Automatic Classification of Athletes With Residual Functional Deficits Following Concussion by Means of EEG Signal Using Support Vector Machine

Cheng Cao, *Student Member, IEEE*, Richard Laurence Tutwiler, and Semyon Slobounov

***Abstract*—There is a growing body of knowledge indicating long-lasting residual electroencephalography (EEG) abnormalities in concussed athletes that may persist up to 10-year postinjury. Most often, these abnormalities are initially overlooked using traditional concussion assessment tools. Accordingly, premature return to sport participation may lead to recurrent episodes of concussion, increasing the risk of recurrent concussions with more severe consequences. Sixty-one athletes at high risk for concussion (i.e., collegiate rugby and football players) were recruited and underwent EEG baseline assessment. Thirty of these athletes suffered from concussion and were retested at day 30 postinjury. A number of task-related EEG recordings were conducted. A novel classification algorithm, the support vector machine (SVM), was applied as a classifier to identify residual functional abnormalities in athletes suffering from concussion using a multichannel EEG data set. The total accuracy of the classifier using the 10 features was 77.1%. The classifier has a high sensitivity of 96.7% (linear SVM), 80.0% (nonlinear SVM), and a relatively lower but acceptable selectivity of 69.1% (linear SVM) and 75.0% (nonlinear SVM). The major findings of this report are as follows: 1) discriminative features were observed at theta, alpha, and beta frequency bands, 2) the minimal redundancy relevance method was identified as being superior to the univariate t-test method in selecting features for the model calculation, 3) the EEG features selected for the classification model are linked to temporal and occipital areas, and 4) postural parameters influence EEG data set and can be used as discriminative features for the classification model. Overall, this report provides sufficient evidence that 10 EEG features selected for final analysis and SVM may be potentially used in clinical practice for automatic classification of athletes with residual brain functional abnormalities following a concussion episode.**

***Index Terms*—Electroencephalography (EEG), feature selection, mild traumatic brain injury (MTBI), support vector machine (SVM).**

I. Introduction

MILD traumatic brain injury (MTBI), commonly known as "concussion," is still one of the most puzzling neurological disorders and least understood injuries. The problem with concussion is that with the exception of the unconscious individual or someone who is severely dazed, it is often very difficult to identify who has sustained a concussion and who has not [1]. Although blunt trauma involved with MTBI appeared to be a relatively simple type of brain injury, the potpourri of symptoms vary from subject to subject and often hint the true complexity of the injury [1]–[4]. The consequence is that attempts to classify concussion as a traumatic event with predictable findings upon initial examination may be erroneous. Most grading scales currently used to diagnose the severity of a head injury are based on loss of consciousness (LOC) or posttraumatic amnesia criteria, both of which occur infrequently in MTBI [3]. It should be noted that common physical and neurological symptoms are usually resolved within a one-week period [5], [6]. Overall, the use of traditional measures of concussion for making athletes' return-to-play (RTP) decisions can have catastrophic results [1].

Despite dramatic advances in understanding the mechanisms of concussion, there are essentially no studies reported aimed at predicting athletes at risk for concussion based upon residual abnormalities after the first incident. This is an important scientific and clinical quest, as undetected signs of a single mild brain injury may be a predisposing factor of concurrent concussions including the risk of permanent brain damage [1].

There are numerous electroencephalography (EEG) studies of MTBI. Early EEG research in 300 patients clearly demonstrated the slowing of major frequency bands and focal abnormalities within 48-h postinjury [7]. A study by McClelland *et al.* [8] has shown that EEG recordings performed during the immediate postconcussion period demonstrated a large amount of "diffusely distributed slow-wave potentials," which were markedly reduced when recordings were performed six weeks later. A shift in the mean frequency in the alpha (8–10 Hz) band toward lower power and an overall decrease of beta (14–18 Hz) power in patients suffering from MTBI were observed by Tebano [9]. The reduction of theta power [10] accompanying a transient increase of alpha-theta ratios [11], [12] was identified as residual symptoms in MTBI patients. The most comprehensive EEG study using a database of 608 MTBI subjects up to eight-year postinjury revealed 1) increased coherence in frontal-temporal regions, 2) decreased power differences between anterior and posterior cortical regions, and 3) reduced alpha power in the posterior cortical region, which were attributed to mechanical head injury [13]. A study by Thornton [14] has shown a similar data trend in addition to demonstrating the attenuation of EEG within the high-frequency gamma cluster (32–64 Hz) in MTBI patients. Recently, the usefulness

Manuscript received June 15, 2007; revised December 31, 2007; accepted January 4, 2008. First published February 15, 2008; last published August 13, 2008 (projected). The work of S. Slobounov was supported by the National Institutes of Health under Grant RO1 NS056227-01A2 "Identification of Athletes at Risk for Traumatic Brain Injury."

C. Cao and R. L. Tutwiler are with the Department of Kinesiology and the Department of Electrical Engineering, Pennsylvania State University, State College, PA 16802 USA (e-mail: cxc687@psu.edu; rlt1@psu.edu).

S. Slobounov is with the Department of Kinesiology, Pennsylvania State University, State College, PA16802 USA (e-mail: sms18@psu.edu).

Digital Object Identifier 10.1109/TNSRE.2008.918422

and high sensitivity of EEG in the assessment of concussion have been demonstrated [15]. Although, it should be noted that a controversial report indicated that no clear EEG features are unique to MTBI, especially after postinjury [16].

The aforementioned studies have provided solid evidence indicating the alteration of EEG patterns in concussed individuals. However, it should be noted that there are still some shortcomings in previous EEG research on concussion both in terms of experiment design and data analysis. First, almost all the previous studies are conducted under the single eyes-closed resting condition. There were few reports about the EEG changes under other conditions, especially those that require more mental activities. In this study, it was expected that the residual functional deficits could be more clearly pronounced under more challenging conditions than under resting and relaxed conditions. Second, the discriminant analyses were mainly conducted on an individual feature basis or on the simple linear combination of several features. The EEG feature selection was based only on known clinical experiences rather than on evidence from statistical learning theory.

A number of recent studies clearly documented the feasibility of supervised and nonsupervised pattern recognition algorithms to classify subjects with various health-related issues. Specifically, a number of neural network studies to detect seizure activity using EEG [17]–[19], characteristically including the Waves of Sleep EEG [20] and brain maturation [21], have been reported. In this paper, a recently developed classification algorithm, the support vector machine (SVM), has been applied as a classifier to identify athletes who suffer from residual functional deficits after a concussion episode. A multichannel EEG data set under multiple conditions has been used as an input for the classification algorithm. The SVM was widely used in the past with a high success rate in many EEG-based brain–computer interface (BCI) applications. As reported, the SVM "probably is the most appropriate classifier to deal with feature vectors of high dimensionality [22]," which is particularly fit for our case (91 features originally). Overall, it is our goal in this report to provide additional evidence that long-lasting residual functional abnormalities in concussed individuals may be overlooked by using current clinical practices to assess concussion.

