

Evaluating the Feasibility of EMG and Bend Sensors for Classifying Hand Gestures

Robert Tidwell, Sarath Akumalla, Sarada Karlaputi, Robert Akl, Krishna Kavi

University of North Texas, Department of Computer science Engineering
Denton, Texas, USA

RobertTidwell@my.unt.edu; SChandraAkumalla@my.unt.edu; SaradaKarlaputi@my.unt.edu;
Robert.Akl@unt.edu; Krishna.Kavi@unt.edu

David Struble

Raytheon
Dallas, Texas, USA
struble@airmail.net

Abstract - This paper presents a feasibility study using Surface Electromyography (EMG) and bend resistive sensors for hand gesture recognition. Using only surface EMG signals, we classified gestures using a neural network, logistic regression, and a Support Vector Machine (SVM) model. The results ranged widely from 40-75% accuracy. To improve accuracy, we included bend resistive sensors that are attached to each of the five fingers of the hand. This approach is sometimes known as a “data glove” and is an alternative to using surface EMG only for gesture recognition. Using the data glove accompanied with an SVM model, we achieved 93.33% accuracy over a 10 hand gesture set. With these sensors, we are able to classify hand gestures, and not arm movements. Future work will be incorporating accelerometers and varied placement of EMG sensors into the model to widen the range of gesture recognition and increase accuracy.

Keywords: Surface EMG sensors, Bend resistive sensors, SVM model, Accelerometers.

1. Introduction

Gestures, such as waving the hand or moving the leg, are commonly used in our daily lives. Gestures are often used as support for our verbal communication (Anton et al., 2013), but they can also be used as a sole, simple and effective way of communication. Gesture recognition seeks to find a way for computers and other processing units to understand human hand and arm gesture movements and thus provide a more convenient way of interaction between computers and humans (Englehart, Hudgins, 2003). To achieve these goals, we create sensor models that can collect raw signal attributes like Electromyography (EMG), Electroencephalography (EEG), etc. and also data acquired from acceleration and orientation measuring systems like gyros and accelerometers (Englehart, McIsaacs, 2006).

The controlling of external peripheral devices can be achieved by identifying the different motion commands from EMG signals. To achieve this, the pattern monikers for each type of motion command are extracted and proper mapping is applied to classify the signals. There are many pattern recognition methods available to categorize the patterns extracted from different features. Using different classification methods like artificial neural networks, using machine-learning techniques like Support Vector Machines (SVM) and logistic regressions, can do pattern matching. Artificial neural networks especially became very popular among researchers in recent times (Ahsan et al., 2011a, b)(Englehart et al., 1995).

Our study was intended to target communication by soldiers involved in close range engagement or combat. Our initial approach to classify these gestures from Fig. 1 was based on the fact that, whenever we perform a gesture an electrical potential is generated in the muscle groups present in the forearm. This electric potential generated in the muscles is called Electromyography potential or simply EMG. In this

way the EMG sensors will act as an interface between human and computer systems used for Human Computer Interaction (HCI) similar to mice, keyboards, joysticks, etc.











Standardized Hand Signals For Close Range Engagement (C.R.E.) Operations				
				
One	Two	Three	Four	Five
				
Six	Seven	Eight	Nine	Ten

Fig. 1. Hand gestures




















				
You	Me	Come	Listen or I Hear	Watch or I See
				
Hurry Up	Stop	Freeze	Cover This Area	Go Here or Move Up
				
Enemy	Hostage	Engage	Stop	Get Leader
				
Column Formation	File Formation	Line Aftrest Formation	Wedge Formation	

Fig. 2. Hand signals

There are two kinds of EMG in wide spread use: surface EMG (SEMG) and intra muscular EMG (invasive EMG). Invasive EMG provides very accurate information about the state of the muscle and its nerves. Invasive EMG is typically used in a clinical environment to determine abnormal muscle conditions. Normal muscles at rest make some “normal” electrical signals when a needle is inserted into them. Abnormal spontaneous activity indicates some muscle damage or muscle activity. Surface EMG is fundamentally noisier than invasive EMG since the motor unit action potentials (MUAP) must pass through body tissues such as fat and skin before a sensor on the surface can capture them.

