ORIGINAL RESEARCH

Human facial neural activities and gesture recognition for machine-interfacing applications

M Hamedi[1]
Sh-Hussain Salleh[2]
TS Tan[2]
K Ismail[2]
J Ali[3]
C Dee-Uam[4]
C Pavaganun[4]
PP Yupapin[5]

[1]Faculty of Biomedical and Health Science Engineering, Department of Biomedical Instrumentation and Signal Processing, University of Technology Malaysia, Skudai, [2]Centre for Biomedical Engineering Transportation Research Alliance, [3]Institute of Advanced Photonics Science, Nanotechnology Research Alliance, University of Technology Malaysia (UTM), Johor Bahru, Malaysia; [4]College of Innovative Management, Valaya Alongkorn Rajabhat University, Pathum Thani, [5]Nanoscale Science and Engineering Research Alliance (N'SERA), Advanced Research Center for Photonics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

Correspondence: PP Yupapin
Nanoscale Science and Engineering Research Alliance (N'SERA), Advanced Research Center for Photonics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
Tel +66 2 329 8414
Fax +66 2 329 8410
Email kypreech@kmitl.ac.th; isoqt@kmitl.ac.th

Abstract: The authors present a new method of recognizing different human facial gestures through their neural activities and muscle movements, which can be used in machine-interfacing applications. Human–machine interface (HMI) technology utilizes human neural activities as input controllers for the machine. Recently, much work has been done on the specific application of facial electromyography (EMG)-based HMI, which have used limited and fixed numbers of facial gestures. In this work, a multipurpose interface is suggested that can support 2–11 control commands that can be applied to various HMI systems. The significance of this work is finding the most accurate facial gestures for any application with a maximum of eleven control commands. Eleven facial gesture EMGs are recorded from ten volunteers. Detected EMGs are passed through a band-pass filter and root mean square features are extracted. Various combinations of gestures with a different number of gestures in each group are made from the existing facial gestures. Finally, all combinations are trained and classified by a Fuzzy c-means classifier. In conclusion, combinations with the highest recognition accuracy in each group are chosen. An average accuracy >90% of chosen combinations proved their ability to be used as command controllers.

Keywords: neural system, neural activity, electromyography, machine learning, muscle activity

Introduction

The interaction between human and computer or machine is of great importance for various fields such as biomedical science and computer, electrical, electronic, or mechanical engineering as well as neuroscience. The most significant reasons of the development of this technology are its efficiency and usability for handicapped and elderly people. The human–machine interface (HMI) is an approach for information transmittal where humans interact with the machine. Designing such interfaces is a challenge and requires a great deal of work to make the interface logical, functional, accessible, and pleasant to use. The most popular HMI modes still rely on keyboard, mouse, and joystick. In recent years, there has been remarkable interest in introducing intuitive interfaces that recognize the user's body movements and translate them into machine commands. The two main methods of designing these interfaces are techniques based on nonbiomedical or biomedical signals. The "tonguepoint" as well as "headmouse" are examples of nonbiomedical signal devices.[1] These approaches are restricted to users with cerebral palsy or spinal vertebrae fusion.[2] Patients with severe multiple sclerosis and spinal cord injuries who have limitations in neck movement are unable to use these devices.[3] An

International Journal of Nanomedicine 2011:6 3461–3472

3461

eye-gaze tracking method has been proposed[4] to provide an accurate cursor. With these systems, a video camera is usually used to record images or videos continuously and image-processing methods are applied to analyze the data. In spite of their usability, there are some disadvantages such as limitation of camera field of view, complexity, computational cost, target size, peripheral conditions, lighting, and image quality. Recently, automatic recognition of facial expressions, which has a direct effect and a vital role in HMI, has been considered. In 1999, image-based facial gesture recognition was proposed[5] for controlling a powered wheelchair. However, the gesture recognition system required costly high-speed image-processing hardware.

As an alternative, biomedical signals have been employed as other kinds of interfaces in many fields of HMI. These signals can be utilized for neural connection with computers and they are obtained from tissues, organs, or the cell system. Electrooculography (EOG), which measures and determines the eye position electrically, was used as a powerful input device. EOG has been applied to control a cursor,[6] but these systems were affected by head and muscle movement, signal drift, and channel crosstalk. Electroencephalography (EEG) has also been considered as a neural linkage to communicate with the computer/machine and is typically called the brain–computer interface (BCI). These systems have been recently developed by removing the requirement for involvement of peripheral nerves and muscles in disabled users[7,8] by applying the modulation of mu and beta rhythms via motor imagery to make BCI control possible. Apart from its benefits, there are some issues with the use of EEGs such as long training time, poor signal-to-noise ratio, and numerous challenges in processing steps.[9,10]

More recently, electromyography (EMG) has become the new basis of biosignal information for designing HMI/human–computer interfaces (HCI). EMG signals are captured from body muscles such as arm or facial muscles and are translated and converted into machine-input control commands. Much work has been done in the field of interface designing in EMG-based HMI systems, which have been applied to upper and lower limb gestures. Two-dimensional pointers are controlled through wrist movements[11] and an EMG-controlled pointing tool was proposed and developed.[12,13] An HMI based on EMG arm gestures within three-dimensional space design was proposed and the final recognition result of this was ~96%.[14] An approach with three hand gestures was investigated to control a mouse cursor and the overall accuracy was 97%.[15] Controlling a hand prosthesis using hand gestures and pattern recognition had around 90% accuracy.[16,17] Control of virtual devices through hand-gesture EMGs[18] and an online EMG mouse-controlling computer cursor with six wrist gestures were selected as control commands and the reported average recognition ratio was 97%.[19] Hand and finger gesture recognition has been explored with an accuracy of almost 93%.[20] More accurate hand-gesture recognition with an accuracy of 99% has been attained.[21] All these interfaces were designed based on hand, finger, wrist, and arm gestures. However, this would be unsuitable for people with both lower and upper extreme impairment, most of whom cannot even move their neck. These can be seen as the most difficult cases, as, for these patients, their only way of communication is by facial expression. For these cases, the interface design solution must therefore be based on facial gestures. Some effort has been made to utilize EMG-based facial gestures to develop proper interfaces. The first was an intelligent robotic wheelchair based on EMG gestures and voice control.[22] The designed interface had three basic facial expressions: happy, angry, and sad.[23] An EMG-based HMI in a mobile robot was tested by asking volunteers to perform ten eye blinks with each eye.[24] Moreover, useful control commands from facial gestures during speech have been provided and investigated for HCI application.[25] In the same year, a hands-free wheelchair controlled through forehead myosignals was suggested.[26] Further, recently, a virtual crane training system controlled through five facial gestures was introduced[27] and a facial EMG-based interface has been proposed and used in maximum medical improvement through three facial EMGs.[28]

The most important facts to be considered throughout the design process are analysis procedure and the flexibility and user-friendliness of these systems. Since real-time control includes processing issues, simplicity, low-cost computation, high speed, and accuracy of analysis are vital factors. Besides, flexibility of these systems mostly relies on the number of control commands; that is, flexibility increases when the number of control commands increases. In addition, the user-friendly systems are more pleasant for the operators since they can choose suitable gestures in each application with a different number of control commands. Therefore, to apply these interfaces in a real HMI system, all of the above issues must be considered to achieve best performance.

