Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Evaluation of surface EMG features for the recognition of American Sign Language gestures

Vasiliki E. Kosmidou *Student Member, IEEE*, Leontios J. Hadjileontiadis, *Member, IEEE*, and Stavros M. Panas, *Member, IEEE*

***Abstract*—In this work, analysis of the surface electromyogram (sEMG) signal is proposed for the recognition of American Sign Language (ASL) gestures. To this purpose, sixteen features are extracted from the sEMG signal acquired from the user's forearm, and evaluated by the Mahalanobis distance criterion. Discriminant analysis is used to reduce the number of features used in the classification of the signed ASL gestures. The proposed features are tested against noise resulting in a further reduced set of features, which are evaluated for their discriminant ability. The classification results reveal that 97.7% of the inspected ASL gestures were correctly recognized using sEMG-based features, providing a promising solution to the automatic ASL gesture recognition problem.**

I. Introduction

SIGN Language (SL) is the native language for the hearing-impaired people and is comprised by a set of specific gestures. There have been many researches that attempted to capture and translate SL gestures employing cameras or sensing gloves [1]-[3]. These techniques either fail to recognize gestures in poor lighting conditions or encumber the user's hands involving expensive devices. In this approach, the use of surface electromyogram (sEMG) signal acquired from the user's forearm is proposed for the recognition of American SL (ASL) gestures. sEMG signals observed at the surface of the skin are the sum of thousands of small potentials generated in the muscle fibers. These signals are bioelectric signals that can range from 0 to 10 mV (peak-to-peak) and the usable energy lies between 0 and 500Hz. The sEMG signals can be easily acquired on the skin above the involved muscles with noninvasive surface electrodes. The use of differential amplification is recommended in order to eliminate the noise-to-signal ratio. sEMG signals from the body intact musculature are currently used to identify motion commands for prosthetic control [4]. The acquisition of EMG signals is followed by the implementation of various statistical and analytical procedures properly designed for the derivation of meaningful conclusions. The crucial point is to choose and compute these features that offer the higher discrimination ability for the recognition of the signed gestures. It has been verified [5] that specific integral absolute and zero-crossing along with frequency domain features can produce the appropriate feature space in order to classify efficiently arm motions.

In this paper, an extended set of features extracted from the acquired sEMG signals are evaluated for their efficiency in correctly classifying ASL gestures using the Mahalanobis distance criterion [6] and discriminant analysis [7]. In this way, a new field of approaching automated ASL gesture recognition is established.

This work was supported by the General Secretariat for Research and Technology (GSRT), Greek Ministry of Development, under the Grant No. 44 (22195/16 Dec. 2005).

The authors are with the Department of Electrical & Computer Engineering, Aristotle University of Thessaloniki, GR 54124, Thessaloniki, Greece. Corresponding author: V. Kosmidou, Tel.: +302310 994180, FAX: +302310 996312, E-mail: vkosm@auth.gr

II. Methodology

A. Feature Extraction

For the efficient identification of the ASL gestures a set of features from the sEMG were extracted. These features are defined in either the time or the frequency domain [8]-[11].

Let $x_i, i = 1,2,...,N$, be a $N-$sample acquired sEMG signal. The features adopted in this research are the following:

Integral of Absolute Value (IAV):

$$IAV = \frac{1}{N}\sum_{i=1}^{N}|x_i| \qquad (1)$$

Difference of Absolute Mean Value ($DAMV$):

$$DAMV = \frac{1}{N-1}\sum_{i=1}^{N-1}|x_{i+1} - x_i| \qquad (2)$$

k-th *order Zero-Crossings* (ZC_k):

$$ZC_k = \sum_{i=1}^{N-1}\mathrm{sgn}(-x_i^{(k)}x_{i+1}^{(k)}), \qquad (3)$$

where $\mathrm{sgn}(x) = \begin{cases} 1 & if \quad x > 0 \\ 0 & otherwise \end{cases}$

After experimenting for the order of the *ZC* that would offer the higher discrimination between gestures, the use of the zero-th, the first, the second and the 8-th order *ZC* is proposed.