Our prototype includes the use of multiple surface EMG sensors called “bioflex” sensors. The manufacturer of these sensors is a company named I-Cube-X. These devices consist of a band carrying the sensors, called gold sensors, and are intended to be attached across the muscle in consideration, as shown in Fig. 3. Proper placement of EMG sensors on the forearm muscles is necessary to obtain noise free quality SEMG signals. Improper placement of surface EMG sensors on forearm muscle groups leads to distorted EMG signals due to noise in EMG readings generated by cross talk signals in the surrounding muscles. In our experiments we modified the sensors to run parallel to forearm and underarm muscles to reduce crosstalk. In this paper we present an overall system concept and prototype implementation and

results of a feasibility study. We also include the use of bend sensors attached to a glove in a separate model, as shown in Fig. 4.



Fig. 3. Bioflex Sensors



Fig. 4. Bend Sensors mounted on a glove

2. Methodology

Fig. 5 shows our modified prototype of the bioflex sensors. The first prototype used only the EMG bioflex sensors and attempted to classify a small subset of the hand gestures in Fig. 1. The subset was: One, Three, Five, Nine, and Ten. The best achieved accuracy for the classification of this subset was approximately 75%. Several methods were used to classify the EMG signals including: time-domain analysis, frequency-domain analysis, time-frequency domain analysis and machine learning techniques (Chu et al., 2005).



Fig. 5. Modified Bioflex Sensors (placed along the upper forearm and upper underarm)

3. Bioflex Sensor Analysis

Our first prototypes initially consisted of dual channel EMG with sensors placed along the upper forearm and upper under arm. After a series of tests yielded only nominal results, we eventually used up to five channels of EMG around the majority of the forearm and underarm to capture the signals (Begg et al., 2008). Placement is shown in Fig. 5.

3.1. Time and Fourier Domain Analysis

Using the EMG sensors the raw time domain signals are sampled at 1000 samples per sec, with the highest frequency detectable at 500 Hz. The five channels showing the raw time domain plots and the

frequency domain plots are given in Fig. 6. This is the minimum recommended sampling rate for EMG. Through our experimentation we found that most of the frequency energy was within 2-200 Hz for our targeted muscle groups. Multiple techniques were tried including frequency normalization, grouping frequencies into buckets (e.g., 0-4 Hz, 4-8 Hz, etc.) and calculating the average frequency energy within each bucket (Englehart et al., 1999)(Bian et al., 2002). The results were then scaled and classified using a multi-class Support Vector Machine. Multiple kernels with the SVM were also tested including a linear, polynomial, sigmoidal, and radial based kernels. Results yielded between 40-75% accuracy.