II. Method: Data Acquisition

A. Subjects

A total of 61 subjects were recruited for this sport-related concussion study. All subjects were Pennsylvania State University athletes at high risk for traumatic brain injury (collegiate rugby and football players), between the ages of 18 and 25 years, male ($n = 27$, mean age $= 19.86$ years) and female ($n = 34$, mean age $= 20.40$ years). None of these subjects had a concussion history according to the subjects' self-report at the time of baseline testing. Thirty of these subjects suffered from grade 1 MTBI (MTBI, Cantu Data Driven Revised Concussion Grading Guidelines, 2006) within six months after baseline testing and were tested on day 30 postinjury basis. It should be noted that all 30 subjects were clinically asymptomatic at day 30 of testing and were cleared for full sport participation based upon neurological and neuropsychological (NS) assessments, as well as clinical symptoms' resolution. Specifically, a commonly accepted assessment in clinical practice medical examination of these subjects was performed at the Athletic Medicine Clinic following the guidelines of the international Co-operative Ataxia Rating Scale [23], including neurological, otologic, and neurotologic testing. The exclusion criteria for the MTBI group tested at day 30 postinjury were the presence of several symptoms such as complaints of loss of concentration, dizziness, fatigue, headache, irritability, visual disturbances, and light sensitivity. Moreover, these subjects were asymptomatic based on NS assessments, including Postconcussion Symptom List, Hopkins Verbal Learning Test, Symbol Digit Modality Test, Stroop Color-Word Test, Trail Making Test, and others. A detailed description of inclusion criteria for MTBI can be found in [24] and [25]. Consent forms approved by the Institutional Review Board of the Pennsylvania State University were obtained from each subject prior to every testing session. The retested EEG data at day 30 postinjury of the 30 injured athletes and the baseline EEG data of the rest of the 31 healthy athletes were used for feature selection and classifier training and validation based on the leave-one-out (LOO) method.

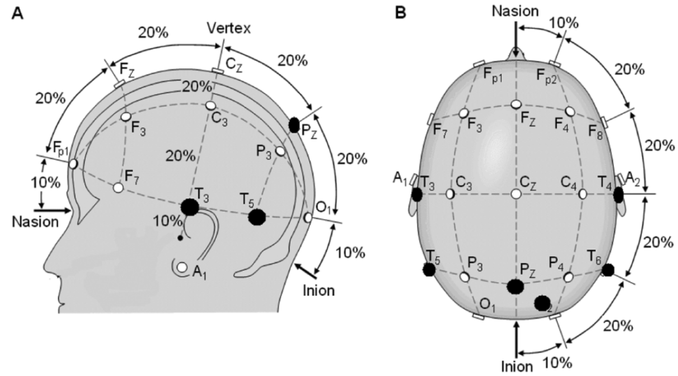


Fig. 1. Position of the 19 electrodes used for data acquisition located according to the 10–20 system. A1 and A2 are referring electrodes. The electrodes filled with black color are where the 10 selected EEG features comes from. (adopted from http://bulter.cc.tut.fi/~ malmivuo/bem/bembook/13/13x/1202ax.htm, and modified by the primary author).

Moreover, to further verify the validity of the features chosen by the algorithm and the discriminant functions acquired by the SVM, and to explore the changes of EEG before and after MTBI within each individual, the baseline data of the 30 concussed subjects, which were exclusive of the training data set, were used as an additional testing data set.

B. Experiment Setup and Posture Conditions

The EEG signals were recorded according to International system 10/20 [26], using 19 electrodes with two reference electrodes located at the earlobes and one ground electrode located at the vertex of the scalp (Fig. 1), in three whole body posture conditions: seated with eyes closed, standing on a firm surface with eyes closed, and standing on a foam (soft) surface with eyes closed. An offline analysis was performed using Neuroguide's 2.3.5 software. Each EEG epoch was visually inspected to eliminate any artifact that was undetected by the software's artifact rejection toolbox. Epochs that contained movement artifacts or any excessive muscle activity at any electrode site were eliminated from further analyses. A minimum of 1 min of artifact-free EEG was used in the analysis of each of the three tasks (seated, firm surface, foam surface). Using Neuroguide's software program, band pass filtering between 0.5 and 30 Hz with zero phase-shift was applied to each of the artifact-free records. A fast Fourier transform (FFT) in the frequency domain was performed on the artifact-free EEG records for each of the 19 recording sites. Frequency bandwidths were divided according to the following division: theta (4–7.5 Hz), alpha (8–12 Hz), beta1 (12–15 Hz), beta2 (15–17.5 Hz), and beta3 (18–25 Hz). Overall, power averages were computed for each bandwidth using Neuroguide's analysis program. In addition to average power within frequency bands defined above, power averages for individual 1 Hz frequencies between 1 and 30 Hz at all 19 electrode locations for each subject and for each of the three posture conditions were computed.

C. Data Preprocessing

According to the definition in Section II-A, the initial data dimension d_0 was large, [$d_0 = 19$ (number of electrodes) $\times$ 35 (number of frequency bands) $\times$ 3 (number of posture conditions) $= 1995$]. To assist in making the sophisticated feature selection algorithms and classification algorithm computationally feasible, a data preprocessing algorithm was performed for dimension reduction in order to select a robust subset of features.

The first stage was an exploratory data analysis (EDA) algorithm using descriptive statistics. Power means and the standard error of those means multiplied by 1.96 were used. Variables that overlap between groups using this analysis were then excluded from the next stage. Stage two was an analysis of Pearson's correlation coefficients, which determine the linear relationship between the response variable (injury) and the features. The features with correlations considered significant at the 0.05 level were kept. After the two stages of preprocessing, 92 features were selected as the candidate subset for further feature selection methods and classification algorithms.

III. Feature Selection Methods and Classification Methods

In many pattern recognition applications, choosing the most robust features from the observed data is very important with respect to empirical accuracy and the general performance of classifiers as well as the computation efficiency [27]–[30]. In this study, two feature selection methods were performed to find the optimal subset of features for classifying concussed and normal subjects. One is based on the univariate t-test, and the other is based on the heuristic minimal-redundancy-maximal-relevance (mRMR) framework. Both methods are independent of type of classifier. In this study, a linear SVM was applied as a classifier to find an optimal feature size. A backward wrapper method was applied to the subset of the features to find the optimal features within the subset. Then, a nonlinear SVM was applied to enhance performance, using the features that were selected from the methods mentioned above. The contents in this part consist of six sections. In Section II-A, the original SVM and one of its variants, referred to as the soft margin SVM, are both introduced. In Sections II-B and II-C, the univariate t-test and the mRMR are described, respectively. In Section II-D, the framework to search for the optimal feature subset is described. In Section II-E, the backward wrapper method, which is used to determine the final feature set, is introduced. The nonlinear SVM is introduced in Section II-F.