Firstly, EMG-recording protocols, preprocessing, and processing of acquired data must be considered the main focus for designing these interfaces. Amongst these, feature extraction and data classification have been always under investigation to optimize the whole procedure and to obtain the best results. As reported previously, due to EMG

characteristics, processing methods are restricted, especially in low-level contraction like in facial muscles, where the firing rate and frequency bandwidth of muscles are almost the same. In the field of EMG feature extraction, time-domain features such as mean absolute value (MAV), root mean square (RMS), and variance (VAR) are chosen and applied in much research since they provide a better estimation of EMGs[29,30] compared with frequency-domain and time frequency-domain techniques, which are mostly applied in muscle fatigue investigations.[31,32] The proper classifier must be accurate and fast enough to attain the real-time prerequisites. There are many tactics available for EMG classification such as neural network, fuzzy, and probabilistic. Suitable classifiers are chosen depending on the patterns that need to be trained, specific applications, and muscle types. The MAV of EMG was extracted and the gestures were discriminated via a threshold with 93% accuracy,[22] and, by combining maximum scatter difference, RMS, as well as power spectrum density features with minimum distance method, 94.44% accuracy was achieved in classification.[25] Facial gestures were classified into two classes via thresholding tactic and obtained 95.71% recognition ratio.[24] Moving RMS features and linear separation technique were used to identify three facial expressions[25] and they achieved 100% recognition accuracy. This 100% accuracy was achieved by employing MAV and support vector machine (SVM) as features and classifier methods, respectively.[26] A control accuracy of 93.2% was reported when RMS features and subtractive fuzzy C-mean classifier were exerted to distinguish five facial gestures as control commands.[27] In the same year, moderate results with accuracies of 61%, 60.71%, and 56.19% were achieved by employing mean, absolute deviation, standard deviation (SD), and VAR features and K-nearest neighbors, SVM, and multilayer perception classifiers.[28] In recent research, RMS features were extracted from eight facial gestures, and the accuracy of two popular classification methods, support vector machine (SVM) and FCM in the field of facial EMG classifications, were compared. Results reported the greater strength of FCM over SVM and confirmed the usefulness of the RMS method in EMG-signal feature extraction.[33,34]

All previously mentioned work was restricted by two issues. Firstly, they used a fixed number of facial gestures for their purpose, which means that their interface was not flexible enough to be employed by other applications. Secondly, the recognition ratio results relied strictly on included facial gestures and there was a possibility of an increase in performance accuracy by applying other gestures. In this paper, the design of a multipurpose interface is proposed based on eleven voluntary facial gestures. This makes the interface much more flexible since each of the included facial gestures can be used as an accurate input control in various HMI applications where 2–11 control commands are needed. Furthermore, this interface became more reliable through creating different combinations with a different number of facial gestures to find the most accurate combinations. Finally, this interface finds the maximum recognition accuracy for each facial combination consisting of 2–11 facial gestures. The design of the proposed interface is described step by step in the following sections.

Materials and methods

The procedure of designing the proposed interface is shown in Figure 1. Figure 1 is divided into six main parts. Figure 1A shows subject preparation, site selection, and electrode placement. System setup, data acquisition, and signal-recording protocol from all participants and signal-filtering processes are demonstrated (Figure 1B). Data segmentation, windowing, and feature extraction are carried out (Figure 1C). The use of threshold lines and achieved features from the previous step are sorted (Figure 1D) to find active features that are more suitable for training and classification. The multipurpose interface is created by including all possible combinations from all existing facial gestures (Figure 1E). All obtained combinations were trained and classified (Figure 1F) to find the most accurate group of facial gestures and a better distribution of them.

Subject preparation, site selection, and electrode placement

Ten healthy volunteers participated as the main data sources for this work. Figure 2 shows all facial gestures, which are performed by each participant. There are some points that must be considered prior to signal recording for all participants. Since the surface EMG method is used in this work, skin preparation is a significant task for getting a clear signal. An alcohol pad is used to clean any dust and sweat from the selected areas for signal recording to avoid any unwanted artifacts and noise. Conductive and adhesive paste or cream is used on the center of the electrodes before placing them on the skin. Since three recording channels are used, three pairs of electrodes are placed on the desired regions in a bipolar configuration. They are placed on the affective muscles involved in chosen facial gestures/expressions.

