

EMG-based Hand Gesture Recognition for Realtime Biosignal Interfacing

Jonghwa Kim, Stephan Mastnik, Elisabeth André
Lehrstuhl für Multimedia Konzepte und ihre Anwendungen
Eichleitnerstr. 30, D-86159 Augsburg, Germany
kim@ieee.org

ABSTRACT

In this paper the development of an electromyogram (EMG) based interface for hand gesture recognition is presented. To recognize control signs in the gestures, we used a single channel EMG sensor positioned on the inside of the forearm. In addition to common statistical features such as variance, mean value, and standard deviation, we also calculated features from the time and frequency domain including Fourier variance, region length, zerocrosses, occurrences, etc. For realizing real-time classification assuring acceptable recognition accuracy, we combined two simple linear classifiers (k-NN and Bayes) in decision level fusion. Overall, a recognition accuracy of 94% was achieved by using the combined classifier with a selected feature set. The performance of the interfacing system was evaluated through 40 test sessions with 30 subjects using an RC Car. Instead of using a remote control unit, the car was controlled by four different gestures performed with one hand. In addition, we conducted a study to investigate the controllability and ease of use of the interface and the employed gestures.

Author Keywords

Biosignal Analysis, Electromyogram, Gesture Recognition, Human-Computer Interaction, Neural Interfacing

ACM Classification Keywords

H5.2. Information interfaces and presentation: Input devices and strategies (e.g., mouse, touchscreen).

INTRODUCTION

Recently, a significant amount of effort in human-computer interaction (HCI) has been dedicated to the development of

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in:

IUI'08, January 13-16, 2008, Maspalomas, Gran Canaria, Spain.

Copyright 2008 ACM 978-1-59593-987-6/ 08/ 0001

user-friendly interfaces employing voice, vision, gesture, and other innovative I/O channels. One of the most challenging approaches in this research field is to link a human's neural signals with computers by exploiting the electrical nature of the human nervous system. For the neural linkage with computers, various biomedical signals (biosignals) can be used, which can be acquired from a specialized tissue, organ, or cell system like the nervous system. Examples include electro-encephalogram (EEG), electrooculogram (EOG), and electromyogram (EMG). Such approaches are extremely valuable to physically disabled persons, as, for instance, unvoiced speech recognition and the hands-free EMG mouse [1].

EMG measures electrical currents that are generated in a muscle during its contraction and represent neuromuscular activities. EMG signals can be used for a variety of applications including clinical applications, human-computer interaction and interactive computer gaming [2, 3]. Moreover, EMG can be used to sense isometric muscular activity which does not translate into movement. This makes it possible to classify subtle motionless gestures and to control interfaces without being noticed and without disrupting the surrounding environment. On the other hand, one of the main difficulties in analyzing the EMG signal is due to its noisy characteristics. Compared to other biosignals, EMG contains complicated types of noise that are caused by, for example, inherent equipment noise, electromagnetic radiation, motion artifacts, and the interaction of different tissues. Hence, preprocessing is needed to filter out the unwanted noises in EMG. This difficulty becomes more critical when resolving a multiclass classifying problem. In most previous works, therefore, multi-channel EMG sensors are used at the same time to detect relevant EMG patterns by a combined signal analysis. In this case, however, users suffer from the inconvenience of carrying many cabled electrodes.

In this work we therefore present an EMG-based controlling interface using a single channel EMG sensor positioned on the inside of the forearm. As a sample application we modified an RC car so that it could be controlled by a user's hand signs, instead of using a remote control unit. The interfacing system first calculates relevant

features in the EMG signal of four hand signs, classifies the hand signs into the four classes, and assigns the result to certain steering commands for the RC car. To be usable, all processing stages must run in real-time without perceptible latency time. At the beginning of this paper, a short overview of existing biosignal based controlling interfaces is given. In the subsequent part, the setup of the developed system including the selection of suitable gestures, features and classification algorithms is described, followed by an evaluation and a discussion of the system.

RELATED WORK

Nowadays many different types of biosignals, such as skin conductance or electrocardiogram, can be measured with many differing procedures. Depending on the respective signal, these biosignals are utilized in industrial applications, such as medicine or entertainment [4]. Some biosignals have also been shown to be suited for the creation of a new communication interface between humans and computers. In this area the use of biosignals offers brand new possibilities when compared to the conventional, mostly audio-visually based human-computer interfaces. Thus, with the help of biosignals, it is today possible to detect emotions [5], make music [6] or develop smart clothes [7]. The famous polygraph is also based on biosignals.

