A Wavelet-Like Filter Based on Neuron Action Potentials for Analysis of Human Scalp Electroencephalographs

Elena L. Glassman, *Student Member, IEEE*

Abstract—This paper describes the development and testing of a wavelet-like filter, named the SNAP, created from a neural activity simulation and used, in place of a wavelet, in a wavelet transform for improving EEG wavelet analysis, intended for brain-computer interfaces. The hypothesis is that an optimal wavelet can be approximated by deriving it from underlying components of the EEG. The SNAP was compared to standard wavelets by measuring Support Vector Machine-based EEG classification accuracy when using different wavelets/filters for EEG analysis. When classifying P300 evoked potentials, the error, as a function of the wavelet/filter used, ranged from 6.92% to 11.99%, almost twofold. Classification using the SNAP was more accurate than that with any of the six standard wavelets tested. Similarly, when differentiating between preparation for left- or right-hand movements, classification using the SNAP was more accurate (10.03% error) than for four out of five of the standard wavelets (9.54% to 12.00% error) and internationally competitive (7% error) on the 2001 NIPS competition test set. Phenomena shown only in maps of discriminatory EEG activity may explain why the SNAP appears to have promise for improving EEG wavelet analysis. It represents the initial exploration of a potential family of EEG-specific wavelets.

Index Terms—Brain-computer interface (BCI), data-specific wavelets, electroencephalography (EEG), pattern classification, pattern recognition, time-frequency representations, wavelet analysis.

I. INTRODUCTION

HUMAN SCALP electroencephalographic measurements (EEGs) are one way to peer into the activity of the brain. There are several distinct neural rhythms found in EEGs which are created by subsystems of different sizes [1]. Accurate interpretation of EEGs, optimally on a single-trial basis, is critical for such applications as brain-computer interfaces (BCIs). When many trials are averaged together, there is often a characteristic waveform that can be recognized. However, on a single-trial basis that characteristic waveform is often too far below the noise floor to be observed.

The original goal of this research was to create an EEG-classification algorithm for eventual use in a BCI. The focus of this paper is an EEG-specific wavelet-like filter that was created with the hope of optimizing EEG signal analysis and thus improving classifier and eventual BCI performance. This paper describes the creation of a filter for use in place of a wavelet in EEG wavelet analysis and documents, through extensive testing, its performance compared to standard wavelets. The platform for comparison, the EEG-classification algorithm that was the original goal, is also described.

A. Context for Wavelet-Like Filter Creation

From the standpoint of developing a classification algorithm, performance depends on the signal analysis, feature selection, and classification methods used. The EEG-classification algorithm used in this paper employs the discrete wavelet transform (DWT) for signal analysis. The wavelet transform's output can be significantly affected by the choice of wavelet (the basic waveshape) with which the signal is analyzed [1]–[3]. As a result, the choice of wavelet can also have a significant impact on the quality of the results with regard to the classifier, which takes the wavelet coefficients as input features. There is no standard method for selecting the best wavelet, and some wavelets are more appropriate than others for particular types of input signals [4]–[7].

For example, the Haar (square wave) wavelet is more appropriate for analyzing a sum of square waves than any other standard wavelet; a single Haar coefficient from wavelet analysis, at the right scale and time, can encode one entire cycle of a square wave by itself [8].

"[L]ike a key designed to fit a lock, the best analyzing function for a neuroelectric event is one that matches that event in shape as closely as possible" [7]. This is also supported in [1]. One must average many trials to visually observe most known neuroelectric events, such as the P300 evoked potential; they are rarely observed in single-trial data. Therefore, instead of designing a filter for single-trial analysis that matches a specific neuroelectric event, the goal of this work is to create a filter that matches the neural activity *underlying* the neuroelectric events. This is consistent with the view of [9] that, when considering which wavelet to use, it is important to take the input signal's underlying structure into account. A wavelet that encapsulates the "underlying component structure" has not yet been documented in the literature [7]. Samar, Swartz, and Raghuveer [1] state that one could potentially improve system performance (i.e., EEG classification) through wavelet shape manipulation. Such a filter could possibly produce superior classification performance across many *different* tasks and their associated neuroelectric events.

The filter developed in this paper is created from a simple model of neural activity. It is not wavelet shape manipulation [1]

but creation, based on a model to give the resulting filter meaningful properties. It is designed to correspond with a general oscillation of neuronal activity, presumed to be a basic underlying component of EEGs. The waveshape generated by the model of neural activity was, after minor adjustments, a wavelet-like filter usable in the Matlab DWT [10].

B. EEG-Classification Algorithm (The Platform)

Before using the algorithm described in this paper as a platform for comparing wavelets, it was necessary to validate its ability to classify EEGs with reasonable accuracy regardless of the wavelet used. By applying the algorithm to the data of the Neural Information Processing Systems (NIPS) BCI Workshop Data Competition [11], it was possible to directly compare its performance to that of algorithms that had been entered into the competition. Its performance was internationally competitive and is documented in the RESULTS.

The algorithm first performs signal analysis on the data using a wavelet transform. Wavelet analysis produces a multiresolution measurement of energy across time and frequency, and is therefore well suited for transient signals like those in EEGs. It essentially finds correspondences between the input signal and a wavelet (a filter with very specific mathematical properties): the wavelet's length is dilated to measure correspondence with different frequency bands and shifted to measure correspondence with the signal in different regions of time.

Once the signals are analyzed using a wavelet transform, a subset of the wavelet coefficients is selected for input into the classifier. In this EEG-classification algorithm, the criterion is discriminability. Discriminability is a standard statistical measure of how well a feature indicates which class the EEG signal belongs to; it measures the relative overlap of the two classes' distributions of values for that feature [12]. It is defined in (1) as

$$D = \frac{|\mu_2 - \mu_1|}{\sqrt{\sigma_2^2 + \sigma_1^2}} \qquad (1)$$

where D refers to discriminability, μ refers to the mean of a distribution, and σ refers to the standard deviation of a distribution. Subscripts identify the class.

The selected features from the signal analysis are then used to train the classifier. In this EEG-classification algorithm, a Support Vector Machine (SVM) with a polynomial kernel was chosen as the classifier. The SVM parameter settings are determined through a grid search, performed by finding the algorithm's tenfold cross-validation error for every combination of parameter values within the ranges considered (see Section I-C below and Section V-D of WAVELET COMPARISON THROUGH PATTERN CLASSIFICATION). The parameter settings that yielded the least cross-validation error were used.

C. Background on Methods Employed in the Platform Algorithm

Wavelet analysis was chosen because of its property of multiresolution. The Fourier Transform (or Fast Fourier Transform, the FFT) is an example of an analysis method without this property [1]. For example, one must use a relatively short window of FFT analysis to clearly observe the relatively short P300 evoked potential in the EEG and a relatively long window of analysis

to clearly observe the relatively long Bereitschaftspotential in the EEG. By separating the signal into "packages" in which different levels of detail (scale/frequency) are prominent, it is possible to look at all levels at once, each at the appropriate resolution [1]. This is done in wavelet analysis.

The classifier employed in the EEG-classification algorithm, a polynomial SVM, combines high-dimensional embedding and maximum margin hyperplane optimization [13]. Maximum margin optimization is a way to optimize a hyperplane: the hyperplane is adjusted to be in between but as far as possible from the training examples that are closest to those of the other class [13]. Some general rules cannot be expressed as a straight line or hyperplane, but it is much harder to optimize a curvy boundary than a straight/planar one. High-dimensional embedding is a way to optimize a curvy boundary using planar boundary optimization techniques; it expands the dimensions of the feature space by adding dimensions that are the products of the original dimensions [13].

This polynomial SVM uses a kernel, which is a similarity measure between the feature vectors of two tokens. It is defined in (2), where i represents the index of the token from class A, j represents the index of the token from class B, and deg indicates the parameter value that sets the amount of high-dimensional embedding [52]

$$\left[1 + \left((Feature\ Vector)_i \bullet (Feature\ Vector)_j \right) \right]^{\deg}. \qquad (2)$$

The total number of dimensions in the new space is determined by the highest degree of the variable, which, in (2), is the result of the dot product after expanding the binomial. The degree parameter sets the number of times that binomial is multiplied by itself to achieve its fully expanded state. In this new feature space, the hyperplane is optimized using maximum margin optimization. When converted back into the original space, it still separates the tokens as it did when optimized, even though the tokens may not be as linearly separable as they were in the higher-dimensional feature space [13]. Though the polynomial SVM can create very nonlinear decision boundaries, it becomes a linear classifier when the degree parameter is set to one.

