Robust outlier detection in high-density surface electromyographic signals

32nd Annual International Conference of the IEEE EMBS
Buenos Aires, Argentina, August 31 - September 4, 2010

Robust Outlier Detection in High-Density Surface Electromyographic Signals

H. R. Marateb, M. Rojas- Martínez, *Member, IEEE*, M. A. Mañanas Villanueva, *Member, IEEE,* R. Merletti, *Senior, IEEE*

***Abstract*— High Density surface Electromyography (HDsEMG) has been applied in both research and clinical applications for non-invasive neuromuscular assessment in several different fields using 2-D array. Proper interpretation of HDsEMG signals requires identifying "good" channels (where there is no short-circuit or bad-contact or major power line interference problem). Recording with many channels usually implies bad-contacts (that introduces large power line interference) and short-circuits (when using gels). In addition to online monitoring the electrode-contact quality, it is necessary to identify "bad" channels, or outliers, prior to the analysis of HDsEMG signal. In this paper we introduce a robust method to identify outliers in a set of monopolar HDsEMG signals recorded from Biceps and Triceps Brachii, Anconeus, Brachioradialis and Pronator Teres. The sensitivity and precision of this method show that this approach is promising.**

I. Introduction

IN the field of pattern recognition, there are several ways to define an outlier. An outlier is a) an observation that deviates so much from the others as to arouse suspicions on the mechanism behind it [1] or b) an observation (or subset of observations) which appear to be inconsistent with the reminder of the dataset [2]. Both definitions refer to some observations that affect the estimation of the general trend of the data.

Outliers affect statistical estimators such as location and scale indicators. In fact, they not only bias the estimator towards them, but also create problems when multiple outliers exist in the data. They can mask each other (masking effect), resulting in the appearance of the bulk of the data as outliers (swamping effect). This is caused by a group of outlying instances that skew the mean and the covariance estimates toward them. Thus, when dealing with outliers, non-parametric (robust) statistics must be used in order to avoid such problems.

Recording HDsEMG signals implies using several channels. When recording with many channels, it is likely to observe low-quality signals due to bad contact conditions caused by poor skin-electrode contact, small electrode displacements during signal recording, power-line interference (especially in monopolar recording) and increasing electrode-skin impedance over time, (Fig.1 as an example).

Human experts can identify outliers with high accuracy but this procedure is time-consuming and depends on the expertise of the operator. Therefore, there is a need for an automatic method to identify "bad" channels. So far, two methods were proposed: one by C. Grönlund et. al. [3] and the other one by H.R. Marateb et. al [4]. The first approach is based on Quelplot, a bivariate extension of Boxplot [5], using two-dimensional features including standard deviation of the signal in short and long epochs for each channel. The second approach is based on a Fuzzy system, and requires tuning membership functions on a training set using particle swarm optimization [6]. In none of these approaches the knowledge of Human Experts was used to identify outliers directly, e.g. in the feature extraction or classification steps. Besides, it is necessary to set some thresholds empirically or based on tuning on training sets. In this paper, we present an approach with the following properties: 1) it is easy to implement and efficient in practice, 2) knowledge of human expert is reproduced in the outlier detection procedure, and 3) the outlier detection threshold is data-dependent and no tuning set is necessary.

The performance of this approach was tested on a set of real monopolar HDsEMG signals recorded from Biceps and Triceps Brachii, Anconeus, Brachioradialis and Pronator Teres and verified by three independent experts.

II. METHODS

A. Feature Extraction

We assume that the number of outlier channels is not more than 40% of the total number of recorded channels. This condition is met in practice since during recording, channels are monitored and signal acquisition can be stopped and repeated if there are too many bad channels.

Besides, considering that the Breaking Point (BP) of the median (which is the best robust location estimator) is around the 50th percentile, the 40th percentile condition allows assuming that the noise is recognizable from the data. In that case it is possible to identify the bulk of the data if appropriate features are used [6].

Hamid R. Marateb and Roberto Merletti are with the Laboratorio di Ingegneria del Sistema Neuromusculare (LISiN), Dipartimento di Elettronica, Politencnico di Torino, Italy (e-mails: hamid.marateb@polito.it and roberto.merletti@polito.it respectively).

