Robust Facial Expression Recognition for MuCI
A Comprehensive Neuromuscular Signal Analysis - Preprint.pdf

Mahyar Hamedi
Sh-Hussain Salleh
Chee-Ming Ting, PhD
Alias M. Noor

Available at: http://works.bepress.com/chee-ming_ting/17/
Robust Facial Expression Recognition for MuCI: A Comprehensive Neuromuscular Signal Analysis

Mahyar Hamedi, Member, EMBS, Sh-Hussain Salleh, Member, IEEE, Chee-Ming Ting, Member, IEEE, Mehdi Astaraki and Alias Mohd Noor

Abstract—This paper presents a comprehensive study on the analysis of neuromuscular signal activities to recognize eleven facial expressions for Muscle Computer Interfacing applications. A robust denoising protocol comprised of Wavelet transform and Kalman filtering is proposed to enhance the electromyogram (EMG) signal-to-noise ratio and improve classification performance. The effectiveness of eight different time-domain facial EMG features on system performance is examined and compared in order to identify the most discriminative one. Fourteen pattern recognition-based algorithms are employed to classify the extracted features. These classifiers are evaluated in terms of classification accuracy and processing time. Finally, the best methods that obtain almost identical system performance are compared through the Normalized Mutual Information (NMI) criterion and a repeated measure analysis of variance (ANOVA) for a statistical significant test. To clarify the impact of signal denoising, all considered EMG features and classifiers are assessed with and without this stage. Results show that: (1) the proposed denoising step significantly improves the system performance; (2) Root Mean Square is the most discriminative facial EMG feature; (3) discriminant analysis when the parameters are estimated by the Maximum Likelihood algorithm achieves the highest classification accuracy and NMI; however, ANOVA reveals no significant difference among the best methods with almost similar performance.

Index Terms—Facial Neuromuscular Activity, Muscle Computer Interaction (MuCI), EMG Denoising, Feature Extraction, Classification, Facial Expression Recognition.

1 INTRODUCTION

Improving life quality has always been one of the priorities amongst scientists, physicians and researchers. In recent years, incorporation of electronic/computer/mechanics with biology/medicine study fields has led to numerous invaluable achievements. One of the most promising results to assist the locked-in patients with crucial disabilities, amputees and the elderly is Human Computer Interaction (HCI) technology which is an approach to transmit the information between humans and a computer [1]. Developing prosthetic limbs, robotic arms and controlling the assistive devices like wheelchairs are counted as important applications in this area. Since the reliability and flexibility of such systems are essential for the users, various techniques and types of interfaces have been suggested and employed. Recognizing the user’s body movements like those of the head, hand and wrist through bioelectrical activities and converting them into computer control commands have been focused recently. However, crucially disabled people cannot even move their neck, and the only existing way of communication for them is through facial expressions.

neuromuscular activities (electromyograms (EMGs)) or brain waves (electroencephalograms (EEGs)). Such HCIs are called Muscle Computer Interaction (MuCI) and Brain Computer Interaction (BCI). It is reported that BCI is only preferred when the use of MuCIs is not feasible [2]. There have been numerous studies on the potential of MuCI systems, including multifunction prosthesis [3-6], power exoskeleton control [7], wheelchairs [8-9], robotic control [10] and grasping control [11].

For affective communication between the user and computer, researchers much consider how facial expressions can be recognized efficiently during MuCI. In order to establish a robust MuCI system based on facial EMGs, several essential factors have to be taken into account, particularly the recording protocol and analysis steps. Since the comfort and friendliness of these systems are vital for users, non-invasive surface electromyography (SEMG) should be considered as the most convenient way to record the underlying activities of muscle contraction. SEMG characteristics, such as amplitude and energy, differ from 0 to 10 mV and 0 to 500 Hz respectively based on the activated muscle and ratio of contraction. These signals are inherently stochastic, non-stationary and noisy, contaminated by a variety of internal (e.g. EEGs, electrooculogram (EOGs)) and external (e.g. recording equipment, ambient noise, motion artifacts) factors [12]. The success of MuCIs depends substantially on classification performance and computational cost consumed during processing to provide a reliable trade-off between accuracy and speed. The significant factors in pattern
recognition that can directly influence the system performance are signal denoising, feature extraction and classification. Thus, accurate and fast techniques of signal processing and machine learning must be considered to design a trustworthy interface for MuCl.

Table 1 shows the previous studies which used facial myoelectric signals to recognize either facial gestures/expressions or human emotions. It reports the number of classes (tasks), recording channels and methods of EMG segmentation, feature extraction and classification as well as their results. The number of classes (facial gestures/expressions, emotions) varied in a range of 2 to 10. For different applications, 1 to 8 recording channels were used, and the three bipolar channels setup was suggested as the most optimum configuration for facial expression recognition [44]. Considering the processing steps, segments with different lengths were adopted. In [29], five time frames between 64 to 1024 msec were inspected on facial myoelectric signals and segment length of 256 msec was recommended, in line with the findings by Englehart et al. [4] for upper limb EMG classification. In terms of feature extraction, time-domain features have been widely used for classification of myoelectric signals as they are easy to compute and can provide better discriminative information than frequency and time-frequency features [45]. Table 1 indicates that different features affect the system performance significantly when being classified by similar classifiers. For example, in [36], Maximum Peak Value (MPV) was identified as the most informative and discriminative feature among ten when classified by Versatile Elliptic Basis Function Neural Network (VEBFNN). Regarding classification, various linear and non-linear techniques have been employed to classify the facial EMG features. Similar to feature extraction, it is obvious that different classifiers influence the system performance. For instance, in [31], it was shown that Fuzzy C-Mean (FCM) obtained higher recognition accuracy compared to Support Vector Machine (SVM) when they classified Root Mean Square (RMS) feature. This behavior of EMG features as well as classifiers emphasizes the important role of processing methods and machine learning algorithms for facial expression recognition. However, in previous studies this behavior has not been widely considered since only very few methods for either feature extraction or classification have been investigated. Classification accuracy has always been the most popular measure to assess the system performance; however, computational cost as another essential metric has only been considered in [29], [36-39] and [42].

According to Table 1, as far as the maximum number of classes (10) is concerned, up to 97.1 % classification accuracy was achieved where Least Square Support Vector Machine (LS-SVM) algorithm classified the fused features although only 87.1% accuracy was obtained. The reported 1.37 seconds processing time seems not to be suitable for real-time applications. The fastest performance was achieved in [36] where only a few milliseconds were consumed by VEBFNN to train the features although only 87.1 % accuracy was obtained.