Table 1 indicates all facial gestures as well as their effective muscle(s)[35] and recording channel(s) applied in the current work. According to Table 1, three pairs of electrodes
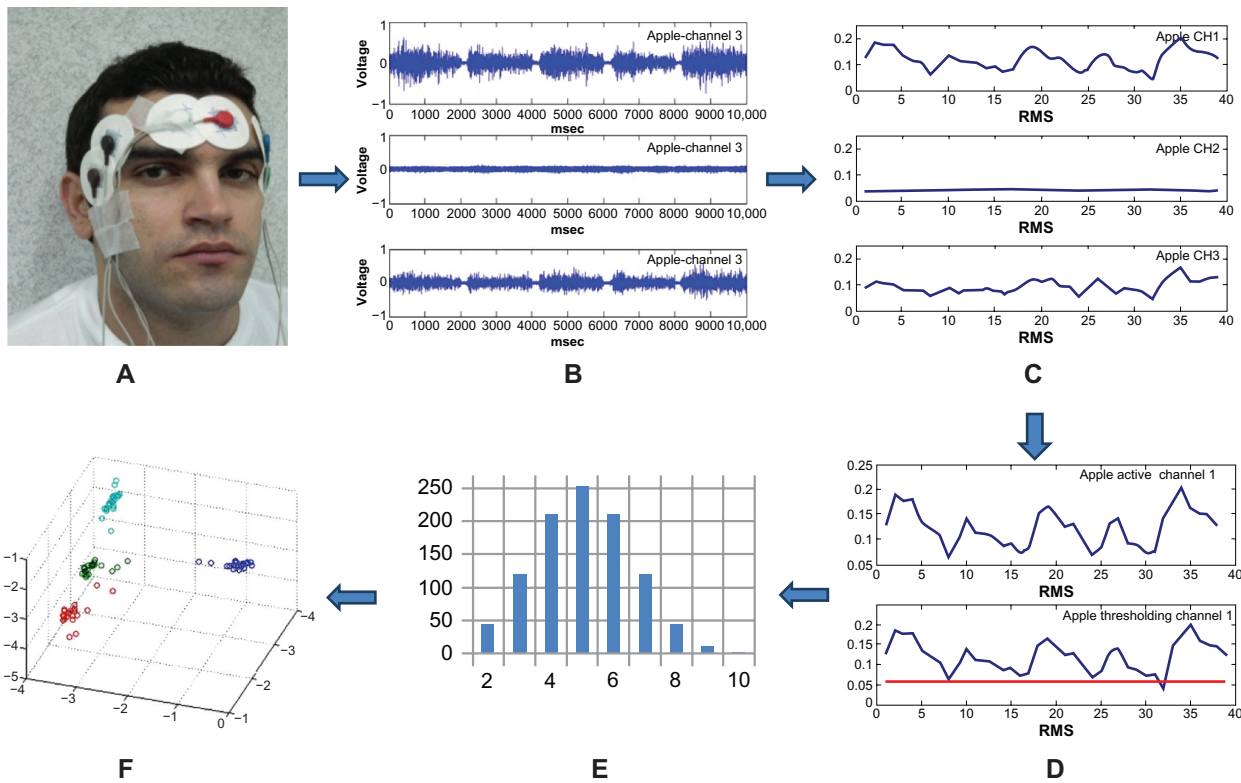
Figure 1 Procedure of designing the proposed interface. **A** shows subject preparation, site selection, and electrode placement. System setup and data acquisition, signal recording protocol from all participants, and signal filtering process are demonstrated in **B**. Data segmentation, windowing, and feature extraction are carried out in **C**. The use of threshold lines to find the active features, which are more suitable for training and classification, is shown in **D**. The multipurpose interface is materialized by making all possible combinations from all existing facial gestures (**E**). All obtained combinations are trained and classified as illustrated in **F** to find the most accurate group of facial gestures and a better distribution of them.

must be placed where all considered gestures could be covered. Two pairs of electrodes (channel 1, channel 3) are located on the left and right temporalis muscle, the other pair is placed on the frontalis muscle, and one ground electrode is sited on the bony part of the left wrist. Each of these pairs is attached within a 2 cm interelectrode distance.[26,27,33,34] After electrode placement and wire attachment, all wires are taped to the face to decrease any unwelcome wire movement and
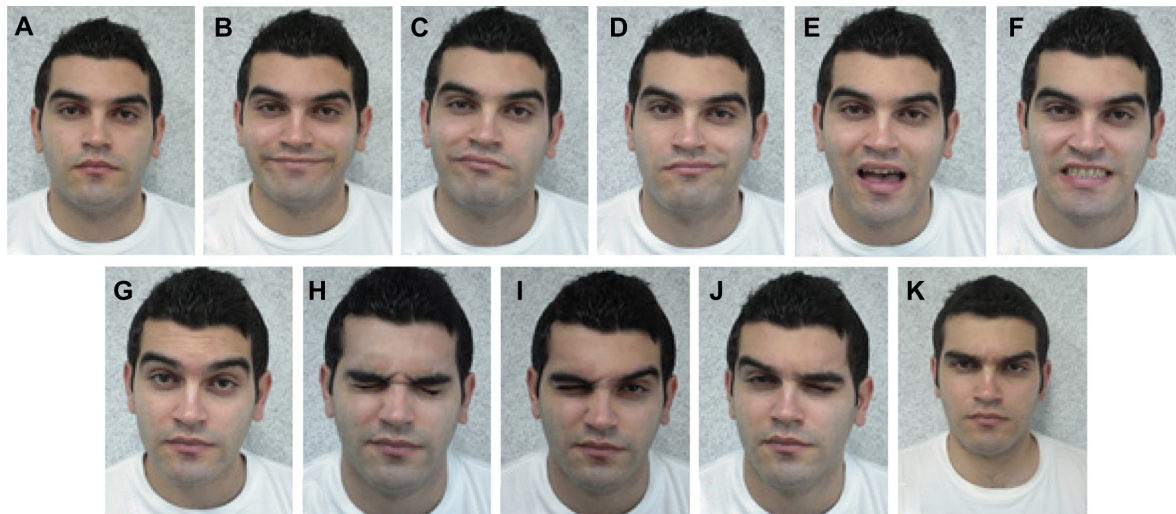

Figure 2 All considered facial gestures in this work: (**A**) natural, (**B**) smiling, (**C**) smiling with right side, (**D**) smiling with left side, (**E**) open the mouth like saying "a" in "apple," (**F**) clenching molar teeth, (**G**) pulling up the eyebrows, (**H**) closing both eyes, (**I**) closing right eye, (**J**) closing left eye, (**K**) frowning.

Table 1 Facial gestures used in this work

Gesture	Gesture name	Effective channel(s)	Effective muscle(s)	
			Major	**Minor**
1	Smiling	1, 3	Zygomaticus major	Levator anguli oris
2	Smiling with right side (pulling up the right corner of lip)	3	Zygomaticus major	Levator anguli oris
3	Smiling with left side (pulling up the left corner of lip)	2	Zygomaticus major	Levator anguli oris
4	Rage (clenching molar teeth)	1, 3	Masseter	Zygomaticus major
5	Gesturing "NO" (pull up the eyebrows)	2	Frontalis, pars lateralis	Pars medialis, Levator palpebrae superioris
6	Opening the mouth like saying "a" in "apple"	1, 3	Pterygoids, digastric	Masseter
7	Closing both eyes	1, 2, 3	Orbicularis oculi, pars palpebralis, orbitalis	Frontalis, temporalis
8	Closing left eye	1, 2	Orbicularis oculi, pars palpebralis, orbitalis	Frontalis, temporalis
9	Closing right eye	2, 3	Orbicularis oculi, pars palpebralis, orbitalis	Frontalis, temporalis
10	Frowning	2	Corrugator supercilii, Depressor supercilii	Frontalis
11	Natural	1, 2, 3	–	–

reduce the number of artifacts. Electrode placement and wire positions are indicated in Figure 3.