Many biosignal based interfaces are used for controlling and communication. For disabled people especially, they offer the possibility of making their lives easier. There have been some promising attempts for the development of a new generation of biosignal controlled prostheses [8], which are much more user-friendly and more easily accepted than customary prostheses. Even people with severe disabilities and whose normal communication channels do not work anymore may receive help by the creation of a new communication interface based on biosignals. This is one major aim of brain computer interface (BCI) research, where communication can take place simply by measuring thoughts. Besides BCIs, which mainly use electroencephalographic signal (EEG) [9], the most important biosignal for controlling interfaces have become the signals received from the EMG. This is due in great part to the fact that most bioelectric signals, such as the EMG or EEG, can be recorded in a comparatively simple and inexpensive manner thanks to the use of electrodes. As it is usually possible to receive the bioelectric signals free of pain by placing the electrodes on the surface of the skin, user acceptance compared to other biosignal measurement methods is also proportionally high. The EMG signal represents the natural electrical activity of the human body, which is used to control the skeletal muscles. Nowadays it is possible to control, in addition to the aforementioned prostheses, robots [10], mobile phones [11] and MP3 players [12] with the help of EMG signals. These systems are usually based on the performance of

several gestures which are recognized through their EMG signal and in this vein trigger specific actions. The type of gesture depends on the number and the positioning of the measuring sensors and varies from nearly motionless arm gestures [13], to hand gestures [12, 14] and movements of single fingers, for example for virtual typing [15]. In 1999, NASA also successfully developed an EMG based controlling interface: It simulates the landing of an aircraft which is solely controlled by gestures resulting from the navigation of a virtual joystick [16].

In this work an EMG based real-time controlling interface was developed to navigate an RC car with the help of four different hand gestures only. Contrary to other existing gesture based controlling interfaces which often require many different EMG channels [10, 12, 14, 16] placed on multiple muscles to be able to distinguish the different gestures, there is only one single EMG channel used in the present case.

METHODOLOGY

Before explaining the setup of the developed system, we first describe the selection of suited gestures and a suited starting or home position:

Gesture selection

For controlling an RC car, four different gestures are required. These gestures should be equally easy to perform and form patterns in the EMG signal which are as discriminative as possible. Before the start and after the finish of each gesture, the hand should be situated in a posture called the home position, in which hardly any signal amplitude occurs. After testing over 20 different gestures the four gestures shown in Fig. 1 were selected as control signs since they seemed to fulfill the mentioned requirements best. The first gesture, *Press* (gesture 1), is a short and very light pressing of the fist. Gesture *Left* (gesture 2) is performed by a quick left motion of the wrist towards the inside of the forearm, followed by a straight return into the home position. The motion directed towards the outside of the arm is called gesture *Right* (gesture 3) and is supposed to be slow and smooth. A fist should be formed in this case. The fourth gesture, *Circling* (gesture 4), is a smooth circling movement of the wrist starting with a swinging movement into the direction of the inside arm. Depending on the way the executive user performs the motion, either a single or two consecutive circle movements are possible. For the home position, a loose, relaxed fist has been proven to be the best suited position.

The choice of the performing hand is not an issue since our experiments showed that both hands generate extremely similar EMG waveforms, except for unremarkable differences in overall amplitude. As most users opt for the right hand, the names and descriptions of the gestures refer to the right hand and have to be mirror-inverted for a left handed usage.

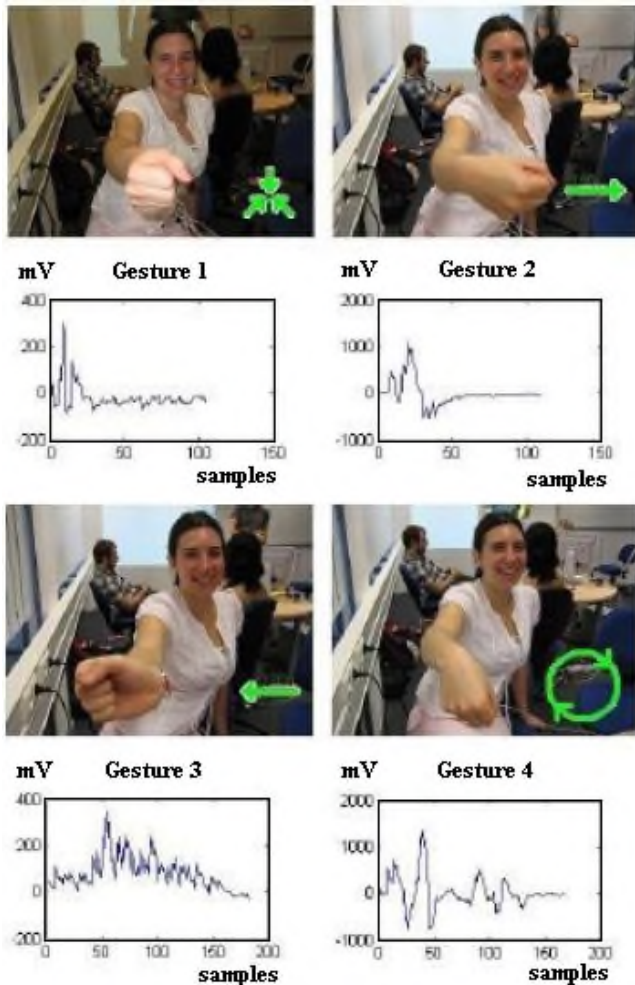


Figure 1. Gesture performance and resulting patterns

System structure

In this section the single processing steps of the system, from the recording of a gesture to the movement of the RC car, are described. An overview is given in Fig. 2.