As can be seen, a promising methodology that can cope with a high number of facial expressions and can deliver a fast and accurate performance has not been proposed yet. This impedes the facial myoelectric-based interfaces to be used in real-time applications. In addition,
efficient facial EMG signal denoising has not been investigated which may significantly increase the system performance. Several techniques have been designed and proposed for signal filtering; however, due to different characteristics of biological signals it is not possible to conclude that they perform well on all types of biosignals. Typically, the filter performance depends on the preceding information of signal statistical characteristics as well as the background noise. Digital filters like finite impulse response filters are time-invariant, and their shortcoming is the overlapping of the biosignals spectrum with background noise. Another denoising approach is the Wiener filter which delivers optimum filtering in the mean-squared error sense [46] only if the time-frequency properties of signals do not vary, and the non-stationary process is considered. Kalman filter (KF) is an appropriate alternative which offers a mean-squared value to minimize the error. Another advantage of KF is its Bayesian aspect [47] that allows possible modeling of prior knowledge about the parameters. This method leads to excellent estimates even for signals with a low signal-to-noise ratio (SNR). KF is a widely used technique for signal denoising [48]; however, to the best of our knowledge there is no study on this filter for facial EMG denoising.

This paper introduces a robust EMG-based facial expression recognition system by proposing the methods that result in almost full recognition accuracy with a very fast processing time. A KF-based denoising protocol in conjunction with Wavelet decomposition is offered to clean the facial EMGs and enhance the SNR. In addition, a comprehensive study on facial neuromuscular signal analysis in different processing stages is carried out. The objective of this paper is fourfold: (1) Developing an MuCI with more degrees of freedom by considering eleven facial expressions, (2) Proposing a robust facial EMG denoising procedure, (3) Introducing the most informative and discriminative time-domain facial EMG feature among eight, (4) Examining fourteen pattern recognition-based classifiers and identifying the best method through extensive assessments. The effectiveness of the signal denoising on system performance is examined while different features and classifiers are applied. This work suggests the most effective, practical and robust combination of methods that delivers the highest performance for recognizing eleven facial states.

The rest of this paper is organized as follows. In the next section, the system overview is sketched and explained. Section 3 describes the materials and methods used for data acquisition, experiments, analysis and evaluation. Following, section 4 presents the experimental results and statistical analyses in which the impact of signal denoising, feature extraction techniques and classification algorithms are inspected. Finally, discussion, conclusion and future works are given in section 5.

2 SYSTEM OVERVIEW

The study structure of this paper is depicted in Fig. 1. Eleven facial expressions are studied to offer a multipurpose interface with eleven degrees of freedom. A robust EMG denoising method comprised of Wavelet transform and Kalman filtering is applied. Eight potentially discriminative time-domain EMG features are considered and compared in order to select the one that leads to the best system performance. Fourteen pattern recognition-based algorithms are employed to classify the features and are compared to identify the best classifier. Final system performance is assessed through three metrics: classification accuracy, processing time and information theory-based method Normalized Mutual Information (NMI). The impact of the EMG denoising stage is investigated not only by SNR measure but also based on the system performance achieved with and without implementing it. The most accurate methods that obtained almost identical system performance are also statistically analyzed with a repeated measure analysis of variance (ANOVA). Eventually, after such extensive analysis and comparisons, the most efficient facial EMG processing pipeline is suggested for designing a reliable and robust multipurpose interface in facial EMG-based MuCIs.

3 METHODS AND MATERIALS

3.1 Data Acquisition

3.1.1 Subject Preparation and Electrode Placement

In this work, facial neuromuscular activities are recorded through three channels from three locations of subjects’ faces. Since superficial electrodes are essential components for measuring the surface EMG signals, they affect the quality of detected signals by reducing the SNR. Accordingly, some precautions are taken into account. The electrodes positions are cleaned with an alcohol pad, and then a conductive gel is used to decrease the electrode-skin impedance. After that, three pairs of rounded pre-gelled Ag-AgCl electrodes are attached on the specified participant’s face in bipolar configuration (with 2 cm inter-electrode distance) to harness the highest energy of EMGs [44]. Several muscles (Zygomaticus major, Masseter, Frontalis, Pars lateralis, Orbicularis oculi, Pars palpebralis, Orbitalis, Corrugator supercili, Depressor supercilii) play a major/minor role in forming the considered facial expressions. The electrodes are placed on the most

---

**Fig. 1. General structure of current study.**
optimum areas in order to cover the muscles and on the longitudinal midline of the major muscles between the muscle fiber innervation zone and the tendon [49]. Therefore, one pair of electrodes (channel 1) is located on the forehead and two other pairs are cited on the left and right lateral face sides (channel 2 and channel 3), as shown in Fig. 1. The ground electrode is placed on the boney part of the left wrist.

3.1.2 System Setup and Recording Protocol
The protocol of this experiment is approved by the University Technology Malaysia Human Ethics Research Committee. In this study, facial EMGs are captured by BioRadio 150 (Clevemed) and they are sampled at ~1000 Hz using a 12 bit A/D converter. Through the activation of a high-pass filter with low cut-off frequency 0.1 Hz and a notch filter 50 Hz, unwanted artifacts from user movements and the power line inference noise are removed by the device software itself. Data are collected from 10 mentally and physically healthy volunteers (who gave informed consent to participate), including five males and five females in the age range of 20-37. The eleven facial expressions considered in this paper are smiling with both sides of lips, smiling with left side of lips, smiling with right side of lips, closing both eyes, closing left eye, closing right eye, clenching the molar, opening the mouth like saying ‘a’ in the word ‘apple’, frowning, raising the eyebrows and keeping the face in a normal condition (neutral state). Subjects are asked to perform each expression for two seconds, five times and with 5 seconds rest between. Since for each expression, 10 (5×2) seconds is informative, and EMGs are captured through three channels simultaneously, a three dimensional data set \([Ch_1; Ch_2; Ch_3]\) is achieved. Thus, for further processing, eleven (number of expressions) data sets of 3×10 seconds (3×10 000 samples) are collected for each subject.