System setup and signal-recording protocol

Surface EMG signals are recorded by the BioRadio 150 (CleveMed, Cleveland, OH). The sampling frequency is adjusted at 1000 Hz. To avoid motion artifacts and power-line interference, the low cut-off frequency of the filter is set at 0.1 Hz and a notch filter is applied to remove the 50 Hz component of the signal. Before signal recording, all volunteers rest for 1 minute. Then, they are asked to perform the facial gestures (see Table 1) for 30 seconds (five trial performances over 2 seconds with a 5-second rest between them to eliminate the effect of exhaustion). This procedure is repeated for all participants to record
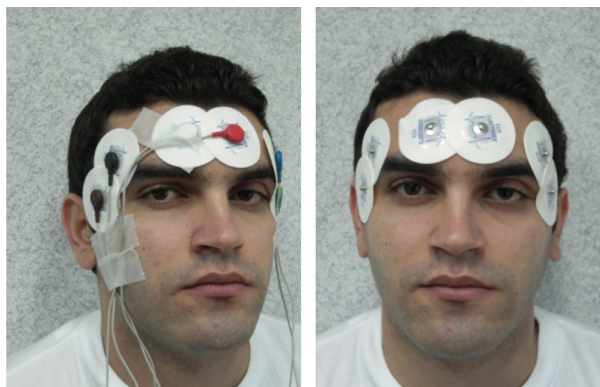
Figure 3 Electrode and wire positions.

each facial gesture. Acquired signals are saved to a personal computer and are prepared for the next stage. An example of the acquired signal of gesturing "a" in "apple" is depicted in Figure 4.

Each recording channel has a different shape based on the activity of the muscles involved in the related gesture (Figure 4). Here, this gesture just affects the recording channels 1 and 3. Channel 2 is inactive and its shape looks like a normal signal. Please note that 20,000 ms parts of all signals that belong to resting stages have been removed and just 2000 ms of active parts are placed next to each other to make continuous signals without rest for a total length of 10,000 ms. All recorded signals are passed through a band-pass filter with a 30-to-450 Hz bandwidth to include the most essential spectrum of EMG signals.[27]

Data segmentation and feature extraction

Filtered signals are segmented into nonoverlapped 256 ms time sections and RMS value is computed by Equation 1 for each section. RMS is a well-known feature to identify the strength of muscle contraction and it offers the extreme probability approximation of amplitude when the signal is modeled as a Gaussian random procedure.

$$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N} X_n^2} \qquad (1)$$

X_n is the achieved raw signal and N is the length of X_n. Therefore, there are 39 features in each channel for each gesture.
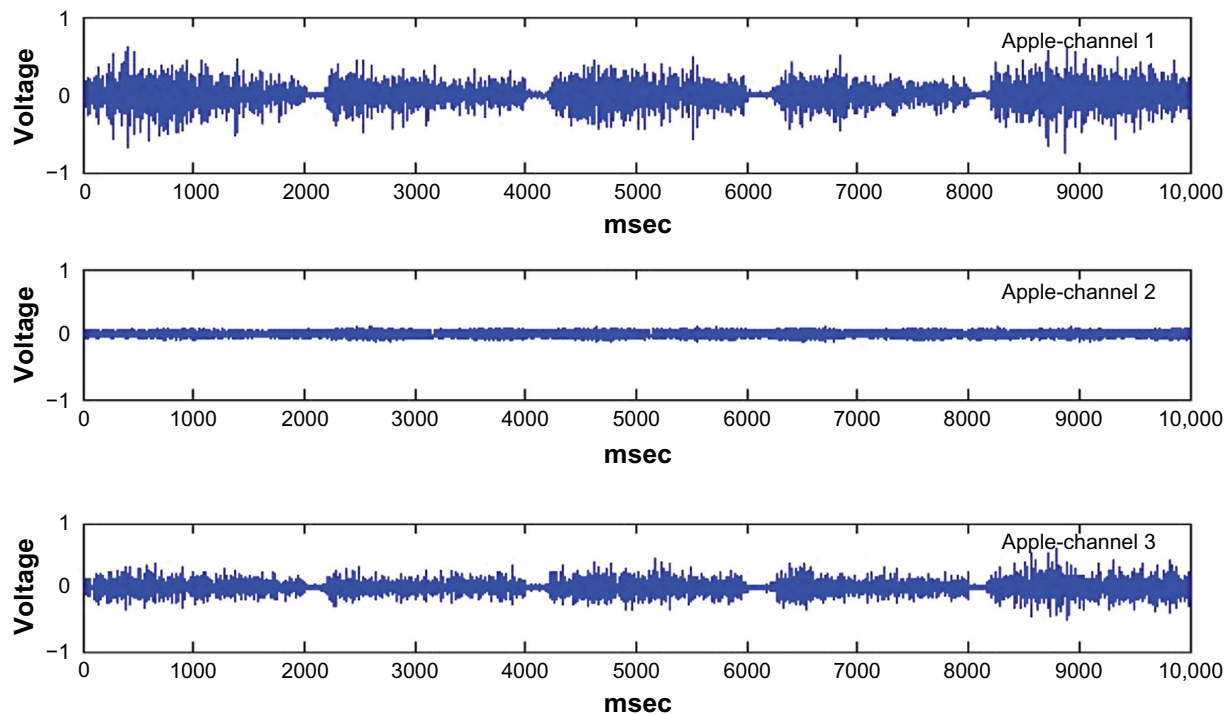
Figure 4 Raw signal of gesturing "a" in apple.

Active feature selection

After feature extraction (RMS), these features are sieved to collect the active ones. Thus, three threshold lines are designed through equation 2 from three normal signal channels (T_1, T_2, T_3).

$$T_x = MEAN\left(RMS\left(X_{normal}\right)\right) + \left(3 \times std\left(X_{normal}\right)\right) \quad (2)$$

Therefore, RMS from each gesture in channel 1, channel 2, and channel 3 are compared with T_1, T_2, and T_3, respectively. The features greater than the threshold are identified as active. As an example, RMS, threshold line, and active features from gesturing "a" in "apple" in channel 1 are shown in Figure 5.

Multipurpose interface design

The main goal of the present work is to design a multipurpose interface for various HMI systems such as controlling assistive devices and artificial limbs. As mentioned earlier, facial gestures are considered an interface in this work. The number of input control commands can be different for controlling various devices, machines, robots, and prostheses. For instance, five input commands can be enough (forward, backward, left, right, and stop) to control a wheelchair. In other devices, the input control commands can be more or less. In this work, eleven facial gestures were chosen and each plays an input command role. Various

devices that need 2–11 input control commands are supported by this interface. In another words, the proposed interface is designed for several applications. This interface becomes practical by including all combinations of facial gestures and assigning them to different groups that include 2–11 gestures. The groups are trained and classified to find the best combination with the highest recognition accuracy in each group. For instance, if the supposed machine needs two movements, the best two facial gestures with the highest recognition accuracy are chosen. All combinations of all facial gestures with different number of gestures are made as shown in equation 3:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (3)$$

where, n is the number of all facial gestures and k is the number of gestures in each group. Based on this equation, the number of all existing combinations in each group for each participant is indicated in Table 2. There are 1013 combinations available to each person.

