FEATURE SELECTION IN MENTAL STRESS ANALYSIS USING MULTIPLE BIOLOGICAL SIGNALS

Maria Goñi
m.goni@edu.uah.es

Inma Mohino
inmaculada.mohino@edu.uah.es

Cosme Llerena
cosme.llerena@uah.es

Roberto Gil-Pita
roberto.gil@uah.es

Manuel Rosa
manuel.rosa@uah.es

University of Alcalá
Signal Theory and Communications Department
28805 Alcalá de Henares (Madrid)
SPAIN

ABSTRACT
Stress is a response of people to face up to daily mental, emotional or physical challenges. Continuous monitoring of stress levels of a subject is of key importance to understand and control personal stress. In this sense, different biological signals can be used, such as, heart rate (HR), respiration, galvanic skin response (GSR) or electric response of the muscles.

In this paper we extract a large number of features from the aforementioned biological signals in order to classify the levels of stress. Once we calculate these features, we use a genetic algorithm combined with a least square linear discriminant (LSLD) in the aim of selecting the most suitable features, considering the error of classification. Results show that respiration is the most useful signal in the classification of stress level and specifically, entropy and recurrence analysis of that signal are the most relevant features. In the case of GSR, we observe that feet are more sensitive to changes of the electrodermal activity than hands. With respect to EMG, it is the less adequate signal to classify stress level.

KEY WORDS
Stress, feature extraction, genetic algorithm, linear discriminant, boostrap techniques.

1 Introduction

Stress is the body's reaction to a change that requires a physical, mental or emotional adjustment or response. During stress periods, many physical changes are produced, like increase of the heart rate, acceleration of the reflex, generation of hormones as adrenaline to intensify the concentration, etc. Under normal conditions, human body returns to its normal state after to face up to the stressors. Unfortunately, sometimes many stressors keep present during the daily activity, producing mental and physical disorders.

A common method to characterize the stress level is to perform an interview with a psychologist, but this method only provides instantaneous values of stress level. Continuous monitoring is essential to understand and control personal stress. Several studies have shown that stress can be assessed analysing biological signals, as for instance, electrocardiogram (ECG) [1][2][3], blood volume pulse [4], pupil diameter [4], skin temperature [4], electromyogram (EMG) [5], GSR [4], respiration [6], etc.

Previous works have demonstrated the feasibility of detecting stress from physiological measurements. For example, there are many papers based on the study and analysis of the heart rate variability (HRV), like, for instance, [1] in which 92 participants carry out a exercise test. In [2], HRV is measured from 23 healthy subjects during periods of rest and arithmetic stress aloud. In the case of [3], HRV is obtained from students during rest and during an exam. Others authors study the stress level combining different biological signals, [4] being an example in which 22 subjects are exposed to stressors, such as, public speaking, arithmetic stressors and cold pressor stressors. This work uses measures of ECG, respiration, GSR and temperature, extracting a short set of features from them. Paper [5] analyses the stress using measurements from HRV, GSR, respiration and EMG, and extracting some features from those biological signals. It uses 30 volunteers who are managed to do calculation tasks, puzzles and memory tasks. One of the most relevant studies in the field of the analysis of the stress was carried out by Healey and Picard [7], in which they show methods for gathering and analysing physiological data during real world driving tasks. They continuously record ECG, GSR, ECG and respiration signals of drivers on an established route through Boston, alternating periods of rest, highway and city, distinguishing three levels of stress. Healey and Picard's study, however, makes a brief extraction of features, and they don't make a separation between train and test in the DLMC, what produces a loss of generalization of the results.

In this work, we use the Healey and Picard's database in order to carry out a complete study of the features extracted from the biological signals and also, in the aim of improving the classification method. For this purpose, we implement a genetic algorithm, which combined with a

LSLD and bootstrapping techniques, will give us information about the most adequate signals and their corresponding features to assess the personal stress.

In Section 2 we describe the database used and the information extracted from drivers. Section 3 develops the extracted features of each of the biological signals. In Section 4 we describe the classification technique used in this article. Section 5 shows the experiments carried out and their corresponding results. Finally, Section 6 summarizes the results presented in this paper, and discusses issues to be explored further in the study of stress detection.