3.2 Signal Denoising using Kalman Filter and Wavelet Transform
As mentioned, facial EMGs have small amplitude, and they are contaminated by external and internal factors like motion artifacts, eye movement, brain activity and recording apparatus itself. Hence, an efficient EMG denoising must be carried out before any further processing. The KF is an efficient adaptive recursive filter that estimates the state of a dynamic system from a series of noisy measurements [50]. In KF, the evolution of the states and the relation between the measurements and states are linear. In addition, the posteriori distribution (state distribution) is Gaussian, so it can be characterized by its mean vector and covariance matrix. The robustness of KF has been already well examined and proved, and it outperformed other adaptive filters like least mean square (LMS) and recursive least squares (RLS) in previous studies [51]. The KF performance in the state space is highly affected by the dimension of the hidden state vector. In order to reduce this dimension, it is recommended [52] to use the discrete wavelet transform (DWT), which decomposes the signal into high and low frequency contained coefficients by incorporating a set of predefined wavelets.

In addition to dimension reduction, DWT also considerably improves the signal denoising [53]. Since EMG is a non-stationary neurophysiological signal, time-variance characteristic of DWT provides excellent joint time-frequency resolution [54]. In the present study, Biorhongonal (bior) 5.5 is selected as the mother wavelet, which is a wavelet where the associated wavelet transform is invertible but not necessarily orthogonal. The important properties of this method are compact support, linear phase, symmetry, regularity and degree of smoothness.

We propose to filter the raw signals by KF in conjunction with DWT. At first, EMGs are band-pass filtered within the range of 60-300 Hz to cover the most significant spectrum of facial EMG signals [55]. Then, EMGs in all channels of each facial expression are separately normalised using the maximum and minimum values of each facial expression. Instead of using raw samples, approximation coefficients of DWT are applied to denote the important frequency spectral components in the signals, which are further smoothed by using KF before reconstruction back to the clean signals. Let \(y_n = [y_{n1},..., y_{nT}]^T\) signifies sequence of \(T\) approximation wavelet coefficients extracted from single-trial measurements at trial \(n\). We assume a state-space model

\[
y_n = x_n + v_n \quad (1)
\]

\[
x_n = x_{n-1} + w_n \quad (2)
\]

The noisy coefficients \(\{y_n\}\) are modeled by a linear additive noise model of the form (1) where \(x_n\) is the state vector of clean wavelet coefficients, hidden in noise \(v_n\) which is \(T \times 1\) i.i.d. Gaussian noise with mean zero and static covariance matrix \(R\), \(v_n \sim N(0, \sigma^2_v I)\). The hidden state \(x_n\) is assumed to follow the first-order Gauss Markov process as in (2) where \(w_n\) is i.i.d. zero mean Gaussian state noise with covariance matrix \(Q\), \(w_n \sim N(0, \sigma^2_w I)\). The denoised \(x_n\) can be extracted as the mean of the filtering density of \(x_n\) conditional on measurement sequence \(y_{1n} = \{y_1, ..., y_n\}\) which can be computed recursively by KF. In this study, the raw signals are parameterized by different levels (to find the optimum level) approximation wavelet coefficients of wavelet transform, which are further denoised using one iteration of KF estimation with \(n=1\) while \(\sigma^2_v = 1\) and \(\sigma^2_w = 0.1\).

3.3 Signal Segmentation
Feeding all signals directly to a classifier is not practical due to the enormous amount of data and some non-informative EMGs. Therefore, signals must be mapped into lower dimension vectors (feature vectors) to highlight the most important properties of EMGs. Accordingly, the data must be segmented into a sequence of time portions for being used in feature extraction. Considering the fact that neuromuscular signals consist of transient and steady states, Hudgins et al. [56] reported that a transient state is capable of classifying the myoelectric features which are extracted from 100 msec after onset with high accuracy. However, a crucial fault of this manner occurs when switching from class to class in real-time control which avoids having multiple degrees of freedom in complex systems. Englehart et al. [3] proved that

steady state data with 128 msec segment length is much more accurate than the transient one for hand gesture classification, and it is more reliable when faster system response is needed. In a generic frame work, long segments result in high computational load and failure in real-time operations, and short segments lead to bias and variance in feature estimation [57]. To address real-time restraints, it is suggested that segment length should be equal or less than 300 mses [56]. It is also stated that the minimum interval between two distinct contractions of motions is almost 200 msec [58] which means the segment longer than 200 msec contains sufficient information for classification. The influence of different segment lengths (50, 100, 150, 200, 300, and 500 msec) on classification of different upper limb EMG features was extensively examined in [45]. They stated that accuracy is not considerably enhanced by increasing the segment length for all single features which supports the idea that a segment with 200 msec length contains enough information. Recently, it was reported that segments with 256 msec length would be the best choice when dealing with facial EMG signals [29]. Thus, in this study EMGs are segmented into non-overlapped windows with 256 msec length prior to feature extraction. Therefore, 39 segments are obtained (10000 ÷ 256 ≈ 39) for each expression in each channel.

### 3.4 Feature Extraction

As noted, classifying large numbers of EMGs is not practical, and it is essential to map them into a lower dimension. In addition, stochastic essence of raw EMG signals necessitates levels of processing in which feature extraction plays the key role to obtain a reliable performance. Feature extraction tries to highlight the significant properties of EMGs and make the raw signals more meaningful for classification. There are numerous methods with a variety of complexity and efficiency in diverse domains (time, frequency, time-scale) which represent different EMG signals characteristics [3], [57]. For the purpose of myoelectric classification, time-domain (TD) features have been widely suggested due to well representation of different motion and gesture characteristics and low computational load in real-time applications [45], [57]. According to Table 1, very few studies investigated and compared facial EMG TD features [35-36]. In [35], six TD features were evaluated only through the obtained classification accuracy while the computational load was neglected. Ten TD features were compared in both terms of accuracy and computational load after classification by VEBFNN [36]; however, the performance was affected not only by the type of feature but also the classifier [45]. Therefore, more extensive evaluation and comparison are required to identify the most discriminative facial EMG feature by considering a range of classifiers.

In this work, the discriminating performance of eight widely used TD EMG features are evaluated and compared. The employed features are integrated EMG (IEMG), modified mean absolute value 1 (MMAV1), modified mean absolute value 2 (MMAV2), willison amplitude (WAMP), simple square integral (SSI), MAV, RMS and MPV. Mathematical definitions of these features are provided in Table 2. Effectiveness of the considered features is evaluated in two scenarios, with and without the proposed signal denoising protocol which also reveals the impact of signal denoising on this stage. Thus, 39 features are extracted from the segmented signals (i.e. for each expression in each channel). By considering three channels, the feature vector is constructed as \([Ch_1; Ch_2; Ch_3]\) for each expression and consequently \([Ch_1; Ch_2; Ch_3]\) for each subject. In order to have more separable feature vectors, log transform is employed to spread the concentrated features when considering the highly scattered ones [59].

