

Robust real-time identification of tongue movement commands from interferences

Khondaker A. Mamun^a, Michael Mace^b, Lalit Gupta^c, Carl A. Verschuur^a, Mark E. Lutman^a, Maria Stokes^d, Ravi Vaidyanathan^{e,f,*}, Shouyan Wang^{a,**}

^a Hearing and Balance Centre, Institute of Sound and Vibration Research, University of Southampton, UK

^b Department of Mechanical Engineering, Faculty of Engineering, University of Bristol, UK

^c Department of Electrical and Computer Engineering, Southern Illinois University, IL, USA

^d Faculty of Health Sciences, University of Southampton, UK

^e Department of Mechanical Engineering, Imperial College London, UK

^f Department of Systems Engineering, US Naval Postgraduate School, USA

ARTICLE INFO

Available online 7 November 2011

Keywords:

Tongue-movement ear pressure signals
Wavelet packet transform
Bayesian classifier
Human machine interface

ABSTRACT

This study aimed to improve the accuracy and robustness of a real-time assistive human machine interface system by classifying between the controlled movements related tongue-movement ear pressure (TMEP) signals and the interfering signals. The controlled movement TMEP signals were collected during left, right, up, down, flicking and pushing tongue motions. The TMEP signals were processed and classified using detection, segmentation, feature extraction and classification. The segmented signals were decomposed into the time-scale domain using a wavelet packet transform. The variance of the wavelet packet coefficients and its ratio between low-to-high scales were defined as features and the intended tongue movement commands and interfering signals were classified using both a Bayesian and support vector machine (SVM) classifiers for comparison. The average classification accuracy for discriminating between the controlled movements and the interfering signals achieved 97.8% (Bayesian) and 98.5% (SVM). The classifiers were robust remaining at a similar performance level when generalised interferences from all subjects were used. It was shown that the Bayesian classifier performed better than the SVM in a real-time environment. The approach of combining the Bayesian classifier and the wavelet packet transform provides a robust and efficient method for a real-time assistive human machine interface based on tongue-movement ear pressure signals.

© 2011 Elsevier B.V. Open access under [CC BY license](http://creativecommons.org/licenses/by/3.0/).

1. Introduction

A wide range of research has been conducted to develop various

HMI systems, current products have not yet fully addressed patient-specific requirements and better interfaces between the patient and peripheral devices are still greatly needed [1,4,5,9]. Recently a novel

and similar papers at core.ac.uk

patients [1–6]. Specific hands free communication and control devices are essential for an individual who has limited mobility or severe motor dysfunctions, for example due to spinal cord injury, congenital limb deformities or arthritis [5,7,8]. In spite of significant progress made in the development of techniques and devices for

tongue-movement ear pressure signals [4,10,11]. Users

express their intention by making impulsive actions of the tongue, which create unique acoustic pressure signals within the ear canal. These pressure signals can be recorded easily using a microphone earpiece positioned non-invasively within the ear canal [4]. The advantage of utilising the tongue is that it has an inherent capability for fine motor control, involving multiple degrees of freedom, as it has evolved to perform sophisticated motions during speech and mastication [1,4]. The system also has the additional benefits of being simple, cheap and non-invasive. Individuals with limited control of their limbs are able to use these prescribed tongue movements to communicate with computers and control assistive devices through the sensing of bio-acoustic pressure signals.

Previously, different types of tongue movements recorded from healthy subjects relating to the controlled (intended) actions

* Corresponding author at: Department of Mechanical Engineering, Imperial College London, London, SW7 2AZ, UK.

** Corresponding author at: Hearing and Balance Centre, Institute of Sound and Vibration Research, University Road, Southampton, SO17 1BJ, UK.
Tel.: +44 23 8059 2310; fax: +44 23 8059 3190.

E-mail addresses: r.vaidyanathan@imperial.ac.uk (R. Vaidyanathan), sy.wang@soton.ac.uk (S. Wang).

URL: <http://www.swanglab.com> (S. Wang).

have been classified using a decision fusion algorithm [4]. The performance of the classifier reached an average of 97% correct accuracy using time domain features and large training sets. The performance of this classifier was shown to be better than three other strategies using time domain information, namely, the matched filter (86%), the parametric autoregressive (AR) Gaussian classifier (85.98%) and the nonlinear alignment classifier (96.27%) [4]. Moreover, to improve the classification performance a single channel independent component analysis (ICA) was used to isolate the critical components of TMEP signals, which are associated with four different tongue motions [10]. This method robustly extracted features and may be more useful when a higher number of movement commands are required or the signals are contaminated with noise. However, the higher computational load makes it unsuitable for real-time applications.

To explore the real-time implementation of a TMEP signal based assistive communication system, several challenges need to be addressed. One significant challenge is the ability of the system to classify TMEP signals in real environments under the influence of interference, including external noise from the surrounding environment (e.g. conversation, road noise), motion artifacts (e.g. head movements), internal noise or artifacts due to natural tongue movements (e.g. speech, mastication). Such interference problems are generally challenging in any human machine interface system. Superior performance of TMEP signal classification has been achieved in the data sets collected in controlled environments [4]. However, significant degradation was experienced when the TMEP signals were contaminated with noise [12,13]. A de-noising algorithm based on discrete wavelet thresholding was applied to improve the quality of signals [11]. Another challenge is that only a limited number of signals are available to train and calibrate the classifier in real environments [12,13]. On the other hand, the accuracy and robustness of a classification algorithm depends highly on its input and therefore optimal selection of its features is very important, especially in noisy environments.

As the TMEP signals of movement actions exhibit transient behaviour in the order of tens of milliseconds, the wavelet packet transform (WPT) should be able to reliably extract features in a multi-scale manner for the classification between movement commands and interferences. The WPT can capture localised time-frequency information of signals and has been implemented widely in signal analysis and modelling [14–18], with significant successful application in diverse fields such as signal detection, classification, compression, noise reduction and image processing [19–21]. To improve the classification accuracy in the presence of external interferences, the WPT was applied to extract features for classification of TMEP controlled actions [12,13,22]. Based on these WPT features, the classification performance has achieved a 97% recognition rate in a simulated noisy environment in comparison to poor performance (88%) with time domain features. In order to further improve the accuracy, reliability and robustness of the assistive HMI system based on TMEP signals, controlled movement related TMEP signals should be discriminated from a wide range of interfering signals that occur in daily life. These interfering signals can be categorised into non-controlled movement or interference related TMEP signals such as speech, swallowing, coughing, eating, drinking, and external artifacts such as the individual's heart beat, remote muscular activity, limb tremor and environmental sounds.

This study aimed to identify controlled movement related TMEP signals from a variety of interferences. The features were extracted using a WPT to capture the transient changes in the TMEP signals and were optimally selected according to statistical distributions of the wavelet packet coefficients so as to maximise the separability between movement commands and interferences. Two types of classifiers, a Bayesian and support vector machine (SVM), were

implemented to perform the classification between two classes of commands and interferences. Their performance was evaluated in both offline and online conditions using both subject specific and generalised interference for training. This work has significantly improved the accuracy and robustness of both offline and online real-time assistive human machine interface systems based on TMEP signals.

2. Experimental paradigm and signal acquisition

2.1. Participants

Ten healthy subjects (6 males, 4 females) ranging from 20–45 years (30.7 ± 6.4 ; mean \pm 1 SD) participated in the experiment. It is noted that within this subject group, five subjects (S6–S10) were well trained to perform the controlled TMEP actions whilst the remaining five subjects (S1–S5) had only half an hour practice prior to data collection. The experiment was approved by the ISVR Human Experimentation Safety and Ethics Committee of the University of Southampton. Participants gave their written informed consent before taking part in the study.