SVMs theoretically determine the hyperplane's position without local minima through the use of convex optimization for a given set of "hyperparameters" [14]. The hyperparameters, hereafter simply referred to as parameters, are the degree and regularization parameters that control the balance between the hyperplane's complexity and its specificity to the given training data [14]. Yet there are local minima in the SVM's performance as those parameters are varied. Generally, a grid search is employed to find the best combination of parameter values [14]. The SVM's performance is recorded in a matrix as each combination of parameters' values is tried, within set limits [14]. Valid statistical methods for assessing SVM performance include cross-validation, bootstrapping, and forward prediction [15]. Tenfold cross-validation was chosen for the described EEG-classification algorithm.

There are more advanced methods for determining the best combination of parameter values. Through approximations or heuristics, these methods attempt to circumvent extensive parameter searches [16], but the variations in SVM performance

can be very hard to predict (see Section V-D of WAVELET COMPARISON THROUGH PATTERN CLASSIFICATION). The parameter settings can have a significant impact on overall performance.

II. RELATED WORK

A. Optimizing EEG Wavelet Analysis

Finding or creating the optimal wavelet for EEG analysis, or any particular type of signal, has been the subject of much investigation. The method that most closely resembles that which is used in this paper is found in [7], [17], [18] and demonstrated in [19]. This method involves the construction of a Meyer wavelet; the difference in spectrum between this constructed wavelet and the neuroelectric waveform of interest is minimized via least-squares [7]. DWT high- and low-pass filters are then derived from the constructed wavelet such that any signal can be analyzed with this matched Meyer wavelet using the DWT [7]. It should be most appropriate for signals with waveforms similar to the waveform the wavelet was originally made to match. The method presented in this paper differs in that, instead of constructing a wavelet that matches a particular signal's waveform, a wavelet is designed to match the *underlying* activity of neuroelectric waveforms (and thus less specific to any one type of waveform).

Others have constructed wavelets using fractal interpolation functions [20] or by creating "super-wavelets" from linear combinations of standard wavelets [21]. It was demonstrated that super-wavelets can be produced with waveshapes very closely matched to those of the signal being analyzed [21]. Note that these super-wavelets do not necessarily conform to the requirements of a filter usable in the DWT.

The matching pursuit technique is another method for decomposing a signal using waveforms that are similar to those of the signal being analyzed [7]. This technique finds the weighted combination of waveforms (or "atoms") that is closest to the input signal using a predefined waveform library [7], [17]–[19], [22]. The flexibility can become a disadvantage, though [7]. Instead of using one waveform that is specifically related to the activity of interest, any waveform in the library may be used [7]. And if the dominant activity in any particular signal is noise, the library waveform that most closely resembles that noise will likely dominate the output of the analysis [7]. If the waveform that most resembled the activity of interest were the only waveform in use, perhaps it would have picked out the activity of interest without being as affected by the noise [7]. More information on library-based methods and adapted waveforms can be found in [23].

Other approaches to the problem of choosing or creating a wavelet include the examination of the wavelet's properties, such as its frequency selectivity, degree of regularity, and how many vanishing moments it has [4].

B. Signal Analysis, Feature Selection, and Classification

There is a wide variety of signal analysis methods currently in use in BCI-oriented EEG-classification algorithms; these methods have a large effect on the overall performance [24]. Autoregressive (AR) model parameters, Principal Component Analysis (PCA), Independent Component Analysis (ICA), common spatial pattern analyses, temporal filtering, power spectral density, temporal and spatial filtering, and wavelet analysis are all documented in the literature [24]–[30]. Once the signals are analyzed using the signal analysis method(s), it is often necessary to select a subset of the analysis's output for input into a classifier. For instance, a learning vector quantizer (LVQ) has been used to select electrodes and frequency bands [29], [31]. Also in the BCI literature, many different methods have been used for classification, such as linear and nonlinear discriminant analysis, supervised neural networks, and SVMs [29], [30], [32]–[34].

C. Current BCI Systems' Accuracies

There are several types of signals or features evident in EEGs that have been used to transmit information through BCIs. They are well represented by the datasets available through the past two BCI dataset competitions, the 2001 Neural Information Processing Systems (NIPS) BCI Workshop Data Competition [11], [35] and the 2003 BCI Competition [36]. EEG-classification algorithms are applied to these datasets with the task of differentiating between classes of signals and/or determining the subjects' intent.

1) Algorithms' Accuracies for EEG Differentiation: Both competitions included differentiation between left- and right-hand movements of some nature, either actual or imaginary. The movements (real or imagined) can be detected by the associated event-related synchronizations and resynchronizations (ERS and ERD) of the mu/beta rhythms [29]. In the first competition, error on actual hand movements ranged between 4% and 46%, with a mean of 16%. For this type of task, the expected error was 50% if classification had been made by chance. When the same task was addressed in the second competition, error ranged between 16% and 49%, with a mean of 32%. For the third dataset, it was the same task but with imagined instead of actual hand movements, and data was available from nine different subjects [11], [37]. Across all nine subjects, the competition entrants' results ranged from 12% to 40%, with a mean of 28% [11], [37].

Another general area of focus has been the self-regulation of slow cortical potentials (SCPs) and mu/beta rhythms. For this task, a cursor's movement is based on these potentials or rhythms, and the subject learns to self-regulate them (e.g., their amplitude) as he/she learns to move the cursor toward specific targets [29]. The algorithm is applied to interpret the EEG and determine which target the subject was intending to move the cursor toward [11], [36]–[38]. Error on SCP differentiation ranged from 11% to 49%, with an average of 22% [36], [38]. This particular task also had an expected error of 50% if classification had been by chance. For the mu/beta rhythm self-regulation, there were four possible targets; therefore the expected error if classification had been by chance was 75% [36], [38]. The entrants' error ranged from 28% to 76%, with a mean of 56% [36], [38].

The final common task for which a dataset was available was determining which letter the subject was focusing on based on his/her P300 evoked potentials to flashes of different columns and rows of a matrix of letters [39]. Since testing on this dataset
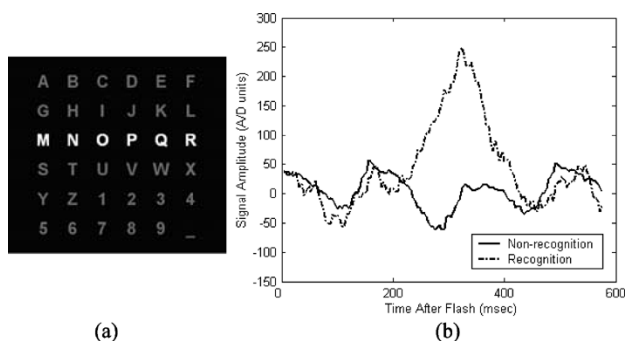
Fig. 1. (a) Matrix of letters to which the subject was responding. In this image, the third row is being flashed. Courtesy of [46]. (b) P300 recognition and nonrecognition responses.

is documented in this paper, there is a more detailed description in the DATA AND SUPPORTING SOFTWARE section. Classification by chance would give an expected error of 97% [36], [38]. All competition entrants combined data from multiple evoked potentials [36], [38]. Five entrants attained zero error and two registered 55% and 65% error [36], [38].

2) Established BCI Systems: The Wadsworth BCI system developed by Wolpaw, McFarland, and their colleagues uses mu/beta rhythm amplitude [29], [40]–[42]. The Graz BCI system utilizes ERD and ERS of the mu/beta rhythms (produced by motor imagery) [29]. The BCI described in [43] uses simple calculation tasks in addition to motor imagery [29].

III. DATA AND SUPPORTING SOFTWARE

Two public domain datasets of actual EEG data were used for experimentation in this study: the EEG Self-Paced Key Typing dataset [44], [45], from the Neural Information Processing Systems (NIPS) BCI Workshop Data Competition [11]; and the P300 Speller Paradigm dataset [46], from the BCI Competition of 2003 [36]. Both datasets enable a researcher to work with data from a laboratory-quality EEG machine. It is also much easier to compare algorithms when they have been tested on the same set of data [15], [24]. These particular datasets are the basis of much research.

All code for processing this data was developed in Matlab [51], supplemented by the Matlab Signal Processing Toolbox [25] and Matlab Wavelet Toolbox [10], Schwaighofer's SVM toolbox [52], and SNNAP, a neuron action potential simulator [53].