Skewness:

$$\gamma_3 = \frac{\mu_3}{\mu_2^{3/2}} \qquad (4)$$

Kurtosis:

$$\gamma_4 = \frac{\mu_4}{\mu_2^2} - 3, \tag{5}$$

where μ_i denotes the i-th central moment

Auto-regressive coefficients (AR): The simplest time-series model is the AR model, in which signal samples are estimated by linear combination of their earlier samples. It has been shown [12] that sEMG spectrum changes with muscle contraction state, resulting in changes in AR coefficients. Various experimental and theoretical approaches have been conducted to define the adequate order of an AR model for sEMG signals. It has been derived that the model order $P = 4$ is usually suitable for EMG signals [12].

Mean Frequency (MNF):

$$f_{mean} = \frac{\int_0^\infty f \cdot P(f)\, df}{\int_0^\infty P(f)\, df}, \tag{6}$$

where $P(f)$ and f denote the power spectrum density and frequency, respectively.

Cepstral parameters: the zero-th, first and second moment of the wavelet spectrum in the fifth scale ($WaveletMoms_{0-2}$):

$$WaveletMoms_0 = \int_0^\infty S_m(f)\, df \tag{7}$$

$$WaveletMoms_1 = \int_0^\infty f \cdot S_m(f)\, df \tag{8}$$

$$WaveletMoms_2 = \int_0^\infty f^2 \cdot S_m(f)\, df, \tag{9}$$

where $S_m(f)$ is the wavelet spectrum in the fifth scale using the bi-orthogonal family wavelets [13]. The $WaveletMoms_1$ is known as the *energy* of the wavelet coefficients.

Wavelet Difference Absolute Mean Value ($WaveletDAMV$):

$$WaveletDAMV = \frac{1}{N-1} \sum_{i=1}^{N-1} \left| ci_{i+1} - ci_i \right|, \tag{10}$$

where ci_i is the i-th wavelet coefficient.

Wavelet Zero-Crossing of 2nd order ($WaveletZC_2$):

$$WaveletZC_2 = \sum_{i=1}^{N-1} \operatorname{sgn}(-ci_i^{(2)} ci_{i+1}^{(2)}) \tag{11}$$

Cepstral Coefficients ($CepCoeff$):

$$c_1 = a_1, c_n = -\sum_{i=1}^{n} \left(1 - \frac{i}{n}\right) \cdot a_i \cdot c_{n-i} - a_n, \tag{12}$$

where c_i is the i-th cepstral coefficient and a_i is the i-th AR coefficient.

The parameters of (1)-(12) consist the analysis feature set applied to the sEMG signal.

B. Discriminant analysis

The clustering and the statistical discrimination of ASL gestures are based on the theory and techniques of discriminant analysis. Discriminant analysis is used to determine which variables discriminate between two or more naturally occurring groups and consequently to reduce the number of the variables considered adequate to fully discriminate the gestures [7]. Moreover, discriminant analysis has prevailed as a useful and efficient tool for building predictive models of group membership based on observed characteristics of each case (in our case, of each gesture). The analytical procedure of discriminant analysis generates a discriminant function (or a set of discriminant functions). These functions are linear combinations of the predictor variables that provide the best discrimination between the groups (gestures). More specifically, the first function maximizes the difference between the values of the dependent variable. The second function is orthogonal (hence, uncorrelated) to it, and maximizes the differences between values of the dependent variable, controlling for the first factor, and so on. Each discriminant function is orthogonal to the others. Obviously the first function represents the most powerful differentiating dimension. The functions are determined observing a sample of cases for which the group membership is known. After generating the functions, they can be applied to new cases with measurement for the predictor variables, but with unknown group membership. Discriminant analysis shares all the usual assumptions of correlation: it requires linear and homoscedastic relationships, and untruncated interval or near interval data.

The discriminant analysis used in this approach is based on the Mahalanobis distances. Mahalanobis distance is the distance between a case and the centroid for each group in attribute space (*n*-dimensional space defined by *n* variables) [6]. There is a Mahalanobis distance for each case and each case is classified as belonging to the group for which the Mahalanobis distance is minimum. The statistical distance or Mahalanobis distance between two points $x = (x_1, x_2, ..., x_n)^t$ and $y = (y_1, y_2, ..., y_n)^t$ in the n-dimensional space $\Re^n$ from the same distribution which has a covariance matrix C is defined as

$$d_s(x, y) = \sqrt{(x - y)^t C^{-1} (x - y)}. \tag{13}$$

Obviously the Mahalanobis distance is the same as the Euclidean distance if the covariance matrix is the identity matrix.