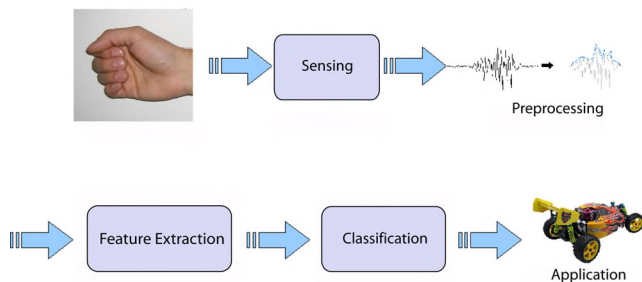


Figure 2. General procedure of a biosignal based recognition system

The EMG signal of the performing arm muscles is detected by electrodes connected to a sensor. In order to qualify the incoming raw signal for further processing the signal is

preprocessed first. Next, the incoming patterns, which represent a gesture movement in the signal, are matched. To be able to distinguish patterns, the significant features of each pattern are first extracted. The resulting feature vector is used for the classification of the movement order. Finally, the control command briefed through the performed gesture is executed by the RC car.

Signal acquisition

We used NeXus-10™ with Myoscan-Pro™ EMG sensor which enables to record EMG signals of up to 1600 μV in an active range of 20 to 500Hz. For the recording of the EMG signal, only one pair of pre-gelled single Ag/AgCl electrodes was fixed on the skin of the system user's inside forearm (Fig. 3). Usually, each pair of electrodes is used to examine mainly one single muscle. Signal interferences of adjacent muscles, known as crosstalk, are normally undesirable. Since we used only one channel sensor for signal acquisition it was necessary to examine several muscles simultaneously with one pair of electrodes. These observed muscles were mainly the flexor carpi radialis and the palmaris longus, both of which are responsible for wrist movements, as well as the flexor digitorum superficialis, which is used for finger movements [17].

All three electrodes are situated in a line in the middle of the forearm parallel to the length of the forearm muscle fibers. By placing the first electrode near the wrist, it is possible to examine the muscles of the forearm between their tendon insertions and their motor points, which seems to be the best location for a constant measurement [18]. The reference electrode is placed in the middle. The sampling rate of the EMG signal in the system was set at 125 samples per second.



Figure 3. Position of the electrodes according to the muscle structure

Preprocessing

Fortunately, in our experiment, the common noises found in EMG signals, such as inherent equipment noise, electromagnetic radiation, or moving artifacts, were hard to find. However there was another problem hindering the raw signal from being subsequently processed. The incoming signal exhibited an extremely unstable baseline which made it difficult to calculate reasonable values for statistical and frequency features. As a consequence, the signal had to be preprocessed first to be suitable for further processing. For

this purpose, the raw signal values x were detrended with the following simple detrending function D :

$$D(x_i) = x_i - \frac{\sum_{j=i-7}^i x_j}{8}$$

Pattern extraction

To capture only relevant patterns in EMG waveforms, we needed to design an adaptive thresholding. Although the system provided a default setting for the concrete threshold values which normally provides comparatively accurate pattern boundaries, the best results were achieved when the threshold values were individually adjusted to the user of the system. This is due to the fact that the EMG signal of each person is slightly different.

An incoming preprocessed value was marked as the beginning of a pattern if a certain defined threshold value was reached. This boundary was found to be quite unproblematic, as there exists a considerable difference in the signal amplitude between staying in the home position and starting to perform a gesture. The upper bound was limited by the *Press* gesture which generally generated the lowest amplitude and absolutely needed to be identified, while the lower bound was limited by noise, for example resulting from small unintended movements which must not be recognized as patterns.

Detecting the ending of a pattern turned out to be more complicated. Here, a trade-off between two different problems had to be found: First, the ending boundary should not be too high, otherwise the pattern might not be caught completely. Second, the ending boundary should not be too low, otherwise the end of a pattern might be detected much too late. A further complication resulted from the fact that some gestures caused a reverberation effect in the signal amplitude, long after the motion had already stopped.

If, because of the characteristic zigzag form of an EMG signal, the examination of an incoming value does not make any sense, a couple of consecutive values have to be observed. The best results for the detection of a pattern ending in the system were achieved by observing the root mean square (RMS) of the last 16 incoming values. If this RMS value fell twice in a row below two (possibly different) boundary values, it was considered as being the end of a pattern.