2 Materials and methods

The database used in this paper is included in *Physionet*, and their authors are Jennifer Healey and Rosalind Picard, whose work is exposed in [7]. This database contains a collection of biological signals corresponding to 18 volunteers. Those volunteers drive through a route along Boston's streets. The route alternates periods of rest, highway and city, simulating three levels of stress. We consider the rest state as the lower level of stress, driving on highway as the medium level, and the periods along the city as the highest level of stress.

The biological signals measured are: ECG, GSR, EMG and respiration. Furthermore, the recordings rely on markers, which indicate the transition between a state to another. Taking into account the position of the electrodes, with respect to the ECG ones, they are placed according to II configuration and the EMG electrodes are placed on the trapezium. Skin conductance is measured on the palm of the left hand, using electrodes placed on the first and middle finger, and on the sole of the left foot, using electrodes placed at both ends of the arch of the foot. Respiration is measured through chest cavity expansion and contraction using an elastic Hall effect sensor attached around the chest with a velcro band.

Not all the information of the database is used, since the information associated with some drivers is not valid. We discard drivers 1 and 3 because they have not data of the markers. We also eliminate the information of driver 4 since it has not data related to the EMG, and finally, driver 13 is eliminated as well, due to it does not capture data for the GSR on the hand. For the rest of them, we carry out a study of the markers, realizing that many of them do not allow us to distinguish in which state they are, due to different reasons. For example, one of them lost himself in the way of coming back. Another one found a closed street of the route, so his route was different to the rest. Another one forgot to press the marker at the beginning of the test, etc. Finally, a total of 8 drivers were selected, labelled as 5, 6, 7, 8, 10, 11, 12 and 15 in the original database.

3 Feature extraction

For each biological signal, we extract a set of features which were suggested to be relevant in related literature [1][3][8]. A total of 260 features are obtained. The features for each biological signal can be divided into the following groups:

- **Temporal domain features:** A set of statistical features have been applied over all the biological signals. These statistical features are the mean, geometric mean, harmonic mean, standard deviation, mean absolute deviation, variance, median, mode, maximum and minimum values, 25th and 75th percentiles, range, interquartile range, variation coefficient, central moment, skewness and kurtosis. Furthermore, other specific temporal features of each signal have been extracted, which we explain in next paragraph.
- **Frequency domain features:** In this paper, we have extracted some frequency domain features in the case of ECG, respiration and EMG.

In this section we include a review of the different measurements and the corresponding features that have been extracted in each case.

3.1 ECG

ECG is the register of the changes of electric potential generated by the cardiac cells. These variations produce the particular waves of the ECG. If we analyze the beating, we conclude that the ECG morphology changes beat to beat, and the distance between beats (measured as the distance between R waves), changes along the time. The last conclusion is based in the heart rate variability (HRV).

Concerning the temporal features for the ECG, we have considered the following measurements:

- Cardiorespiratory arrythmia index RSA_{index} [9], expressed in equation(1).

$$RSA_{index} = \frac{RR_{max} - RR_{min}}{\overline{RR}} \quad (1)$$

  where RR is the RR interval of the cardiac signal, calculated as the difference between two consecutive R waves. RR_{max} and RR_{min} are the maximum and minimum RR intervals respectively, and $\overline{RR}$ is the mean of the RR intervals.

- Difference in deep breathing DBD [10], calculated as the difference between the maximum RR interval and the minimum RR interval.
- Difference between exhalation and inspiration EI [11], which is the ratio between the RR_{max} and the RR_{min}.
- pNN50 parameter, defined as the percentage of heartbeat intervals with a difference in successive heartbeat intervals greater than 50 ms.