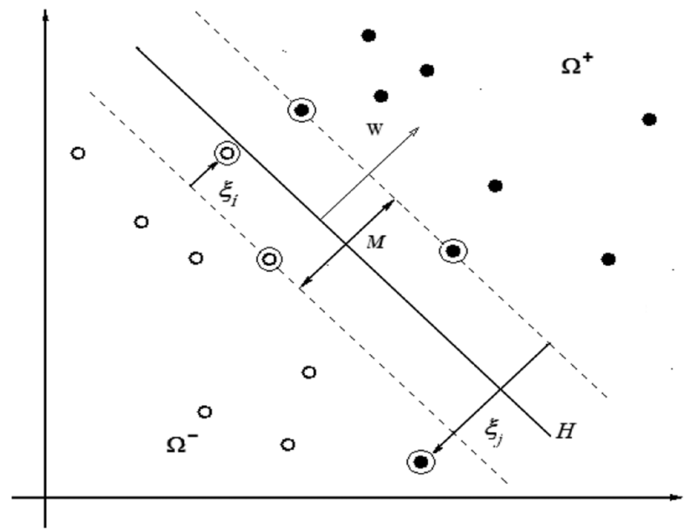


Fig. 2. Sketch map of linear SVM. For nonseparable data sets, slack variables Ξ_i are introduced. Thick points are called SVs. The solution for the hyperplane H can be written in terms of the SVs.

A. Support Vector Machines (SVMs)

The SVM is a novel classification technique developed by Vapnik [31] whose theoretical foundation is derived from VC dimension theory [31]–[35]. The central idea of SVM is to minimize the VC dimension of the classifier by separating the data, $X \in \mathbf{R}^d$ from two classes by finding a weight vector $w \in \mathbf{R}^d$ and an offset $b \in \mathbf{R}^d$ of the hyperplane

$$\begin{aligned} H : \mathbf{R}^d &\to \{-1, 1\} \\ x &\mapsto \operatorname{sign}(x \cdot w + b) \end{aligned} \tag{1}$$

with the largest possible margin, $M = 2/\|w\|$, which is the distance between the closest members from the two classes. One variant of the original SVM algorithm, referred to as soft margin SVM, was developed to deal with the situation in which the data cannot be linearly separated. The soft margin SVM solves the following optimal problem

$$\begin{aligned} &\min_{w \in R^d} \ \|w\|^2 + C \sum_{i=1}^{n} \xi_i^2 \\ &\text{s.t. } y_i (w \cdot x_i + b) \geq 1 - \xi_i \ (i = 1, \ldots, n). \end{aligned} \tag{2}$$

In this minimization problem, n is the number of samples in the training data set and $y_i \in \{-1, 1\}$ is the target value of the training data set. In this study, $y = 1$ denotes the concussed subject and $y = -1$ denotes the normal subject (Fig. 2).

The parameters ζ_i are called slack variables. The trade-off between the maximal margin and the minimal training error $\sum_{i=1}^{n} \zeta_i^2$ is controlled by the regulation parameter C. Choosing a proper value for C is important to the performance of the classifier. If C is set to be too large, the classifier is overfitted and the generalization performance suffers. Particularly, when C goes to infinity, the soft margin SVM algorithm performance is degraded into the original SVM algorithm. On the other hand, if C is set to be too small, the classifier will have a large training error, which means that the classifier is underfitted. In practical applications, C is always estimated from the training data by using a cross-validation procedure or other methods. In this study, C was estimated from the training data after preprocessing. Further feature selection methods were performed by the LOO method due to the relatively small size of the input data set.

B. Univariate t-test

The univariate t-test evaluates the probability (denoted as p) that a feature has different population means across two groups. The smaller p is, the more different a feature appears across two groups, indicating a stronger correlation between the feature and the group label. The features are ordered according to the p values from the feature that yields the smallest p to the feature that yields the largest p [36], [37]. The univariate t-test method is often used in practical applications because of the independence of classifiers and the computational simplicity; however, it suffers from the dependence between features.

C. mRMR Method

Selecting optimal features based on mutual information is a popular method. The mRMR method developed by Peng *et al.* [30], [38] is an improvement over one of the most popular methods, the Max-Relevance approach, and yields a better performance. Both the mRMR and Max-Relevance methods are referred to as "filter" methods. In practice, the "filter" method always has low computational complexity compared to the wrapper method, and chooses the features that lead to comparable classification accuracies on different classifiers [38]. Given two random variables x and y, their mutual information is defined in terms of their probability density functions, $p(x)$ and $p(y)$ and $p(x,y)$

$$I(x,y) = \int\int p(x,y)\log\frac{p(x,y)}{p(x)p(y)}dxdy. \tag{3}$$

The Max-Relevance method searches for features that satisfy the following conditions:

$$\mathbf{max}D(S,c), \quad D = \frac{1}{|S|}\sum_{x_i\in S} I(x_i,c) \tag{4}$$

where S is a feature set, x_i is the individual feature that belongs to S, and c is the class label.

Features selected according to the Max-Relevance method could have large redundancy. When two features have a high dependence on each other, the respective class-discriminative power will not be reduced much if one of them was removed. Therefore, the following minimal redundancy condition can be added to select mutually exclusive features [38]:

$$\mathbf{min}_R(S), \quad R = \frac{1}{|S|^2}\sum_{x_i,x_j\in S} I(x_i,x_j). \tag{5}$$

The criterion combining the above two constraints is called mRMR. By maximizing $\mathbf{\Phi}(\cdot)$ defined by the following equation, R and D are optimized simultaneously:

$$\mathbf{max}\mathbf{\Phi}(D,R), \quad \mathbf{\Phi} = D - R. \tag{6}$$

In practice, the following incremental search method is used to find the near-optimal features:

$$\mathbf{max}_{x_j\in X-S_{m-1}}\left[I(x_j,c) - \frac{1}{m-1}\sum_{x_i\in S_{m-1}} I(x_j,x_i)\right] \tag{7}$$

where X is the set of all the features and S_{m-1} is the feature set with $m-1$ features. The incremental search method finds the mth feature that maximize $\mathbf{\Phi}(\cdot)$. The computational complexity of the incremental research method is $O(|S|\cdot n)$, where n is the total number of the features. The mutual information $I(x,y)$ between two continuous variables x and y was estimated by the density estimation method using Parzen Windows [38].

D. Selection of the Feature Subset

Suppose that $S_1^i, S_2^i, \ldots, S_n^i$ $(S_1^i \subset S_2^i \subset \cdots \subset S_n^i)$, $i = 1$, 2 are n sequential feature sets selected by the univariate t-test $(i=1)$ or the mRMR method $(i=2)$, and e_k^i is the classification error estimated by the LOO method corresponding to each specific S_k^i; then, the optimal n^{*i} is determined by the minimal classification error $e^{*i} = \min e_k^i$.

In this section, the classifier associated with the classification is a linear SVM utilizing regulation parameter C determined previously. The performance curves e_k^i generated by the univariate t-test and the mRMR methods were compared to find a better method that leads to smaller n^*, lower e^*, and a smoother performance curve.