#### 3.5 Classification

Extracted features need to be classified into distinct classes for the recognition of facial expressions. Different factors, such as electrode position, muscle fatigue and sweat, cause variations in the EMG pattern over time and result in misclassification; hence, a classifier should be able to cope with these flaws. Moreover, the classifier must be fast and proficient enough to meet real-time restraints and classify the novel patterns during the online training. In the following experiment, different pattern recognition-based classification algorithms are applied on extracted features to classify the facial expressions. In order to select the classifiers, several criteria are taken into account, such as: evaluating different families of pattern recognition algorithms by choosing the widely used technique(s) in each family; capability to cope with the aforementioned prerequisites for EMG classification; high performance reported in previous works and studying new methods in this field. The selected classifiers are employed to evaluate the effectiveness of signal denoising step as well as the discriminating power of TD features, and they are compared to examine their capability for

---

**TABLE 2**

**Mathematical Definition of Features, \( N \) is the length of the segment, \( k \) is the current segment, \( x_i \) is the current point of signal and \( i \) is the index of the current point**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Mathematical Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAV</td>
<td>[ MAV_k = \frac{1}{N} \sum_{i=1}^{N} x_i ]</td>
</tr>
<tr>
<td>MMAV1</td>
<td>[ \text{MMAV}<em>1 = \frac{1}{N} \sum</em>{i=1}^{N} o_x</td>
</tr>
<tr>
<td>MMAV2</td>
<td>[ \text{MMAV}<em>2 = \frac{1}{N} \sum</em>{i=1}^{N} o_x</td>
</tr>
<tr>
<td>RMS</td>
<td>[ \text{RMS}<em>k = \frac{1}{N} \sum</em>{i=1}^{N} x_i^2 ]</td>
</tr>
<tr>
<td>IEMG</td>
<td>[ \text{IEMG}<em>k = \sum</em>{i=1}^{N}</td>
</tr>
<tr>
<td>SSI</td>
<td>[ \text{SSI}<em>k = \sum</em>{i=1}^{N}</td>
</tr>
<tr>
<td>MPV</td>
<td>[ \text{MPV}_k = \max</td>
</tr>
</tbody>
</table>

\( \omega(i) = \begin{cases} 1, & 0.25N \leq i \leq 0.75N \\ 0.5, & \text{otherwise} \end{cases} \)

\( f(x) = \begin{cases} 1, & x > T \\ 0, & \text{otherwise} \end{cases} \)
recognizing the facial expressions. These classifiers are grouped into parametric and nonparametric Discriminant Analysis (DA), Neural Networks, Kernel Machines and Decision Tree method. It must be noted that a wide range of parameter values are examined in order to find the ones that result in the best performance and form the optimum model of each classifier. For this purpose, a five-fold random cross-validation scheme is used to test the parameters and evaluate the classifier performance. In this study, library for support vector machine (LIB-SVM) [60], least square support vector machine laboratory (LS-SVM Lab) [61] and Classification toolbox [62] are utilized.

3.5.1 Parametric Discriminant Analysis

Parametric DA is a classification method that assumes the data from each class follows a multivariate Gaussian distribution (i.e. forms of underlying density functions are known). For the purpose of training a classifier, the parameters of a Gaussian distribution for each class are estimated by the fitting function. Then, the trained classifier chooses the classes of unknown data based on the lowest misclassification cost. This type of classifier can be grouped based on the prior knowledge on probabilistic structure of the problem and distribution parameters estimation routines. In this paper, some popular types of such methods are considered and examined on classification of facial EMG TD features for the first time.

Normal Density Discriminant Function (NDDF). In this method the normal distribution is determined by two parameters, mean and variance, which can be estimated directly from each class pattern [63]. This technique has been successfully used in different applications (e.g. [64]).

Linear Discriminant Analysis (LDA). Linear discrimination fits a multivariate normal density to each class while the model includes identical covariance matrix and different means for each class. Generally, it intends to represent one dependent categorical variable (the class label) as a linear mixture of other features and to model the difference between the classes of data [65]. The success of this technique was clarified on upper limb myoelectric signal classification and control [4].

Maximum-Likelihood (ML) Estimation. In this technique parametric distribution of patterns are classified when the ML algorithm estimates the parameters of Gaussian distribution or mixture of Gaussian distributions for each class. ML chooses parameter values that maximize the probability of the data [66]. ML has many optimal properties in estimation: sufficiency, consistency, efficiency, and parameterization invariance [67]. In the present work, we model the pattern of each class using only one Gaussian distribution.

Expectation Maximization (EM) Estimation. In this method parameters of Gaussian distribution for each class are estimated using the EM algorithm [68]. It has been found in most instances to have the advantages of reliable global convergence, low cost per iteration, economy of storage and ease of programming as well as a certain heuristic appeal [66]. In our case, only one Gaussian distribution is used to model the pattern of each class.

Regularized Discriminant Analysis (RDA). In this method, the parameters of Gaussian distributions are estimated by the Friedman Shrinkage algorithm by making use of a regularized covariance matrix [69]. The purpose of the regularization is to reduce the variance related to the sample-based estimates at the expense of potentially increased bias. The high potential of this classifier for face recognition is reported in [70]. Here, one Gaussian distribution is considered, and the regularization parameter is tuned at $\gamma = 0.1$.

3.5.2 Nonparametric Discriminant Analysis

In most pattern recognition real applications, the common parametric forms rarely fit the densities actually encountered in practice [63]. Nonparametric techniques do not assume a particular form of density function but usually estimate one. In this study, two different types of such classifiers are investigated in the context of facial EMG classification for the first time.

Parzen Windows (PW). Using a given kernel function, this technique approximates a training set distribution via a linear combination of kernels centered on the observed points [71]. In this study, the multiclass PW classifier introduced in [72] is formed by a window width of 0.5.

K-Nearest Neighbor (KNN). KNN is a statistical instance-based technique which supposes that all samples relate to points in the space $\mathbb{R}^n$ [73]. In this method features are classified based on their nearest neighbors’ class, and it rates a pattern by regarding the most similar labeled training samples. Peleg et al. [74] employed KNN as a classifier for EMG signals in finger activation in order to be used for a robotic prosthesis arm. The number of adjacent samples which are taken into account is defined by the parameter $k$ (number of nearest neighbors), and in this study it is set to five.