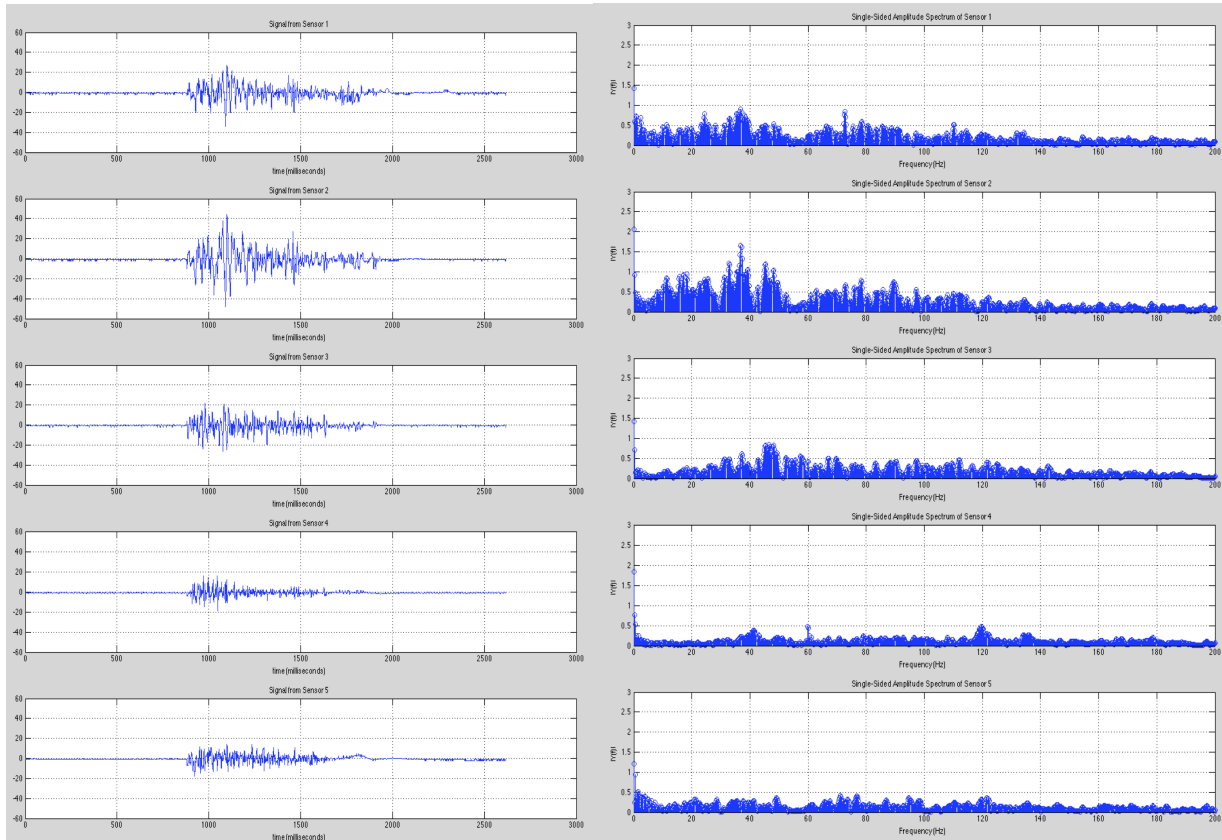


Fig. 6. Five channels showing raw time domains plots (left), Five channels showing frequency domain plots (right)

3.2. Support Vector Machine (SVM)

Two-channel System

We built a two-channel EMG sensor system with one sensor placed on top of the Flexor Carpi Radialis muscle, which runs in the upper forearm and the second one was placed on the Flexor Carpi Ulnaris muscle, which runs in the lower forearm. The classification was performed by taking samples of all the ten hand gestures (one to ten). Then the total sample set was divided into two sets; one for the training set to build a training matrix and the other a test set to build a testing matrix. The training matrix was built using the samples from the training set and the corresponding rows of the training matrix were labeled using the class labels of the gestures. The class labels of the gestures are the first ten positive integers. The class labels were placed in a column matrix with the number of rows equal to the number of rows in the training matrix. Each sample was placed in a row in the training matrix. In a similar way, the test matrix and test labels were created.

LIBSVM is an integrated software for support vector classification, regression and distribution estimation. It also supports multi class classification. LIBSVM is used for classification of the samples.

The training matrix and the training label were given as inputs to LIBSVM to create a training model. Then the model was used to predict the labels of the testing samples. The signal analysis was done in 3 domains: time domain, frequency domain, and Cepstrum domain.

Time domain is the raw time signals sampled at 1000 samples/sec. Frequency domain analysis is done by taking the Fast Fourier Transforms (FFT) of the signals in the time domain. A Cepstrum is obtained by taking the Inverse Fourier Transform of the logarithm of the estimated spectrum of the signal. Cepstrum coefficients are widely used in speech recognition, generally combined with a perceptual auditory scale.

A Cepstrum is used to separate two or more signals from a mixture. Fig. 7 shows the Cepstrum plot for gesture one in addition to the time and frequency domain plots for channel-1 and channel-2. The Cepstrum plots look similar for the different gestures. This similarity in waveforms is due to the crosstalk between different muscle pairs, which impaired localization of the sensor nodes. The bioflex sensors have circular gold sensor plates with large surface area. The signals captured by these sensors contain some noise, which makes it difficult to separate the signals from the noise.