E. Backward Wrapper Method

The backward selection tries to exclude one redundant feature at a time from the current feature set S_k (initially, k is set to be n^* selected in Section II-D), to satisfy the constraint that the reduced feature set, denoted as S_{k-1}, leads to a classification error e_{k-1} not worse than e_k. As every feature in S_k is considered to be a possible feature to be removed, there are k different configurations of S_{k-1}. For each possible configuration, the respective classification error e_{k-1} is calculated. The backward selection is terminated if, for every configuration, the corresponding e_{k-1} is larger than e_k. Otherwise, among the k configurations of S_{k-1}, the feature that yields the largest error reduction is chosen as the new feature set. If the number of configurations leading to the largest error reduction is more than one, the new feature set is randomly chosen from one of these configurations. This decremental selection procedure is repeated until the termination condition is satisfied.

The wrapper method usually exhibits high classification accuracy at the cost of large computational complexity and less generalization of the features selected on other classifiers [38]–[40]. In this study, as the wrapper method is applied on a relatively small subset of features, the computational complexity is acceptable. Moreover, as the feature subset was chosen by the feature selection methods (the univariate-test and the mRMR) that are not dependent on the classifier, the performance of the features chosen by the wrapper method

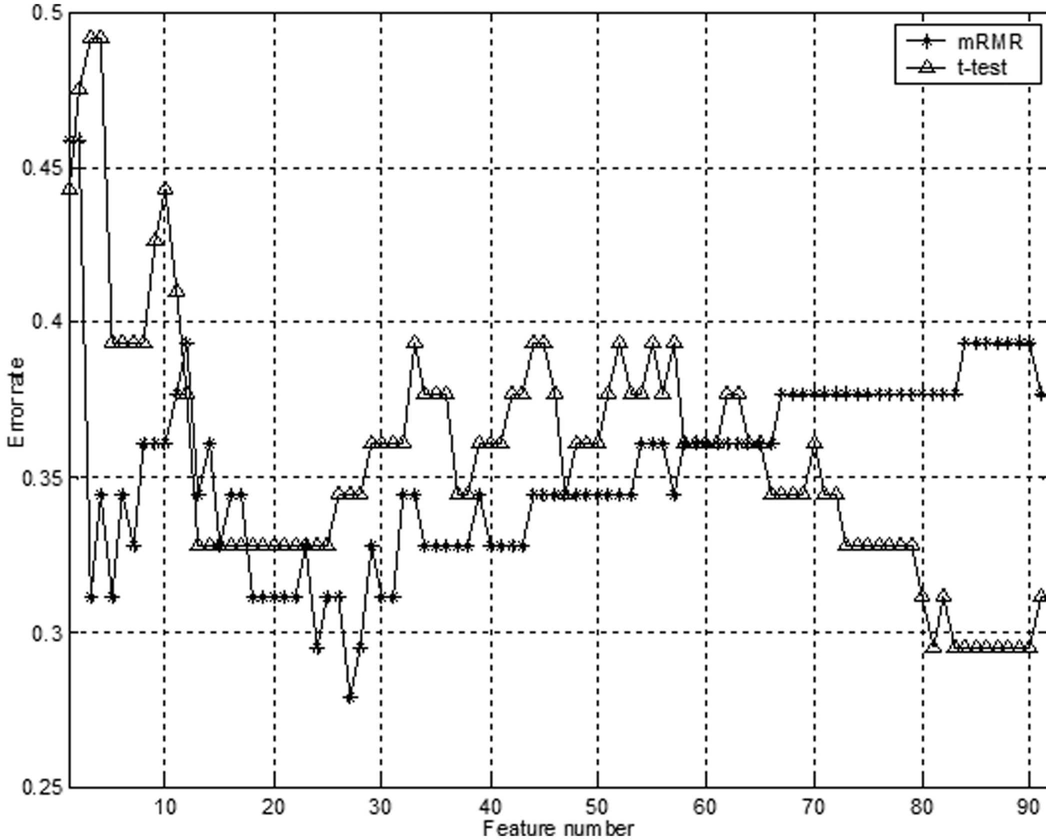


Fig. 3. Comparison of classification accuracies based on the features selected by the mRMR method and the t-test. Liner SVM was used. LOO CV errors were calculated.

from this subset will be less dependent on the classifiers than from the total feature set. A nonlinear SVM was utilized at the end of this analysis to justify the feature selection performance of our scheme and to enhance the classification performance.

F. Nonlinear SVMs

The basic idea of nonlinear SVMs is to map the input vecto $X \in R^d$ into the vectors $\Psi(X)$ in a higher dimensional feature space F, where Ψ represents mapping $R^n \to R^f$, and also to perform a linear classification in this feature space. In practice, instead of performing a mapping $\Psi(X)$, a kernel function $K(x_i, x_j)$ is used. By choosing the kernel $K(x_i, x_j)$ an SVM that operates in an infinite dimensional space can be constructed. Refer to [33] for details.

In this study, a Gaussian kernel was used. The Gaussian kernel is defined as follows:

$$K(x_i, x_j) = e^{1/2\left(\|x_i - x_j\|^2\right)/\gamma}. \tag{8}$$

The kernel parameter γ is referred to as the kernel width. The regulation parameter C of the SVM defined in (2) and the kernel width γ were determined by the training data after preprocessing prior to feature selection. The optimal C and γ were searched for on a 13×19 grid where the two parameters take the following values:

$$C = \left\{10^2, 10^{2.5}, 10^3 \ldots 10^7, 10^{7.5}, 10^8\right\},$$
$$\gamma = \left\{10^{-5}, 10^{-4.5}, 10^{-4} \ldots 10^3, 10^{3.5}, 10^4\right\}.$$

For each pair (C, γ), the generalization error was estimated by the LOO method. The pair that yields the smallest generalization error was chosen [33].