Feature extraction

To be able to classify a performed gesture some distinctive features have to be found and taken from each matched pattern [19, 20]. Therefore several features were extracted, including common statistical features like *maximum*, *minimum*, *mean value*, *variance*, *signal length* and *root mean square*.

In the frequency domain obtained by using typical fast Fourier transform (FFT), we calculated *fundamental*

frequency (F0) and *Fourier variance* of the spectrum. Given the spectrum of signal we also extracted the *region length* which is defined as a partial length of the spectrum containing greater magnitude than the mean value of total Fourier coefficients. This feature should be an indicator for how periodic a signal is: the smaller the region the more periodic the signal. In the cases where there is more than one region in the spectrum, the lengths of these regions are added.

Furthermore we attached some particular features which appeared to be suited for the distinction of the four gestures through their EMG signals. First, we added the *positions of the maximum and the minimum* which are defined by the relative position (as a percentage) of the max. and min. values within the length of the entire pattern. Next, we calculated the *zerocrosses*, which are defined by the number of times the zero line is crossed or touched in relation to the length of the signal. The feature *number of occurrences* results from the number of vertices existing in the pattern graph.

Classification using decision-level classifier fusion

In general, for classification, feature vectors, which represent a pattern, are assigned to specific classes. For solving this task various classification algorithms exist [21, 22]. Many of these supervised classifiers have to be trained at first. During training, the classifier learns to arrange the patterns. Training results can be very different, with variations ranging from the setting of some parameters or weights to the creation of a new grammar.

As the duration of the classification process is an essential factor for the efficiency of a realtime system, we decided to apply two comparatively simple and thus fast algorithms: the K-nearest neighbor (KNN) classifier and the Bayes classifier. Despite their simplicity these algorithms generally provide proportionally good results.

The K-Nearest Neighbor classifier, which belongs to the non-parametric statistical classifiers, rates a pattern by regarding the most similar labeled training samples. For this purpose, the distances (e.g. Euclidean distance) between the feature vector of the current pattern and the feature vectors of each training sample are calculated. Beforehand, all vectors are generally normalized. The number of adjacent samples which are taken into account is defined by the parameter k . In our pattern recognition system, we considered the five nearest neighbors.

The parametric statistical Bayes classifier decides based on probabilities. Each pattern is attached to the class to which it belongs to through the highest probability. The calculation of the probabilities is derived from the Bayes theorem. A cost function is required first for the classification. This function calculates the costs of all possible classification decisions, whereas the decision is made in favour of the class that produces the lowest costs.

Thus, the possibility of a wrong decision is minimized, if wrong decisions produce high costs. A precondition for a classification with the Bayes algorithm is that the probabilities for all possible feature vectors have to be known a priori. As in practice not all possibilities are normally known, they are usually estimated with a Gaussian distribution or a Gaussian mixture model.

In addition to using the single classifiers separately, we designed two novel combination methods using both classifiers. Both methods are based on a decision tree structure as illustrated in Fig. 4 and 5. As splitting attribute we defined a prediction rate (PR) in per cent, the ratio of the number of predicted neighbors to the number of all considered neighbors (corresponding to the parameter k of the KNN classifier). For example, PR of KNN in Fig. 4 means the ratio of the number of considered neighbors in the KNN algorithm which are labeled as predicted by the KNN classifier to the number of all considered neighbors, while the PR of Bayes in Fig. 5 denotes the number of considered neighbors in the KNN algorithm which are labeled as predicted by the Bayes classifier to the number of all considered neighbors. This means that both combination methods are designed based on the assumption that the KNN algorithm only delivers reliable results when the PR of KNN is high enough.

The first method (Fig. 4) first checks whether the PR of Bayes is at least 40%. In this case, the Bayes classification result is immediately taken. Otherwise, if the PR of KNN is at least 60%, the KNN classification result is chosen. When both PRs are too small, the Bayes classification result is preferred.

The second method (Fig. 5) is less complicated. It simply depends on the PR of KNN. If the PR is at least 60%, the KNN classification result is taken, otherwise the Bayes result is selected.

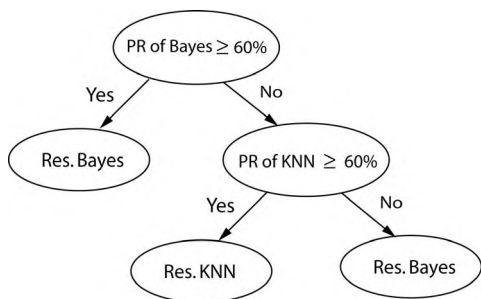


Figure 4. Decision tree of classifier combination 1

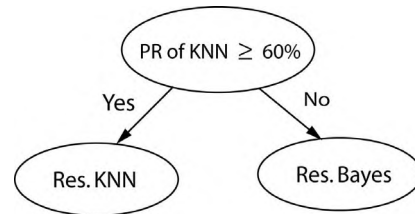


Figure 5. Decision tree of classifier combination 2

Classifiers usually need a certain number of features to possess enough information for a reliable classification. Due to the curse of dimensionality, however, an increase of the number of features does not always produce better results. Thus, a feature selection has to take place before classification. In the section below on offline analysis, we discuss the effectiveness of feature selection.