2.2. Experiment and signal recording

The oral cavity is connected to the ear via the Eustachian tube. Tongue movements cause pressure changes within the ear canal, which can be detected by a sensor. The sensor includes a shielded housing plug and an internal microphone. The microphone was inserted into the ear canal and connected to an amplifier. The pressure change was picked up by the microphone and digitised and stored in a computer similarly as in [4]. The distinct movement related actions can be differentiated from signatures of the recorded ear pressure signals.

In the present study the classification was performed between controlled movement commands and interference related TMEP signals. TMEP signals were recorded when subjects performed six types of controlled tongue movement: moving the tongue from the neutral position to the top/front centre of the roof of the mouth ('up'), touching the tongue to the bottom/front centre of the mouth ('down'), the front/right side of the mouth ('right'), the front/left side of the mouth ('left'), flicking the tongue up and down once ('flicking') and moving the tongue to the outside of the oral cavity in a straight manner with closed lips ('pushing'). TMEP signals during these six intended tongue actions were defined as controlled or intended movement related TMEP signals.

In contrast, non-controlled movement or interference related TMEP signals were collected while subjects were speaking, coughing, drinking or resting. The speech activity included utterances of words consisting of numbers from 0 to 9, and words 'start', 'stop', 'open', 'close', 'on' and 'off'. The drinking activity was to drink 15 ml of water from a glass, whilst the resting activity was recorded during normal relaxation. This set of words represents a wide range of tongue movement patterns.

Each subject was seated in a comfortable armchair with a recording microphone sensor inserted into the ear canal. Prior to the experiment, the selection of ear (left or right) to insert the earpiece was made by the participants according to individual preference. The signals were recorded using custom made software written in Microsoft C# running on a laptop computer.

A visual cue was presented on a computer screen to instruct the subjects to perform a specific tongue movement action. Subjects were instructed to move their tongue in the respective direction as much as possible, so as to perform each action correctly. The cues were represented by text as well as direction, via a moving circle on the screen. Before making each movement,

the participant was instructed to always place the tongue back into its neutral position. Each action was randomly repeated every 5 s, to minimise the possible effects of fatigue or learning. Each controlled (six) and non-controlled (nineteen) movement were repeated 100 and 20 times, respectively. After each movement, the direction or type of movement was labelled by the subject and indexed in customised software for classification analysis. Signals were sampled at 8000 Hz and then digitally down-sampled to 2000 Hz for further analysis.

3. Signal analysis methods

3.1. Discrete WPT for feature extraction

The discrete WPT represents a generalisation of multi-resolution analysis to decompose a signal into sub-bands and presents both approximation and detail spaces in a binary tree [23,24]. The wavelet packet coefficients at one scale can be recursively decomposed into the coefficients at the next scale using a low-pass and high-pass analysing filter. To compute the WPT coefficients of scale levels $j=1, \dots, J$, it filters the WPT coefficients recursively at the previous stage. Let $W_{j,p}(k)$, $p=0, \dots, 2^j-1$ represent the WPT coefficients at level j . Then the following two wavelet packet orthogonal bases equations are used to compute the wavelet packet coefficients:

$$W_{j,2p}(k) = \sum_{l=0}^{L-1} h(l)W_{j-1,p}(2k+1-l \bmod N_{p-1}) \quad (1)$$

$$W_{j,2p+1}(k) = \sum_{l=0}^{L-1} g(l)W_{j-1,p}(2k+1-l \bmod N_{p-1}) \quad (2)$$

where $k=1, \dots, N$ and $N_p=N/2^p$. $h(l)$ and $g(l)$ are the impulse responses of scaling and wavelet filters, which represents low-pass and high-pass filters, respectively. They are quadrature mirror filters and have only finite non-zero filter coefficients, which results in an efficient way to compute the WPT coefficients.

3.2. Wavelet filter selection

The efficacy of the wavelet packet transformation is dependent on the wavelet basis or filter. One common approach to specifying the wavelet filter is to select one with minimum reconstruction error according to an entropy cost function [24–26]. This is considered optimal for signal compression, but may be inappropriate for signal classification. A modified algorithm was proposed to maximise the discriminant ability of the WPT using a class separability cost function [27]. More often the wavelet filter selection was performed empirically according to the above criteria [25]. In this study the selection of wavelet filter was made with the criteria (1) properties of the wavelet filter and (2) a class separability based objective function for evaluation amongst all possible wavelets in the following families: Daubechies, Coiflets and Symlets. These families of wavelets were considered due to their properties of (1) orthogonal transform, (2) compact support and (3) optimal number of vanishing moments. A Symlet wavelet filter of order seven (Sym7) was selected as it gave the best classification performance based on a Euclidean distance measure among the available wavelet families [12]. A few other wavelets, i.e., Daubechies with order 5, Coiflet with order 4 and Symlet with order 5 also achieved comparable performance.

3.3. Bayesian classification

The naive Bayesian classifier classifies a pattern into one of a set of classes by maximising the posterior probability $p(c_i | \mathbf{x}) = p(\mathbf{x} | c_i)$

$p(c_i)/p(\mathbf{x})$, where c_i is the i^{th} class, and \mathbf{x} is a test pattern defined by features from the measured signals. If the features are independent given the class, the likelihood $p(\mathbf{x} | c_i)$ can be decomposed into the product of $p(x_1 | c_i), \dots, p(x_n | c_i)$ and the posterior becomes

$$p(c_i | \mathbf{x}) = \frac{p(c_i)}{p(\mathbf{x})} \prod_{k=1}^n p(x_k | c_i) \quad (3)$$

In this study, the discriminant function of the Bayesian classifier was derived under multivariate Gaussian assumptions [21,28].

3.4. SVM classification

The SVM estimates the optimal boundary in the feature space by combining a maximal margin strategy with a kernel method. The machine is trained according to the structural risk minimisation criterion [21,29]. The decision boundaries are directly derived from the training data set by learning. The SVM maps the inputs into a high-dimensional feature space through a selected kernel function. It then constructs an optimal separating hyper-plane in the feature space. To obtain optimal performance of the SVM classifier, selection of a proper kernel function is essential [29]. The optimal kernel function is dependent on the specific data and linear or radial basis function (RBF) kernel is generally used in bio-signal classification [30]. In this study, a RBF kernel was selected as it performed much better than a linear kernel. The hyperparameters of a SVM classifier, i.e., the regularisation parameter C and the RBF kernel parameter γ , were estimated during training to optimise classification performance.

3.5. TMEP signal classification

The classification of controlled actions and interferences related TMEP signals consisted of signal activity detection,

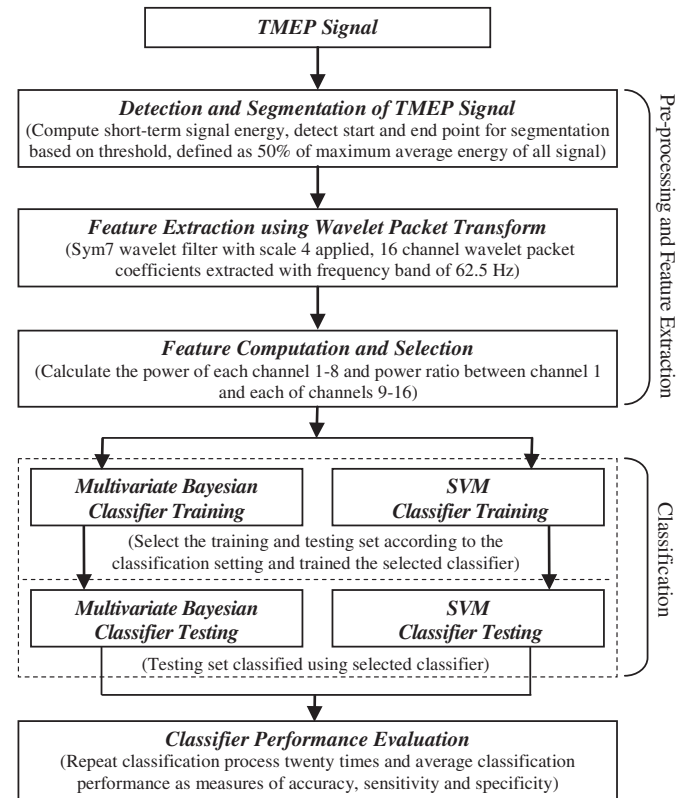


Fig. 1. Flowchart for controlled movement commands and interferences related TMEP signal classification.