3.5.3 Neural Networks

Probabilistic Neural Network (PNN). The PW method implemented as a neural network is known as PNN [63]. It is similar in structure to back-propagation and differs primarily in the sigmoid activation function that is replaced by a statistically derived one. PNN has many advantages [75], such as its training speed which is much faster than a BP network and its robustness against noise. This method has proven to be efficient for bioelectric signals [76]. It includes three layers: input (equal to dimension), pattern (features) and category (number of classes) which are 3, 429 and 11 neurons respectively for each subject in this paper.

Reduced Coulomb Energy Neural Network (RCENN). In PW a fixed window is used all over the feature space, and KNN adjusts the region according to the data density. An in-between technique is to adjust the window size during training based on the distance to the nearest datum of a different class. Such region adjustment can be implemented in a neural network called as RCENN [63]. In RCENN, each neuron has an adjustable parameter $\lambda$ that corresponds to the radius of a d-dimensional sphere in the input space which is $\lambda = 0.5$ in this paper.

Radial Basis Function Neural Network (RBFNN). This method consists of instance-based and PNN methods.
It is a three layer neural network with radial basis functions (RBFs) as activation functions in the hidden layer. The network output is a linear mixture of RBFs of the inputs and neuron parameters [77]. According to [38], RBF can provide more satisfactory trade-off between accuracy and speed compared to Multilayer Perceptron (MLP) for EMG-based facial expression recognition. In this study, the input layer contains 3 neurons (equal to feature vector dimensions). The number of neurons in the hidden layer is determined manually in each run in order to obtain the best performance. The output layer includes eleven neurons (number of classes in the training data set) and the neuron parameters known as spread of RBFs is tuned at 0.05.

**Versatile Elliptic Basis Function Neural Network (VEBFNN).** VEBFNN is an extension of RBF and elliptical basis function (EBF) neural networks [78]. The significant benefit of this method is that only one epoch is enough to train the data, and once a datum is learnt it can be discarded which makes VEBFNN very fast and reasonable for real-time applications [78]. In our study, the layers of this network contain: three neurons in the input layer, equal to the dimension of the input dataset; eleven neurons in the output layer, equal to the number of classes in the training dataset; and the hidden layer neurons are not defined in advance and they are formed during the training process. Moreover, a sphere with a radius of 0.5 is considered as a simple hyper-ellipsoidal function.

### 3.5.4 Kernel Machines

**Support Vector Machines (SVMs).** SVM is a nonparametric machine learning algorithm which is originally designed for binary classification [79]. It aims to find discriminant hyperplanes that classify the data to different groups with the maximum possible margin to increase the classifier generalization ability. SVM employs a regularization parameter that allows adaptation to outliers and error on the training set. This classifier has been widely used for EMG classification (e.g. [45]), and its advantages are addressed in [80]. In this paper, a multiclass SVM with a one-versus-one scheme is used along with a kernel function, and the penalty factor is tuned at 1000.

**Least-Square Support Vector Machines (LS-SVMs).** LS-SVM is the modified version of SVM, and it tackles the high degree of computational cost during the training stage in SVM. In this method, equality constraints are used to find the optimal solution by solving a set of linear equations instead of solving a quadratic optimization problem [81]. This classifier has already been examined for facial EMG classification and promising results were reported in [34], [39] and [42]. The LS-SVM model used in this paper is formed using RBF kernel where the regularization parameter is defined as the rate of correct classification in a test set and the latter is used to determine the computational cost for signal denoising and the time consumed to train the classifiers. In addition, an information theory based measure is proposed to be used for more investigation on classification problems [83]. The NMI criterion is an entropy type quantity which provides a measure of the statistical relationship between variables and contains all the statistics of the related distributions. Additionally, it compares the amount of information carried by the classifier output and by the true class labels. It is a good internal measure when classification accuracies of models are equivalent, and it assesses the correctness of classification.

NMI is defined as equation (3) where \( m \) is a total class number, \( c_i \) represents the sample number of the \( i \)-th class that is classified as the \( j \)-th class, \( n \) is a total sample number and \( c_{ij} \) is the total number for the \( i \)-th class.

\[
NMI = -\frac{1}{\sum_{i=1}^{m} C_i \log_2 \left( \frac{C_{ij}}{C_i \sum_{i=1}^{m} c_{ij}} \right)}
\]

where 
\[
C_i \sum_{i=1}^{m} c_{ij} \leq n
\]

Finally, a statistical significant test by a repeated measure ANOVA is carried out to evaluate the most accurate methods that obtained almost identical system performance.

### 4 Experimental Result

#### 4.1 Signal Denoising Evaluation

This experiment explores the impact of signal denoising on the system performance by considering different criteria. As mentioned earlier, for the observation model, the wavelet coefficient is used instead of raw signals. Therefore, wavelet decomposition level plays a key role for signal denoising and has direct influence on the final performance. To find the optimum decomposition level, the SNR of all facial expressions EMGs averaged over all subjects is measured after signal denoising at different levels.

As can be seen in Fig. 2, increase in this level leads to
higher SNR for all signals. This trend is more considerable from level 2 to 3, and it remains almost unchanged from level 4 to 5. Accordingly, level 4 is selected as the optimum level to extract the wavelet coefficients. Fig. 3 illustrates raw and denoised EMG signals in one trial of gesturing ‘a’ in apple for all three channels. It is obvious that the amplitude and frequency of denoised EMGs are smoothed and the background noise is eliminated after denoising. In order to compare the system performance with and without the denoising step, the classification accuracy obtained by all classifiers using different features averaged over all subjects is computed and shown in Fig. 4. Regardless of the type of classifier, feature and exact classification accuracy values, it is very clear that in all cases the performance achieved by denoised signals is much higher than that by the raw signals. The interesting point is that the average system performance when applying different classifiers and features improves considerably from 83.05% to 97.27% after denoising. Moreover, it can be seen that the highest classification accuracy provided by raw features (91.15%) is still lower than the minimum result obtained by denoised features (94.18%). Unlike the raw features, the differences between the classification performances reached by various algorithms over denoised features are not substantial and are in a smaller range of 94.18%-99.31%. Table 3 presents the best classification accuracy provided by all features averaged over all subjects before and after denoising. It is indicated that the results of denoised features are enhanced by at least about 10% with respect to the raw ones. Obviously, the highest accuracy for both raw and denoised features is offered by RMS, with about 8% difference. To clarify this difference, Fig. 5 illustrates raw and denoised RMS feature distribution of one subject in feature space. As can be seen, denoised RMS provides more separable boundaries within facial EMG patterns which results in better discrimination within different classes (Fig. 5(b)), whereas raw RMS features are mixed and difficult to recognize in Fig. 5(a). Accordingly, in this experiment raw features are outperformed, and the robustness of the proposed methodology is proven for facial EMG denoising. Hereafter, for further evaluation, we only focus on the performances provided by denoised features.