Mónica Rojas- Martínez (phone: +34934016971 ; fax: +34934017045; e-mail: monica.rojas@upc.edu) and Miguel Ángel Mañanas Villanueva (e-mail: Miguel.angel.mananas@upc.edu) are with the Biomedical Research Networking Center in Bioengineering, Biomaterials and Nanomedicine, CIBER-BBN, the Biomedical Research Centre, CREB and the Department ESAII at Technical University of Catalonia, UPC, Spain.

978-1-4244-4124-2/10/$25.00 ©2010 IEEE

Experts identify “bad” channels as those that differ from or that are not similar to “good” channels. To implement this approach, it is necessary to quantify similarity. We may introduce the first feature F for channel i as $F(i) = med\{CC(x_i, x_j) \quad j = 1 \dots n, i \neq j\}$, where n is the total number of recorded channels, x_i are the temporal samples of channel i in a 250 ms epoch and $CC(x_i, x_j)CC$ is the cross-correlation coefficient between channels i and j. Consequently, F is defined as a measure of similarity that reproduces expert’s outcome.

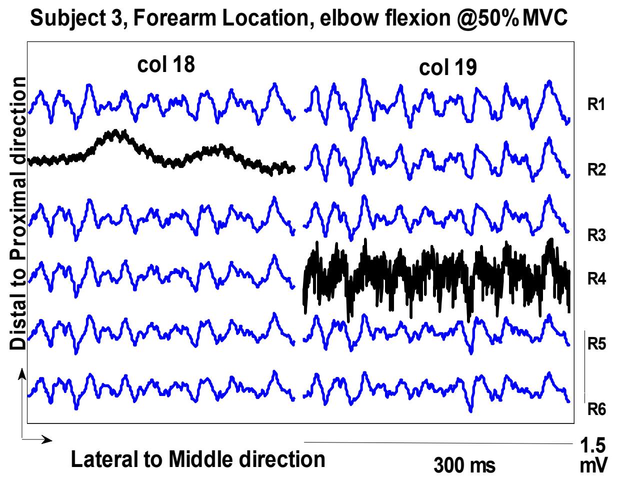

Figure 1. Monopolar HDsEMG signal set distributed in six rows (r1,r2,,…,r6) and two columns. R2 (col. 18) and R4 (col. 19) were recognized as outliers according to three experts’ opinions.

The test set consisted of 19 signal sets, each of them had at least 114 monopolar channels (refer to section C for information on the experimental protocol). The effect of multi-outliers on the extracted F feature in one of the sets is shown in Fig.2 Three independent experts agreed that outlier channels were 1, 9 ,10 ,17 ,24 ,28 ,40 ,48 ,53 ,94 ,96 ,105 , and 108 that can be discriminated using feature F without masking and/or swamping.

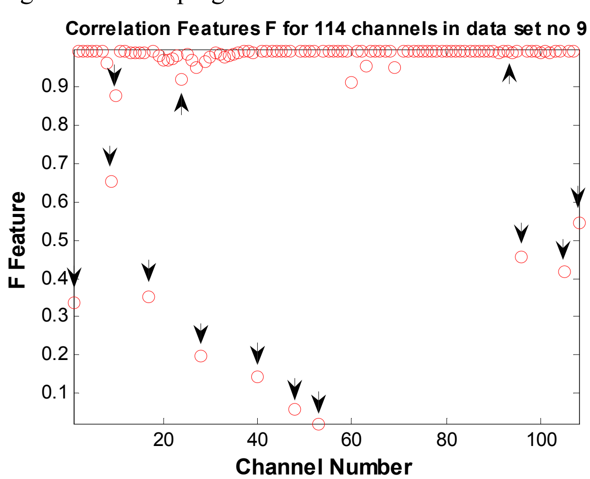

Figure 2. The value of F features for 114 channels in the ninth signal set. In this set there are 13 outliers according to the agreement between three experts. These are channels no. 1, 9 ,10 ,17 ,24 ,28 ,40

The second Bivariate feature, $P = [P_{l/t} \ \ P_{50/t}]$, is defined in the frequency domain by estimating the relative power of each channel in the low and power line frequencies (with its harmonics) with respect to the total power of the signal as:

$$P_{l/t} = \frac{P_{low}}{P_{Tot}} \quad P_{50/t} = \frac{P_{50}}{P_{Tot}}, \text{ Where}$$

$$P_{low} = \sum_{k=0}^{12} P_k \quad P_{50} = \sum_{k=1}^{4} P_{50 \cdot k} \text{ and } P_{Tot} = \sum_{k=0}^{400} P_k$$

And samples P_k are 1 Hz apart (1 (s) epochs). These features are shown in Fig. 3 for a different signal set (120 channels) after de-correlation using Principal Component Analysis (PCA). According to Fig. 3, these features can be considered as compact representations of good and bad channels.