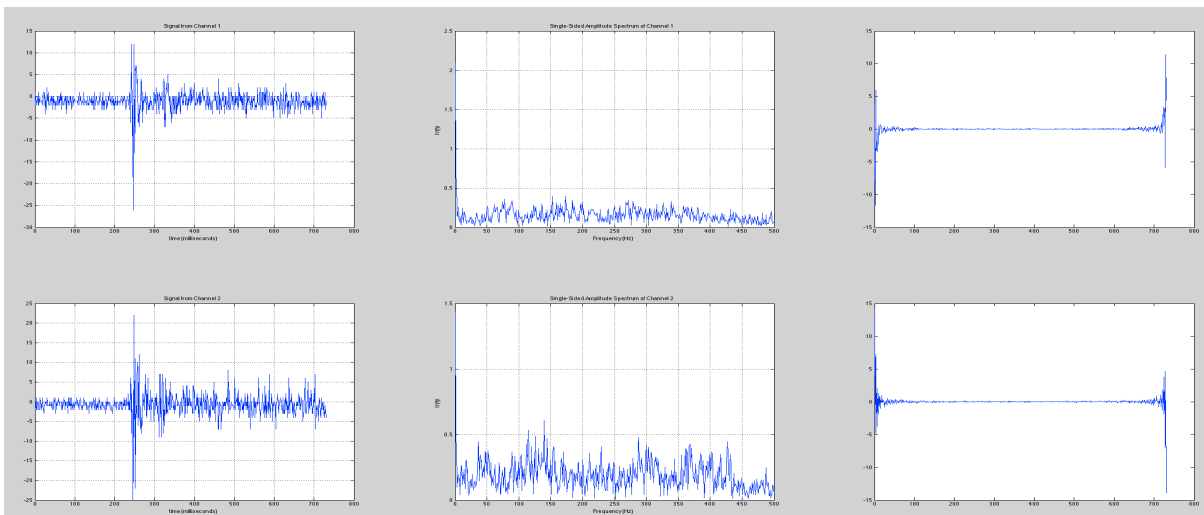


Fig. 7. Finger gesture one 2-channel signals plots. The top figures show the time, frequency and Cepstrum plots of channel-1. The bottom figures show the time, frequency and Cepstrum plots of channel-2.

Other machine learning techniques like Neural Networks and Logistic Regressions were also used to see if they could produce better classification of these samples than the support vector machines (Bishop 2006). Neural network functions can be found in the neural network toolbox of MATLAB. In a similar way there are many different types of regression functions provided in MATLAB toolboxes. Out of the techniques just mentioned, SVM gave the best classification results.

Five-channel System

The 2-channel system was expanded to 5-channel EMG sensors system. We were able to achieve classification accuracy up to 75% using the samples of the finger gestures for training and testing purposes. To increase accuracy, five gestures were selected from the set of ten gestures. The gesture subset consists of the following 5 gestures:

- Gesture 1
- Gesture 3
- Gesture 5
- Gesture 9
- Freeze (fist position of the hand)

This set provided the most distinction between the finger positions required in making the gestures. The subset selection goal was to determine the narrowly distributed frequency bandwidths, which can be used for more precise classification of these gestures.

Fresh round of samples were taken for each of the five gestures involving three different individuals. Each subject took 50 samples for each of the five gestures. On the whole each gesture has 150 samples. Each sample was broken down into two groups: data from upper forearm group of sensors and the data from lower forearm group of sensors. The primary and the secondary sensors were selected for each sample. The 5-channel data was reduced down to a 2-channel data with the selection of the two best sensor data from each group. After the selection of the primary and the secondary sensors the Fourier Transform of these samples was calculated using the FFT algorithm in MATLAB to convert the data in time domain to the frequency domain. The Fourier data for each of the five gestures were placed in separate data matrices. Mean, variance and standard deviations were calculated at each frequency for each gesture. The mean, variance and standard deviation of each frequency were plotted for each gesture separately (Chen et al., 2007).