- RMSSD paramater [8], given by equation(2).

$$RMSSD = \sqrt{\frac{\sum_{n=1}^{N-1}(DARR_n - \overline{DARR})^2}{N-2}} \quad (2)$$

being N the total numer of RR intervals and $DARR = |RR_{n+1} - RR_n|$

- MDARR parameter, defined as median of the DARR sequence.
- Approximate entropy [2][3]: measures the complexity or irregularity of the RR series.
- Poincare Plot [12][13]: is a graphical illustration of each RR interval against the following RR interval. That is, Poincare Plot represents $RR(n)$ on the x-axis, and $RR(n+1)$ on the y-axis. The Poincare descriptors are represented using equations(3) and (4).

$$SD1^2 = Var\left(\frac{1}{\sqrt{2}}RR_n - \frac{1}{\sqrt{2}}RR_{n+1}\right) \quad (3)$$

$$SD2^2 = Var\left(\frac{1}{\sqrt{2}}RR_n + \frac{1}{\sqrt{2}}RR_{n+1}\right) \quad (4)$$

where SD is standard deviation, $Var(\cdot)$ is variance and RR_n is the n-th RR interval.

On the other hand, for the extraction of the frequency domain features, HRV spectrum was divided in a very low frequency component VLF (frequencies below 0.04 Hz), low frequency LF (0.04-0.08Hz), medium frequency MF (0.08-0.15 Hz) and high frequency HF (0.15-0.5 Hz), taking into account previous works as [1][2][7][8]. From this band, the following features were extracted:

- Integration over the power VLF, LF, MF and HF band, and its corresponding natural logarithms.
- Sum of energy in 0-0.1 Hz, 0.1-0.2 Hz, 0.2-0.3 Hz and 0.3-0.4 Hz.
- $HF/(HF+LF)$, LF/HF, $(LF+MF)/HF$ and $HF/(LF+MF+HF)$ ratios [7][8][14].
- Coefficient of component variance in LF and HF band [2]. Those components are calculated as the square root of LF or HF power divided by mean RR.

3.2 Respiration

Four series have been extracted from the respiration signal: T_i, T_e, T_{tot} and V_t. The inspiration time T_i is defined as the difference of time between the end and the beginning of an inspiration cycle. The exhalation time T_e is the difference between the end and the beginning of an exhalation cycle. The breath duration T_{tot} is calculated as the sum of T_i plus T_e. Finally, the tidal volume V_t represents the total volume of air of the lungs, at the end of an inspiration cycle.

From those time series we have extracted several features.

With regard to temporal domain, those are the extracted features:

- Changes of slope of the signal.
- VTE index, calculated as the standard deviation divided by the mean of the respiratory signal.
- Respiratory frequency, calculated as the number of inspirations during a certain time.
- Inspiration index, defined as the ratio between T_i and T_{tot} [15].
- Mean inspiratory flux, evaluated as the ratio between V_t and T_i [15].
- Relation between frequency and tidal volume f/V_t.
- Coefficients of variation of T_e, T_i, T_{tot}, V_t, T_i/T_{tot}, V_t/T_i and f/V_t [16].
- Mean of the three first derivatives of the respiration signal [16].
- Entropy of the respiration signal, its 3 first derivatives, T_e, T_i, T_{tot} and V_t, to quantify the variability of the respiration signal [16].
- Recurrence analysis. This technique allow the study of the time series through the analysis of the created recurrences on its phase space. To quantify the recurrence, Webber and Zbilut [17] proposed 3 measurements:

$$Percent\,recurrence = \frac{N_R}{N^2} \quad (5)$$

$$Percent\,determinism = \frac{N_L}{N_R} \quad (6)$$

$$Entropy = -\sum P_i log P_i \quad (7)$$

where N is the number of points in the considered time series, N_R is the number of dots to consider on the recurrence plot, N_L is the number of dots which make up diagonal lines of 3 or more consecutive points, and P_i is the fraction of the line segments which have length "i".

Regarding the frequency domain features, the selected frequency bands for the respiration are the same as the ECG, because the respiration and the heart rate are closely connected. The frequency domain features extracted are:

- Integration over the power VLF, LF, MF and HF bands.
- Sum of energy in 0-0.1 Hz, 0.1-0.2 Hz, 0.2-0.3 Hz and 0.3-0.4 Hz [7][18].

