. Discriminating mental tasks using EEG represented by AR models. In *Proceedings of the 17th IEEE Annual Conference on Engineering in Medicine and Biology Society*, Montreal, Que., Canada, Sept 1995, 875-876.

Anderson, C.W., Stolz, E.A., and Shamsunder, S. 1998. Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks. *Biomedical Engineering*, *45*, 277–286.

Bhatia, K.P., and Marsden, C.D. 1994. The behavioural and motor consequences of focal lesions of the basal ganglia in man. *Brain*, *117*, 859–876.

Birbaumer, N., Kubler, A., Ghanayim, N., Hinterberger, T., Perelmouter, J., Kaiser, J., Iversen, I., Kotchoubey, B., Neumann, N., and Flor, H. 2000. The Thought Translation Device (TTD) for completely paralyzed patients. *Rehabilitation Engineering, 8*, 190-193.

Bishop, C.M. 2006. *Pattern recognition and machine learning.* Springer, New York, NY.

Chan, M., Estève, D., Escriba, C., and Campo, E. 2008. A review of smart homes – Present state and future challenges. *Computer Methods and Programs in Biomedicine, 91*, 55-81.

Christoforou, C., Haralick, R.M., Sajda, P., and Parra, L.C. 2010. The Bilinear Brain: Towards Subject-Invariant Analysis. In *Proceedings of 4th International Conference on Communications, Control and Signal Processing,* Limassol, March 2010, 1-6.

Cover, T.M., and Hart, P.E. 1967. Nearest Neighbor Pattern Classification. *Information Theory*, *13*, 21-27.

Coyle, S., Ward, T., Markham, C., and McDarby, G. 2004. On the suitability of near-infrared (NIR) systems for next-generation brain-computer interfaces. *Physiological Measurement, 25*, 815-822.

Coyle, S.M., Ward, T.E., and Markham, C.M. 2007. Brain-computer interface using simplified functional near-infrared spectroscopy system. *Journal of Neural Engineering, 4*, 219-226.

Curran, E.A., and Stokes, M.J. 2003. Learning to control brain activity: A review of the production and control of EEG components for driving brain-computer interface (BCI) systems. *Brain and Cognition, 51,*326-336.

Delorme, A., and Makeig, S. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Neuroscience Methods, 134*, 9–21.

Doyle, T.E., Kucerovsky, Z., and Ieta, A. 2006. Affective State Control for Neuroprostheses. In *Proceedings of 28th Annual Innternational Conference on Engineering in Medicine and Biology Society*, New York, NY, August-September 2006, 1248-1251.

Doynov, P., Sherwood, J., and Derakhshani, R. 2008. Classification of Imagined Motor Tasks, In *IEEE Region 5 Conference*, Kansas City, MO, April 2008, 1-6.

Duda, R.O., Hart, P.E., and Stork, D.G. 2001. *Pattern classification*. Wiley, New York, NY.

Fahn, S. 1987. Systemic therapy of dystonia. *The Canadian Journal of Neurological Sciences, 14,* 528–532.

Fawcett, T. 2006. An Introduction to ROC analysis. *Pattern Recognition Letters*, *27*, 861-874.

Felton, E., Lewis, N.L., Willis, S.A., and Radwin, R.G. 2007. Neural Signal Based Control of the Dasher Writing System. In *Proceedings of 3rd International IEEE/EMBS Conference on Neural Engineering,* Kohala Coast, HI, May 2007, 366-370.

Feve, A., Bathien, N., and Rondot, P. 1994. Abnormal movement related potentials in patients with lesions of basal ganglia and anterior thalamus. *Neurology, Neurosurgery & Psychiatry, 57,* 100-104.

Fross, R.D., Martin, W.R., Li, D., Stoessl, A.J., Adam, M.J., Ruth, T.J., Pate, B.D., Burton, K., and Calne, D.B. 1987. Lesions of the putamen: their relevance to dystonia. *Neurology, 37,* 1125-1129.

Fukui, Y., Ajichi, Y., and Okada, E. 2003. Monte Carlo prediction of near-infrared light propagation in realistic adult and neonatal head models. *Applied Optics, 42*, 2881–2886.

Gonon F.G. 2003. Nonlinear relationship between impulse flow and dopamine released by rat midbrain dopaminergic neurons as studied by in viva electrochemistry. *Neuroscience, 24*, 19-28.

Guehl, D., Cuny, E., Ghorayeb, I., Michelet, T., Bioulac, B., and Burbaud, P. 2009. Primate models of dystonia. *Progress in Neurobiology, 87*, 118-131.

Gwet, K. 2002. Inter-rater reliability: dependency on trait prevalence and marginal homogeneity. *Statistical Methods for Inter-Rater Reliability Assessment, 2,* 1-9.

Heeger, D.J., and Ress, D. 2002. What does fMRI tell us about neuronal activity? *Nature Reviews Neuroscience, 3*, 142-151.

Herman, P., Prasad, G., McGinnity, T.M., and Coyle, D. 2008. Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification. *Neural Systems and Rehabilitation Engineering, 16,* 317-326.

Huppert, T.J., Hoge, R.D., Diamond, S.G., Franceschini, M.A., and Boas, D.A. 2006. A temporal comparison of BOLD, ASL and NIRS hemodynamic responses to motor stimuli in adult humans. *NeuroImage, 29,* 368-382.

Jain, A.K., Duin, R.P.W., and Mao, J. 2000. Statistical pattern recognition: a review. *Pattern Analysis and Machine Intelligence, 22,* 4-37.