A. P300 Letter Recognition Task

The P300 Speller Paradigm dataset [46] was designed for development of a system that can interpret which letter the subject is concentrating on, based on his/her P300 evoked responses to flashed letters [39]. The subject's EEG was recorded while looking at a display of a 6×6 matrix of characters (see Fig. 1) [39], [47], [48]. The subject concentrated on one specific letter while rows and columns of this matrix were rapidly flashed one at a time. The EEG reflects a detectable recognition response when a flashed row or column contains the letter the subject is concentrating on. Potentially, the subject could type by concentrating on letters in the flashing matrix one at a time. Each of the six rows and six columns (the stimuli) were flashed in a random
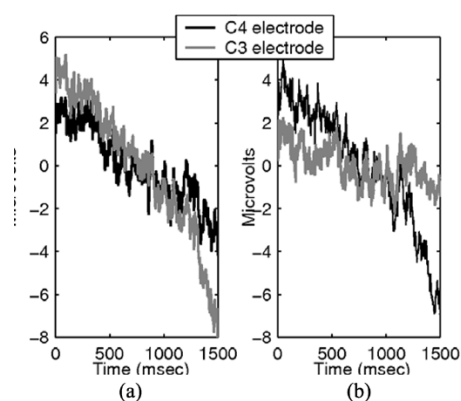

Fig. 2. Average potentials on the C3 and C4 electrodes during (a) left-hand movement and (b) right-hand movement.

order while the subject concentrated on one letter. To ensure there was enough data for determining which letter the subject was concentrating on, the twelve stimuli were each flashed, in random order, a total of fifteen times before the subject moved on to concentrate on the next letter. Therefore, for each letter, there were fifteen responses to each stimulus. The letter the subject was concentrating on can be found by determining which column and row elicited recognition responses.

To reduce the time necessary to train and test the algorithm on the data, for each letter, the 15 responses to each stimulus were averaged together; the classifier's task was to differentiate between these averaged recognition and nonrecognition responses.

B. Left- versus Right-Hand Movements

The Self-paced Typing dataset [44] consists of EEG potentials recorded from one subject, who pressed keys with either the left or right hand in self-chosen order and timing. Only data from the C3 and C4 electrodes were used; those two electrodes are over the left and right hemispherical primary motor cortex, where the largest Bereitschaftspotentials (BPs) occur (see Fig. 2). The BP, also known as the readiness potential, is a slowly increasing negative voltage [49], [50]. Its intensity is greatest on the side of the brain that controls the hand about to move. It precedes the voluntary initiation of movement and is the known distinguishing characteristic between left- and right-hand movements. The classifier's task was single-trial differentiation.

IV. CREATION OF A WAVELET-LIKE FILTER FOR EEG ANALYSIS

A. Developing the Model of Underlying Neural Activity

A model was built to approximate the simplest underlying activity in the brain, based on the hypothesis that this most basic underlying neural activity was a propagating wave of depolarization through a population of neurons. The model was constructed with consideration for how such activity would be measured by EEG electrodes.

While large ensembles of neurons bristle with voltage potentials and chemical transmissions, the majority of the recorded EEG signals are produced by potentials of one type of cell,
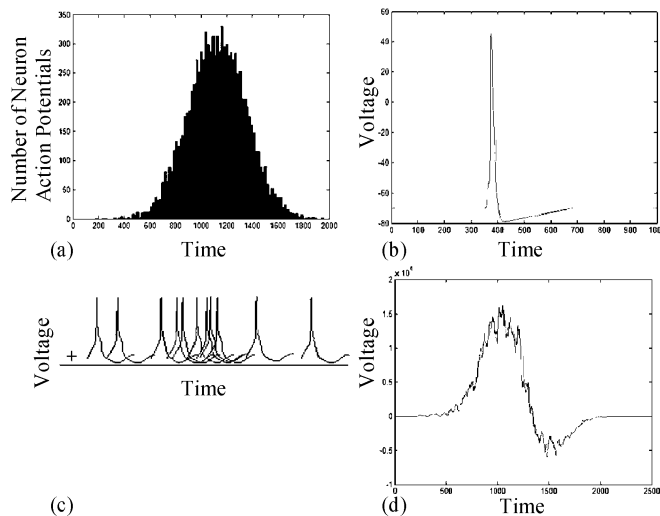
Fig. 3. (a) Histogram of 10 000 normally distributed random numbers, simulating the density of neuron firings over time. (b) The simulated neuron action potential, showing the starting resting state, depolarization, natural overshoot, and return to the resting state. (c) Time placement of neuron action potentials using random numbers. (d) Sum of neuron action potentials after shifting the resting level to zero.

the pyramidal cells [54]. These cells have apical dendrites that are locally aligned and perpendicular to the brain's surface, generating electromotive forces that the electrode can pick up [54]–[56]. Many neurons are interfaced with the pyramidal cells by synapses along these apical dendrites [54]–[56]. As the presynaptic neurons fire, postsynaptic potentials are generated within local regions of the apical dendrite, creating differences in voltage along its length and an electromotive field that an EEG electrode above can pick up [54]–[56]. These apical dendrites have the ability within themselves to generate action potentials; the action potentials sustain synaptic currents that would otherwise diminish into nonexistence [54]–[56]. Yet most of the EEG is produced by pyramidal cells' summated synaptic potentials [54]–[56]. The shorter the interval between presynaptic potentials, the more likely it is that the postsynaptic potentials overlap in time, elevate the membrane potential past the threshold, and initiate an action potential in the postsynaptic cell [54].

It was hypothesized that, from the EEG electrode's perspective, a wave of activity would first appear when the summation of postsynaptic potentials created dipoles and triggered action potentials from the apical dendrites on the outer range of the electrode's region of sensitivity. This wave would stimulate the surrounding neurons in the network, which would stimulate even more neighbors, until it grew into a peak of activity and then eventually died away, the wave of activity having passed on to another region of the brain.

To model this, a neuron action potential was obtained from a simulator, and ten thousand normally distributed random numbers were generated (see Fig. 3). The normal distribution was chosen because the actual distribution was not known, and the normal distribution is often assumed when that is the case. For each random number, a simulated neuron action potential was added to an array, its position in time determined by the random number's value. The result was a sum of ten thousand simulated

neuron action potentials; the density of firings was normally distributed over time.

Because of the possibility of using different distributions and action potential types, there is the potential for a family of wavelet-like filters that may have varying degrees of success depending in part on the accuracy of the approximation (combination of distribution and action potential shape) and the nature of the cortex or mental/physical task to which it is applied.

B. Conforming to the Requirements of a DWT Filter

The resulting array of simulated voltage samples over time begins with a few action potentials, builds up to a peak, and dies away according to a normal curve. In order to use this sum of neuron action potentials as a filter in the Matlab DWT [10], certain requirements had to be met [57]. The DWT uses two filters to decompose signals: a low-pass (the scaling function) and a high-pass (the wavelet).

The majority of the energy of the sum of neuron action potentials belonged to the lower half of the spectrum, so it was transformed into the low-pass filter, from which the high-pass filter (its quadrature mirror filter) was calculated. To transform the sum of neuron action potentials into a DWT low-pass filter, the waveshape's properties were adjusted to conform to the DWT low-pass filter spectral, sum, norm, and length requirements.

The Daubechies 4 (Db4) low-pass filter's waveshape was similar to the sum of neuron action potentials, so it was chosen as a guide during the transformation process [8]. The similarity can be observed by comparing the overall shape of the sum of neuron action potentials in Fig. 3(d) and the Db4 scaling function (in the time domain) in Fig. 4(a). They both have heavily damped single sine wave-like shapes.

Since the Db4 wavelet in the Matlab Wavelet Toolbox [10] is eight samples long, the length of the sum of neuron action potentials was also set to eight samples. The Db4 low-pass filter has no energy in the highest-frequency bin of its FFT; therefore the energy in the highest-frequency bin of the eight-sample sum of neuron action potentials was removed by taking its complex FFT and setting the energy in the highest-frequency bin of both the real and imaginary portions to zero. The inverse FFT was applied to the modified spectrum to reconstruct the waveform of the resulting low-pass filter. Its magnitude FFT spectrum is shown in Fig. 5. By adjusting the filter's amplitude and sum, it was made to conform to the Matlab DWT low-pass filter requirements of a sum of 1.0 and a norm of $\sqrt{2}/2$. The final step before using the filter in wavelet analysis was installing it as a scaling function (low-pass filter) in the Matlab Wavelet Toolbox [10], [57]. Once installed, the Wavelet Toolbox calculates the filter's quadrature mirror filter to obtain its corresponding wavelet (high-pass). This pair of wavelet- and scaling-function-like filters was named the Sum of Neuron Action Potentials or SNAP filter (see Figs. 4 and 6).