IV. Results

A. Feature Selection

Two feature selection methods, the univariate t-test and the mRMR frame work, were performed on the data set to select the compact feature sets. The generalization error curves computed by the univariate t-test and mRMR methods were compared to find a better method that leads to a comparatively small feature set size n^* and low e^* with a smooth error curve. The generalization errors were estimated by the LOO cross-validation procedure. The estimated error curves are depicted in Fig. 3. The mRMR procedure generated a smoother performance curve with a smaller n^* and lower minimum error e^*. Note that the generalization error curve of mRMR drops rapidly up to the first 27 features and then rises gradually as the number of features increases. This trend is consistent with the pattern of the idealized generalization error curve when some optimal feature selection method is used in the presence of irrelevant and redundant features (Fig. 4). This consistency indicates that the mRMR method is capable of finding relevant concussion features. In comparison, the generalization error curve generated by the univariate t-test does not follow this pattern. The generalization error initially descends up to the first 30 features, then increases with oscillation between the 30th and the 70th features, and finally drops

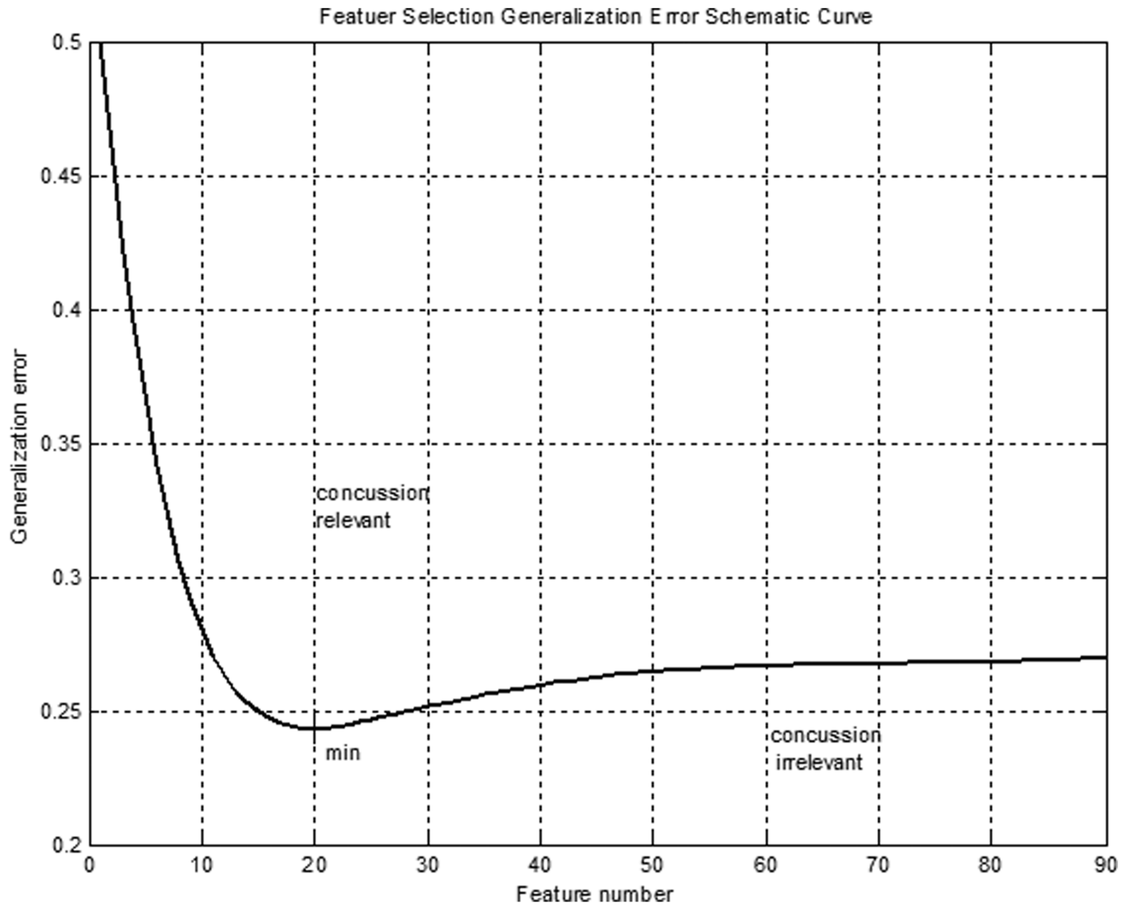


Fig. 4. Idealized error curve using a feature selection method in the presence of irrelevant and redundant features. As the number of included relevant features increases, the classification error drops until the remained features are either irrelevant or redundant.

TABLE I
FINALLY CHOSEN 10 FEATURES BY THE WRAPPER METHOD AND THEIR mRMR RANKS

	Features	mRMR rank
1	O2 10Hz foam	1
2	T5 beta2 foam	3
3	T4 beta2 foam	5
4	T3 theta sec	6
5	T4 beta2 sec	9
6	T5 beta3 stec	13
7	Pz 10Hz foam	15
8	T5 beta2stec	18
9	T6 10Hzfoam	19
10	T5 beta3foam	27

TABLE II
P-VALUES OF THE CHOSEN FEATURES AND THE MUTUAL INFORMATION BETWEEN EACH FEATURE AND CLASS LABEL

	Features	p-value of t-test	MI(bit)
1	O2 10Hz foam	0.0259	0.2682
2	T5 beta2 foam	0.0259	0.1738
3	T4 beta2 foam	0.0670	0.2183
4	T3 theta sec	0.8100	0.1139
5	T4 beta2 sec	0.3558	0.1791
6	T5 beta3 stec	0.0168	0.1573
7	Pz 10Hz foam	0.0188	0.2301
8	T5 beta2stec	0.0222	0.1731
9	T6 10Hzfoam	0.0331	0.2159
10	T5 beta3foam	0.0374	0.1317

to the minimum at the end of the curve. This pattern indicates that the t-test method is not capable of finding the relevant features associated with concussion. Therefore, the compact feature set selected by the mRMR method was used to further search for the best combination of features using the backward wrapper method that leads to the highest classification accuracy.

Table I shows the 10 features selected by the backward wrapper method. The p-value of each feature and the mutual information between each feature and the class label, denoted as MI, are listed in Table II. The p-values and MIs are highly negatively correlated, but not completely consistent with each other. This is because the computation of the p-value is based on the assumption that each feature has a normal distribution, which is not necessarily true in practice, whereas the mutual information is calculated from the estimation of the actual distribution of the features. It is reasonable to believe that the mutual information is more reliable than the p-value in terms of calculating the dependences as well as the independences of the individual feature and the class label (i.e., the subject's condition)

B. Classification

Forty-seven of 61 subjects were correctly classified by the linear SVM using the 10 selected features yielding a classification accuracy of 77.1%. In practice, the classification accuracy alone is not enough to describe the performance of the classifier. Two additional parameters, sensitivity and selectivity, are introduced in this paper. The sensitivity parameter is defined as the probability that the subject has suffered a concussion within one month and is successfully identified by the classifier. The selectivity parameter is defined as the probability that the subject identified by the classifier truly suffered a concussion within one month. The accuracy, sensitivity, and selectivity were estimated by

$$\begin{aligned} A &= \frac{n_{\text{correctly\_classified\_concuss}} + n_{\text{correctly\_classified\_normal}}}{n_{\text{concuss}} + n_{\text{control}}} \\ S_{en} &= \frac{n_{\text{correctly\_classified\_concuss}}}{n_{\text{concuss}}} \\ S_{el} &= \frac{n_{\text{correctly\_classified\_concuss}}}{n_{\text{concuss\_as\_"concuss"}}} \end{aligned} \tag{9}$$

where A stands for total accuracy, S_{en} stands for sensitivity, and S_{el} stands for selectivity.