Controlling the RC car

The purpose of the developed interface is to control an RC car. To be able to control the car via a PC, the remote control of the RC car was connected with a computer in the following way (see also [22]): First, we opened the remote control unit and removed the old board by cutting off the wires. To insert a new board, we soldered four wires onto the control points and a further wire onto the negative pole of the board. We then fixed four transistors on the new board and connected them with the four wires soldered onto the control points. Another wire was set from the transistor to the grounding. The remaining transistor base was connected with a male parallel port connector. Finally the board was reconnected with the cut wires. After attaching a 9-volt-battery and connecting the remote control and a PC by attaching a common printer cable to the parallel port connector, the RC car could be controlled based on the classification result by directly addressing the ports through the software.

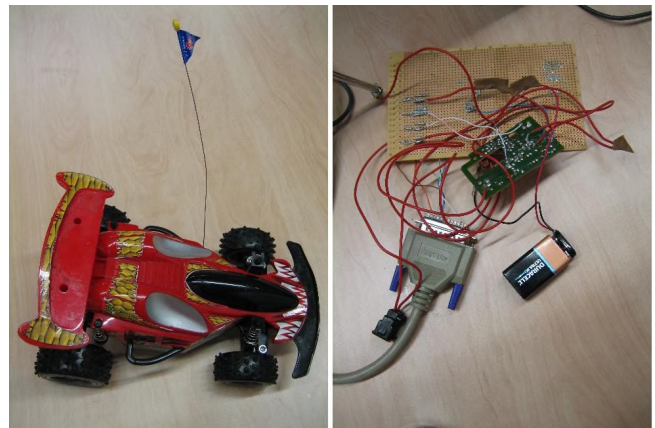


Figure 6. RC Car and modified remote control

The control of the car with the four gestures works as follows: When a control command has been given by

performing a gesture, this command is executed by the car as long as the next control command is provided. A control command can be canceled by giving the opposite command. Thus, for example, if the car is moving forwards and towards the left and a backward control command is given, the car will first stop. With a further backward command the car will move backwards and left. As the left movement order has not been canceled, it is still activated. The control of the left, straight and right movements works in an analogous way.

Calibration

To adjust the system to the current user at the beginning of each driving session a calibration had to be performed. Here, training samples of the user are taken for each gesture. These samples are used for the training of the classifiers. Usually we recorded 10 or 20 samples of each gesture per user, as the recording of more samples is very time-consuming and uncomfortable for a user. Furthermore no significant improvement of the classification rate was obtained by training the classifiers with more samples. Although it is possible to use an old, previously recorded training dataset when the user is not dealing with the interface for the first time, it is recommendable to perform a new calibration for each session. The threshold values for the pattern extraction should also be adapted to the user at the beginning of a session.

A calibration phase is necessary since the EMG signal is very sensible. First, the EMG signal of every user is naturally slightly different, as all humans are different. A different muscle mass for example is one reason for the influencing differences. But the signal does not only vary between different users. There are also variations for single users, since the signal depends on the current condition of the user. Even during one session, sometimes, a noticeable alteration of the EMG could be observed as a result of increasing muscle exhaustion. In addition, many users have individually slight differences in the performance of the gestures and it was noticed that even small variations in gesture performance could cause quite diverse signal shapes. That is the reason why an exact performance of the gestures became an important factor for a successful system run. Furthermore it was observed that other conditions such as the temperature can influence the EMG signal.

The major difference was often found in the signal amplitude, but there were also variations in the shape and length of the extracted patterns. Moreover, the amplitude of the dithering of the signal while the user was in the home position was sometimes quite diverse. For a few users, a delayed dithering after the performance of a gesture could also be perceived, which was probably due to the operating muscles.

In spite of these differences in the EMG signal, our long-term goal is to build up a user-independent system beside

the user dependent usage. To achieve this goal, we decided to train the classifiers with a huge sample dataset consisting of the samples of many different users. Furthermore the threshold values have to be set in such a way that all users perform well. The required samples for the creation of the user-independent training dataset were collected during an evaluation after the completion of the controlling interface.

Evaluation of the system

Preceding tests

To collect first feedback from users, we presented our interface to the RC car to school girls during an annual information day called “Girls Day” at our university. In general, the school girls did well with the EMG-based interface even though some training was required until they were able to perform gestures in a sufficiently constant manner. Fortunately, this problem could be resolved by using electrodes with better adhesive properties.