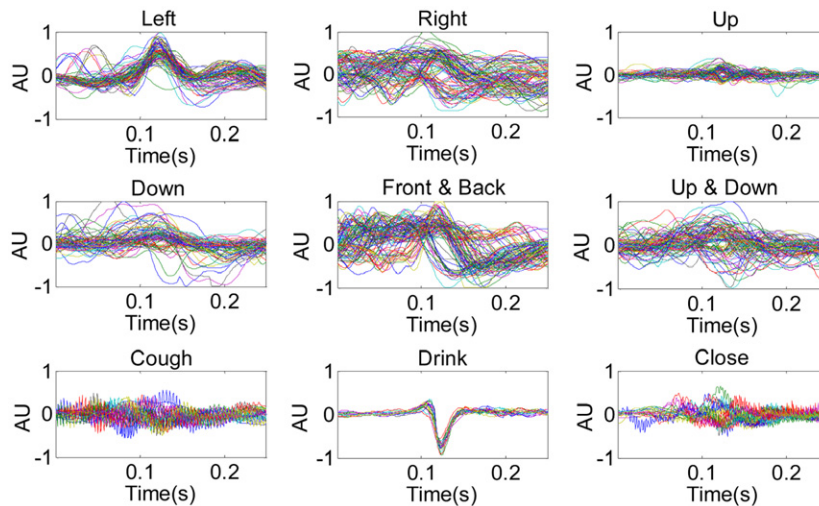


Fig. 2. Superimposed segmented TMEP traces for controlled movements (Left, Right, Up, Down, Flicking and Pushing) and interferences (Cough, Drink and Speech-‘Close’) related signals for Subject #5. The duration of each segment is 0.25 s and each activity repeated 100 and 20 times for controlled movement and interference, respectively.

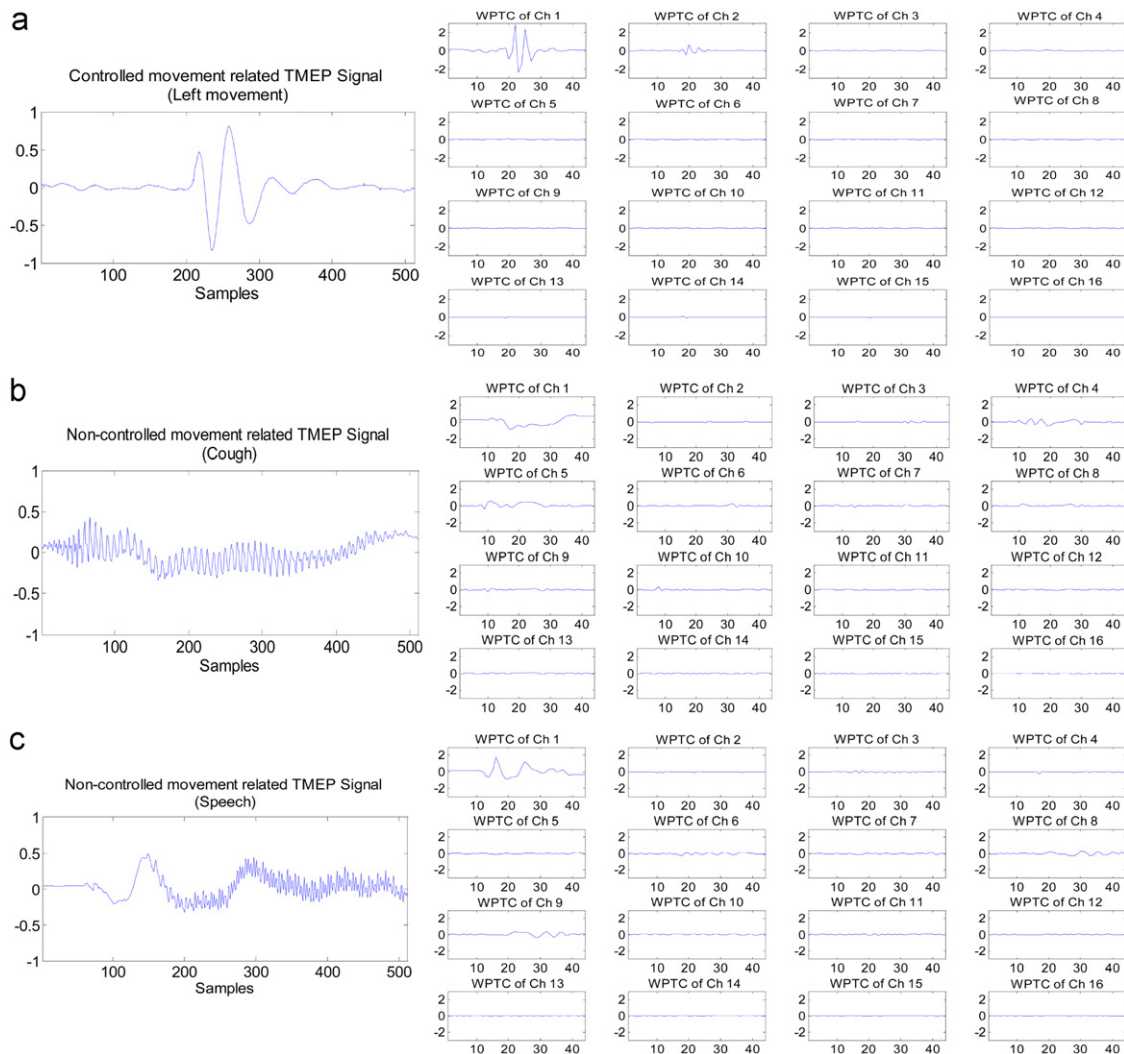


Fig. 3. Controlled movement (a, moving left action) and interferences (b, cough and c, speech of ‘close’) related segmented TMEP signals and their 16 channels wavelet packet transform coefficients at level 4 for subject #5.

segmentation, feature extraction, feature selection and classification. The flowchart of these stages is shown in Fig. 1. The TMEP signals during tongue movement actions need to be detected and

segmented appropriately. The detection method is similar to that in automatic speech recognition systems by setting a threshold on the short-term energy of the incoming signal. The threshold was

determined for each subject as 50% of the maximum average peak energy across training TMEP signals during tongue movement commands calibration. The signal is then segmented to a section with 512 samples, which is slightly longer than the typical 0.2 s duration of the TMEP signal during a controlled movement. Details of detection and segmentation methods are available in [4,31]. The superimposed segmented TMEP traces of controlled movement and interferences have large variation in shape, duration and frequency (Fig. 2).