### 4.2 Feature Evaluation

It has been shown that various TD features extracted from either upper limb [45] or facial [35-36] EMGs directly affect the system performance. However, in previous studies, their effectiveness has only been examined through the performance achieved by one classifier which is not a fair assessment. In this experiment, the performance of each single denoised feature is assessed by considering different classifiers which reveals a brighter perspective of features efficiency on system performance. Fig. 6(a) depicts that all TD features have high discriminating capability for facial expressions classification since they deliver almost 98% of accuracy by at least one classifier. The maximum (99.31%) and minimum (94.18%) classification accuracy is achieved when RMS and IEMG features are classified by ML and C4.5 respectively. MPV, MAV and MMAV1 also exceed 99% accuracy when in order classified by LSSVM, ML and NDDF.

MAV is the only feature which leads to accuracy of about 99% through three classifiers ML, NDDF and RDA. The highest average performance across all classifiers is delivered by RMS with 97.94 ± 0.87% of accuracy which implies that the EMG pattern generated by this feature

---

**Table 3**

<table>
<thead>
<tr>
<th>Feature</th>
<th>IEMG</th>
<th>MPV</th>
<th>MAV</th>
<th>MMAV1</th>
<th>MMAV2</th>
<th>RMS</th>
<th>WAMP</th>
<th>SSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acc. (%)</td>
<td>R</td>
<td>D</td>
<td>R</td>
<td>D</td>
<td>R</td>
<td>D</td>
<td>R</td>
<td>D</td>
</tr>
<tr>
<td>ML</td>
<td>86.9</td>
<td>98.8</td>
<td>90.9</td>
<td>99.1</td>
<td>87.3</td>
<td>99.2</td>
<td>85.1</td>
<td>99.2</td>
</tr>
<tr>
<td>ML</td>
<td>89.7</td>
<td>99.3</td>
<td>90.2</td>
<td>99.8</td>
<td>90.5</td>
<td>99.9</td>
<td>86.7</td>
<td>98.8</td>
</tr>
<tr>
<td>RDA</td>
<td>86.9</td>
<td>98.8</td>
<td>90.9</td>
<td>99.1</td>
<td>87.3</td>
<td>99.2</td>
<td>85.1</td>
<td>99.2</td>
</tr>
<tr>
<td>LSSVM</td>
<td>89.7</td>
<td>99.3</td>
<td>90.2</td>
<td>99.8</td>
<td>90.5</td>
<td>99.9</td>
<td>86.7</td>
<td>98.8</td>
</tr>
<tr>
<td>ML</td>
<td>86.9</td>
<td>98.8</td>
<td>90.9</td>
<td>99.1</td>
<td>87.3</td>
<td>99.2</td>
<td>85.1</td>
<td>99.2</td>
</tr>
<tr>
<td>ML</td>
<td>89.7</td>
<td>99.3</td>
<td>90.2</td>
<td>99.8</td>
<td>90.5</td>
<td>99.9</td>
<td>86.7</td>
<td>98.8</td>
</tr>
<tr>
<td>NDDF</td>
<td>86.9</td>
<td>98.8</td>
<td>90.9</td>
<td>99.1</td>
<td>87.3</td>
<td>99.2</td>
<td>85.1</td>
<td>99.2</td>
</tr>
<tr>
<td>ML</td>
<td>89.7</td>
<td>99.3</td>
<td>90.2</td>
<td>99.8</td>
<td>90.5</td>
<td>99.9</td>
<td>86.7</td>
<td>98.8</td>
</tr>
<tr>
<td>SVM</td>
<td>86.9</td>
<td>98.8</td>
<td>90.9</td>
<td>99.1</td>
<td>87.3</td>
<td>99.2</td>
<td>85.1</td>
<td>99.2</td>
</tr>
<tr>
<td>ML</td>
<td>89.7</td>
<td>99.3</td>
<td>90.2</td>
<td>99.8</td>
<td>90.5</td>
<td>99.9</td>
<td>86.7</td>
<td>98.8</td>
</tr>
<tr>
<td>NDDF</td>
<td>86.9</td>
<td>98.8</td>
<td>90.9</td>
<td>99.1</td>
<td>87.3</td>
<td>99.2</td>
<td>85.1</td>
<td>99.2</td>
</tr>
<tr>
<td>SVM</td>
<td>89.7</td>
<td>99.3</td>
<td>90.2</td>
<td>99.8</td>
<td>90.5</td>
<td>99.9</td>
<td>86.7</td>
<td>98.8</td>
</tr>
<tr>
<td>LSSVM</td>
<td>86.9</td>
<td>98.8</td>
<td>90.9</td>
<td>99.1</td>
<td>87.3</td>
<td>99.2</td>
<td>85.1</td>
<td>99.2</td>
</tr>
<tr>
<td>RDA</td>
<td>89.7</td>
<td>99.3</td>
<td>90.2</td>
<td>99.8</td>
<td>90.5</td>
<td>99.9</td>
<td>86.7</td>
<td>98.8</td>
</tr>
</tbody>
</table>

---

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2016.2569098, IEEE Transactions on Affective Computing
provides higher discriminative information than others. Fig. 6(b) indicates the distributional characteristics of classification accuracies delivered by each denoised feature across all classifiers in the form of box-plots. This figure illustrates that MPV, MAV, MMAV1 and RMS have no substantial difference in the performance, and they provide more discriminating information than other features since the overall range of classification accuracy is meaningfully higher. In addition, the interquartile ranges of MAV and WAMP are reasonably similar which indicates their same degree of dispersion when being classified by all methods. This equality is also obvious among IEMG, MPV, RMS and SSI but within a smaller range. SSI is shaped in a short box and it suggests that all classifiers obtain almost similar rates of accuracy using this feature. On the contrary, the long spread of accuracy for MMAV2 emphasizes the low accordance of different algorithms to classify this feature. This figure also shows that RMS delivers the highest classification accuracy since its upper quartile peaks the maximum.