Furthermore, we used this first test to decide which features should be selected as input for the classification algorithms. For this purpose, we performed a sequential backward search (SBS) with one dataset consisting of 10 samples per class, i.e. altogether 40 samples. Based on this result the features *number of occurrences*, *root mean square*, *signal length*, *region length* and *zero crosses* were selected. Based on several preceding tests the threshold value for the beginning of a pattern was set at 90 mV as default value, while the two values for pattern ending were set at 30 mV.

Experimental setting

After the preceding tests, we conducted a more comprehensive experiment with a total of 30 subjects at an Open Lab Party at our University. First of all, we collected personal information about the subjects, including age (average: 25.63 years), gender (16 females and 14 males), size (average: 1.75 m), weight (average: 72.3 kg) and performing hand (28 right handers, 2 left handers).

The experiment consisted of two phases. Twenty subjects participated in the first phase. Ten of them also took part in the second phase, in addition to ten new subjects. That is, altogether, we obtained 40 test sets. For each subject, we recorded 20 training samples and 20 test samples. The experiment consisted of the following steps:

- 1) The subjects were asked to provide us with demographic data.
- 2) The subjects got an individual introduction to the EMG-based interface.
- 3) The performed arm was disinfected and electrodes were attached to it as shown in Fig. 3.
- 4) The participants practiced the single gestures until they succeeded in performing them with sufficiently little variation. If necessary, the single thresholds were adjusted

to the participants.

5) For each participant, 20 training samples were recorded per gesture.

6) For each participant, 20 test samples were taken to test the online behavior of the system.

If necessary, practice of the gestures was repeated after recording the training and test samples.

For the 10 participants who were tested twice, the second phase was identical except for the explanations that were no longer needed. Furthermore, training times were significantly shorter.

Results

The following table shows the average classification rates of the classifiers and classifier combinations for all 40 test sets (Table 1).

Classifier	Average classification rate
Combination 1	94.38%
Combination 2	93.94%
KNN	93.84%
Bayes	91.84%

Table 1. Average classification rates of all test sets

Among the four classification methods, the more complex classifier combination 1 achieved the best result with a classification rate of 94.38%, followed by the classifier combination 2 and the KNN classifier. The Bayes classification result on its own achieved with a classification rate of 91.84%, the worst result.

When regarding the gesture classification rates separately, it appears that the gestures *Press* (gesture 1) and *Circling* (gesture 4) could be identified best (Table 2).

Classifier	G 1	G 2	G 3	G 4
Combination 1	96.88%	92.63%	92.75%	95.25%
Combination 2	96.75%	91.75%	91.88%	95.38%
KNN	96.75%	91.63%	91.75%	95.25%
Bayes	93.38%	90%	87.88%	96.13%

Table 2. Average classification rates for each gesture

The classification rates obtained in phase 1 and phase 2 for the ten subjects who participated in both phases were quite similar, while the result of the best classification method, i.e. classifier combination 1, was slightly improved (Table 3).

Classifier	Phase 1	Phase 2
Combination 1	95.25%	95.63%
Combination 2	95.13%	94.88%
KNN	95.13%	94.63%
Bayes	92%	93.75%

Table 3. Comparison of phase 1 and phase 2

Classifier	Males	Females
Combination 1	94.20%	93.67%
Combination 2	93.84%	93.28%
KNN	93.84%	93.20%
Bayes	91.07%	88.56%

Table 4. Comparison of male versus female subjects

When analyzing the classification results of the first test of all 30 subjects according to gender (Table 4), size (Table 5) and weight (Table 6), we found that taller than average participants had better rates than shorter subjects.

Classifier	Taller	Smaller
Combination 1	95.73%	92.71%
Combination 2	95.42%	92.29%
KNN	95.42%	92.22%
Bayes	93.02%	90.21%

Table 5. Comparison of subjects taller than the average height versus smaller subjects

Classifier	More Heavy	Lighter
Combination 1	92.73%	94.51%
Combination 2	92.39%	94.03%
KNN	92.39%	93.96%
Bayes	89.32%	92.08%

Table 6. Comparison of subjects which are heavier than the average weight versus lighter subjects

Although nearly all subjects did quite well with the default setting of the threshold values, the value for the detection of the beginning of a pattern was increased to 110 mV during

the evaluation. The reason for this change was that we wanted to avoid small signal amplitudes, for example resulting from small unintended moves or signal dithering in the home position, from being detected as a pattern. For this reason, the threshold value even had to be increased to 130 mV for one subject. The threshold values for pattern ending only had to be slightly adjusted in a few cases due to delayed dithering in the signal after a gesture had been performed.

Offline analysis

As mentioned earlier, not all features we calculated are relevant for the 4-class gesture classification. To improve classification accuracy, we performed an offline analysis focusing on feature selection. The phase 1 training samples of every subject in the evaluation were taken for the significance analysis of features. The sequential backward search (SBS) method was performed to select significant features by using Leave-One-Out cross validation based on classifier combination 1.