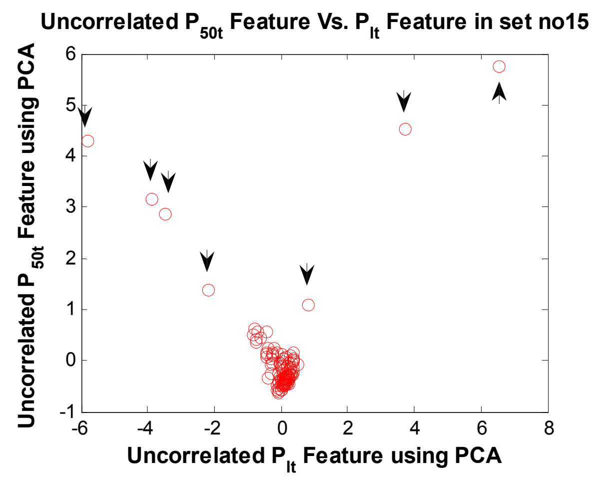

Figure 3. 2-D representation of Bivariate frequency domain features (uncorrelated $P_{l/t}$ in x -axis and $P_{50/t}$ in y-axis) for 120 channels in signal set no15. The expected detected outliers (indicated by arrow) can be recognized from the “good” channels if an appropriate area is chosen to represent the bulk of the data.

B. Classifier

There are different classifiers to identify outliers, for example, distance-based methods that identify outliers based on distance measures with respect to the bulk of the data. One of them is the Mahalanobis Distance (D_M) between each sample (y) and the rest of the data which is defined as:

$$D_M(y) = \sqrt[2]{(y-\mu)^T S^{-1} (y-\mu)}$$

Where $y = (y_1, y_2, .., y_N)^T$ is a multivariate feature vector from a group of values with mean $\mu = (\mu_1, \mu_2, .., \mu_N)^T$) and Covariance Matrix S.

However, D_M is affected by the presence of multiple outliers. Robust D_M estimators, like those proposed in [7], solve this problem but there is still the need to set a threshold to separate outliers from the rest of the signals. This threshold should be robust enough to be suitable for different data sets. Besides, in some cases the global outlier definition is not efficient for localized outliers that do not deviate that much from the main data trend. In this work, a different local-distance-based outlier detection approach is considered [8]. The output of the algorithm is the Local Distance-based Outlier Factor (LDOF) for each sample.

LDOF is calculated by using the mean of the samples as location estimator but in our case, we used the median ($\tilde{g}$) because it is more robust in the presence of outliers than the mean. Suppose that k, the number of nearest samples, is 40% of the total number of channels in each 2-D array signal, g is the feature and N_p is a set including k nearest neighbors (K-NN) of the measured feature g_p set, then

$$\bar{d}_{gp} = \|g_p - \tilde{g}\|^2 + \frac{1}{k}\sum_{g_i \in N_p} \|g_i - \tilde{g}\|^2,$$

$$\bar{D}_{gp} = \frac{2}{k-1}\sum_{g_i \in N_p} \|g_i - \tilde{g}\|^2$$

Then LDOF for each sample g_p can be calculated using:

$$LDOF_k(g_p) = \frac{\bar{d}_{gp}}{\bar{D}_{gp}}$$

The interesting property of LDOF is that when LDOF$\approx$ 0.5, the point is lying in a uniform cloud of objects [8]. Instead of setting a fixed threshold empirically or tuning it based on a training set, a data-dependent threshold was used according to the separation of clusters created after non-parametric density estimation of LDOF outputs [9]. Those channels whose LDOF output was higher than that threshold were considered as outliers. The output of LDOF classifier using feature F for the same channels as in Fig. 2 is shown in Fig.4. In this case, the obtained threshold was 0.786.

LDOF is used sequentially twice for the classification procedure: First, some channels are detected as outliers using feature F. Then the remaining channels are evaluated by using LDOF for feature P in order to find more outliers. The second step is important for identifying a number of channels that can be similar with each other but contain components induced by high power line interference and (or) baseline noise .