Overall, it is obvious that all denoised TD features contain high discriminative information as they can provide high degree of accuracy for classification of facial EMGs. The lowest accuracy reported in this study (94.18%) is higher than the best result reported for single TD features in previous works with ten facial expressions (93.1%) [42]. According to the results, it can be concluded that RMS is the best feature in terms of classification accuracy and contains invaluable information to discriminate different facial expressions EMGs. The classification performance delivered by the denoised RMS feature in form of confusion matrices for all classifiers is provided in appendix B.

4.3 Classifier Evaluation

4.3.1 Accuracy

Classification accuracy is the most important criterion to measure the system performance. Fig. 7 depicts the accuracy obtained by each classifier when different features are used. As mentioned before, the performances of classifiers are significantly improved by applying denoised features rather than the raw ones. Fig. 7(a) shows that the lowest classification accuracy is obtained by C4.5 (94.18%) when IEMG is fed. There is no significant difference among the best performance of PW, RBF, RCE and VEBF

since the maximum accuracies vary between 97.51% and 97.86%. This phenomenon is also observed amongst SVM, EM, KNN, LDA and PNN where the maximum accuracy is around 98.5%. As can be observed, the majority of classifiers including PNN, LDA, RCE, C4.5, SVM and ML obtain the highest accuracy when RMS is used. KNN, RBF, VEBF, EM and LSSVM reach the maximum performance by MPV. PW, RDA and NDDF gain this level of accuracy using SSI, MAV and MMAV1 respectively. The most promising results are acquired by ML when classifying RMS (99.31%) and MAV (99.25%), followed by LSSVM (99.17%) and NDDF (99.19%) through MPV and MMAV1 respectively. This experiment also shows that the average performance of the classifiers on all features is in the range of 95.63%–98.53% where the minimum is obtained by C4.5 and the maximum by ML. Distribution of classification accuracies attained by each classifier across all features is illustrated in Fig. 7(b). It shows the robustness of LSSVM, RDA, NDDF and ML compared to other classifiers as they provide a much higher overall range of accuracy. However, no considerable difference is seen among the performance of ML and NDDF. Unlike PNN and SVM where the performances are scattered in a long range, short boxes of LSSVM and RDA indicate low discrepancy in accuracy which suggests stability of these classifiers across different features. According to the position of upper whiskers and medians, this figure once more confirms that ML is the most accurate method for classification of facial expressions. Its superiority is emphasized as it obtains the highest performance for both raw (shown in Fig. 4) and denoised features. It can be interpreted that facial EMG patterns are likely distributed as Gaussian forms where parameters can be optimally estimated by the ML algorithm.

### 4.3.2 Processing Time

Processing time is counted as another essential criterion for evaluating the system performance. Due to the simplicity of TD features, their computational load throughout the processing is negligible. In this experiment, processing time is comprised of the time needed for denoising the EMGs (white bars) and training the classifiers (colored bars), shown in Fig. 8. As can be seen, all classifiers need less than a second to be trained. For DA methods, training time is less than 0.1 second where PW and ML are the fastest nonparametric and parametric techniques respectively. In contrast, according to this bar chart, SVM is the slowest classifier for training. As described before, the denoising stage on all facial EMGs is carried out in one run for each subject. In this figure, white bars show that denoising of facial EMGs averaged over all subjects takes about 0.09 seconds, which is constant for each classifier. There is no remarkable difference in processing time (about 0.12 ± 0.01 sec) when PW, PNN, KNN, ML, RCE and RDA classifiers are employed, regardless of the type of feature. For NDDF, LDA, EM and LSSVM, the processing time grows from about 0.16 to 0.21 seconds. As demonstrated, C4.5, RBF, VEBF and SVM require higher computational cost for processing, and the maximum time is almost 0.6 seconds for SVM. It must also be noted that the training process is faster when denoised features are used compared to the raw ones. This is because clean features provide more separable boundaries within classes which facilitate the training procedure for classifiers. For example, training time for SVM decreases from about 0.9 to 0.5 seconds when denoised features are utilized. Finally, it can be concluded that although the total processing time is slightly affected by the denoising stage, it is still quite a fast procedure as the time for most of the classifiers is below 0.2 seconds.

### 4.3.3 NMI

NMI is used to measure and validate the classification correctness of different algorithms. As mentioned, it is a robust tool to evaluate the internal model efficiency of classifiers, especially when the recognition accuracy ratio by different algorithms is identical or relatively close. Therefore, this experiment evaluates the classifiers with
average accuracy higher than 99% which are LSSVM, NDDF and ML. According to Table 4, it may be considered that LSSVM and NDDF are equivalent in performance since they result in almost the same accuracy when using denoised features MPV and MMAV1 respectively. However, the higher NMI for NDDF indicates that its misclassified samples are assigned into a small-number label whereas NMI value of LSSVM receives more impact since its misclassified data are assigned into a large-number label. Therefore, the former is preferred in practice. Based on this theory, the model combination of ML and RMS performs better than ML and MAV as its NMI is bigger. The considerable point is that when NDDF classifies MAV, it reaches a higher NMI than LSSVM despite its lower classification accuracy. This clarifies that the samples which cause reduction in classification accuracy for NDDF are not misclassified but are assigned into the rejected classes which does not affect NMI at all. Finally, according to NMI values, the robustness of ML is approved again as the superior classifier in our study.

### 4.3.4 Statistical Analysis

According to classification accuracy, processing time and NMI, the ML method offers the best performance when classifying RMS facial EMG features. To investigate whether this method provides significantly better performance compared to others presented in Table 4, a one-way repeat measure ANOVA is examined where the factor is the classification accuracy and the five levels are the methods in Table 4. A significant level is set at $P = 0.05$. This test reveals no significant differences for the main factor as it returns an $F$ value of 1.359 and $P = 0.267$. Therefore, based on ANOVA we may conclude that all methods provided in Table 4 perform similarly.