It is easy to prove that the total accuracy A, the sensitivity S_{en}, and the selectivity S_{el} have the following relation:

$$A = 1 - P_1 \left[1 + S_{en}\left(\frac{1}{S_{el}} - 2\right)\right] \tag{10}$$

where

$$P_1 = \frac{n_{\text{concuss}}}{n_{\text{concuss}} + n_{\text{control}}}. \tag{11}$$

Twenty-nine of the 30 subjects with concussion were correctly classified by the linear SVM using the 10 features and there were 13 false positives. The sensitivity and selectivity of the linear SVM associated with the 10 features are 96.7% and 69.1%, respectively.

A nonlinear Gaussian kernel SVM was applied to the 10 selected features, yielding a comparable performance. The total accuracy, sensitivity, and selectivity are 77.1%, 80.0%, and 75.0%, respectively (i.e., 47 of the 61 subjects were correctly classified, 24 of the 30 subjects with concussion were correctly identified and there were eight false positives). Although the total accuracy of nonlinear SVM is the same as that of the linear SVM, the sensitivity and selectivity of the nonlinear SVM are more balanced than those of the linear SVM.

With regard to the additional verification using the baseline data of the 30 subjects who suffered from concussion after the baseline test, 66.7% (linear SVM) and 73.3% (nonlinear SVM) of the data were correctly classified as "normal."

V. Discussion

To the authors' knowledge, this study is the first report that attempts to use the SVM to detect residual functional abnormalities at 30-day postinjury using a multichannel EEG data set under multiple conditions. We believe our approach may be complementary to numerous EEG-based discriminant analyses of MTBI patients. There are several findings of interest. First, the locations of the 10 features are concentrated on the temporal (T3, T4 and T5, T6) and occipital (O2) areas. It should be noted that the subjects' reports indicated that concussion accidents were associated with impact to the side of the head during collision. A recent report by Delaney *et al.* [41] has also indicated that the side/temporal area of the head or helmet is the most probable area to be struck, resulting in concussion for both football and soccer players. Biomechanical events set up by the concussive blow (such as amount of head movement about the axis of the neck at the time of impact, the site of impact, etc.) ultimately result in concussion [42], and their analysis may contribute to a more accurate assessment of the degree of damage and potential for recovery. Our findings are also in agreement with numerous imaging studies indicating that MTBI most likely affects and possibly damages the prefrontal and baso-temporal areas of the brain with the possibility of superficial contusions and acute sharing injury [43]–[48]. Finally, the most comprehensive EEG study using a database of 608 MTBI subjects up to 8-year postinjury revealed 1) increased coherence in frontal-temporal regions and 2) decreased power differences between anterior and posterior cortical regions. Overall, the identified locations of the 10 features (i.e., temporal area of the brain) are in agreement with previous MTBI brain imaging and biomechanical research.

Also, five out of 10 features that classified concussed individuals were derived from EEGs when subjects were tested standing on foam. This is consistent with our previous research indicating that the alteration of postural control measures was observed in concussed individuals only during the more challenging dynamic postural tasks [25]. For example, the residual deficits were most prominent when concussed individuals were exposed to more demanding, conflicting visual scene information induced by virtual reality graphics [49], [50]. Also, previous research has shown not only erroneous motor responses but also prolonged reaction time (RT) values especially during the choice RT tests, but not simple reaction time tests between normal control and concussed individuals [51]. Similarly, abnormalities in isometric force production in concussed individuals have been documented, but only when the complexity of the motor task was systematically increased [24]. Moreover, a series of recent studies have clearly demonstrated motor impairments, including abnormal eye movement and visuomotor/occulomotor arm movement deficits in MTBI patients in the absence of traditional neuropsychological markers of concussion [52]–[54].

Finally, as can be seen from Tables I and II, the frequency band configuration of the features is also an important factor in classifying concussed individuals. Specifically, discriminative features were observed at theta, alpha, and beta frequency bands. It should be noted that similar frequency bands were targeted in a number of previous EEG studies of concussion. Specifically, a shift in the mean frequency in the alpha (8–10 Hz) band toward lower power and an overall decrease of beta (14–18 Hz) power in patients suffering from MTBI were observed by Tebano *et al.* [9]. The reduction of theta power [10] accompanying a transient increase of alpha-theta ratios [11],

[12] was identified as residual symptoms in MTBI patients. Reduced alpha power in the posterior cortical region, which was attributed to mechanical head injury, was also observed by Thatcher [13].

The testing result on the additional independent data set verified that the features and discriminant functions are robust. It should be noted that the sensitivity corresponding to the EEG data of the same 30 subjects after MTBI is 96.7% and 80.0% for linear and nonlinear SVMs, respectively. This proves that there are significant differences detectable by SVM in the EEG signals between baseline and MTBI within each individual.

Overall, this report demonstrated that a feature selection scheme utilizing the t-test or mRMR to find the candidate compact features followed by the application of the forward wrapper method to find the best combination of the compact feature set yields the lowest generalization error. Due to the relatively small sample size of this study, a LOO method was used to estimate the generalization error. The estimated error curves illustrate that the mRMR method yields both a good compact feature selection and a good generalization performance.

The total accuracy of the classifier using the 10 features is 77.1%. The classifier has a high sensitivity of 96.7% (linear SVM) and 80.0% (nonlinear SVM), and a relatively lower but still acceptable selectivity of 9.1% (linear SVM) and 75.0% (nonlinear SVM). The relatively lower selectivity may be due to the relatively small sample size of brain-injured subjects. However, this report provides sufficient evidence that the 10 EEG features from multiple conditions selected for final analysis and the SVM may be potentially used in clinical practice for automatic classification of athletes suffering from residual functional abnormalities after concussion episodes. The proposed approach, in conjunction with other modern approaches (e.g., analysis of biomechanical impact, brain imaging studies, duration of symptoms resolution, etc.) can be potentially used by clinicians for more accurate return-to-sport participation criteria after concussive episodes.

References

[1] R. Cantu, "Concussion classification: Ongoing controversy," in *Foundations of Sport-Related Brain Injuries*. New York: Springer, 2006, pp. 87–111.

[2] J. T. Barth, J. R. Freeman, D. K. Boshek, and R. N. Varney, "Acceleration-deceleration sport-related concussion: The gravity of it all," *J. Athletic Training*, vol. 36, no. 3, pp. 253–256, 2001.

[3] K. Guskiewicz, "Assessment of postural stability following sport-related concussion," *Current Sports Medicine Rep.*, vol. 2, no. 1, pp. 24–30, 2003.

[4] M. Lovell, "Ancillary test for concussion. Neurotrauma and sport medicine review," *J. Neurosurgery*, vol. 98, pp. 296–301, 2003.

[5] R. J. Echemendia, M. Putukien, R. S. Mackin, L. Julian, and N. Shoss, "Neuropsychological test performance prior to and following sports-related mild traumatic brain injury," *Clin. J. Sports Medicine*, vol. 11, pp. 23–31, 2002.