Reference [58] describes wavelet-like filters as having a constant Q, the proportion of filter bandwidth to center frequency. During wavelet analysis, the wavelet is dilated to measure energy at different scales (frequency bands), but by definition, the dilation does not change the shape of the wavelets/filters. Therefore, they all retain a constant Q.
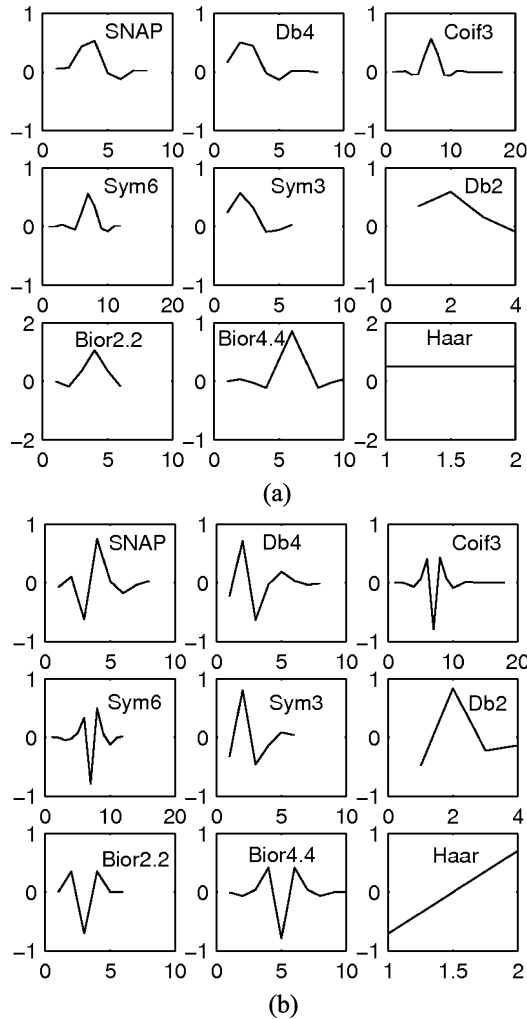
Fig. 4. Wavelets in the time domain, including the final version of the new filter labeled as "SNAP." (a) Low-pass and (b) high-pass. Note: The Haar at the scale shown has just two points.
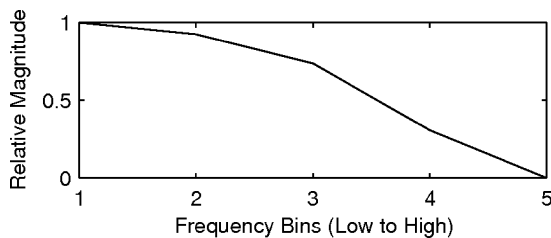

Fig. 5. SNAP filter magnitude FFT spectrum.

The SNAP filter also conforms to all the DWT filter requirements specified by the Matlab Wavelet Toolbox [10], notably, having the appropriate spectral properties to be used as a high-pass/low-pass pair for DWT or CWT decompositions, even though it does not have other complex mathematical properties that, for instance, enable standard wavelets to perfectly reconstruct the input signal. Therefore, though it is not a wavelet, the SNAP filter is a wavelet-like filter that can be used in a wavelet transform to analyze signals and produce coefficients for use as features in a classifier.
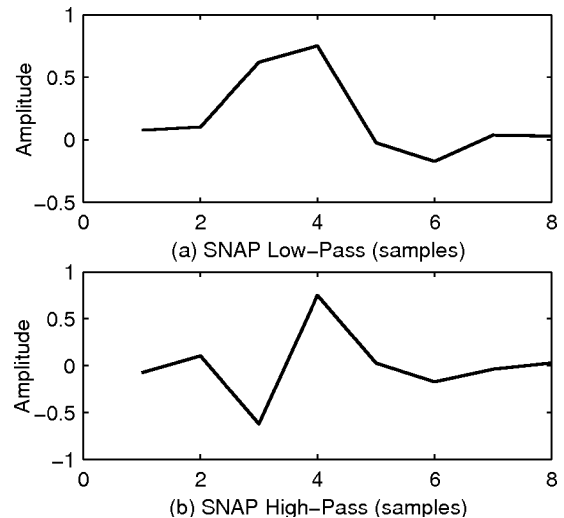

Fig. 6. SNAP filter in the time domain.

C. Relevance of SNAP Filter to Different Scales of EEG Wavelet Analysis

The SNAP filter is designed to correspond with a general oscillation of neuronal activity. Since there are several distinct neural rhythms found in EEGs that are created by subsystems of different sizes [1], this filter may perform well regardless of scale. The possibility of developing scale-specific EEG filters is addressed in Section VII.

V. WAVELET COMPARISON THROUGH PATTERN CLASSIFICATION

There are many ways of evaluating a wavelet's appropriateness for a specific type of signal, and a very direct method was used in this study. A classifier was trained and tested to differentiate between EEG signals from two different classes, such as left- and right-hand movements. The classifier identified the class of an EEG signal based upon its wavelet coefficients. By training and testing the classifier using a different wavelet to generate the coefficients each time, it is possible to compare how effective the wavelets are, relative to each other. The process comprises wavelet analysis, feature selection, and SVM classification.

A. Preprocessing and Signal Analysis for P300 Letter Recognition Data

The P300 Speller Paradigm dataset [46] is a continuous stream of EEG data from a large array of electrodes. The Cz electrode was chosen after reviewing the dataset's documentation, in which a chart displays high correlations between signal variance at the Cz electrode and whether or not the flashed row/column included the letter being concentrated on. The documentation also includes a demonstration analysis that employed data from just the Cz electrode.

Segments of data were extracted from the Cz electrode stream; each segment was 600 ms long, beginning at the same time as the onset of the flash of the stimulus. The segments were separated into two groups, recognition and nonrecognition responses, based on the labels provided with the data. As explained in DATA AND SUPPORTING SOFTWARE, each set of
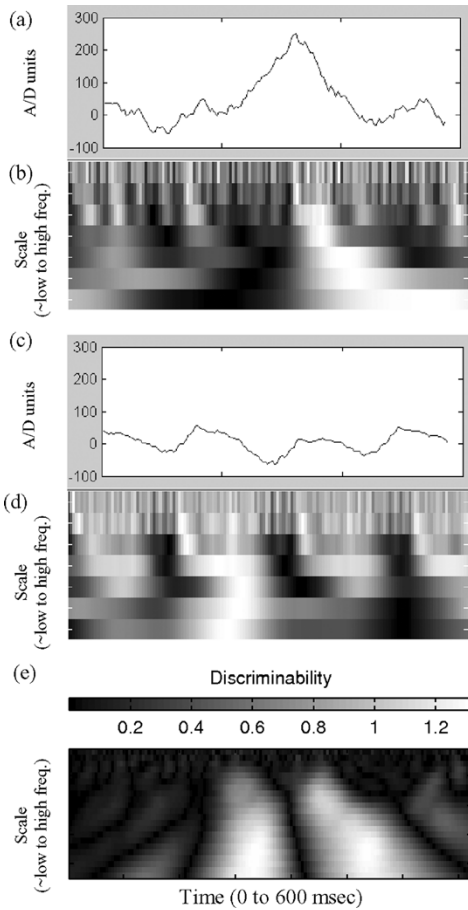
Fig. 7. P300 recognition and nonrecognition responses: (a) Average of all recognition responses. (b) CWT analysis of the averaged signal above it. (c) Average of all nonrecognition responses. (d) CWT analysis of the averaged signal above it. (e) Discriminability map produced from the dataset. Note: Discriminability is calculated from the wavelet coefficients of the individual trials, not an averaged signal.
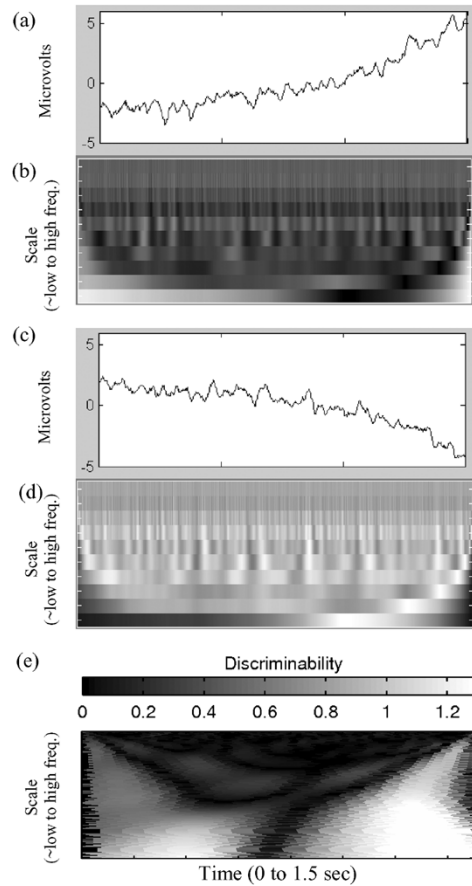
Fig. 8. Left- and right-hand movement data: (a) Average of all right-hand movements. (b) CWT analysis of the averaged signal above it. (c) Average of all left-hand movements. (d) CWT analysis of the averaged signal above it. (e) Discriminability map produced from the dataset. Note: Discriminability is calculated from the wavelet coefficients of the individual trials, not an averaged signal.