### 5 Discussion and Conclusion

A comprehensive offline study on facial neuromuscular signal is carried out in order to find a procedure that can provide the best performance to recognize eleven facial expressions to be applied in MuCI systems. The major goal is to propose an efficient facial EMG processing pipeline to be implemented easily by providing fast and accurate performance. In addition, a robust methodology is proposed for facial EMG denoising including Wavelet transform and Kalman filtering, and it is shown that this method significantly enhances the classification performance though minor computational cost is imposed on the system. In the meantime, the optimum decomposition level of wavelet transform which directly influences the quality of facial EMGs is found. As a key component of MuCI systems, this study offers an improved interface with eleven degrees of freedom. This is fulfilled through the activation of neuromuscular signals generated from nine facial muscles. Investigation on eight TD features clarifies that denoised features provide much better boundaries within different facial expression patterns and raw features are completely outperformed. Among all TD features, RMS is the most discriminative and informative feature as it leads to 99.31% classification accuracy which is in accordance with the results in [28-32], [35] and [42].

Comparison on fourteen classifiers reveals that parametric DA techniques have the highest capability to classify facial TD features, especially when the parameters of Gaussian distributions are estimated by the ML algorithm. This result is also confirmed from information entropy point of view by means of NMI measure. In terms of processing time, DA methods, either parametric or nonparametric, perform very fast as they only need less than 0.1 second to be trained by denoised features. ANOVA statistical analysis shows that there is no significant difference among the system performance reported for the top five methods in our study.

In comparison with the previous studies that used single features, the best result of this paper outperforms them all. Compared to [39], classification accuracy is improved by 2.2%, and processing time is reduced about 0.6 second which again emphasizes the usefulness of EMG denoising rather than just using multi-feature sets. It must be highlighted that in this study the number of classes is also increased to eleven which affects the system performance. As expected, this phenomenon is observed for raw features; for instance, in [36], classification accuracy obtained by VEBF when classifying MPV was 87.1%; whereas here, the result of this combination is reduced to 85.27% which is due to the increment in class numbers. The methodology proposed in this paper can also be employed when interfaces with lower numbers of control commands are needed, and better results can be expected as reported in [32].

Compared to image/video based facial expression recognition systems, more promising performances can be expected from EMG-based recognition systems. Several problems are observed during experiments using images and video, such as undetectable weak or moderate affective facial responses [84]. For instance, for subjects with insufficient muscles control or for the movements with small changes in facial expression (like clenching the molars), the camera-based methods are ineffective because of their low sensitivity. Using EMG techniques, even the weakest responses can be detected as most facial muscles are very close to the surface electrodes. Another problem is that small dynamic transitions in activity may be less well observed since they may be masked by the stiffness of overlying cutaneous and subcutaneous tissues [85]. While due to great sensitivity in amplitude domain, EMGs also have a good time resolution so rapid changes in activity can be reliably measured. In addition, video-based systems require a high speed image processing hardware; thus, the overall cost of system is much higher.
than EMG-based technique. Other drawbacks are slow communication rate, detecting small number of facial gestures, being sensitive to background, camera position and lightening.

Our study shows how efficiently signal denoising can improve the system performance with a negligible increase in the computational cost. Besides, RMS has high discriminating power to provide clear boundaries within different facial expressions EMG patterns. The DA method, when the parameters are estimated through ML, plays a pivotal role for classification in facial neuromuscular-based control systems, especially the complex ones. As a limitation of our study, because of the low number of subjects along with the absence of patients, we cannot safely generalize our findings and conclude an effective MuCl approach. Small interindividual variability can be expected when analyzing other subjects facial EMGs which is partly due to differences in the morphology of the facial musculature [85]. In the future, the proposed method will be employed on the data from more subjects, including locked-in patients who suffer from crucial disabilities. In addition, an extensive statistical analysis is needed to study the differences among the repetitions of the facial expressions. This requires recordings with more repetitions of each facial expression. Besides, the homogeneity of each facial expression in the within and between subject analysis must be examined. This technology, if confirmed in a larger population of normal subjects as well as patients, could pave the way for the use in a clinical context. Besides, since the ultimate goal of MuCl is to provide an efficient online communication between the human muscle and computer, the proposed interface must be assessed for real-time processing and online applications, such as controlling assistive devices like wheelchairs.

ACKNOWLEDGMENT
This research project is supported by Center of Biomedical Engineering (CBE), Universiti Teknologi Malaysia research university grant (RJ130000.7845.4L840) and funded by Ministry of Higher Education (MOHE).

REFERENCES
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAFFC.2016.2569098, IEEE Transactions on Affective Computing


Mahyar Hamedi received the BSc degree in biomedical engineering from Science and Research Branch, Islamic Azad University (SRIAU), Tehran, Iran, in 2008, and the MSc degree in biomedical engineering from Universiti Teknologi Malaysia (UTM), Malaysia, in 2012. He is now a Ph.D. Candidate at the Center for Biomedical Engineering (CBE), UTM, and working on the advanced biomedical signal processing for human/brain machine interfacing applications. Particularly, EEG and EMG signal modelling, denoising, feature extraction and classification. His current research interests include BCI, neuro-engineering, affective computing, computational neuroscience, HMI and pattern recognition.

Sh-Hussain Salleh received the PhD degree in speaker recognition from University of Edinburgh, U.K. in 1997. He had been a lecturer with Faculty of Electrical Engineering, UTM where he was appointed as Head of Microelectronic and Computer Engineering Department (1999-2004). He was later appointed as Dean of Faculty of Biomedical Engineering and Health Science, UTM (2007-2009).

He currently serves as a professor of biomedical engineering and the Deputy Director of CBE, UTM. His research interests include biomedical signal processing and instrumentation, biometrics and speech synthesis.

Chee-Ming Ting received the PhD degree in Mathematics from the Universiti Teknologi Malaysia (UTM), Malaysia, in 2012. He was a Postdoctoral Fellow at the CBE, UTM, where he is currently a Research Fellow. He is also a Senior Lecturer at the Faculty of Biosciences and Medical Engineering, UTM. His research interests include statistical signal processing, spatiotemporal modeling, state-space methods, time series analysis, and high-dimensional statistics with applications to speech and biomedical signal processing and classification.

Mehdi Astaraki received the BSc and MSc degrees in biomedical engineering from Science and Research Branch of Islamic Azad University (SRIAU), Tehran, Iran, in 2009 and 2013, respectively. His research interests include medical image segmentation and registration, biomedical signal processing and pattern classification.

Alias Mohd Noor received the PhD degree at Bath University in 1990. He was appointed various academic administrators at Faculty level such as Head Department, Deputy Dean (Academic) and Dean of Mechanical Engineering Faculty. He was further appointed as Deputy Vice Chancellor (2003-2009), Dean of Research Alliance (2009-2012). Currently he is the Director of CBE at UTM. His research activities are on the areas of mechanical and BioMedical Engineering.