As a result of the SBS feature selection, the following significance rating of features was obtained: *signal length, root mean square, mean value, fundamental frequency, zerocrosses, number of occurrences, region length, position of maximum, position of minimum, Fourier variance* (in downward rating).

In the case of user-dependent classification, when considering only phase 1 for each participant, the average classification rate of 93.92% achieved in online classification was improved to an average rate of 94.46% by using the new set of features in offline classification. For user-independent offline classification, the training datasets of all 30 subjects were mixed together and the same classification was performed. In this case, a remarkable average classification rate of 93.17% could be achieved. When looking at the gestures separately (Table 7), the classification rate of gesture 1 even reached 98%, even though this was at the expense of the classification rate of gesture 2.

Classifier	G 1	G 2	G 3	G 4
Combination 1	98%	87%	93.34%	94.34%

Table 7. User independent gesture classification rates

Subjective evaluation of the system

We did not only evaluate the robustness of the system; we were also interested in the users’ subjective assessment of the interface. We therefore asked twenty persons that had used our system to fill in a questionnaire. The questionnaire used 11 attitude statements with a 5-ary rating scale ranging from “do not agree at all” to “completely agree” in order to evaluate the user’s attitude towards the system as a whole (3 questions addressing fun, effort of use and

controllability) and towards the single gestures (2 questions for each of the four gestures referring to naturalness and difficulty of use).

To show that the mean value of a rating was significantly above or below the neutral value of 3, we applied t-tests for one sample. Overall, the interface was perceived as positive. The users thought it was fun to use the interface with a mean value of 4.10 ($t(19)=6.85, p<0.001$) and they had the feeling that the interface was responding in the intended manner with a mean value of 3.85 ($t(19)=4.344, p<0.001$). Nevertheless, the gesture-based interface was also perceived by a few subjects as tiring. For the corresponding attitude statement, we obtained a mean value of 2.9 only, which did not significantly deviate from the neutral value of 3.0.

Apart from *Circle* (gesture 4), all gestures were perceived as natural and not difficult to use. The best results were achieved for *Press* (gesture 1) with a mean value of 4.65 ($t(19)=11.0, p<0.001$) for naturalness and a mean value of 1.45 ($t(19)=-7.815, p<0.001$) for difficulty of use. In comparison, we only obtained a mean value of 3.35 for naturalness and a mean value of 2.45 for difficulty of use when evaluating gesture 4. None of the mean values for gesture 4 significantly deviated from the neutral value of 3.0.

Interestingly, the fact that a gesture is recognized more easily does not necessarily mean that users also perceive it more positively. As we have shown in the previous section, gesture 1 and gesture 4 achieved the highest recognition rates. Nevertheless, gesture 4 scored worst than any other gesture in the subjective evaluation, while gesture 1 got the best subjective rating.

Discussion

Although the classification rates achieved by the system are comparatively high, and it was observed that nearly all participating subjects could familiarize themselves quite fast with the controlling interface, the developed controlling interface is still more of a prototype than a finalized system. We achieved a rather good controllability of the interface with the four selected gestures. Yet this is conditioned on a precise performance of the gestures. Hence, before a user is able to handle the interface, he first has to practice the performance of each gesture in a constant manner. Of course the interface would be more user-friendly if no gesture training was required, but unfortunately even small discrepancies can result in major differences in the EMG signal. Therefore, one possible approach for improving the interface would be to find more amenable command gestures that make the system less sensitive to user- and time-dependent performance variations. With gesture *Right* such an improvement could already be made by having the subjects to form loose fist during the motion.

Overall, it should be kept in mind that the gestures should

be as fast, easy and natural to perform as possible. Indeed, it turned out that a rather anticipatory driving style was necessary to control the RC efficiently, as the time that is needed to perform the gestures has to be taken into account.

Another approach to improve the system might be an EMG-specific analysis for extracting more relevant features. The quality of the system may also be improved by capturing the pattern boundaries more precisely. One way to achieve this is to very precisely adjust the threshold value to the user. Unfortunately this procedure runs counter to another desirable development goal, i.e. the advancement to a completely user-independent interface. The performed offline test has already created promising expectations in relation to such a development.

Conclusion

This paper has shown that an EMG signal can be effectively employed in human-machine neural interface. The presented EMG-based controlling interface is able to reliably recognize various hand gestures with a positive classification rate of over 94% even though we used only one single EMG sensor, in contrast to related work which is based on multiple EMG sensors. For real-time condition, different decision tree approaches were proposed to combine several classifiers at the decision level by employing an empirical splitting attribute. Despite its simple structure, classifier fusion was shown to improve the accuracy of single classification. The good performance of the system was also reflected by the users' subjective ratings of the system's usability.