Normal signal is not counted as a facial gesture and is excluded from the combinations. In each application, normal signal needs to be considered as a command input and there is no need to classify it apart from other facial gestures. The normal signal or rest condition does not produce any signals compared with other signals and it is defined as a
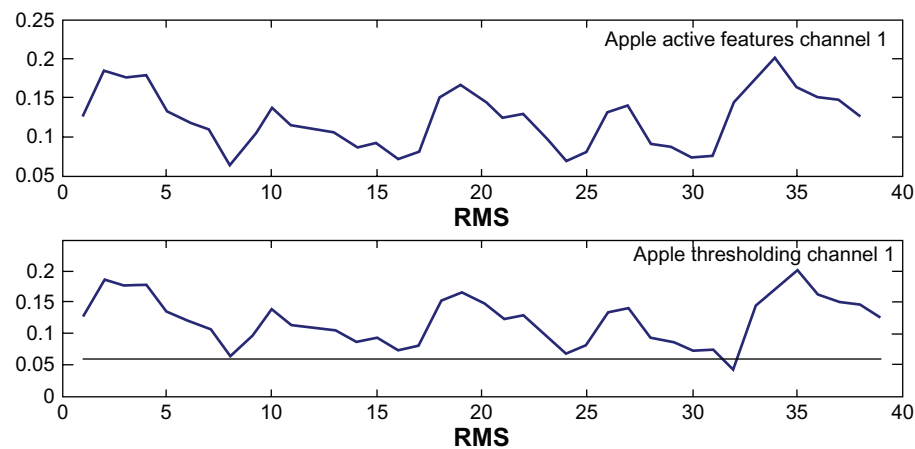
Figure 5 The bottom plot determines the root mean squares (RMS) of gesturing "a" in apple word and the threshold line and above plot indicates the active RMS.

neutral-input command, which is usually used to disable the devices. For an interface with nine control commands, there are 10 combinations of facial gestures available, and one of them has the highest classification accuracy (Table 2). Suppose that each facial gesture is represented by a number (Table 1). Table 3 shows all combinations of nine facial gestures.

Classification

In this step, all combinations must be trained and classified for each person (Table 2). The best combination with the highest recognition accuracy in each group is selected. Each combination contains a different number of facial gestures' active RMS. These features must be classified and separated from each other as much as possible. There are some concerns with EMG classification such as electrode-dislocation responsiveness and noticeable interruption in real-time control and the classifier must cope with them all. In this work, Fuzzy c-means clustering (FCM) is employed for its capability, simplicity, importance, and its wide use in EMG processing and classification.[27,33,34] The most significant characteristic of FCM clustering is that it permits data to belong to two or more clusters and that makes it more flexible than other methods. Class labels always deliver a convenient direction throughout the training procedure, as they do in all supervised training methods. This idea led to the development of a new mode of FCM called "supervised Fuzzy c-means." In this method, the numbers of cluster centers are

given before training as well. By utilizing the obtained active features as input data and adjusting other initial parameters, FCM can compute the location of each cluster center (V_i) and the membership value of each data toward each cluster (U_{ik}). This procedure is repeated until the optimized value of V_i and U_{ik} are found. This technique has already proven its ability to recognize eight facial gestures[33] and it is applied here to design the proposed interface.

Results

To design a multipurpose interface, flexible sets of control commands are required for various applications. The best combinations of facial gestures are assigned to different usages where 2–10 control commands are needed. All 1013 combinations obtained from each participant are trained and classified by FCM in the last stage, and the best combinations containing 2–10 facial gestures with the highest recognition accuracy are chosen. Two criteria are considered to find the best combinations: the first is the classification and recognition performances and the other is the way they are distributed and discriminated in feature space. Each combination of facial gesture can be recognized by its gesture's labels (Table 1). For instance, combination code 7652 includes the four gestures: closing eyes, gesturing "a" in "apple"; "No," pulling up eyebrows; and smiling with right side. All classification results, average, and SDs of all ten participants with different numbers of facial gestures were collected (Table 4).

Table 2 Number of combinations in each group of facial gestures per participant

Number of facial gestures in each combination	2	3	4	5	6	7	8	9	10	Total
Number of combination(s)	45	120	210	252	210	120	45	10	1	1013

Table 3 All combinations of nine facial gestures

Combination	1	2	3	4	5
Facial gestures	1, 2, 3, 4, 5, 6, 7, 8, 9	1, 2, 3, 4, 5, 6, 7, 8, 10	1, 2, 3, 4, 5, 6, 7, 9, 10	1, 2, 3, 4, 5, 6, 8, 9, 10	1, 2, 3, 4, 5, 7, 8, 9, 10
Combination	6	7	8	9	10
Facial gestures	1, 2, 3, 4, 6, 7, 8, 9, 10	1, 2, 3, 5, 6, 7, 8, 9, 10	1, 2, 4, 5, 6, 7, 8, 9, 10	1, 3, 4, 5, 6, 7, 8, 9, 10	2, 3, 4, 5, 6, 7, 8, 9, 10

Combinations of two facial gestures

According to Table 2, 45 combinations can be made by just two facial gestures for each participant. After training, all of these combinations are distinguished perfectly with 100% recognition accuracy in classification. In other words, results show that the robustness as well as accuracy of all combinations for all participants is identical and the SD is 0. Therefore, based on the application, each of these 45 combinations can be considered as input commands. Table 4 provides the results of the best combinations for all participants. Most of these combinations are well distributed in feature space as well. Figure 6A illustrates the distribution of the combination code 56 of the first participant.

Combinations of three facial gestures

One hundred and twenty existing combinations for each participant were trained and classified. More than 95% of these combinations for each participant reached 100% classification accuracy. The average result and SD were 100% and 0, respectively. Nevertheless, the work was accomplished with the best combinations where the three classes were well discriminated. For the first participant, combination code 532 is chosen as the best among all and it is illustrated in Figure 6B. Table 4 reports the results of the best combinations for all participants.

Combinations of four facial gestures

Two hundred and ten combinations of facial gestures were classified for all participants and once again 95% achieved 100% recognition accuracy. Average results and SD of all were 100% and 0, respectively. Combination code 53210 was chosen as the best one. Table 4 and Figure 6C represent the results of the best combination for all participants and the distribution of the best four facial gestures of the first volunteer, respectively.