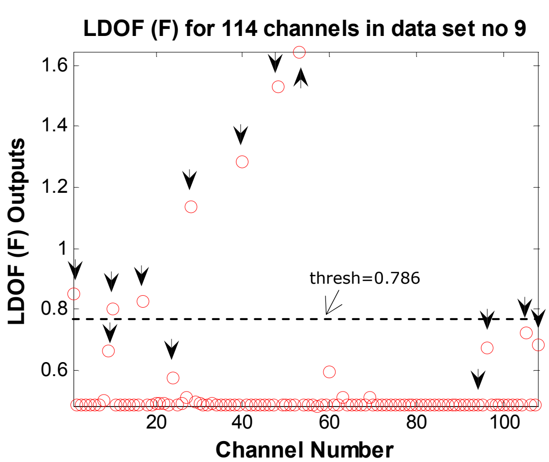

Figure 4. The output of LDOF classifier using the correlation-based *F* feature for the same 114 channels shown in Fig 2.. Outliers identified by experts are indicated by arrow. The estimated threshold to identify outliers is also shown.

C. HDsEMG recording

A protocol for the evaluation of contractions involving the elbow joint was designed. Five different muscles were included: Triceps and Biceps Brachii in the upper arm and Brachioradialis, Anconeus and Pronator Teres in the forearm. Surface EMG signals were recorded with 2-D electrode arrays. Two 15x8-channel arrays were used to cover the surface of biceps and triceps brachii on the anterior and posterior regions of the upper arm respectively, and a third array composed of 19 x 6 channels was used for the assessment of the three forearm muscles. Inter-electrode distance was 10 mm in both directions. Midpoints of the arrays were aligned with anatomical landmarks recommended by SENIAM project [10].

Monopolar HDsEMG signals were recorded simultaneously using three synchronized amplifiers (EMG-USB,3dB bandwidth 10-750 Hz, LISiN-OT Bioelettronica). A Driven Right Leg (DRL) circuit was used for reduction of power line interference. Twelve healthy male volunteers (age, 28.3 ± 5.5 years; height: 177.8 ± 6.0 cm; weight: 75.7± 8.7kg) participated in the experiment. No subject had known symptoms of neuromuscular disorders. Subjects sat with the dominant arm placed in a mechanical brace with the elbow joint flexed at 45º, shoulder abducted at 90º (parallel to the Sagittal plane), and forearm twisted 90º (midway between supination and pronation). This position was selected in order to maximize the action of the flexors and extensors of the elbow. The experiment consisted of series of flexion, extension, supination and pronation contractions in isometric condition at 10%, 30% and 50% of maximum voluntary contraction (MVC).

D. Validation of the outlier detection method

Outliers were manually detected by three different experts in 19 signal sets selected from different arrays (triceps, biceps brachii or forearm), during different contractions and force levels. The "MODE" operator (the most frequently occurring value) was used to combine three experts' opinions to identify artifacts for each channel of each set. Reliability of agreement between experts was assessed using Fleiss' Kappa index [11] and scored 88.96 % pointing to "almost perfect agreement". The next step consisted of the accuracy assessment of the algorithm in comparison with that of expert's opinions. Performance of the algorithm was assessed in terms of Sensitivity(S), Specificity (SP), Precision (P) and Accuracy (ACC) defined as (see TABLE I):

$$S = \frac{TP}{TP+FN} \qquad SP = \frac{TN}{TN+FP}$$

$$P = \frac{TP}{TP+FP} \qquad ACC = \frac{TP+TN}{TP+FP+TN+FN}$$

TP, TN, FP and FN are defined in Table I.

Note that by defining TP as the number of true outliers that were detected using the presented method, it is possible to avoid overestimation of performance indices. Since the number of "good" channels is usually much higher than the number of "bad" ones, performance indices will bias to higher values when TP is calculated as the number of good channels detected as non-outliers. The two most important performance measures are Sensitivity, the capability of the algorithm to identify outliers, and Precision, the capability of the method to preserve "good" channels.

Classical cross-validation with training and testing sets is not necessary to be considered because the proposed outlier detection approach requires no tuning and the thresholds are all data-dependent and based on theoretical concepts of robust statistics.

TABLE I
THE DEFINITON OF PERFORMANCE MEASURES

		Automatic Detection	
		Artifact	**Non artifact**
Experts' Knowledge	**Artifact**	TP	FN
	Non Artifact	FP	TN

III. RESULTS

The overall performance measures of the 19 analyzed sets are listed in TABLE II. . Additionally, the performance of the algorithm for the set displayed in Fig.2 (set no. 9) is presented as an example. Most of the outliers were detected in channels in outermost columns. The performance of the method was independent of the required tasks (flexion, extension, supination or pronation) and of the effort level (10%, 30% or 50% MVC).