From the analysis of the plots of the mean, variance and standard deviations of each gesture separately, it was found that all of them have one common feature. The samples were found to be very randomly distributed from their means (Ajiboye et al., 2005). In other words they had high variances and standard deviations. This analysis implied that every person had their own unique gesture performing styles. Also it was difficult to get the subjects to perform the gestures with the same intensity. One assumption from this analysis was that the samples of the individual subjects would have very narrow distribution from their respective means. The samples of just one subject were taken and similar analysis was performed. The plots showed very low variances and standard deviations. So to improve the classification accuracy the 5 channels prototype has to be calibrated from subject to subject before classification.

To determine the dominant frequency bandwidths of each gesture the difference between the samples of every combination of gesture pairs were calculated. Before taking the differences each sample was normalized to even out the variations in the samples, and to make each sample uniform. It was found that each gesture has higher amplitude spikes compared to the other gestures in a very narrow frequency bandwidth of 1-1.5 Hz. These bandwidths are very narrowly distributed in the range of 33-44 Hz.

Probability distribution tables were constructed using the dominant frequency bands of each gesture. The contents of the probability distribution tables are as follows:

- Columns of the tables consist of the dominant frequency bands for each gesture.
- Rows consists of multiple thresholds defined. A threshold is defined as the difference between the amplitude values of two gestures.
- Elements of the table consists of the probability estimation or percentage estimation of the number of different sample combinations which have a difference equal to the threshold value between the gesture pair.

Hence, all tables were constructed in a similar way. A new sample was taken and tested against these tables to estimate the probability contribution of each gesture to the sample. Which ever had the highest contribution or the highest probability then the sample was determined to belong to that gesture. So many test matrices were created by taking multiple samples and tested against these tables. The overall test results yielded on average around 70-75% accuracy.

Finally other supervised machine learning algorithms Naïve Bayes and K – nearest neighbors were used to test these samples. MATLAB provides inbuilt functions for both the Naïve Bayes and K-nearest neighbors' classification. The Naïve Bayes classifiers are based on so called Bayes Theorem and are particularly suited when the dimensionality of the input is very high. A Naïve Bayes classifier assigns a new observation to the most probable class, assuming the features are conditionally independent given the class value. The K-nearest neighbor is a method for classifying objects based on the closest training examples in the feature space. It is a type of instance-based learning where the function is only

approximated locally and all computation is postponed until classification. The results were similar to the SVM classification but LIBSVM computes at a much faster pace and thus was preferable.

4. Bend Sensor Analysis

The best achieved accuracy for the classification of the finger gestures of the standardized hand signals for close range engagement operations used by the military and swat teams using the EMG sensor affixed to the upper and lower parts of the forearm was approximately 75%. These gestures have higher voltage components in the frequency spectrum compared to the other gestures in narrowly distributed frequency bands such that it becomes very difficult to attain much higher classification accuracies. But still we can clearly distinguish the wrist movements to a much higher accuracy using EMG sensors.

So we decided to use other sensor types to classify the finger gestures to a higher accuracy. The other sensor types used are the flex sensors or the bend sensors (Anton et al., 2013). These sensors, when bent increase the resistance across the ends of the carbon element embedded in a thin substrate. This resistance can be converted to an electrical voltage when connected to a voltage divider circuit. These bend sensors use a piezo-resistive method of detecting the bend.

The piezo-resistive method describes the change in electrical resistivity of a semiconductor when a mechanical stress is applied. The sensor output is determined by both the angle between the ends of the sensors and the flex radius. Bend sensors are attached to a glove that is worn on the hand. The sensors are located over the first knuckle of each finger and can determine if the finger is bent partially, fully or straight. If the finger is bent then the signal would increase and be measured against a threshold value. If the finger is straight then the signal would remain close to zero.