fifteen segments (responses) was averaged into one segment to reduce the time necessary to train and test the classifier. These averaged segments were then analyzed with a DWT (at the maximum possible decomposition of seven scales), resulting in a set of wavelet coefficients to be used as inputs for the classifier.

B. Preprocessing and Signal Analysis for Left- versus Right-Hand Movements

The Self-paced Typing dataset [44] is a collection of segments, each 1500 ms long and ending 120 ms before a keypress. For each segment, the difference between the C3 and C4 electrodes was calculated. This is referred to as the bipolar derivation (the first spatial derivative) [29]. It emphasizes the lateral dissimilarity between EEG signals from the left and right motor cortices, over which the C3 and C4 electrodes are positioned [29]. The difference signals were then analyzed with a DWT (at the maximum possible decomposition of ten scales), resulting in a set of wavelet coefficients to be used as inputs for the classifier. The wavelet analysis of these difference signals did yield coefficients with greater discriminability, defined in (1), than the coefficients from wavelet analysis of the C3 and C4 electrodes individually.

C. Feature Selection

A subset of the wavelet coefficients (the coefficients produced by wavelet analysis) was selected as input (features) for the classifier. Computational limitations affecting SVM training constrained the maximum number of coefficients used to approximately 100 (5% of the total number produced by wavelet analysis of the Self-paced Typing dataset and 50% of the total produced by wavelet analysis of the P300 dataset). The coefficients were rank-ordered by discriminability, and the classifier was trained on the n coefficients with highest discriminability, where n was determined by the grid search method described in Section V-D. There was no threshold for the discriminability of selected coefficients.

In order to visualize the discriminability values, signals were analyzed with the continuous wavelet transform (CWT). In contrast to the DWT, the CWT produces the same number of coefficients for each scale. The resulting coefficients can be arranged in a matrix and plotted as an image in which the coefficients' discriminability values are mapped to a color map. It produces a "discriminability map" across frequency and time. The discriminability maps in Figs. 7(e) and 8(e) show the regions of frequency and time where coefficients with high discriminability were located. Due to the clarity of CWT discriminability maps

and the ease of their creation, DWT discriminability maps were not made. Yet it is possible to create them.

Note: During cross-validation, as shown in the pseudo-code in Section V-D, the features' discriminabilities were calculated from the cross-validation training set data only. This helped prevent overfitting.

D. Pattern Classification

Once the features were ranked by discriminability, the classifier, a polynomial SVM, was trained and tested using tenfold cross-validation on the data. To find the best parameter values, a grid search, which is a standard method, was used. Cross-validation [13] was performed for every combination of values within set ranges for the following parameters: the polynomial SVM kernel's degree, which sets the level of high-dimensional embedding; the SVM cost parameter, which controls the classifier's specificity to the training data; and the number of features used. For each combination of parameter values, 90% of the signals were randomly chosen to serve as the training set. The SVM was trained on the training set and tested on the remaining 10% of the signals (those not included in the training set). This was repeated ten times for each combination of parameter values; the average error rate across all ten test sets was recorded. Finally, the combination of parameter values with which the SVM yielded the lowest cross-validation error rate was recorded, along with the error rate itself.

The following pseudo-code outlines this SVM parameter grid search.

```
• For degree = 1 through 5
  • For cost setting A through E
    • For number of coefficients = 1 through
      100, skipping by 10
      • For 10 cross-validation (CV) sets
        • Create CV training and test set
        • Calculate discriminability values
          from CV training set
        • Select features
        • Train SVM on CV training set
        • Forward CV test set through trained
          SVM
        • Calculate accuracy on CV test set
      • End
      • Average accuracy across all 10 CV
        test sets
    • End
  • End
• End
```

Note that the SVM degree parameter's range includes the degree of one, at which the SVM is a linear classifier. Therefore, the grid search method has the opportunity to select, based on cross-validation error, either a linear decision boundary or the most appropriate degree of nonlinearity.

With the goal of reducing the number of settings tried for each parameter and thus reducing computation time, a series of plots was made to observe whether some parameter values were clearly superior to others. As demonstrated in Fig. 9, there is no
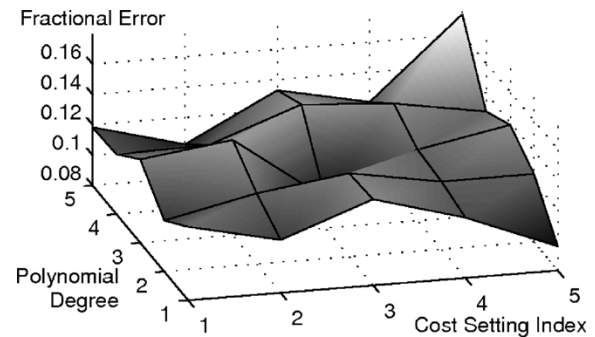

Fig. 9. Cross-validation error over SVM polynomial degree and cost settings, using the SNAP wavelet and 21 coefficients on the self-paced typing dataset.

apparent trend toward better or worse cross-validation error for any values of degree or cost. There were also no apparent trends for the number of features used.

The process of calculating cross-validation error for each combination of parameter settings was repeated for each wavelet. Total computational time per wavelet was approximately eight hours on a Dell Inspiron 8200. The majority of this time was consumed by the repeated training of the SVM. It may be reasonable to assume that, after SVM training, which would be done off-line anyway, this process could be applied in a real-time system. The DWT, an algorithm with an efficiency of $O(n)$, would be applied to the EEG data, and the fraction of its output coefficients that was selected during training as discriminable features would then be inputs to the trained SVM [59]. The SVM would determine each new signal's class by its relationship to the SVM hyperplane, a relatively fast process [32].

E. The Continuous Wavelet Transform (CWT) Experiment

Since the SNAP filter's waveshape is most similar to that of the Db4, and the Db4 was used as a guide during the SNAP filter's transformation into a DWT low-pass, an additional experiment was conducted to test whether the process of creating a filter based on action potentials had not, in effect, re-created the Db4 wavelet or a variant of it.

The left- and right-hand movement data were preprocessed as before, but CWT, not DWT, analysis was conducted. The discriminability values for all the CWT coefficients were sorted by value and plotted (see Section V-C). This was performed once using the SNAP filter and once using the Db4. (The CWT uses the high-pass counterpart of the filter/wavelet by definition.) The CWT was chosen because it measures the correspondences between the wavelet/filter and the signal at many more points in time and scale than the DWT. This increases the probability that the wavelet/filter will be compared with the signal at the time and scale at which it matches the signal most closely. If the discriminabilities of the Db4's and SNAP's CWT coefficients are very similar, then this would indicate that they themselves are very similar.