III. Dataset Characteristics

The dataset used in this work consists of a set of two-channel sEMG signals acquired from an electromyography device in collaboration with the Myoelectrical Activity

Laboratory of Aristotle University of Thessaloniki, Thessaloniki, Greece. The sEMG signals were acquired with a sampling frequency of 1 kHz and amplified by a factor of 1000. The ASL gestures performed during the sEMG acquisition correspond to the words: *come, communicate, dream, goodbye, father, hello, how, yes* and *you*. Each gesture was repeated twenty times. The duration of each recorded sEMG signal was fixed to 3.5 sec After experimentation on the exact placement and type of the sEMG electrodes, their optimum position was identified, in terms of ensuring high signal quality and discrimination for the performed motions. In this work, we used only the right hand for the identification of the ASL gestures and the muscles *Flexor Carpi Radialis* and *Flexor Carpi Radialis Brevis* presented the most appropriate performance. The electrodes used were wet disposable electrodes that stick to the skin. An example of the two-channel EMG signals from the aforementioned muscles, when the words *come* and *yes* are performed, is depicted in Fig. 1. The features described in Section II-A were extracted within an overlapping hamming window of 512 samples length.

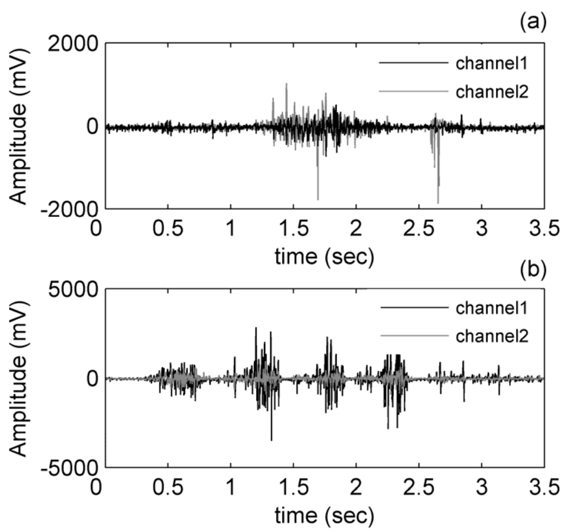

Fig. 1. Two-channel EMG signals for the words (a) *come* and (b) *yes*.

IV. Results and Discussion

The estimated features constitute the data which were statistically processed using the SPSS 13. Discriminant analysis with the Mahalanobis distance was performed and the features were evaluated by the Wilks' Lambda (WL) parameter (Fig. 2) [14]. The bar-plot of Fig. 2 represents the contribution of each feature (variable) to the discriminant function. The smaller the variability of WL for an independent feature, the more that feature contributes to the discriminant function. WL varies from 0 to 1, with 0 indicating significant difference in group means and 1 indicating equality of all group means. Obviously, by considering the parameter $(1-WL)$, the bar-plot can be interpreted as follows: the higher a bar in Fig. 2 is, the more the corresponding feature contributes to the discriminant function. The bar-plot offers a visual evaluation of the most significant features; these are the *IAV, DAMV, WaveletDAMV,* ZC_0*,* ZC_1*,* ZC_8*,* $WaveletZC_2$*, Kurtosis, MNF,* $WaveletMoms_1$, and the *AR coefficients.*

Using these eleven features, the discriminant analysis produces eight significant functions known as discriminant

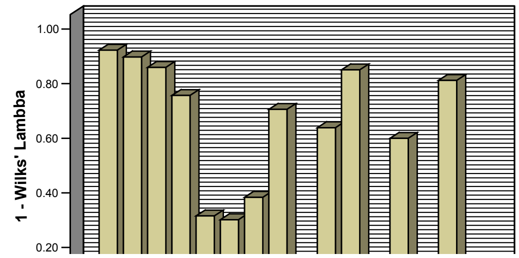

TABLE I
STANDARDIZED CANONICAL DISCRIMINANT FUNCTION COEFFICIENTS

Feature	Function 1	Function 2	Function 3	Function 4
IAV	-3.966	1.360	-0.374	-0.418
DAMV	3.329	-1.006	0.449	0.805
WaveletDAMV	-0.468	-0.001	0.101	-0.404
ZC_0	-0.381	0.687	0.026	-0.442
ZC_1	0.470	0.593	-0.196	0.801
ZC_8	0.510	0.773	1.970	0.363
$WaveletZC_2$	0.181	-0.043	-0.397	0.301
Kurtosis	0.113	-0.081	0.204	0.090
MNF	0.575	-1.071	0.851	0.074
$WaveletMoms_1$	0.182	2.046	1.281	1.462
AR	0.029	-0.101	1.290	-0.677