To extract features, segmented TMEP signals were decomposed into the WPT domain using a Sym7 wavelet filter at scale 4. The selection of the decomposition scale was made by comparing two other scales (3 and 5) based on their capability to localise discriminative information. After transformation, 16 channels of WPT coefficients at scale 4 were obtained, each with a frequency bandwidth of approximately 62.5 Hz. An illustration of the WPT coefficients of a controlled movement (left) and two interferences (coughing and speech of ‘close’) related TMEP signal is presented in Fig. 3. It has been found that the controlled movement related

TMEP signals have the majority of their energy located in the low frequency band (0–62.5 Hz) [4,12], whereas the energy of interfering TMEP signals is distributed at low and/or high frequency. There are large variations of the energy distributions associated with interferences. Some only have signal energy at low frequency, such as drinking, and some have signal energy at low and high frequencies. Therefore all channels of the WPT coefficients were considered for feature definition. Sixteen discriminative features (x_1, \dots, x_{16}) were computed as the absolute power and low-to-high frequency power ratios. The power p_i of each WPT channel was calculated as the variance of the wavelet coefficients. The first eight features ($x_1 = p_1, \dots, x_8 = p_8$) were computed as the absolute power of the WPT channels 1–8 (frequency range 1–500 Hz). To isolate the discriminative information content between controlled movement and interfering TMEP signals by utilising the very low (0–62.5 Hz) and high (500Hz or more) frequency WPT channels, the power ratios between channel 1 and each of channels 9–16 were computed as the remaining eight features ($x_9 = p_1/p_9, \dots, x_{16} = p_1/p_{16}$) for each signal. These features

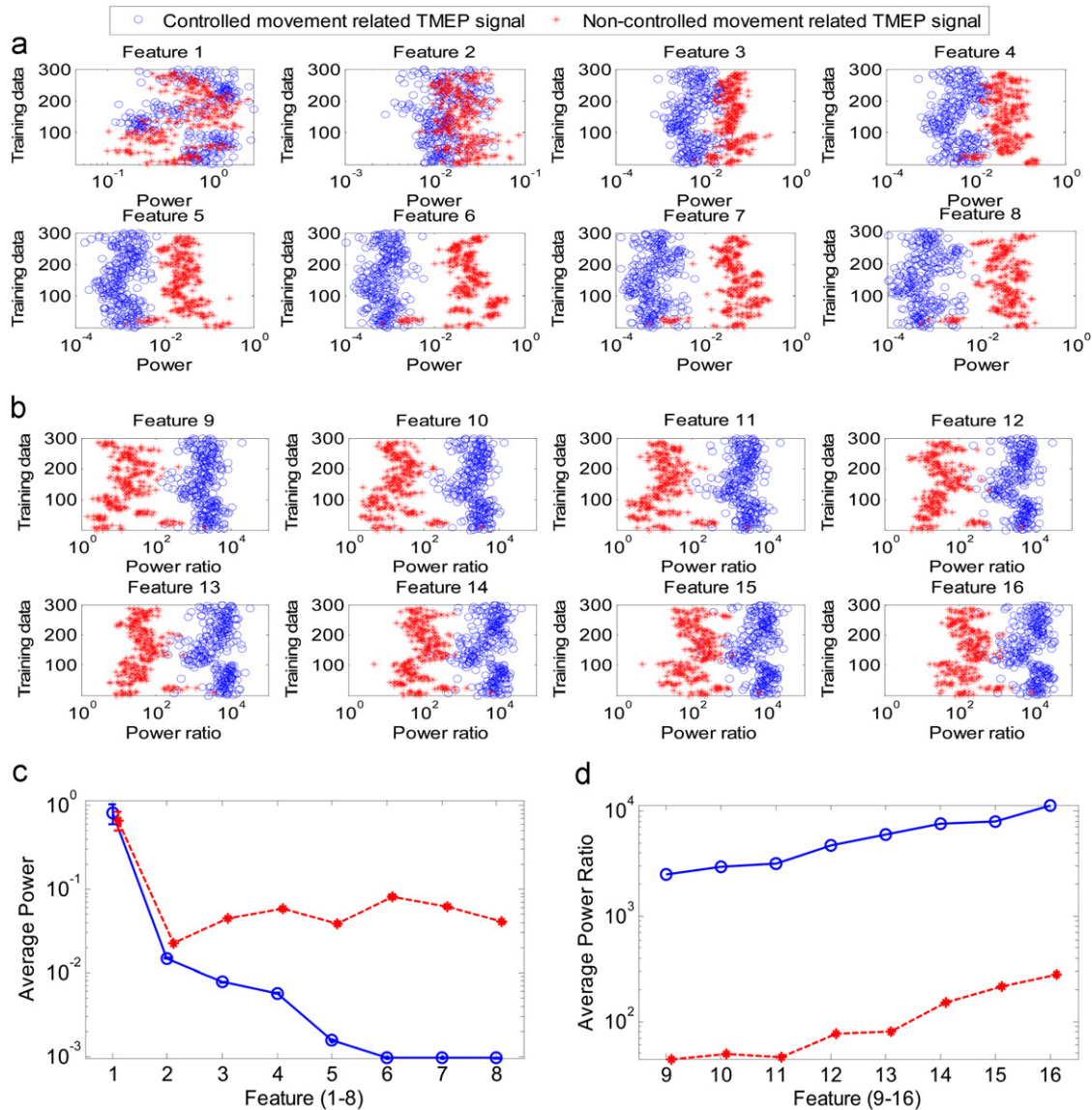


Fig. 4. Feature distribution of controlled movements (circle, blue) and interferences (star, red) related TMEP signals among all training data set in Subject #5. Feature 1–8 computed as power of each WPT channel 1–8 (a) and feature 9–16 computed as power ratio between the channel 1 and each of channel 9–16 (b). The average power for feature 1–8 (c) and average power ratio for feature 9–16 (d) of controlled movements (blue, circle) and interferences (red, star) related TMEP signal were computed and are presented on a logarithmic scale. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

were determined based on maximised class separability to provide optimal classification performance after comparison with other channels and various combinations of their ratios and quantities. The distribution of the 16 features related to TMEP controlled movement and interferences for Subject #5 is shown in Fig. 4. In this case, it was noted that the classes were almost linearly separable by features 5 to 16, while features 1 and 2 overlapped significantly and features 3 and 4 were only partially overlapping. However feature 1 (channel 1) is important as it carries most of the power of the signal in both TMEP controlled movements and interferences (Figs. 3, 4c).

Based on the extracted features, a multivariate Bayesian classifier and a SVM classifier were designed to classify the controlled movement related signals from interferences. Each classifier was constructed in specific and generalised interference situations. In specific interference situation, the classifier was trained and tested using each subject's specific TMEP controlled movements and interference related features. The training and testing data sets of the specific interference were selected randomly, with 60% from each type of signal assigned to train the classifier and the rest (40%) used to test the classifier. The training and testing data were mutually exclusive. In the generalised interference situation, the classifier was further extrapolated to be more robust to address a wide variety of interferences from all subjects. As the characteristics of the TMEP controlled movements are unique to each subject and the types of the interference are not limited to them, the classifier was constructed with subject specific TMEP controlled movements and generalised interference related features. The training and testing data sets of the generalised interference were selected similarly to the specific interference for controlled movements, and using a leave-one-subject-out cross-validation procedure (i.e. data from one subject used for testing and the remaining subjects used to train the classifier) for interferences.

The discriminant functions were used to separate different classes in a Bayesian classifier [32]. Fig. 5 shows the discriminant functions for classifying controlled movement TMEP signals from the specific interferences situation. The discriminant function for the controlled movement class was much higher than the interference class when the input was features of controlled movement related TMEP signals in most cases. The opposite occurred when the input was features of interferences related TMEP

signals. It indicates a separation boundary existing between the controlled movements and interferences related TMEP signals.