Table 4 Final classification results of all participants with different number of facial gestures

Number of facial gestures: 2											
Participants	1	2	3	4	5	6	7	8	9	10	Average	SD
Results (%)	100	100	100	100	100	100	100	100	100	100	100	0
Number of facial gestures: 3											
Participants	1	2	3	4	5	6	7	8	9	10	Average	SD
Results (%)	100	100	100	100	100	100	100	100	100	100	100	0
Number of facial gestures: 4											
Participants	1	2	3	4	5	6	7	8	9	10	Average	SD
Results (%)	100	100	100	100	100	100	100	100	100	100	100	0
Number of facial gestures: 5	5										
Participants	1	2	3	4	5	6	7	8	9	10	Average	SD
Results (%)	100	100	100	100	100	100	100	100	100	100	100	0
Number of facial gestures: 6											
Participants	1	2	3	4	5	6	7	8	9	10	Average	SD
Results (%)	98.21	97.68	97.2	98.1	97.7	99.1	98.45	97.9	97.34	98.03	98.03	0.5515
Number of facial gestures: 7											
Participants	1	2	3	4	5	6	7	8	9	10	Average	SD
Results (%)	97.24	96.39	96.85	97.24	96.98	98.43	97.62	96.81	96.92	97.14	97.16	0.5517
Number of facial gestures: 8											
Participants	1	2	3	4	5	6	7	8	9	10	Average	SD
Results (%)	95.5	95.21	94.45	95.78	94.65	96.13	95.89	94.82	94.37	95.08	95.18	0.6217
Number of facial gestures: 9											
Participants	1	2	3	4	5	6	7	8	9	10	Average	SD
Results (%)	93.1	93.74	91.95	92.23	92.7	94.76	93.15	92.37	92.87	93.07	93.02	0.8084
Number of facial gestures: 10											
Participants	1	2	3	4	5	6	7	8	9	10	Average	SD
Results (%)	90.63	88.45	89.52	90.11	86.71	91.02	89.15	88.97	87.83	91.27	90.41	3.1270

Abbreviation: SD, standard deviation.

Table 5 All combination results for first participant

Combination	Results	Combination	Results
123456789	88.42%	1234678910	81.56%
1234567810	90.04%	1235678910	93.10%
1234567910	76.30%	1245678910	78.07%
1234568910	83.32%	1345678910	76.98%
1234578910	78.19%	2345678910	84.60%

Combinations of five facial gestures

Two hundred and fifty-two combinations made by five facial gestures were classified and combination code 45732 was picked as the best with 100% recognition accuracy. Table 4 indicates the results of the best combination for all participants, average result, and SD. Figure 6D shows the distribution of the best five facial gestures of the first volunteer.

Combinations of six facial gestures

Among all 210 existing combinations made by six facial gestures for each participant, the combination code 4578910 was chosen as the best with 98.03% average recognition accuracy. An SD of 0.5515 was achieved for all volunteers (Table 4). Figure 6E demonstrates best distribution of the first participant with the combination code 4578910.

Combinations of seven facial gestures

One hundred and twenty different combinations were constructed with seven facial gestures for each volunteer. The average result reported that the combination code 64578910 provided 97.16% recognition accuracy with an SD of 0.5517 (Table 4). Figure 6F shows this combination code for the first participant.

Combinations of eight facial gestures

Forty-five combinations were trained and classified. The average result of the best combination reached 95.18% recognition accuracy with an SD of 0.6217 (Table 4) with code 234578910. Figure 6G shows the results for the first volunteer.

Combinations of nine facial gestures

Ten combinations were made by nine facial gestures for each volunteer. Table 5 shows the results of all combinations for the first participant. Combination code 1235678910 is the best combination with 93.1% recognition accuracy. Table 4

reports the final results for all participants with an average result of 93.02% recognition accuracy and an SD of 0.8084. Figure 6H shows the distribution of this combination for the first participant.

Combinations of ten facial gestures

There is just one combination of all facial gestures (without normal signal) for each person. This group set achieved 90.41% recognition accuracy with an SD of 3.1270 for all participants. Table 4 describes all results attained from all volunteers and Figure 6I depicts the distribution of the first participant's data distribution. To distribute the concentrated features better while condensing the highly scattered points, a nonlinear transformation logarithm is applied to the features. By increasing the number of facial gestures, the variation between participants' recognition results increases, which causes an increase in SD. Table 6 shows the average of the best results from all facial gesture groups of all participants. When the number of facial gestures increases, the classification accuracy becomes lower because when the number of facial gestures is increased, the amount of data as well as the number of clusters becomes higher. The probability of data overlapping in each cluster increases.

Discussion

The main goal of this work is to design a multipurpose interface based on facial EMG gestures for HMI systems. It is accomplished and developed by employing eleven facial gestures and expressions. The obtained results show that this method produces an accurate interface in HMI systems. The main factors that differentiate this interface from others are: it contains more facial gestures, offers variety in choosing suitable facial gestures, is accurate, and can be used in various applications.

Table 7 compares the research on applied facial EMG gestures as the interface. The comparison of the results shows that the current work has achieved the best performance. In Ang et al, recognition of three facial gestures[23] obtained just 94.44% recognition accuracy while this work achieved 100%. Rezazadeh et al recognized five facial gestures,[27] but attained 93.2% while 100% classification accuracy was achieved in this work. Furthermore, its easy implementation, simplicity, high speed and user-