TABLE II
FIRST 8 CANONICAL DISCRIMINANT FUNCTIONS WERE USED IN THE ANALYSIS.

Function	Eigenvalue	% of Variance	Cumulative %
1	19.106	50.7	50.7
2	6.751	17.9	68.6
3	5.137	13.6	82.2
4	3.922	10.4	92.6
5	1.582	4.2	96.8
6	0.694	1.8	98.7
7	0.402	1.1	99.7
8	0.102	0.3	100.0

functions. The standardized discriminant coefficients for the first four discriminant functions presented in Table I are used to compare the relative importance of the independent variables (features). The eigenvalue of each discriminant function (often encountered as characteristic root) is indicative of the ratio of importance of the dimensions which classify cases of the dependent variable. The role of eigenvalues is to assess relative importance of each function, reflecting the percents of variance explained in the dependent variable. Table II presents the importance of each one of the eight functions, indicating that the first four explain the significant percent 92.6%. We can visualize how the two first functions discriminate between groups by plotting the individual scores for the two discriminant functions (Fig. 3). The group centroids are the mean discriminant scores for each of the dependent variable categories for each of the discriminant functions. The closer the means, the more errors of classification there likely will be. As it is depicted in Fig. 3, not all gestures are completely discriminated by the first two discriminant functions (in Table II it can be observed that the first two functions offer a classification score of 68.6% only). With the use of all the eight discriminant functions the gestures are recognized with a percentage of correct classification, known as *hit ratio*, of 99.4%, as denoted in Table III.

The features were tested against white noise of up to 20dB, and the noise sensitivity was evaluated, accordingly. Results from this noise stress test are depicted in Fig. 4. The features *IAV* and ZC_0 that were noted as significant at the discriminant analysis appear to have low tolerance to noise. Excluding these two features from the classification analysis, the *hit ratio* derived by using the rest nine features (i,e., *DAMV, WaveletDAMV,* ZC_1, ZC_8, *WaveletZC*$_2$, *Kurtosis*, *MNF, WaveletMoms*$_1$ and *AR coefficients*) decreases to the from 99.4% to 97.7% (Table III).

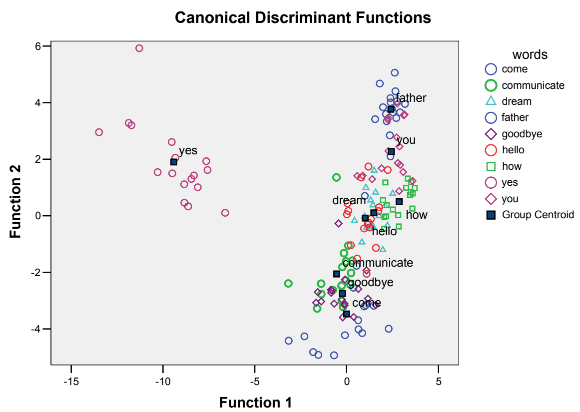

Fig. 3. Classification of the gestures using the first two discriminant functions.