The classification between the controlled movements and interferences related TMEP signals was further explored using a SVM classifier in both specific and generalised interference situations. The optimal selection of SVM parameters (C and γ) was performed through a 5-fold cross-validation procedure [33]. The SVM classifier was implemented using LIBSVM [34]. To statistically compare the performances among the classification methods in specific and generalised interference situations, as well as trained and un-trained groups, a Student's t -test was performed using SPSS (Ver. 15, Chicago, Illinois).

4. Results

The classification performance was evaluated with averaged accuracy, sensitivity and specificity, by repeating the classification process twenty times with random selection of the training and testing data. Accuracy is defined as the percentage of correctly classified instances. Sensitivity is defined as the ratio of the number of true positives classified to the number of actual total positive cases. Specificity is defined as the ratio of the number of true negative classified to the number of actual total negative cases. The Bayesian classifier and SVM classifier were evaluated on the ten subjects in both the specific and generalised situations.

The classification accuracy of the multivariate Bayesian classifier was $97.8 \pm 2.1\%$ (mean ± 1 SD) across all 10 subjects when subject specific interferences were used. The sensitivity and specificity were $98.8 \pm 1.7\%$ and $96.1 \pm 3.5\%$, respectively. The SVM classifier achieved slightly better performance based on the classification accuracy, sensitivity and specificity of $98.5 \pm 1.9\%$, $99.2 \pm 1.0\%$ and $99.3 \pm 3.7\%$, respectively. In the generalised interferences situation, the performance remained at a similar level. The accuracy, sensitivity and specificity were $96.4 \pm 3.8\%$, $98.7 \pm 1.5\%$ and $94.5 \pm 6.0\%$ for the Bayesian classifier, and $96.6 \pm 3.6\%$, $95.4 \pm 5.0\%$ and $97.1 \pm 3.2\%$ for the SVM classifier, respectively (Figs. 6a,b). In the specific interference situation, the SVM classifier performed significantly better than Bayesian in terms of accuracy ($98.5 \pm 1.9\%$ vs. $97.8 \pm 2.1\%$ ($t(9) = -4.1$, $p < 0.05$)) and specificity ($99.3 \pm 3.7\%$ vs. $96.1 \pm 3.5\%$ ($t(9) = -2.3$, $p < 0.05$)). In the generalised interference situation, the Bayesian classifier performed significantly better than the SVM in terms of sensitivity ($98.7 \pm 1.5\%$ vs. $95.4 \pm 5.0\%$ ($t(9) = 3$, $p < 0.05$)) although the SVM performed better in terms of specificity ($97.1 \pm 3.2\%$ vs. $94.5 \pm 6.0\%$ ($t(9) = -2.8$, $p < 0.05$)). Overall these two classifiers achieved similar level of performance.

The effect of training was further investigated. Among 10 subjects, half (S1–S5) had a short practice before the experiment (un-trained group) and the other half (S6–S10) had intensive training to adequately make tongue movement commands (trained group). The trained group had significantly better performance than the un-trained group: accuracy $99.3 \pm 0.3\%$ vs. $96.3 \pm 2.0\%$ ($t(8) = -3.4$, $p < 0.05$), sensitivity $100.0 \pm 0.0\%$ vs. $97.6 \pm 1.7\%$ ($t(8) = -3.1$, $p < 0.05$) and specificity $98.8 \pm 1.01\%$ vs. $93.4 \pm 2.9\%$ ($t(8) = -3.9$, $p < 0.05$) in the Bayesian classifier (Fig. 6c), and accuracy $99.8 \pm 0.2\%$ vs. $97.1 \pm 1.8\%$ ($t(8) = -3.3$, $p < 0.05$), sensitivity $99.9 \pm 0.1\%$ vs. $98.5 \pm 0.9\%$ ($t(8) = -3.7$, $p < 0.05$) and specificity $99.8 \pm 0.2\%$ vs. $94.8 \pm 3.9\%$ ($t(8) = -2.8$, $p < 0.05$) in the SVM classifier (Fig. 6d).

5. Real-time evaluation

In the above offline experiments various interferences of speaking, coughing or drinking were investigated. In a real world

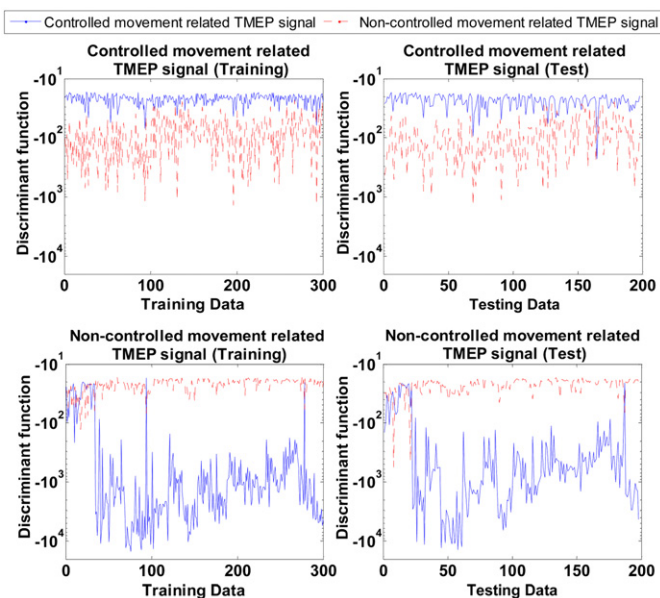


Fig. 5. Discriminant functions of Bayesian classifier for the training and testing data to identify the state of controlled movements (solid, blue) and interferences (dotted, red) related TMEP signals in subject #5.