Table 6 The average of all results of all participants

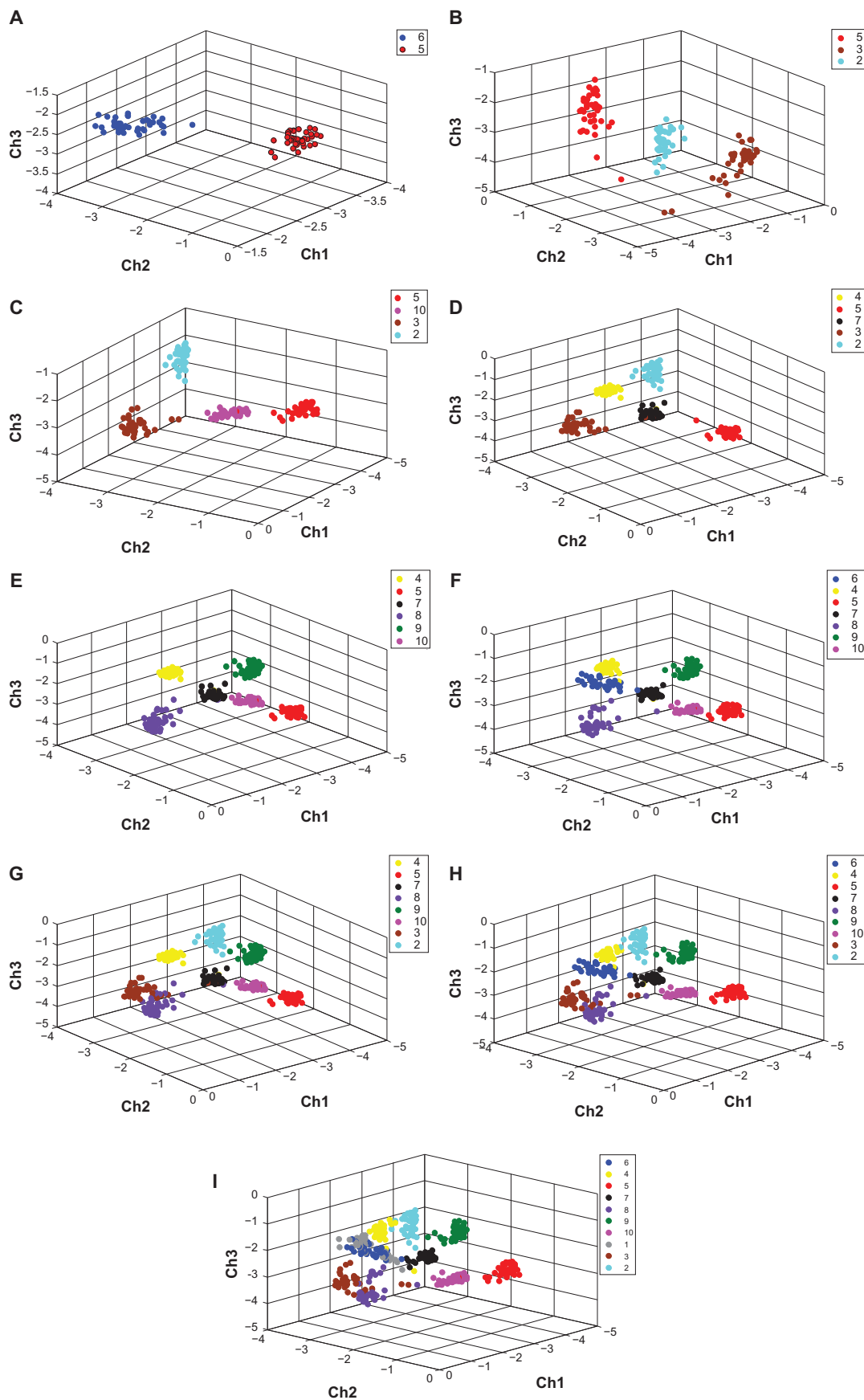| Number of facial gestures in each combination	2	3	4	5	6	7	8	9	10
Average of the best combinations	100%	100%	100%	100%	98.03%	97.16%	95.18%	93.02%	90.41%

Figure 6 Data distribution of the best combinations with different numbers of facial gestures for the first participant.

Table 7 Comparison of research in this field of study

Reference	Participants	Classes	Feature	Classifier	Results	Application
Present work	10	11	RMS	FCM	90.41%	Interface design
Moon et al[22]	1	5	MAV	Thresholding	93%	Intelligent robotic wheelchair
Ang et al[23]	2	3	MSD, RMS, PDS	Minimum Distance	94.44%	Interface design via three expressions
Ferreira et al[24]	8	2	–	Thresholding	95.71%	Mobile robot
Arjunan et al[25]	–	3	MRMS	Linear separation	100%	Unvoiced speech interface
Firoozabadi et al[26]	3	5	MAV	SVM	100%	Control a virtual wheelchair
Guillaume et al[36]	1	6	Absolute value	Gaussian models	92%	Enhancement of HCI via facial EMGs
Rezazadeh et al[27]	10	5	RMS	SFCM	93.2%	Virtual brain control
Egon et al[28]	21	4	Mean AD, SD, VAR	K-NN, SVM, MLP	61%, 60.71%, 56.19%	Facial EMG-based interface used in MMI

Abbreviations: AD, absolute deviation; EMGs, electromyograms; FCM, fuzzy C-mean; HCI, human–computer interface; K-NN, K-nearest neighbors; MAV, mean absolute value; MLP, multilayer perception; MMI, man–machine interface; MRMS, moving root mean square; MSD, maximum scatter difference; PDS, power spectrum density; RMS, root mean squares; SD, standard deviation; SFCM, subtractive fuzzy C-mean; SVM, support vector machine; VAR, variance.

friendliness are the advantages of this work compared to other methods. Apart from the above-mentioned points, the best combinations of facial gestures were found for each person in this work. Except for their general utility, various combinations of facial gestures are available to each person and can be applied depending on the user's condition. In this approach, normal signal is not considered as a gesture and it is not involved in declared combinations. However, it can still be used as a control command in this interface where the system needs to be disabled or put into a standby mode.

While this interface can be used by everyone, it is specifically targeted to people with both lower and higher impairments. It is one of the last opportunities for communication for these people, but, fortunately, it can help to improve their quality of life.

Conclusion

A multipurpose interface based on facial expression and gesture recognition was designed and applied in different HMI systems. Eleven facial gestures were chosen and their EMG signals recorded in three bipolar configuration channels. This interface can be used to control any prosthesis, rehabilitation device, or robots where 2–11 control commands are needed. This work developed the earlier approaches by making all combinations of facial gestures for all participants and identifying the one with the highest recognition accuracy and the best data distribution. Facial EMG analysis methods applied in this work provide enough accuracy, speed, and simplicity for utility; therefore, sufficient conditions for device control are prepared. In real-time and online control, especially in HMI systems where there is a user on one side of this system (a majority are elderly or handicapped people), the given interface can cope with all existing limitations.

Acknowledgments

Disclosure

The authors report no conflicts of interest in this work.

References

1. Salem C, Zhai S. An isometric tongue pointing device. Proceedings of CHI 97; March 22–27, 1997; Atlanta, GA. URL: http://www.sigchi.org/chi97/proceedings/tech-note/cs.htm. Accessed August 25, 2011
2. Barreto AB, Scargle SD, Adjouadi M. A practical EMG-based human-computer interface for users with motor disabilities. *J Rehabil Res Dev*. 2000;37(1):53–63.
3. LoPresti EF, Brienza DM, Angelo J, Gilbertson L. Neck range of motion and use of computer head control. *J Rehabil Res Dev*. 2003;40(3):199–212.
4. Mason KA, inventor. Control apparatus sensitive to eye movement 1969. US Patent 3-462-604-1969.
5. Nakanishi S, Kuno Y, Shimada N. Shirai Y. Robotic wheelchair based on observations of both user and environment. *Proc IROS*. 1999;99:912–917.
6. Krueger TB, Stieglitz T. A naïve and fast human computer interface controllable for the inexperienced – a performance study. *Conf Proc IEEE Eng Med Biol Soc*. 2007:2508–2511.
7. Palaniappan R. Brain computer interface design using band powers extracted during mental tasks, neural engineering. 2nd International IEEE EMBS Conference. March 16–19, 2005; Arlington, VA. 2005:321–324.
8. Rasmussen RG, Acharya S, Thakor NV. Accuracy of a brain computer interface in subjects with minimal training. Bioengineering Conference Proceeding of the IEEE 32nd Annual Northeast. May 15, 2006; Easton, PA. 2006:167–168. IEEE 0-7803-9563-8.
9. Wolpaw J R, Loeb GE, Allison BZ, et al. BCI meeting 2005 – Workshop on signals and recording methods. *IEEE Trans Neural Syst Rehabil Eng*. 2006;14(2):138–141.