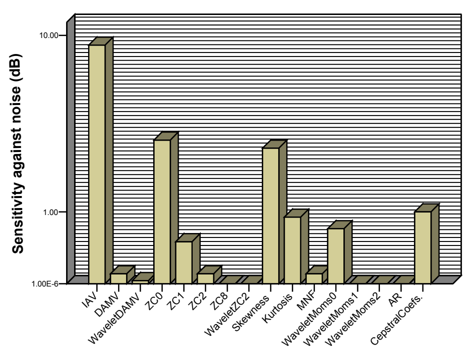

Fig. 4. Evaluation of the features sensitivity to noise

V. Conclusion

In this work, sixteen features extracted from two-channel sEMG signals were evaluated for their recognition power of nine ASL gestures. After statistical process involving discriminant analysis based on the criterion of Mahalanobis distance, the features with the adequate discrimination ability were reduced to eleven (*IAV, DAMV, WaveletDAMV,* ZC_0, ZC_1, ZC_8, *WaveletZC*$_2$, *Kurtosis*, *MNF, WaveletMoms*$_1$ and the *AR coefficients)*. These can be equally represented by the eight resulting discriminant functions for further feature reduction in classification process. Testing of these features against noise revealed that two features, i.e., *IAV* and ZC_0, were noise susceptible and they were excluded from the feature set, decreasing the hit ratio down to 97.7%. Nevertheless, the classification performance of the proposed features could be used as a set-bed for the recognition of additional ASL gestures, where the use of more than two channels sEMG signals is needed.

TABLE III
Classification Results

Words	Hit ratio from 11 features (%)	Hit ratio from 9 features (%)
come	100	100
communicate	100	100
dream	95	90
father	100	100
goodbye	100	95
hello	100	95
how	100	100
yes	100	100
you	100	100
Total	99.4	97.7

VI. Acknowledgments

The authors would like to acknowledge Ms. Katerina Dikaiou for performing the ASL gestures during the acquisition and archiving the sEMG signals, Associate Professor Christos Kontzamanidis for providing the electromyography devices, and Mr. Mavroudis Eleftheriou for his assistance with SPSS.

References

[1] T. Starner, J. Weaver, and A. Pentland, “Real-time American Sign Language recognition using desk and wearable computer based video,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, pp.1371–1375, 1998.

[2] A. D. Wexelblat, *A feature-based approach to continuous gesture analysis*. M.S. thesis, MIT, 1993. Available at: ftp://media.mit.edu/wex/MS-Thesis.ps.gz.

[3] C. Vogler and D. Metaxas, “A Framework for Recognizing the simultaneous Aspects of American Sign Language,” *Computer Vision and Image Understanding* vol. 81, pp. 358–384, 2001.

[4] E.Park and S. Meek, “Fatigue compensation of the electromyogram signal for prosthetic control and forced estimation,” *IEEE Trans. on Biomedical Engineering*, vol. 40, pp. 1019-1023, 1993.

[5] G. Saridis and T.Gootee, “EMG pattern analysis and classification for a prosthetic ARMA,” *IEEE Trans. on Biomedical Engineering*, vol 30, pp. 18-29, 1983.

[6] Hongyu Xu, “Mahalanobis Distance Based ARTMAP Networks,” M.S. thesis, Dept. Computer Science, San Diego State Univ., California, USA, Oct. 2003.

[7] B. Tabachnick, and L. Fidell, *Using multivariate statistics (4-th Edition)*, Boston: Allyn and Bacon, 2001, ch 11.

[8] Sijiang Du, M. Vuskovic, “Temporal vs. Spectral Approach to Feature Extraction from Prehensile EMG Signals,” in *Conf. Rec. 2004 IEEE Int. Conf. Information Reuse and Integration,* pp. 344–350.

[9] B. Hannaford, S. Lehman, “Short time Fourier analysis of the electromyogram: fast movements and constant contraction,” *IEEE Trans. on Biomedical Engineering*, vol. 33, pp. 1173-1181, Dec. 1986.

[10] C. Paolini, “Feature Extraction of Myoelectric Data,” *Technical Report*, Department of Mathematical Sciences, SDSU, 1994.

[11] K. Wheeler, C. Jorgensen, “Gestures as Input: Neuroelectric Joysticks and Keyboards,” *IEEE Computer Science*, pp. 56-61, June 2003.

[12] J. Graupe and Kk. Cine, “Fuctional separation of EMG signals via ARMA identification method for prosthesis control purpose,” *IEEE Trans. Syst. Man Cybern.*, vol. 5, pp.252-259, 1985.

[13] R. Boostani and M. H. Moradi, “Evaluation of the forearm EMG signal features for the control of a prosthetic hand,” *Physiological Measurements*, vol. 24, pp.309-319, March 2003.

[14] J. N. Mrija, *SPSS 11 Guide to Data Analysis*, Prentice Hall, 2002.