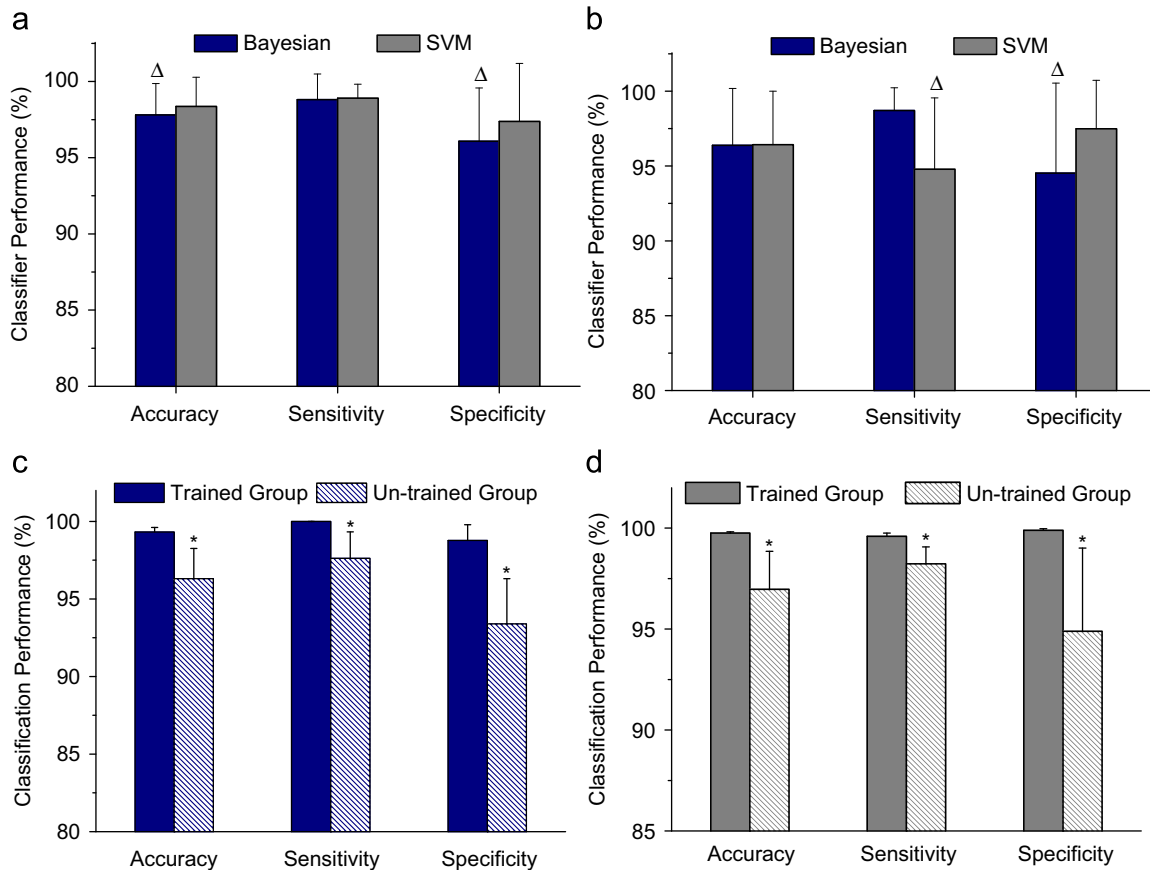


Fig. 6. Average offline classification performance of accuracy, sensitivity and specificity across all subjects based on specific interferences (a) and generalised interference (b) situations. The training can significantly improve the performance of both Bayesian (c) and SVM classifiers (d). $\Delta p < 0.05$, Bayesian vs. SVM, and $* p < 0.05$, trained vs. untrained group, mean ± 1 SD.

implementation there will be a lot of other interferences, for instance, free speech, and the algorithm needs to be incorporated with the successive inter-command classification scheme. Therefore the algorithm was further evaluated in a real-time environment. Both the Bayesian and SVM classifiers were trained with generalised interferences from three randomly selected subjects with a small training set and tested on an additional two subjects. The training set has only 120 trials of controlled movements and 162 trials of interferences. The test subjects performed controlled movements of 'up', 'down', 'right' and 'left' actions and interferences segmented during a 5-min newspaper reading, 1-min conversation, swallowing, coughing and drinking. The testing was carried out in a normal office environment and the signals were detected, segmented and classified in real-time. The results showed that the Bayesian classifier achieved 88.1% in accuracy, 95.0% in sensitivity and 85.9% in specificity, while the SVM achieved only 68.6% in accuracy, 97.5% in sensitivity and 59.4% in specificity. The SVM performed considerably worse than the Bayesian classifier. It may be due to the fact that the classifier parameters optimisation is only based on a small training set with large variability. The SVM is more sensitive to the size of training set than the Bayesian classifier. The small training size tends to cause the SVM classifier to over-fit to the training data and therefore have poor generalisation during testing [35,36]. After rejecting interferences, the movement commands were further identified and used to control a simulated wheelchair on a computer screen [31]. The wheelchair was well controlled with only a few false actions. In contrast, the wheelchair went quickly out of control when no interference rejection procedure was

utilised. A demonstration video of the system is available at <http://www.swanglab.com/software.htm>.

6. Discussion

Interference is one of the major challenges in developing human machine interfaces, including brain computer interfaces, due to its variety and uncertain sources [7,30]. In the present study, robust identification of tongue movement commands from interferences was explored using Bayesian and SVM classifiers with features extracted by a WPT. The robustness of the classification was also tested in a real-time environment. Both classifiers performed better offline, although the multivariate Bayesian classifier achieved higher accuracy than the SVM in the real-time system.

Previously the wavelet packet transform was used to extract features for classifying tongue movement actions and achieved better performance than time domain features in both clean and noisy environments [12,13,32]. The WPT has been widely used for feature extraction from bio-signals, such as ECG [37], MES [38] or EEG [39,40], and has provided better performance for pattern recognition over time and spectral domain features compared to the wavelet transform and short-time Fourier transform [38]. This may be owing to its capability to precisely localise the time-scale information in non-stationary signal dynamics. To achieve optimal performance, the wavelet filter and decomposition scale were empirically determined by comparing several wavelet families

and scales. It should be noted that some wavelet filters achieved comparable performance to the selected Sym7.

Two classifiers were investigated in this study. The Bayesian classifier is simple and computationally efficient, [28,30] while the SVM is more complex due to its optimisation characteristics [35,36]. The SVM performed slightly better than the Bayesian classifier during offline classification of controlled movement related TMEP signals and interferences (Figs. 6a,b). Both classifiers were robust in both the subject specific and generalised interference situations. In the generalised interference situation, the classifier requires no previous information about interfering signals from the test subject. This implies that as it is trained with other subjects' signals, it does not require collecting a personalised training set for interferences. The performance in the generalised interference situation is slightly lower than in the specific interference case (Figs. 6a,b). Training for the subject associated with the execution of the controlled actions can significantly improve the classification performance (Figs. 6c,d). The classifiers were further implemented in a real-time system using generalised interference data for training. The Bayesian classifier performed much better than SVM in a real-time system environment. The performance was worse than that in an offline situation but the classifier made it possible to control a simulated wheelchair precisely in real-time. As a problem-specific application, it has been demonstrated that the Bayesian classifier with features extracted using the wavelet packet transform is suitable for a real-time system of identifying movement commands from interferences, and potentially can be used within the command classification algorithm as well. The reduction of the performance in a real-time environment may be related to the wide feature variation of the TMEP signals. The variance of features of interfering signals is higher than those of the controlled movement related signals. Such asymmetric distribution might be a contributing factor to the higher specificity error than the sensitivity error in the Bayesian classifier in both offline and online classification. Another possible cause of the high specificity error may be that some interfering signals have energy concentrated at very low frequencies, such as drinking, and they have comparable signatures to the actual controlled actions (Fig. 2). Taking temporal patterns into account could potentially reduce the error and improve the overall performance.

This work provides an efficient approach for distinguishing TMEP signals from varied interferences in both offline and online systems. The Bayesian classifier is computationally efficient to be applied to HMI or BCI systems (for instance, deep brain local field potentials [41,42] or EEG signals [30,43]) to improve the accuracy of the decision making. It may be the case that more features are required when complex tasks are involved. In future work, dynamic feature selection methods, such as principal component analysis (PCA), singular value decomposition (SVD) or discriminative common vector (DCV) techniques will be applied for feature reduction and selection. This is expected to provide better performance in more challenging tasks, such as movement classification based on deep brain local field potentials [33,42].

In summary, the notable technical contributions that were introduced in this paper are as follows:

- A wavelet packet based feature extraction and selection approach developed to identify TMEP actions and interferences.
- The robustness of the identification method was evaluated with various types of interference signals in subject specific and generalised interference setting, and achieved high accuracy.
- Feasibility of robust identification of tongue movement commands from interferences was also evaluated in a real-time setting while considering a wide range of potentially interfering factors, for example free speech and was still able to maintain a good performance level.
- Typical results from ten subjects in offline and two subjects in real-time have demonstrated the success of the method.
- The effect of training to perform the task was also investigated as a way of improving classification performance.

The Bayesian classifier with features extracted by the wavelet packet transform can reliably distinguish controlled movement related TMEP command signals from the interference signals both offline and online. The rejection of various interfering signals has significantly improved the robustness of the assistive HMI based on TMEP signals and makes the real-time implementation and application in real living environments possible.