10. Guger C, Edlinger G, Harkam W, Niedermayer I, Pfurtscheller G. How many people are able to operate an EEG-based brain-computer interface (BCI)? *IEEE Trans Neural Syst Rehabil Eng*. 2003;11(2):145–147.

11. Rosenberg R. The biofeedback pointer: EMG control of a two dimensional pointer. Digest of Papers. Second International Symposium on Wearable Computers. October 13–20, 2003; Pittsburg, PA: IEEE; 1998:162–163. IEEE 0-8186-9074-7.

12. Fukuda O, Tsuji T, Kaneko M. An EMG controlled pointing device using a neural network. IEEE International Conference on Systems, Man, and Cybernetics. October 12–15, 1999; Tokyo, Japan; IEEE SMC '99 Conference Proceedings. 1999;4:63–68.

13. Fukuda O, Arita J, Tsuji T. An EMG-controlled omnidirectional pointing device using a HMM-based neural network. Proceedings of the International Joint Conference on Neural Networks; July 20–24, 2003. 2003;4:3195–3200.

14. Alsayegh OA. EMG-based human-machine interface system. Proceedings of IEEE International Conference on Multimedia and Expo 2000; July 30–August 2, 2000; New York, NY. Vol 2. 2000;2:925–928. IEEE 0-7803-6536-4.

15. Itou T, Terao M, Nagata J, Yoshida M. Mouse cursor control system using EMG. Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2001; Istanbul, Turkey. Vol 2. 2001:1368–1369. IEEE 0-7803-7211-5.

16. Weir R. Design of artificial arms and hands for prosthetic applications. In: Kutz M, editor. *Standard Handbook of Biomedical Engineering and Design*. New York, NY: McGraw-Hill; 2003:32.1–32.61.

17. Chen X, Zhang X, Zhao ZY, Yang JH, Lantz V, Wang KQ. Multiple hand gesture recognition based on surface EMG signal. ICBBE 1st International Conference on Bioinformatics and Biomedical Engineering; July 6–8, 2007; Wuhan, China. 2007:506–509.

18. Wheeler KR. Device control using gestures sensed from EMG. SMCia/03. Proceedings of IEEE International Workshop on Soft Computing in Industrial Applications, 2003; June 23–25, 2003. 2003:21–26.

19. Kim J-S, Jeong H, Son W. A new means of HCI: EMG-MOUSE. IEEE International Conference on Systems, Man and Cybernetics; The Hague, The Netherlands. 2004;1:100–104.

20. Tsenov G, Zeghbib AH, Palis F, Shoylev N, Mladenov V. Neural networks for online classification of hand and finger movements using surface EMG signals. 8th Seminar on Neural Network Applications in Electrical Engineering, 2006. NEUREL, 2006; September 25–27, 2006; Belgrade, Serbia and Montenegro. 2006:167–171.

21. Naik GR, Kumar DK, Palaniswami M. Multi run ICA and surface EMG based signal processing system for recognising hand gestures. CIT 2008. 8th IEEE International Conference on Computer and Information Technology, 2008; July 8–11, 2008; Sydney, Australia. 2008:700–705.

22. Moon I, Lee M, Ryu J, Mun M. Intelligent robotic wheelchair with EMG-, gesture-, and voice-based interfaces. Proceedings of the 2003 IEEE/RSJ International Conference on intelligent Robots and Systems. (IROS 2003); October 27–31, 2003; Las Vegas, NV. 2003;3: 2453–3458.

23. Ang LBP, Belen EF, Bernardo RA, Boongaling ER, Briones GH, Coronel JB. Facial expression recognition through pattern analysis of facial muscle movements utilizing electromyogram sensors. TENCON 2004. IEEE Region 10 Conference; November 21–23, 2004, Chiang Mai, Thailand. Vol 3. 2004:600–603. IEEE 0-7803-8560-8.

24. Ferreira A, Celeste CW, Cheein AF, Bastos-Filho FF, Sarcinelli-Filho M, Carelli R. Human-machine interfaces based on EMG and EEG applied to robotic systems. *J Neuroeng Rehabil*. 2008;5:10.

25. Arjunan SP, Kumar DK. Recognition of facial movements and hand gestures using surface electromyogram (sEMG) for HCI based applications. 9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications; December 3–5, 2008; Glenelg, Australia. 2008:1–6.

26. Firoozabadi SMP, Asghari Oskoei MR, Hu H. A human-computer interface based on forehead multi-channel bio-signals to control a virtual wheelchair. Proceedings of the 14th ICBME, Tehran, Iran, 13–14 February. 2008:272–277.

27. Rezazadeh IM, Wang X, Firoozabadi M, Hashemi Golpayegani MR. Using affective human machine interface to increase the operation performance in virtual construction crane training system: A novel approach. *Journal of Automation in Construction*. 2010;20:289–298.

28. van den Broek EL, Lis´y V, Janssen JH, Westerink JHDM, Schut MH, Tuinenbreijer K. Affective man-machine interface: Unveiling human emotions through biosignals. *Biomedical Engineering Systems and Technologies*. Berlin, Germany: Communications in Computer and Information Science, Springer Verlag; 2010:52(Pt 1):21–47.

29. Clancy EA, Hogan N. Probability density of the surface electromyogram and its relation to amplitude detectors. IEEE Transactions on Biomedical Engineering, Jun 1999. Vol 46. 1999;46(6):730–739. IEEE 0018-9294.

30. Oskoei MA, Hu H. Support vector machine-based classification scheme for myoelectric control applied to upper limb. *IEEE Trans Biomed Eng*. 2008;55(8):1956–1965.

31. Subasi A, Kiymik MK. Muscle fatigue detection in EMG using time–frequency methods, ICA and neural networks. *J Med Syst*. 2010;34(4): 775–785.

32. Knaflitz M, Bonato P. Time-frequency methods applied to muscle fatigue assessment during dynamic contractions. *J Electromyogr Kinesiol*. 1999; 9(5):337–350.

33. Hamedi M, Rezazadeh IM, Firoozabadi SMP. Facial gesture recognition using two-channel biosensors configuration and fuzzy classifier: A pilot study. International Conference on Electrical, Control and Computer Engineering (INECCE), 2011; June 21–22, 2011; Pahang, Malaysia. 2011:338–343.

34. Hamedi M, Sheikh HS, Tan TS, Kamarul A. Surface electromyography-based facial expression recognition in bipolar configuration. *J Comp Sci*. 2011;7(9):1407–1415.

35. Ekman P, Friesen WV. *Facial Action Coding System*. Palo Alto, CA: Consulting Psychologist Press; 1978.