To classify the gestures using the flex sensors there's no need of any specialized machine learning algorithms like support vector machines, neural networks or Naïve Bayes algorithm etc. but a simple logical function like a Boolean function is sufficient enough to correctly classify gestures. If flex sensor is bent then we can assign a variable to 1 otherwise we can assign it to 0. Depending on the gesture type there can be different combinations of 1s and 0s, which can be used to accurately classify the gestures. Based on this idea we designed a simple custom algorithm for the classification purpose. The results for all 10 hand gestures from Fig. 1 were 93.33%.

5. Conclusion

Surface EMG sensors can possibly be effective at classifying hand gestures. Several attributes such as sensor placement and sensor type and size could be refined to enhance this system. The methods used to analyze the time series signals of surface EMG are also of great importance to an overall classification model. Refining these methods as well as developing the machine learning algorithms to do the eventual classification is very important to increasing accuracy. Finally, augmentation with various other types of sensors could greatly increase the accuracy and range of possible gestures classified in the system. Future work will be incorporating accelerometers and varied placement of EMG sensors into the model to widen the range of gesture recognition and increase accuracy.

Acknowledgements

This work was sponsored in part by the National Science Foundation Industry & University Cooperative Research Center for Net-Centric Software Systems (NSF Net-Centric I/UCRC), and member companies.

References

Ahsan R., Muhammad Ibrahimy, Othman Khalifa. (2011a) Hand motion Detection from EMG signals by using ANN based classifier for Human Computer Interaction, "Modeling, Simulation and Applied Optimization (ICMSAO), 2011 4th International Conference on," April 19-21, pp. 1-6.

- Ahsan R., Muhammad Ibrahimy, Othman Khalifa (2011b). Electromyography Signal based hand gesture recognition using Artificial Neural Network, "Mechatronics (ICOM), 4th International Conference On," pp. 1-6.
- Anton S., Posternikov Anton, Osika Maxim, Yasakov Valeriy, (2013) "EnableTalk", Microsoft Imagine Cup, Sydney. <http://enabletalk.com/> consulted February 2013
- Ajiboye A., Richard F. Weir, (2005). A heuristic fuzzy logic approach to EMG pattern recognition for multifunctional prosthesis control, "IEEE Transactions on Neural Systems & Rehabilitation Engineering," 13(3), pp. 280-291.
- Begg R., D. Lai, and M. Palaniswami, (2008) "Computational Intelligence in Biomedical Engineering," CRC Press.
- Bian Z., Gongxue Zhang, (2002). "Pattern Recognition 2nd edition," Beijing: Tsinghua University Press. (In Chinese)
- Bishop C., (2006) "Pattern Recognition and Machine Learning," Springer.
- Chen X., Xu Zhang, Zhang Yan Jhao, Ji- Hai Yang, Vuokko Lantz, Kong Qaio Wang (2007) Multiple Hand Gesture Recognition based on Surface EMG Signal, "Bioinformatics and Biomedical Engineering. ICBBE 2007. The 1st International Conference on," July 6-8, pp. 506-509.
- Chu J., Inhyuk Moon, and Mu-Seong Mun, (2005). A real-time EMG pattern recognition based on linear-nonlinear feature projection for multifunction myoelectric hand, "Proceedings of the IEEE 9th International Conference on Rehabilitation Robotics," Chicago, IL, USA, pp. 295-298.
- Englehart K., B. Hudgins, M. Stevenson, and P.A. Parker, (1995). A dynamic feedforward neural network for subset classification of myoelectric signal patterns, "Engineering in Medicine and Biology Society, IEEE 17th Annual Conference," pp. 819-820.
- Englehart K., B. Hudgins, P.A. Parker, M. Stevenson, (1999). Classification of the myoelectric signal using time-frequency based representations, "Medical Engineering & Physics," 21 (6-7), pp. 431-438.
- Englehart K., B. Hudgins, (2003). A robust, real-time control scheme for multifunction myoelectric control, "Biomedical Engineering, IEEE Transactions on," vol. 50, pp. 848-854.
- Englehart K., and D. McIsaacs, (2006). The science in science fiction's artificial men, "Crosstalk: The Journal of Defense Software Engineering," 19(10), 4-8.