F. NIPS 2001 Brain-Computer Interface Workshop: Post Workshop Data Competition

The EEG-classification algorithm was used to compare the tested filter/wavelets' relative effectiveness. To validate its use

TABLE I
P300 LETTER RECOGNITION TASK

Wavelet Utilized	Cross-Validation Error
SNAP	6.92%
Coiflet 3 (Coif3)	7.14%
Biorthogonal 2.2 (Bior2.2)	8.55%
Biorthogonal 4.4 (Bior4.4)	9.83%
Symlet 3 (Sym3)	9.83%
Daubechies 2 (Db2)	11.77%
Daubechies 4 (Db4)	11.99%
Haar*	50.00%

*The Haar's associated error rate is not included in the range of errors referenced in the abstract because it is such an extreme outlier.
Note: The set of wavelets tested on this dataset was expanded and altered after testing was first performed on the left- and right-hand data. Specifically, the Daubechies 2 (Db2) and Haar wavelets were added, and the Symlet 3 (Sym3) wavelet replaced the Symlet 6 wavelet.

TABLE II
LEFT- VERSUS RIGHT-HAND MOVEMENT TASK

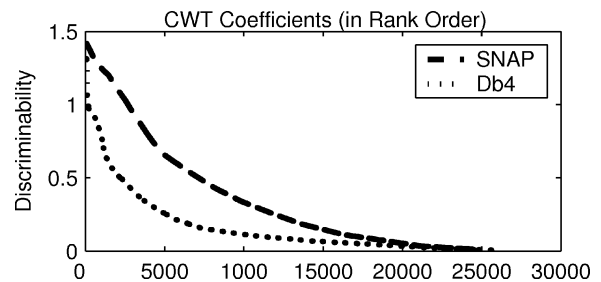| Wavelet Utilized	Cross-Validation Error
Coiflet 3 (Coif3)	9.54%
SNAP	10.03%
Biorthogonal 4.4 (Bior4.4)	10.56%
Symlet 6 (Sym6)	10.89%
Biorthogonal 2.2 (Bior2.2)	11.34%
Daubechies 4 (Db4)	12.00%

Fig. 10. Discriminability values of the Db4 CWT coefficients are plotted with a dotted line. The discriminability values of the SNAP CWT coefficients are plotted with a dashed line.

as a platform for wavelet comparison, its own efficacy was compared to that of the six entries in the 2001 NIPS EEG dataset competition [11]. The algorithm's filter/wavelet was set to the SNAP, and the SVM-related parameters (degree, cost, and number of features) were set to the values recorded when its lowest cross-validation error was attained. It was trained on all the data (which had previously been split up into training and test sets for cross-validation). The trained algorithm was then tested on the official competition test set, data the algorithm had never previously been trained or tested on.

VI. RESULTS

A. P300 Letter Recognition

The level of performance of each wavelet on the P300 letter recognition task is summarized in Table I. The wavelets tested are in order from best algorithm performance to worst.

The algorithm's error rate, as a function of the wavelet/filter used, ranged from 6.92% to 11.99%, almost twofold, when differentiating between P300 recognition and nonrecognition responses to flashed letters. Using the SNAP filter, the algorithm attained 6.92% error, better than that for any of the six standard wavelets tested (7.14% to 11.99% error, excluding the seventh wavelet, the Haar, at 50% error).

B. Left- versus Right-Hand Movement

The level of performance of each wavelet on the left- versus right-hand movement task is summarized in Table II. The SNAP filter, with which the algorithm attained 10.03% error, is ranked second out of six standard wavelets, which ranged from 9.54% to 12.00% error. The range of accuracy for this task is not as great as for the P300 letter recognition task, but it is still possible to see the relative success of each wavelet. Across both datasets a general trend may be observed; the SNAP filter or Coif3 wavelet performed best, followed by the Biorthogonal, Symlet, and Daubechies wavelets, in that order.

C. A CWT Coefficient Discriminability Value Comparison of the Left- versus Right-Hand Movement Data

The discriminability values for the CWT coefficients were sorted by value and plotted in Fig. 10, which shows that the SNAP coefficients were consistently more discriminable than

the Db4 coefficients. The slope of the SNAP sorted-discriminabilities graph was gentler than the Db4's since the trend of greater discriminability continued across virtually all SNAP coefficients. The comparison of discriminability values of the SNAP and Db4 CWT coefficients is another confirmation that they are indeed distinct filters, despite the fact that the Db4 was used as a guide during the SNAP filter's transformation.

D. NIPS 2001 Brain-Computer Interface Workshop Post Workshop Data Competition Results [11], [37]

When ranked by performance on the competition test set for the left- versus right-hand movement dataset, the EEG-classification algorithm described in this paper (which, for this test, was using the SNAP filter) yielded an error rate of 7% (see Table III). It was not a formal entry in the competition, but its performance was only 2 percentage points behind the entries from the NASA Ames Research Center and Tsinghua University and 3 percentage points behind the competition winner. It was also 6 percentage points more accurate than the next most accurate entry.

VII. DISCUSSION

The wavelet/wavelet-like filter utilized in the EEG-classification algorithm had a marked effect on overall performance on the datasets tested. The algorithm's legitimacy as a platform for wavelet comparison was validated by its internationally competitive performance on the NIPS competition test set [11]. Of the nine that were tested, the SNAP filter and the Coif3 wavelet performed best and appear to be most appropriate for EEG analysis. The results are encouraging because the SNAP and Coif3 performed the best on two datasets (two different tasks and two different subjects). One task is a voluntary movement and the other involves the recognition response to a flashed column/row of letters.

These two different tasks have substantially different types of distinguishing characteristics that the SNAP and Coif3 were

TABLE III
NIPS 2001 BCI POST WORKSHOP DATA COMPETITION
RESULTS ON THE COMPETITION TEST SET

Algorithm	Error	University/Center	Entrant
Recurrent Neural Network Optimized by Dynamic Noise Annealing	4%	Ecole Polytechnique Federale de Lausanne (EPFL)	P.-E. Sottas
Common Spatial Subspace Decomposition (CSSD) and Multiple Electrode Activity Subtraction (MEAS) Linear Discriminant Analysis	5%	Tsinghua University	S. Gao
Slow Potential Shift nu-Support Vector Classifier with CV Criterion Parameter Tuning	5%	NASA Ames Research Center	R. Rosipal, L. Trejo, K. Wheeler
SVM trained on coefficients of wavelet analysis with SNAP	**7%**		**Author**
Decision Tree and Neural Network Classifier	13%	University of Sydney	M. Dam, K. Tosevski, K. Belista, M. El-Ali
AutoRegressive Model with eXogenous input (ARX) Linear Discriminant Analysis	22%	National University of Ireland, Dublin	D. Burke, S. Kelly, P. de Chazal, R. Reilly
Not specified	46%	Israel Institute of Technology	E. Yom-Tov

used to encode into wavelet coefficients. The most obvious distinguishing characteristic for the left- versus right-hand movement task is the BP, a steadily increasing difference in voltage (nearly DC) across the two sides of the motor cortex. When the P300 letter recognition task signals are averaged over many trials, the prominent distinguishing characteristic is a positive spike in voltage 300 ms after the flash.

Yet the discriminability maps [Figs. 7(e) and 8(e)] indicate that there may be more complex phenomena that could not be observed by other means of visualization. This is most evident for the P300 data: unlike the single positive spike visible when P300 data are averaged together, the discriminability map shows two very distinct regions of discriminability from 250 to 340 ms and from 430 to 520 ms after the flash. The discriminability map for the left- versus right-hand movement data also shows regions of discriminability other than that of the BP. The map shows increasing discriminability in the lowest frequencies beginning 0.75 s before the onset of finger movement, which agrees with the definition of the BP, but it also shows regions of discriminability between 1.5 and 1.3 s before the onset of finger movement. These additional phenomena, though visible in the discriminability maps, may be too subtle to pick out in raw data. They may also be averaged out of the averaged signals, even though other distinguishing characteristics are more visible after averaging (see Figs. 1(b) and 2). They may be still visible in the discriminability maps because discriminability is calculated from the wavelet coefficients of all the individual trials, not averaged signals.

The SNAP and Coif3's superior performances on both EEG datasets may indicate that they are particularly suited to encoding, not specific distinguishing characteristics like P300 recognition responses or BPs, but the underlying activity of which they are composed. Wavelets and wavelet-like filters that are well suited to encoding the underlying activity of EEGs are of value because such filters would then, presumably, be effective at encoding any distinguishing characteristic made up of such underlying activity. This general applicability may be much more desirable than having a particular filter or wavelet that is only appropriate for encoding, for instance, BPs, especially since the discriminability maps indicate that there are more distinguishing characteristics than those that are already well-known. Encoding all the distinguishing characteristics is critical since they are what the classifier relies on to determine to which class an EEG signal belongs. EEG data may generally be made up of underlying activity that the SNAP and Coif3 may be particularly well suited to encode.