Koles, Z.J., Lind, J.C., Soong, A.C.K. 1995. Spatio-temporal decomposition of the EEG: a general approach to the isolation and localization of sources. *Electroencephalography and Clinical Neurophysiology, 95*, 219–230.

Li, L., Umbach, D.M., Terry, P., and Taylor, J.A. 2004. Application of the GA/KNN method to SELDI proteomics data. *Bioinformatics, 20,* 1638-1640.

Mallat, S. 2001. *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, USA.

Marsden, C.D., Harrison, M.J., and Bundey, S. 1976. Natural history of idiopathic torsion dystonia. *Advances in Neurology, 14,* 177–187.

Martinez, P., Bakardjian, H., and Cichocki, A. 2007. Fully Online Multicommand Brain-Computer Interface with Visual Neurofeedback Using SSVEP Paradigm. *Computational Intelligence and Neuroscience, 2007,* doi:10.1155/2007/94561.

Mason, S.G., Bashashati, A., Fatourechi, M., Navarro, K.F., and Brich, G.E. 2007. A Comprehensive Survey of Brain Interface Technology Designs. *Annals of Biomedical Engineering,35,*137-169.

McFarland, D.J., Miner, L.A., Vaughan, T.M., and Wolpaw, J.R. 2000. Mu and beta rhythm topographies during motor imagery and actual movement. *Brain Topography, 12*, 177–186.

Millan, J.R. 2003. Adaptive Brain Interfaces. In *Communications of the ACM*, 46, 75-80.

Mitra, S.K. 2006. Digital *Signal Processing: A Computer Based Approach*. *McGraw Hill, New York, NY*.

Niide, W., Tsubone, T., and Wada, Y. 2009. Identification of moving limb using near infrared spectroscopy signals of brain activation. In *Proceedings of International Joint Conference on Neural Networks*, Atlanta, GA, June 2009, 2264-2271.

Oppenheim, A.V., and Schafer, R.W. 1975. *Digital Signal Processing*. Prentice Hall, Englewood Cliffs, NJ.

Ortiz, S., Jr. 2007. Brain-Computer Interfaces: Where Human and Machine meet. In *Computer*, Lee Garber, IEEE Computer society, 34-36.

Ozelius, L.J., Hewett, J., Page, C.E., Bressman, S.B., Kramer, P.L., Shalish, C., Leon, D., Mitchell, F.B., Deborah, R., David, P.C., Stanley, F., Neil, J.R., Alan, J.B., James, F.G., and Xandra, O.B. 1997. The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. *Nature Genetics, 17*, 40–48.

Parra, L.C., and Sajda, P. 2003. Blind source separation via generalized eigen value decomposition. *The Journal of Machine Learning Research, 4,* 1261–1269.

Roberts, S.J., and Penny, W.D. 2000. Real-time brain-computer interfacing: a preliminary study using Bayesian learning. *Medical and Biological Engineering and Computing,38*, 56-61.

Schlögl, A., Kronegg, J., Huggins, J.E., and Mason, S.G. 2007. Evaluation Criteria for BCI Research. In *Towards Brain-computer Interfacing,* MIT press, Cambridge, MA.

Schlögl, A., Lee, F., Bischof, H., and Pfurtscheller, G. 2005. Characterization of Four-Class Motor Imagery EEG Data for the BCI-Competition. *Journal of Neural Engineering, 2,* L14-L22.

Schmitz, C.H., and Koch, S.P. 2009. Ultra-compact, EEG-compatible NIRS system. *Poster C 05 at the Berlin Brain Computer Interface Workshop*, Berlin, Germany, July 2009.

Sebald, D.J., and Bucklew, J.A. 2000. Support Vector Machine Techniques for Nonlinear Equalization. *Signal Processing, 48,* 3217-3226.

Sherwood, J., and Derakhshani, R. 2009. On classifiability of wavelet features for EEG-based brain-computer interfaces. *International Joint Conference on Neural Networks,* Atlanta, GA, June 2009, 2895-2902.

Sitaram, R., Zhang, H., Guan, C., Thulasidas, M., Hoshi, Y., Ishikawa, A., Shimizu, K., and Birbaumer, N. 2007. Temporal Classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface. *NeuroImage, 34*, 1416-1427.

Theodoridis, S., and Koutroumbas, K. 2009. *Pattern Recognition*. Academic Press, Burlington, MA.

Ting, W., Guo-zheng, Y., Bang-hua, Y., and Hong, S. 2008. EEG feature extraction based on wavelet packet decomposition for brain computer interface. *Measurement, 41*, 618-625.

Valsan, G., Grychtol, B., Lakany, H., and Conway, B.A. 2009. *The Strathclyde Brain Computer Interface.* In *Proceedings of Annual Innternational Conference on*

Engineering in Medicine and Biology Society, Minneapolis, MN, September 2009, 606-609.

Vetterli, M., and Herley, C. 1992. Wavelets and Filter Banks: Theory and Design. *Signal Processing, 40,* 2207-2231.

Waske, B., Schiefer, S., and Braun, M. 2006. Random feature selection for decision tree classification of multi-temporal SAR data. In *Proceeding of IEEE International conference on Geosciences and Remote Sensing Symposium*, Denver, CO, July-August 2006, 168-171.

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

VITA

Vikas Gottemukkula, received B.Tech. in Electronics and Communication engineering from Scient Institute of Technology, Hyderabad, India. He is pursuing M.S. in Electrical Engineering at University of Missouri Kansas City. He is a member of the Eta Kappa Nu society. His research interests include machine learning, biomedical signal analysis, and brain computer interfacing.