One important consideration must be noted when dealing with channels lying over non active muscles. These channels are correctly classified as non-outliers. That is because this approach is local-based and these good channels are preserved even if their underlying muscles are not active.

Finally, the high standard deviation of sensitivity in our approach is due to the presence of outlier channels that are not identified using the proposed method in sets with only few artifacts that dramatically reduces their sensitivity. It can be improved using robust statistical methods to combine features and (or) better classifying strategies in comparison with the simple serial classifier used in this paper.

TABLE II
PERFORMANCE MEASURES OF THE PROPOSED OUTLIER DETECTION METHOD (MEAN (± STANDARD DEVIATION))

SET	**S (%)**	**P (%)**	**SP (%)**	**ACC (%)**
1-19	87.2 (±20.2)	90.4 (±12.7)	99.5 (±1.0)	98.9 (±1.1)
9	76.9	100.0	100.0	95.4

NOTE : S,P,SP AND ACC STAND FOR SENSITIVITY, PRECISION ,SPECIFICITY, AND ACCURACY.

IV. CONCLUSION/DISCUSSION

In this work, we presented a method to identify "bad" channels in HDsEMG signals. According to the results, this method has high precision that indicates its capability to preserve "good" channels and acceptable sensitivity to identify outliers. The definition of TP based on TABLE 1 avoids overestimation of the performance of the algorithm because the number of bad channels is much lower than the number of good ones in a standard recording. When dealing with HDsEMG, it is also very important to preserve good channels in order to evaluate Motor Unit Action Potentials morphology, conduction velocity, direction of propagation, and the global activation in the 2-D space.

In this work, an adaptive method was used to set the classifier threshold to detect outliers. This threshold depends on the distribution of the data and it is not necessary to be fitted by experimental data or obtained by cross-validation.

Future work will focus on combining appropriate classifiers to increase the overall performance of the method. Additionally, it is necessary to extend identification of interferences in the frequency domain to artifacts other than 50/60 Hz power line interferences.

ACKNOWLEDGMENT

The first author is grateful to Marjan Mansourian for fruitful discussions in this project. This work was supported by the Spanish Government (TEC2008-02754) and the Doctoral School of Politecnico di Torino, Italy.

REFERENCES

[1] D. Hawkins, "Identification of outliers," Chapman and Hall, London, 1980.

[2] V. Barnet, T. Lewis, "Outliers in statistical data," John Wiley, ISBN 0-471-93094-6, Chichester, 1994.

[3] C. Grönlund, K. Roeleveld, A. Holtermann, JS Karlsson," On-line signal quality estimation of multichannel surface electromyograms," Med Biol Eng Comput. 2005 May,43(3):357-64.

[4] H.R. Marateb, A.F. Soares, M. Rojas, R. Merletti " An expert-based fuzzy system for automatically identifying the location of muscle innervation zones in surface electromyography," Proceedings of the International Xth Quantitative EMG conference, 2009 May 6-10, Venice, Italy.

[5] H.R. Marateb, K.C. McGill, "Resolving Superimposed MUAPs using Particle Swarm Optimization," IEEE Trans. on Biomed. Eng., 2009; 56: .916-919.

[6] F.R. Hamptel, "A general qualitative measure of robustness," Annals of mathematics Statistics, 42, 1887-1896, 1971.

[7] Ali S. Hadi, "Identifying Multiple Outliers in Multivariate data," Journal of the Royal Statistical Society. Series B (Methodological), Vol. 54, No.3, (1992), pp.761-771.

[8] M.M. Breunig, et. al. , "LOF : Identifying Density-Based Local Outlier," Proc. ACM SIGMOD 2000 Int. Conf. On Management of Data, Dallas, TX,2000.

[9] A.W. Bowman, and A. Azzalini, "Applied Smoothing Techniques for Data Analysis," Oxford University Press,1997.

[10] H. Hermens et al., "European Recommendations for Surface Electromyography," SENIAM. 1999.

[11] J.L. Fleiss,"Measuring nominal scale agreement among many raters," Psychological Bulletin, Vol. 76, No. 5 pp. 378–382. 1971.