The one complicating factor is that the two datasets were preprocessed such that the voluntary movement data were actually the difference between two electrodes, and the P300 data were actually the average of several responses to flashes. Therefore, the data that were analyzed still included the underlying activity, but not in its original form. Despite this, the SNAP and Coif3's performances were consistently superior to those of the other tested wavelets. The SNAP and Coif3's performances were also too close for meaningful comparison to each other.

Because of the SNAP filter's effectiveness on both datasets, the data does not disprove the hypothesis that, due to its neuronal-model-derived shape, the SNAP filter may have a connection to the underlying neural activity that produces the potentials associated with both voluntary motor tasks and visual tasks. Furthermore, this model of neural activity was not optimized in any way to maximize the SNAP filter's performance. One could perform a grid search (similar to finding the best settings for the SVM parameters in this study) across a range of neuron action potential types and distributions to optimize the model and produce more effective versions of this filter. There is a possibility of finding versions that are particularly well suited for certain cortices, time scales, and/or mental tasks. The model of neural activity used in this study was not designed for any specific size or type of network of neurons.

The results are even more intriguing when one compares the original intention and the end result. The original intention was to create a filter that could be used in place of a wavelet in wavelet analysis and that matched the underlying activity of the brain. For DWT analysis, two filters are needed, a high-pass (wavelet) and a low-pass (scaling function). And though both filters are integral to the analysis, most of the resulting DWT coefficients are calculated directly using the high-pass (wavelet) and only indirectly from the low-pass (scaling function). Therefore, one would ideally make the high-pass filter into the filter that matched the underlying neural activity. Yet the sum of neuron action potentials could only serve as the low-pass filter (scaling function) due to its spectral properties. Therefore, when using the SNAP filter pair, the majority of the coefficients are produced not by the low-pass filter derived from the sum of neuron action potentials, but by its quadrature mirror filter, the SNAP high-pass.

Despite this reversal, the data indicated that the SNAP filter pair was one of the most appropriate of those tested for EEG analysis. One possible explanation is that when the signal is decomposed into the approximations and details, the sums of neuron action potentials are preserved in the approximations and, since the detail filter (wavelet) is a high-pass filter, the on- and offsets of the sums (looking at the summations essentially as pulses) are encoded, though less efficiently, in two or more detail coefficients. Since wavelet analysis is inherently a multiresolution analysis, the sums of neuron action potentials occurring on different time scales could be isolated at different scales of the wavelet analysis.

Potentially, the SNAP filter could be the first of an entire family. It may be possible to construct filters from a variety of neuronal models. There may even be models that could serve as the high-pass filter for the creation of a high-pass matched filter pair, in which the high-pass is the filter that matches the underlying EEG components.

Interestingly, the wavelet packet decomposition algorithm uses both the high- and low-pass filters to directly calculate output coefficients [57]. Therefore, regardless of whether the sum of neuron action potentials is transformed into the scaling function or wavelet, the wavelet packet algorithm would be able to utilize the direct correspondence between the signals and the EEG-component-matching filter of the pair. Employing the SNAP filter in wavelet packet analysis would then make the frequency-inverted relationship no longer an issue. Even if a neuronal model produces a waveform best suited for a bandpass filter but which cannot be utilized in the DWT, like the Mexican Hat wavelet [8], it could potentially be employed in the continuous wavelet transform.

Additionally, even though wavelets are constrained to having the same shape for every scale, variations of the neuronal model may be particularly applicable to certain scales of EEG decomposition and analysis. One could construct a filter bank in which each scale is analyzed using a filter specifically designed for that time scale. The successful creation of new filters and filter banks may be a way to further probe the neuronal mechanisms behind the EEG, explore data-specific wavelet-like filters, and yield better EEG analyses for pattern classification in applications like BCIs.

REFERENCES

[1] V. J. Samar, K. P. Swartz, and M. R. Raghuveer, "Multiresolution analysis of event-related potentials by wavelet decomposition," *Brain Cogn.*, vol. 27, pp. 398–438, 1995.

[2] V. J. Samar, "Wavelet analysis of neuroelectric waveforms," *Brain Lang.*, vol. 66, no. 1, pp. 1–6, Jan. 1999.

[3] S. Kadambe and G. Boudreaux-Bartels, "A comparison of wavelet functions for pitch detection of speech signals," in *Proc. ICASSP*, 1991, pp. 449–452.

[4] B. B. Hubbard, *The World According to Wavelets: The Story of a Mathematical Technique in the Making*. Wellesley, MA: Peters, 1996, p. 196.

[5] G. Lemaur, "On the Choice of the Wavelet Basis Function for Image Processing," Ph.D., Université de Mons-Hainaut, Belgium, 2003.

[6] Y. F. Low and R. Besar, "Optimal wavelet filters for medical image compression," *Int. J. Wavelets, Multiresolution, Inform. Process.*, vol. 1, no. 2, pp. 179–197, 2003.

[7] V. J. Samar, A. Bopardikar, R. Raghuveer, and K. Swartz, "Wavelet analysis of neuroelectric waveforms: A conceptual tutorial," *Brain Lang.*, vol. 66, no. 1, pp. 7–60, Jan. 1999.

[8] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[9] S. Kadambe, "On the choice of a wavelet, and the energy and the phase distributions of the wavelet transform," in *Proc. IEEE-SP Int. Symp. Time-Frequency and Time-Scale Analysis*, 1992, pp. 339–382.

[10] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi. (2002) Matlab Wavelet Toolbox. MathWorks. [Online]. Available: http://www.mathworks.com

[11] P. Sajda, A. Gerson, K.-R. Müller, B. Blankertz, and L. Parra, "A data analysis competition to evaluate machine learning algorithms for use in brain-computer interfaces," *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 11, no. 2, pp. 184–185, Feb. 2003.

[12] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2001, pp. 48–49.

[13] N. Cristianini and B. Schölkopf, "Support vector machines and Kernel methods: The new generation of learning machines," *AI Mag.*, vol. 23, no. 3, pp. 31–41, 483, 484, 2002.

[14] F. Friedrichs and C. Igel, "Evolutionary tuning of multiple SVM parameters," *Neurocomput.*, to be published.

[15] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: A review of the first international meeting," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, Jun. 2000.

[16] C.-W. Hsu, C.-C. Chang, and C.-J. Lin. (2003) A Practical Guide to Support Vector Classification. National Taiwan Univ., Dept. Comput. Sci. Inf. Eng.. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

[17] J. O. Chapa and M. R. Raghuveer, "Optimal matched wavelet construction and its application to image pattern recognition," in *Wavelet Applications: II—Proc. Soc. Photo Instrumentation Engineering*, vol. 2491, 1995, pp. 518–529.

[18] M. R. Raghuveer and J. O. Chapa, "Object detection through matched wavelet transforms," in *Wavelet applications: III—Proceedings of the Society for Photo Instrumentation Engineering*, vol. 2762, 1996, pp. 45–50.

[19] V. J. Samar, H. Begleiter, J. O. Chapa, M. R. Raghuveer, M. Orlando, and D. Chorlian, "Matched Meyer neural wavelets for clinical and experimental analysis of auditory and visual evoked potentials," in *Proc. EU-SIPCO-96, Signal Processing: VIII. Theories and Applications*, 1996, pp. 496–507.

[20] G. C. Donovan, J. S. Geronimo, D. P. Hardin, and P. R. Massopust, "Construction of orthogonal wavelets using fractal interpolation functions," in *Proc. Soc. Industrial and Applied Mathematics Journal on Mathematical Analysis*, vol. 27, pp. 1158–1192.

[21] C. J. A. Tollig and A. J. Hoffman, "Wavelet neural network for classification of transient signals," in *Proc. 1997 South African Symp. Communications and Signal Processing (COMSIG '97)*, 1997, pp. 161–166.

[22] S. G. Mallat and Z. Zhong, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[23] R. R. Coifman and M. V. Wickerhauser, "Wavelets, adapted waveforms and de-noising," *Electroencephalogr. Clin. Neurophysiol. Suppl.*, vol. 45, pp. 57–78, 1996.

[24] T. M. Vaughan, "Guest editorial brain-computer interface technology: A review of the second international meeting," *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 11, no. 2, pp. 94–109, Jun. 2003.