Acknowledgement

We would like to thank reviewers for their constructive comments and subjects for their participation in the experiment. This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC; Grant number EP/F01869X/1) and an ISVR Rayleigh Scholarship, University of Southampton, UK.

References

- [1] X.L. Huo, J. Wang, M. Ghovanloo, Introduction and preliminary evaluation of the tongue drive system: wireless tongue-operated assistive technology for people with little or no upper-limb function, *J. Rehabil. Res. Dev.* 45 (6) (2008) 921–930.
- [2] T.E. Hutchinson, K.P. White, W.N. Martin, K.C. Reichert, L.A. Frey, Human-computer interaction using eye-gaze input, *IEEE Trans. Syst. Man Cybern.* 19 (6) (1989) 1527–1534.
- [3] C.K.H. Law, M.Y.Y. Leung, Y. Xu, S. K. Tso, A cap as interface for wheelchair control, in: Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2002), Lausanne, Switzerland, 2002, pp. 1439–1444.
- [4] R. Vaidyanathan, B. Chung, L. Gupta, H. Kook, S. Kota, J.D. West, Tongue-movement communication and control concept for hands-free human-machine interfaces, *IEEE Trans. Syst. Man Cybern. A Syst. Hum.* 37 (4) (2007) 533–546.
- [5] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, T.M. Vaughan, Brain-computer interfaces for communication and control, *Clin. Neurophysiol.* 113 (6) (2002) 767–791.
- [6] Y.-L. Chen, F.-T. Tang, W.H. Chang, M.-K. Wong, Y.-Y. Shih, T.-S. Kuo, The new design of an infrared-controlled human-computer interface for the disabled, *IEEE Trans. Rehabil. Eng.* 7 (4) (1999) 474–481.
- [7] M.K. O'Malley, Principles of human-machine interfaces and interactions, in: M. Zhang, B. Nelson, R. Felder (Eds.), *Life Science Automation: Fundamentals and Applications*, Artech House Publishers, 2007, pp. 101–125.
- [8] S.G. Mason, A. Bashashati, M. Fatourehchi, K.F. Navarro, G.E. Birch, A comprehensive survey of brain interface technology designs, *Ann. Biomed. Eng.* 35 (2) (2007) 137–169.
- [9] M.A. Lebedev, M.A.L. Nicolelis, Brain-machine interfaces: past, present and future, *Trends Neurosci.* 29 (9) (2006) 536–546.
- [10] R. Vaidyanathan, C.J. James, Independent component analysis for extraction of critical features from tongue movement ear pressure signals, in: Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2007), Lyon, France, 2007, pp. 5481–5484.
- [11] R. Vaidyanathan, S. Wang, L. Gupta, A wavelet denoising approach for signal action isolation in the ear canal, in: Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2008), Vancouver, Canada, 2008, pp. 2677–2680.
- [12] K.A. Mamun, M. Mace, M.E. Lutman, R. Vaidyanathan, S. Wang, Pattern classification of tongue movement ear pressure signal based on wavelet packet feature extraction, in: Proceedings of the 5th UK & RI Postgraduate Conference in Biomedical Engineering & Medical Physics (PGBioMed 2009), Oxford, UK, 2009, pp. 33–34.
- [13] K.A. Mamun, M. Mace, M.E. Lutman, R. Vaidyanathan, S. Wang, Bayesian classification of tongue movement based on wavelet packet transformation, in: Proceedings of the Network on multiScale Information RePresentation and Estimation (INSPIRE 2009), London, UK, 2009.
- [14] H. Dickhaus, H. Heinrich, Classifying biosignals with wavelet networks, *IEEE Eng. Med. Biol. Mag.* 15 (5) (1996) 103–111.
- [15] M. Cocchi, R. Seeber, A. Ulrici, WPTER: wavelet packet transform for efficient pattern recognition of signals, *Chemometr. Intell. Lab. Syst.* 57 (2) (2001) 97–119.
- [16] R.E. Learned, A.S. Willisky, A wavelet packet approach to transient signal classification, *Appl. Comput. Harmon. Anal.* 2 (3) (1995) 265–278.
- [17] A. Quinquis, A few practical applications of wavelet packets, *Digit. Signal Process.* 8 (1) (1998) 49–60.

- [18] M. Unser, A. Aldroubi, A review of wavelets in biomedical applications, in: Proceedings of the IEEE 84, no. 4, 1996, pp. 626–638.
- [19] G. Antonini, A. Orlandi, Wavelet packet-based EMI signal processing and source identification, IEEE Trans. Electromagn. Compat. 43 (2) (2001) 140–148.
- [20] H.B. Xie, Y.P. Zheng, J.Y. Guo, Classification of the mechanomyogram signal using a wavelet packet transform and singular value decomposition for multifunction prosthesis control, Physiol. Meas. 30 (5) (2009) 441–457.
- [21] R. Duda, P. Hart, D. Stork, Pattern Classification, 2nd ed., John-Wiley, 2001.
- [22] K.A. Mamun, M. Mace, R. Craig, M.E. Lutman, R. Vaidyanathan, S. Wang, Tongue movement ear pressure signal classification using wavelet packet transform, Int. J. Audiol. 49 (9) (2010) 701.
- [23] D.B. Percival, A.T. Walden, Wavelet Methods for Time Series Analysis, Cambridge University Press, 2000.
- [24] A.T. Walden, Wavelet analysis of discrete time series, in: Proceedings of the European Congress of Mathematics, vol. 202, 2001, pp. 627–641.
- [25] S. Wang, X.G. Liu, J. Yianni, T.Z. Aziz, J.F. Stein, Extracting burst and tonic components from surface electromyograms in dystonia using adaptive wavelet shrinkage, J. Neurosci. Methods 139 (2) (2004) 177–184.
- [26] R.R. Coifman, M.V. Wickerhauser, Entropy-based algorithms for best basis selection, IEEE Trans. Inf. Theory 38 (2) (1992) 713–718.
- [27] N. Saito, R.R. Coifman, Local discriminant bases and their applications, J. Math. Imaging Vis. 5 (4) (1995) 337–358.
- [28] D.G.T. Denison, C.C. Holmes, B.K. Mallick, A.F.M. Smith, Bayesian Methods for Nonlinear Classification and Regression, Wiley Series in Probability and Statistics, 2002.
- [29] S.R. Gunn, Support Vector Machines for Classification and Regression, Technical Report, ECS, University of Southampton, UK, 1998.
- [30] A. Bashashati, M. Fatourehchi, R.K. Ward, G.E. Birch, A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals, J. Neural Eng. 4 (2) (2007) R32–R57.
- [31] M. Mace, K.A. Mamun, R. Vaidyanathan, S. Wang, L. Gupta, Real-time implementation of a non-invasive tongue-based human-robot interface, in: Proceedings of the 2010 IEEE/RJS International Conference on Intelligent Robots and Systems (IROS 2010), Taipei, Taiwan, 2010, pp. 5846–5852.
- [32] K.A. Mamun, M. Mace, M. Lutman, R. Vaidyanathan, L. Gupta, S. Wang, Multivariate Bayesian classification of tongue movement ear pressure signals based on the wavelet packet transform, in: Proceedings of the 2010 IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2010), Kittila, Finland, 2010, pp. 208–213.
- [33] K.A. Mamun, M. Mace, M.E. Lutman, R. Vaidyanathan, S. Wang, SVM classification of tongue movement ear pressure signals for human machine interface, in: Proceedings of the Network on multiScale Information Re-Presentation and Estimation (INSPIRE 2010), London, UK, 2010.
- [34] C.-C. Chang, C.-J. Lin, LIBSVM: A Library for Support Vector Machines, <<http://www.csie.ntu.edu.tw/~cjlin/libsvm>>, 2001.
- [35] F. Lotte, M. Congedo, A. Lecuyer, F. Lamarche, B. Arnaldi, A review of classification algorithms for EEG based brain-computer interfaces, J. Neural Eng. 4 (2) (2007) R1–R13.
- [36] D. Gacquer, V. Delcroix, F. Delmtte, S. Piechowiak, Comparative study of supervised classification algorithms for the detection of atmospheric pollution, Eng. Appl. Artif. Intell. 24 (2011) 1070–1083.
- [37] E.D. Ubeyli, ECG beats classification using multiclass support vector machines with error correcting output codes, Digit. Signal Process. 17 (3) (2007) 675–684.
- [38] K. Englehart, B. Hudgins, P.A. Parker, A wavelet-based continuous classification scheme for multifunction myoelectric control, IEEE Trans. Biomed. Eng. 48 (3) (2001) 302–311.
- [39] W.Y. Hsu, Y.N. Sun, EEG-based motor imagery analysis using weighted wavelet transform features, J. Neurosci. Methods 176 (2) (2009) 310–318.
- [40] V.J. Samar, A. Bopardikar, R. Rao, K. Swartz, Wavelet analysis of neuroelectric waveforms: a conceptual tutorial, Brain Lang. 66 (1) (1999) 7–60.
- [41] C. Loukas, P. Brown, Online prediction of self-paced hand-movements from subthalamic activity using neural networks in Parkinson's disease, J. Neurosci. Methods 137 (2) (2004) 193–205.
- [42] K.A. Mamun, R. Vaidyanathan, M.E. Lutman, J. Stein, X. Liu, T. Aziz, S. Wang, Decoding movement and laterality from local field potentials in the subthalamic nucleus, in: Proceedings of the 5th International IEEE/EMBS Conference on Neural Engineering (NER 2011), Cancun, Mexico, 2011.
- [43] M. Thulasidas, C. Guan, J.K. Wu, Robust classification of EEG signal for brain-computer interface, IEEE Trans. Neural Syst. Rehabil. Eng. 14 (1) (2006) 24–29.