[6] S. T. Macciocchi, J. T. Barth, W. Alves, R. W. Rimel, and J. Jane, "Neuropsychological functioning and recovery after mild head injury in collegiate athletes," *Neurosurgery*, vol. 3, pp. 510–513, 1996.

[7] W. Geets and N. Louette, "Early EEG in 300 cerebral concussions," *Electromyogr. Clin. Neurophysiol.*, vol. 14, no. 4, pp. 333–338, 1985.

[8] R. J. McClelland, G. W. Fenton, and W. Rutherford, "The postconcussional syndrome revisited," *J. Roy. Soc. Med.*, vol. 87, pp. 508–510, 1994.

[9] T. M. Tebano, M. Cameroni, G. G. A. Loizzo, G. Palazzino, G. Pessizi, and G. F. Ricci, "EEG spectral analysis after minor head injury in man," *Electroenceph. Clin. Neurophysiol.*, vol. 70, pp. 185–189, 1988.

[10] A. Montgomery, G. W. Fenton, R. J. McCLelland, G. MacFlyn, and W. H. Rutherford, "The psychobiology of minor head injury," *Psychol. Medicine*, vol. 21, pp. 375–384, 1991.

[11] R. Pratar-Chand, M. Sinniah, and F. A. Salem, "Cognitive evoked potential (P300): A metric for cerebral concussion," *Acta Neurol. Scand.*, vol. 78, pp. 185–189, 1988.

[12] W. R. Watson, R. J. Fenton, J. McClelland, J. Lumbsden, M. HeadleyRutherford, and W. H. Rutherford, "The post-concussional state: Neurophysiological aspects," *Br. J. Psychiatr.*, vol. 167, pp. 514–521, 1995.

[13] R. W. Thatcher, R. A. Walker, I. Gerson, and F. H. Geisler, "EEG discrimination of mild head injury," *Electroencephalogr. Clin. Neurophysiol.*, vol. 73, pp. 94–106, 1989.

[14] K. E. Thornton, "Exploratory investigation into mild brain injury and discriminant analysis with high frequency bands (32–64 hz)," *Brain Injury*, vol. 13, no. 7, pp. 477–488, 1999.

[15] J. Duff, "The usefulness of quantitative EEG (QEEG) and neurotherapy in the assessment and treatment of post-concussion syndrome," *Clin. EEG Neurosci.*, vol. 35, no. 4, pp. 198–209, 2004.

[16] M. Nuwer, D. Novda, M. Schrader, and M. Vespa, "Routine and quantitative EEG in mild traumatic brain injury," *Clin. Neurophysiol.*, vol. 116, no. 9, pp. 2001–2025, 2005.

[17] O. Ozdamar, I. Yaylali, P. Jayaker, and C. N. Lopez, "Multilevel neural network system for EEG spike detection," in *Computer-Based Med. Syst.*, May 4, 1991, vol. 12, pp. 272–279.

[18] N. Acir, İ. Öztura, M. Kuntalp, B. Baklan, and C. Güzeliş, "Automatic detection of epileptiform events in EEG by a three-stage procedure based on artificial neural networks," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 1, pp. 30–40, Jan. 2005.

[19] I. N. Bankman, V. G. Sigillito, R. A. Wise, and P. L. Smith, "Feature-based detection of the k-complex wave in the human electroencephalogram using neural networks," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 12, pp. 1305–1310, Dec. 1992.

[20] T. Shimada, T. Shiina, and Y. Saito, "Detection of characteristic waves of sleep EEG by neural network analysis," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 3, pp. 369–379, Mar. 2000.

[21] L. Moreno, J. D. Pineiro, J. L. Sanchez, S. Manas, J. Merino, L. Acosta, and A. Hamilton, "Brain maturation estimation using neural classifier," *IEEE Trans. Biomed. Eng.*, vol. 42, pp. 428–432, Apr. 1995.

[22] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for eeg-based brain-computer interfaces," *J. Neural Eng.*, vol. 4, pp. R1–R13, 2007.

[23] P. Trouillas, T. Takayanagi, M. Hallett, D. Currier, S. Subramony, K. Wessel, A. Bryer, H. Diener, S. Massaquoi, C. Gomez, P. Coutinho, M. Ben Hamida, G. Campanella, A. Filla, L. Schut, J. Honnorat, N. Nighoghossian, and B. Manyam, "International cooperative ataxia rating scale for pharmacological assessment of the cerebellar syndrome," *J. Neurol. Sci.*, vol. 145, pp. 205–211, 1997.

[24] S. Slobounov, J. Johnston, H. Chiang, and W. Ray, "Movement-related EEG potentials are force or end-effector dependent: Evidence from a multi-finger experiment," *Clin. Neurophysiol.*, vol. 113, pp. 1125–1135, 2002.

[25] S. Slobounov, W. Sebastianelli, R. Tutwiler, and E. Slobounov, "Alteration of postural responses to visual field motion in mild traumatic brain injury," *Neurosurgery*, vol. 59, no. 1, pp. 134–139, 2006.

[26] H. H. Jasper, "The ten-twenty system of the international federation," *Electroencephalogr. Clin. Neurophysiol.*, vol. 10, pp. 371–375, 1958.

[27] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, Jan. 2000.

[28] N. Kwak and C. H. Choi, "Input feature selection by mutual information based on parzen window," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 12, pp. 1667–1671, Dec. 2002.

[29] F. J. Iannarilli and P. A. Rubin, "Feature selection for multiclass discrimination via mixed-integer linear programming," *IEEE Trans. Pattern. Anal. Mach. Intell*, vol. 25, no. 6, pp. 779–783, Jun. 2003.

[30] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," in *Proc. Comput. Syst. Bioinformat.*, 2003, pp. 523–528.

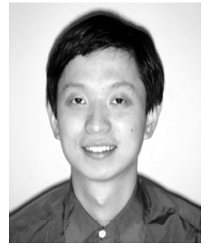
[31] V. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.

[32] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: Wiley, 2002.

[33] L. Wang, *Support Vector Machines: Theory and Applications*. Berlin, Germany: Springer, 2005.