[25] Matlab Signal Processing Toolbox. MathWorks. [Online]. Available: http://www.mathworks.com

[26] F. H. L. da Silva and N. J. I. Mars, "Parametric methods in EEG analysis," in *Handbook of Electroencephalography and Clinical Neurophysiology*, 4th ed, A. S. Gevins and A. Remond, Eds. Amsterdam, The Netherlands: Elsevier, 1987, vol. 1, Methods of analysis of brain electrical and magnetic signals, pp. 243–260.

[27] J. Parday, S. Roberts, and L. Tarassenko, "A review of parametric techniques for EEG analysis," *Med. Eng. Phys.*, vol. 18, pp. 2–11, 1996.

[28] F. L. da Silva, "Event-related potentials: Methodology and quantification," in *Electroencephalography: Basic Principles, Clinical Applications and Related Fields*, 4th ed, E. Niedermeyer and F. Lopez da Silva, Eds. Baltimore, MD: Williams and Wilkins, 1999, pp. 947–957.

[29] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, pp. 767–791, 2002.

[30] F. Beverina, G. Palmas, S. Silvoni, F. Piccione, and S. Giove, "User adaptive BCIs: SSVEP and P300 based interfaces," *Psych. J.*, vol. 1, no. 4, pp. 331–354, 2003.

[31] M. Pregenzer, G. Pfurtscheller, and D. Flotzinger, "Automated feature selection with a distinction sensitive learning vector quantizier," *Neurocomp.*, vol. 11, pp. 19–29, 1996.

[32] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification," *IEEE Trans. Neural Sys. Rehab. Eng.*, vol. 11, no. 2, pp. 141–144, Jun. 2003.

[33] B. Blankertz, G. Curio, and K.-R. Müller, "Classifying single trial EEG: Toward brain computer interfacing," in *Advances in Neural Information Processing Systems (NIPS 01)*, T. G. Diettrich, S. Becker, and Z. Ghahramani, Eds. Cambridge, MA: MIT Press, 2002, vol. 14.

[34] D. Burke, S. Kelly, P. de Chazal, R. B. Reilly, and C. Finucane, "A parametric feature extraction and classification strategy for brain–computer interfacing," *IEEE Trans. Neural Syst. Rehab. Eng.*, to be published.

[35] Laboratory for Intelligent Imaging and Neural Computing, Dept. Biomed. Eng., Columbia Univ., New York. (Accessed: 2002, Sep.) NIPS 2001 Brain Computer Interface Workshop Post Workshop Data Competition. [Online]. Available: http://newton.bme.columbia.edu/competition.htm

[36] B. Blankertz, K.-R. Müller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schloegl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schroeder, and N. Birbaumer, "The BCI competition 2003: Progress and perspectives in detection and discrimination of EEG single trials," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1044–1051, Jun. 2004.

[37] Laboratory for Intelligent Imaging and Neural Computing, Dept. Biomed. Eng., Columbia Univ. , New York. NIPS 2001 Brain Computer Interface Workshop Post Workshop Data Competition Results. [Online]. Available: http://newton.bme.columbia.edu/competitionresults.htm

[38] B. Blankertz, T. M. Vaughan, G. Schalk, J. R. Wolpaw, G. Curio, K.-R. Müller, A. Schlögl, C. Neuper, G. Müller, B. Graimann, G. Pfurtscheller, T. Hinterberger, M. Schröder, and N. Birbaumer. BCI Competition 2003 Final Results. Fraunhofer FIRST IDA, Berlin, Germany. [Online]. Available: http://ida.first.fraunhofer.de/projects/bci/competition/results/index.html

[39] J. R. Wolpaw. Documentation: 2nd Wadsworth BCI Dataset (P300 Evoked Potentials). [Online]. Available: http://ida.first.fraunhofer.de/projects/bci/competition/albany_desc/albany_desc_ii.html

[40] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, "An EEG-based brain–computer interface for cursor control," *Electroenceph. Clin. Neurophysiol.*, vol. 78, pp. 252–259, 1991.

[41] J. R. Wolpaw, D. J. McFarland, and T. M. Vaughan, "Brain–computer interface research at the Wadsworth Center," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 222–225, 2000.

[42] D. J. McFarland, A. T. Lefkowicz, and J. R. Wolpaw, "Design and operation of an EEG-based brain–computer interface (BCI) with digital signal processing technology," *Behav. Res. Meth. Instrum. Comput.*, vol. 29, pp. 337–345, 1997.

[43] W. D. Penny, S. J. Roberts, E. A. Curran, and M. J. Stokes, "EEG-based communication: A pattern recognition approach," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 214–215, Jun. 2000.

[44] B. Blankertz, G. Curio, and K.-R. Müller. EEG Self-Paced Key Typing Dataset. NIPS 2001 Brain Computer Interface Workshop Post Workshop Data Competition. [Online]. Available: http://ida.first.fraunhofer.de/projects/bci/competition/results/index.html

[45] ——, "Classifying single trial EEG: Toward brain computer interfacing," in *Advances in Neural Information Processing Systems (NIPS 01)*, T. G. Diettrich, S. Becker, and Z. Ghahramani, Eds. Cambridge, MA: MIT Press, 2002, vol. 14.

[46] J. R. Wolpaw. P300 Speller Paradigm. BCI Competition 2003. [Online]. Available: http://ida.first.fraunhofer.de/projects/bci/competition/

[47] L. A. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroenceph. Clin. Neurophysiol.*, vol. 70, no. 6, pp. 510–523, 1988.

[48] E. Donchin, K. M. Spencer, and R. Wijensinghe, "The mental prosthesis: Assessing the speed of a P300-based brain-computer interface," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 174–179, Jun. 2000.

[49] H. H. Kornhuber and L. Deecke, "Hirnpotentialänderungen bei willkürbewegungen und passiven bewegungen des menschen: Bereitschaftspotential und reafferente potentiale," *Pflügers Arch. Ges. Physiol.*, vol. 284, pp. 1–17, 1965.

[50] L. Gliden, H. G. Vaughan Jr., and L. D. Costa, "Summated human EEG potentials with voluntary movement," *Electroenceph. Clin. Neurophysiol.*, vol. 20, pp. 433–438, 1966.

[51] *Matlab Student Version.* [Software], MathWorks, Natick, MA, 2002.

[52] A. Schwaighofer. SVM Toolbox for Matlab. [Online]. Available: http://www.cis.tugraz.at/igi/aschwaig/software.html

[53] SNNAP: Simulator for Neural Networks and Action Potentials. Univ. Texas Health Sci. Center Houston, Dept. Neurobiol. Anatomy. [Online]. Available: http://snnap.uth.tmc.edu

[54] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, *Principles of Neural Science*, 3rd ed. New York: Elsevier, 1991, pp. 27, 98, 778–779, 784, 785.

[55] F. H. Duffy, V. G. Iyer, and W. W. Surwillo, *Clinical Electroencephalography and Topographic Brain Mapping*. New York: Springer-Verlag, 1989, pp. 73, 74, 84–86.

[56] M. F. Bear, B. W. Connors, and M. A. Paradiso, *Neuroscience: Exploring the Brain*, 2nd ed. Baltimore, MD: Lippincott, Williams, & Wilkins, 2001, pp. 119–123.

[57] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, *Wavelet Toolbox User's Guide*. Natick, MA: MathWorks, 1997, pp. 5-2–5-7, 6-21.

[58] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995, p. 6.

[59] H. Guo and C. S. Burrus, "Fast approximate Fourier transform via wavelets transform," presented at the SPIE Conf. Mathematical Imaging, Denver, CO, 1996.

Elena L. Glassman (S'04) attends the Massachusetts Institute of Technology, Cambridge, in the class of 2008. Her intended major is electrical engineering and computer science.

She has done much of her work outside of the lab setting, but conducted research at the Princeton University Cognitive Electrophysiology Lab, Princeton, NJ, in the spring and summer of 2004. She is now an undergraduate researcher in the Computer Science and Artificial Intelligence Lab (CSAIL) Networks and Mobile Systems Group. Previous publications include abstracts in the 52nd, 53rd, 54th, and 55th Intel International Science and Engineering Fair *Abstracts of the Finalists* (Washington, DC: Science Service, 2001–2004). She is currently interested in signal analysis, feature selection, pattern classification, artificial intelligence, and the brain.

Ms. Glassman was awarded the IEEE President's Scholarship for her work presented at the 2004 Intel International Science and Engineering Fair and in 2003 won the highest honor given at the Intel International Science and Engineering Fair, the Intel Foundation Young Scientist Award. She is also a member of the American Association for Artificial Intelligence.