Khondaker A. Mamun received the B.Sc. degree in Computer Science & Engineering from Ahsanullah University of Science and Technology, Dhaka, Bangladesh, in 2002, and M.Sc. degree in Computer Science & Engineering from Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 2007. He is currently a Ph.D. candidate in Institute of Sound and Vibration Research, University of Southampton, UK. His research interests include biomedical signal processing and machine learning, rehabilitation systems, human machine interfaces, brain computer interfaces and neuroinformatics. He has published a number of papers in international journals and conferences. He is a student member of IEEE.



Michael Mace received the M.Eng. degree in mechanical engineering from the University of Bristol, in 2008. He is currently a Ph.D. candidate in the Department of Mechanical Engineering at the University of Bristol, Bristol, UK. His research interests are in pattern recognition and physiological signal processing with emphasis on bio-acoustic signal classification and assistive human-machine (robotic) interfaces.



Lalit Gupta received the B.E. (Hons.) degree in electrical engineering from Birla Institute of Technology and Science, Pilani, India, in 1976, the M.S. degree in digital systems from Brunel University, Middlesex, U.K., in 1981, and the Ph.D. degree in electrical engineering from Southern Methodist University, Dallas, TX, in 1986. He is currently a Professor of electrical and computer engineering at Southern Illinois University, Carbondale. He has been awarded contracts from the Army Research office to conduct research in the development of smart munitions, from Seagate Technology on image compression research, Cleveland Medical Devices on National Institute of Health (NIH) funded projects related to brain wave form analysis and classification, Think-A-Move, Inc., on an NIH funded project related to human-machine interfacing, Neuronetrix on an NIH funded project on detecting neurological disorders from evoked potentials, and the Naval Postgraduate School on research related to human-machine interfacing. His current research interests include pattern recognition, neuroinformatics, neural networks and signal processing. He has numerous publications in the areas of neural networks, evoked potential analyses and classification and multichannel/sensor information fusion strategies. Dr. Gupta is an Associate Editor of the *Pattern Recognition Journal*.



Carl A. Verschuur is a Lecturer in Audiology at the Sound and Vibration Research. He has a background in clinical audiology as well as theoretical and practical applications of speech sciences. His primary research interest is in the relationship between speech processing in hearing aids and cochlear implants and speech perception. He has published work in speech perception in cochlear implant users, auditory localisation in bilateral cochlear implant users and speech perception in hearing aid users. His current work focuses on the role of low frequency speech cues in determining success with auditory prosthetic devices, and also the role inflammatory processes in determining loss of hearing.



Mark E. Lutman is Professor of Audiology and Head of the Hearing and Balance Centre at the Institute of Sound and Vibration Research of the University of Southampton. His research interests include: signal processing for hearing aids and cochlear implants, measurement of cochlear function from otoacoustic emissions, neonatal and adult hearing screening, cochlear implantation, epidemiology of hearing impairment and balance disorder, evaluation of benefit from hearing instruments and noise-induced hearing loss. He was editor of the British Journal of Audiology from 1991 to 1995 and President of the British Academy of Audiology from 2007 to 2008.



Maria Stokes is Professor of Neuromusculoskeletal Rehabilitation in the Faculty of Health Sciences, University of Southampton, and is a physiotherapist by background, with a PhD in Neuromuscular Physiology. Her research interests cover active living and healthy ageing and developing health technologies. Themes: 1) mechanisms of musculoskeletal function, dysfunction and recovery (across sporting elite to frail); 2) prevention and rehabilitation of musculoskeletal conditions. Development of technologies as assessment tools and assistive devices includes: rehabilitative ultrasound imaging (RUSI) to evaluate and re-educate muscle function; brain-computer interfacing (BCI) for communication and investigating neuromuscular mechanisms; mechanomyography (muscle sounds) and myotonometry to investigate muscle mechanical properties. A key feature is multidisciplinary collaboration, primarily with engineering scientists.



Ravi Vaidyanathan is a Senior Lecturer in Mechatronic Systems at Imperial College London, UK and a Research Assistant Professor at the US Naval Postgraduate School, USA. He earned his Ph.D. in biologically inspired systems at Case Western Reserve University in 2001, and worked in industry through 2004, holding directorships in control systems and medical engineering. His current research interests include biologically inspired robotics, human-machine interface and complex adaptive systems. He has led more than 20 research separate program in U.S.A., Singapore and U.K, authored over 90 refereed publications, and two (pending) patents. Dr. Vaidyanathan has been the recipient of international awards from organizations

such as the IEEE Robotics and Automation Society, the Robotics Society of Japan (RSJ) and the American Institute of Aeronautics and Astronautics (AIAA), including Best Paper in Conference at the IEEE International Conference on Intelligent Robots and Systems (IROS) and a Finalist for the New Technology Foundation (NTF) Award on Entertainment and Robotic Systems celebrating top innovations in robotics from 1987 to 2007. He also holds honorary academic posts at the University of Bristol (UK), Case Western Reserve University (USA) and Anna University (India).



Shouyan Wang is a Lecturer at the Hearing and Balance Centre at the Institute of Sound and Vibration Research of the University of Southampton. His research interests include: neural signal processing, brain plasticity after neural prostheses stimulation, brain connectivity modelling, neural engineering for neural prostheses, signal processing for hearing aids and cochlear implants.