[34] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[35] B. Dong, C. Cao, and S. E. Lee, "Applying support vector machines to predict building energy consumption in tropical region," *Energy Build.*, vol. 37, no. 5, pp. 545–553, 2005.
[36] Y. Xiao, J. P. Hua, and E. R. Dougherty, "Quantification of the impact of feature selection and the variance of cross-validation error estimation," *EURASIP J. Bioinformatics Syst. Biol.*, 2007.
[37] W. S. Nobel, "Support vector machine applications in computational biology," in *Kernel Methods in Computational Biology*. Cambridge, MA: MIT Press, 2004, pp. 71–92.
[38] H. C. Peng, F. H. Long, and C. Ding, "Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern. Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.
[39] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artif. Intell.*, vol. 97, no. 1-2, pp. 273–324, 1997.
[40] T. N. Lal, M. S. Schroder, T. Hinterberger, J. Weston, M. Bogdan, N. Birbaumer, and B. Schölkopf, "Support vector channel selection in BCI," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1003–1010, Jun. 2004.
[41] S. Delaney, J. Puni, and F. Rouah, "Mechanisms of injury for concussions in university football, ice hockey, and soccer: A pilot study," *Clin. J. Sport. Med.*, vol. 16, no. 2, pp. 162–165, 2006.
[42] N. Shaw, "The neurophysiology of concussion," *Progr. Neurobiol.*, vol. 67, pp. 281–344, 2002.
[43] H. H. Abu-Judeh, R. Parker, M. Singh, H. el-Zeftawy, S. Atay, and M. Kumar, "SPET brain perfusion imaging in mild traumatic brain injury without loss of consciousness and normal computed tomography," *Nucl. Med. Comm.*, vol. 20, pp. 505–510, 1999.
[44] I. Bicik, B. P. Radanov, N. Schafer, J. Dvorak, B. Blum, B. Weber, C. Burger, G. K. v. Schulthess, and A. Buck, "PET with 18 fluorodeoxyglucose and hexamethylpropylene amine oxime SPECT in late whiplash syndrome," *Neurology*, vol. 51, pp. 345–350, 1988.
[45] P. A. Hofman, F. R. Verhey, J. T. Wilmink, N. Rozendaal, and J. Jolles, "Brain lesions in patients visiting a memory clinic with postconcussional sequelae after mild to moderate brain injury," *J. Neuropsych.*, vol. 14, pp. 176–184, 2002.
[46] R. Kant, L. Smith-Seemiller, G. Isaac, and J. Duffy, "Tc-HMPAO SPECT inpersistent post-concussion syndrome after mild head injury: Comparison with MRI/CT," *Brain Injury*, vol. 11, pp. 115–124, 1997.
[47] M. Lorberboym, Y. Lampl, I. Gerzon, and M. Sadeh, "Brain SPECT evaluation of amnestic ED patients after mild head trauma," *Am. J. Emergency Medicine*, vol. 20, pp. 310–313, 2002.
[48] T. W. McAllister, M. B. Sparling, L. A. Flashman, and A. J. Saykin, "Neuroimaging findings in mild traumatic brain injury," *J. Clin. Exp. Neuropsychol.*, vol. 23, pp. 775–791, 2001.
[49] S. Slobounov, T. Wu, and M. Hallett, "Neural basis subserving the detection of postural instability: An FMRI study," *Motor Control*, vol. 10, no. 1, pp. 69–89, 2006.
[50] S. Slobounov, M. Hallett, T. Wu, H. Shibasaki, and K. Newell, "Neural underpinning of postural responses to visual field motion," *Biol. Psychol.*, vol. 72, pp. 188–197, 2006.
[51] H. Hugenholtz, D. T. Stuss, L. L. Stethen, and M. T. Richards, "How long does it take to recover from a mild concussion?," *Neurosurgery*, vol. 22, no. 5, pp. 853–857, 1988.
[52] M. H. Heitger, T. J. Anderson, R. D. Jones, J. C. Dalrymple-Alford, C. M. Frampton, and M. W. Ardagh, "Eye movement and visuomotor arm movement deficits following mild closed head injury," *Brain*, vol. 127, pp. 575–590, 2004.
[53] M. H. Heitger, R. D. Jones, J. C. Dalrymple-Alford, C. M. Frampton, M. W. Ardagh, and T. J. Anderson, "Mild head injury–a close relationship between motor function at one-week post-injury and overall recovery at three and six months," *J. Neurol. Sci.*, vol. 253, pp. 34–47, 2007.
[54] M. H. Heitger, R. D. Jones, C. M. Frampton, M. W. Ardagh, and T. J. Anderson, "Recovery in the first year after mild head injury: Divergence of symptom status and self-perceived quality of life," *J. Rehabil. Medicine*, vol. 39, pp. 612–621, 2007.



Cheng Cao (S'08) received the B.S.E.E degree form the University of Science and Technology of China, Heifei, China, in 2003 and the M.E degree from the National University of Singapore, Singapore, in 2005. He is currently working toward the Ph.D. degree in the Department of Kinesiology and the M.S. degree in the Department of Electrical Engineering at Pennsylvania State University, State College.

His research interesting involves brain–computer interface, time-frequency analysis, signal separation, and the application of modern signal methods in quantitative analysis of EEG.

Mr. Cao won the Presidential Scholarship of the Pennsylvania State University in 2003 and the third prize of excellent student of University of Science and Technology of China in 1999–2002.

Richard Laurence Tutwiler received the A.S.E.T. degree from the Community College of Allegheny County, PA, Pittsburgh, PA, in 1976, the B.S.E.E. degree from the University of Pittsburgh, Pittsburgh, PA, in 1979, the M.S.E.E. degree from The Pennsylvania State University, University Park, in 1985, and the Ph.D. degree in electrical engineering from The Pennsylvania State University, in 1992.

He is currently the principal investigator involved in machine vision research at ARL. This research involves the application of differential geometry theory to develop invariant feature sets for image analysis and classifications of underwater and hyperspectral remote sensed images. His research interests in ultrasound consist of high-frequency analog and digital design, ultrasonic beamforming architectures, 3-D image reconstruction, parallel processing, image processing/analysis, and pattern recognition. Prior to joining ARL, he worked for HRB-Singer, Inc. (now Raytheon), State College, PA. His responsibilities included optimal signal processing modeling and simulation studies. Prior to joining HRB-Singer, Inc., he was a member of an advanced development laboratory with Motorola, Inc., Communications Division. He was a member of a hardware group responsible for the design and development of the first fully frequency synthesized portable communications product (VHF to GHZ) that Motorola marketed. He is currently a Senior Research Associate and Department Head of the Image Processing Group at the Applied Research Laboratory of The Pennsylvania State University. He is a member of the Graduate Faculty in the Department if Electrical Engineering, the Department of Acoustics, and the Department of Kinesiology.

Semyon Slobounov received the Ph.D. degree from the Department of Psychology, University of Leningrad, Leningrad, USSR (now St. Petersburg State University, St. Petersburg, Russia), in 1978 and the Ph.D. degree from the Department of Kinesiology, University of Illinois, Urbana-Champaign, in 1994.

He is Professor in the Department of Kinesiology, College of Health of Human Development and Adjunct Professor of Orthopaedics and Medical Rehabilitation with Hershey Medical College at Pennsylvania State University, with primary responsibilities to teach undergraduate and graduate courses in the areas of psychology of injury, neural basis of motor behavior, and psychophysiology. He is an adjunct investigator with the National Institute of Heath, National Institute of Neurological Disorders and Stroke. He also is an Adjunct Professor of the Neuroscience Program, Life Science Consortium, and an Affiliate Professor of Gerontology Center at Pennsylvania State. His research focused on neural basis of human movements with special emphasis on rehabilitation medicine, psychology, and neurophysiology, including traumatic brain injuries.