The EMG signal carries valuable information regarding the nerve system. It would be quite easy to transfer its use to applications other than the RC car. Moreover, since the EMG signal can be used to sense isometric muscular activity, it is possible to detect motionless gesture or intention in the EMG signal. Consequently, there is a wide range of potential applications using EMG signal in human-machine interfacing. However, to realize advanced applications, many issues still need to be resolved, including the development of algorithms for EMG-specific analysis, the extraction of relevant features, and the design of real-time classifiers with guaranteed accuracy, as discussed in this paper.

ACKNOWLEDGMENTS

This work was partially supported by the European Community (EC) within the network of excellence Humaine IST-507422.

REFERENCES

1. Kim, J.-S., Jeong, and H., Son, W., A new means of HCI: EMG-MOUSE, 2004 IEEE International Conference on Systems, Man and Cybernetics, Volume 1, 10-13 Oct. 2004, 100 – 104.

2. Kim, J., Bee, N. Wagner, J., and André, E. Emote to Win: Affective Interactions with a Computer Game Agent, Lecture Notes in Informatics (LNI), P-50, 2004, 159-164.
3. Park DG, and Kim HC. Muscleman: Wireless input device for a fighting action game based on the EMG signal and acceleration of the human forearm. [http://melab.snu.ac.kr/Research/melab/doc/HCI/muscleman_paper.pdf].
4. Sakurazawa, S., Yoshida, N., and Munekata, N. Entertainment feature of a game using skin conductance response, in ACE '04: Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology (2004), 181–186.
5. Wagner, J., Kim, J., and André, E. From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extraction and Classification. In IEEE International Conference on Multimedia & Expo (ICME 2005), 2005, 940-943.
6. Dubost, G. and Tanaka, A. A wireless, network-based biosensor interface for music, In Proceedings of International Computer Music Conference (ICMC), 2002.
7. Troester, G. The agenda of wearable healthcare, in IMIA Yearbook of Medical Informatics (2005), 125–138.
8. Reischl, M., Mikut, R., Pylatiuk, C. and Schulz, S. Erweiterung der Ansteuerungsmöglichkeiten myoelektrischer Handprothesen, Biomedizinische Technik 47, Ergänzungsband 1, 2002, 868-870.
9. Ebrahimi, T., Vesin, J. and Garcia, G. Brain-computer interface in multimedia communication, in IEEE Signal Processing Magazine, Volume: 20, Issue: 1, 2003, 14-24.
10. Crawford, B., Miller, K., Shenoy, P. and Rao, R. Real-time classification of electromyographic signals for robotic control, in Proceedings of the National Conference on Artificial Intelligence, Vol. 20, 2005, 523-528.
11. Costanza, E., Perdomo, A., Inverso, S. A. and Allen, R. EMG as a subtle input interface for mobile computing, Lecture Notes in Computer Science, Issue 3160, 2004, 426-430.
12. Fistre, J. and Tanaka, A. Real time EMG gesture recognition for consumer electronics device control, Presented at Sony CSL Paris Open House 10/2002, <http://www.csl.sony.fr/~atau/gesture/>
13. Costanza, E., Inverso, S. A., Allen, R. and Maes, P. Intimate interfaces in action: Assessing the usability and subtlety of emg-based motionless gestures, In: Proceedings of ACM CHI 2007, Conference on Human Factors in Computing Systems, 2007, 819-828.
14. Naik, G.R., Kumar, D.K., Singh, V.P., & Palaniswami, M. Hand gestures for HCI using ICA of EMG. HCSNet Workshop on the Use of Vision in HCI, 2006, 67-72.

15. Wheeler, K. R. Device control using gestures sensed from EMG, in IEEE International Conference on Soft Computing in Industrial Applications, 2003.
16. Jorgensen, C., Wheeler, K. and Stepniewski, S. Bioelectric control of a 757 class high fidelity aircraft simulation, in Proc. World Automation Congr., Wailea, Maui, HI, June 11–16, 2000.
17. Lippert, H. Anatomie, Urban & Fischer Verlag, 2000.
18. DeLuca, C. J. Surface electromyography: Detection and recording, MA: Neuro-Muscular Research Center, Boston Univ, 2002.
19. Ciaccio, E., Dunn and S., Akay, M. Biosignal pattern recognition and interpretation systems, part 2 of 4: Methods for feature extraction and selection, in IEEE Engineering in Medicine and Biology (1993), 106–113.
20. Niemann, H. Klassifikation von Mustern, (2nd Edition). Berlin: Springer, 2003.
21. Duda, R., Hart, P. and Stork, D. Pattern Classification (2nd Edition). John Wiley Sons, Inc., USA, 2001.
22. Bennett, J. How-to build guide: PC Controlled RC Car. <http://www.jbprojects.net/articles/rc/>