- LF/HF ratio.
- First coefficient of an autorregresive model of first order. This model has been applied to the respiration signal, the series T_e, T_i, T_{tot}, V_t, and to the relations T_i/T_{tot}, V_t/T_i and f/V_t.
- Welch periodogram [19], extracting the highest power of each analysed time frame, and the frequency corresponding to that maximum values.

3.3 Galvanic skin response

Galvanic skin response refers to the variation of the electrical properties of the skin in response to sweat secretion. This electrodermal response is, therefore, a sudden rise of the conductance of the skin.

From this response, four temporal series are defined. The amplitude of the response S_m is the difference between the peak of the response and the baseline [7][8][14]. Duration of the response S_d is the difference between the time of the response onset and the time of the peak [7][8][14]. Area of the responses S_a, calculated as $Sm \cdot Sd/2$ [4][7][14][20]. Finally, the half recovery time T_m, defined as the difference between the peak of the response and the baseline [7].

As well as the statistic features common to all the biological signals, the following temporal domain features has been calculated with the time series extracted from the GSR.

- Sum of the time series S_m, S_d, S_a and T_m [4][7][8][14].
- Standard deviation of S_m [8].
- Mean absolute deviation of S_m [4].
- Total number of the responses in the evaluated segment [5][8][14][20].
- Occurrence frequency, calculated as the frequency which more electrodermal responses appear in the assessed segment [7].

The GSR has been measured in both in the hand and in the foot.

3.4 Electromyograph

Electromyography is the study of the electrical signals generated by muscles when those are contracted. With the common extracted statistical features, we have used other temporal domain features.

- 10th and 90th percentiles [5].
- Moments of 1-5th order.
- Mean absolute value [9][21].
- Zero crossings [9][21]. With two consecutive samples x_n and x_{n+1}, the number of zero crossings is increased if:

$$\{x_n > 0\} \ \ and \ \ \{x_{n+1} < 0\} \tag{8}$$

or

$$\{x_n < 0\} \ \ and \ \ \{x_{n+1} > 0\} \tag{9}$$

- Slope sign changes SSC [9][21]. With three consecutive samples x_{n-1}, x_n and x_{n+1}, the number of SSC is increased if:

$$\{x_n > x_{n-1}\} \ \ and \ \ \{x_n > x_{n+1}\} \tag{10}$$

or

$$\{x_n < x_{n-1}\} \ \ and \ \ \{x_n > x_{n+1}\} \tag{11}$$

- Entropy [10]. Returns a non-negative number. As lower the result is, more regularity has the EMG, and vice versa.

Furthermore, we have included the following frequency domain features.

- Autorregresive model [21]. Three first coefficients of a fifth order autorregresive model were extracted.
- Periodogram. Is a form of evaluate the spectral density power of the EMG, calculating the DFT and squaring the absolute value of the result.
- Cepstral coefficients [22]. We used the first 20 MFCCs.

4 Classification method

In this section we explain the classification method used in this study (LSLD), and the selection method of features (genetic algorithm). Furthermore, we explain the validation method of bootstrap.

4.1 Least square linear discriminant

The process of assigning a certain observation to one class is known as classification. A classifier is linear when the decision rule $\mathbf{g}$ is a function of a linear combination of the observation components. This can be expressed in equation (12) as a matrix way:

$$\mathbf{g} = f\left(\mathbf{Vz}\right) \tag{12}$$

The input vector is $\mathbf{z} = [x_1 \ \ldots \ x_L \ 1]^T$ and the vector of weights is expressed in equation (13).

$$\mathbf{V} = \begin{bmatrix} w_{11} & w_{12} & \ldots & w_{1L} & b_1 \\ w_{12} & w_{22} & \ldots & w_{2L} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{C1} & w_{C2} & \ldots & w_{CL} & b_C \end{bmatrix} \tag{13}$$

There are made C different linear combinations, one per class.The decision is obtained, checking which of the outputs (linear combinations) is maximum so, the decision frontiers consist of a serial of hyperplanes:

$$\mathrm{y} = \sum_{n=1}^{L} w_n x_n + b = \begin{cases} -1, & \text{si } y < 0 \\ +1, & \text{si } y \geq 0 \end{cases} \quad (14)$$

In order to implement feature selection algorithms, we need classification algorithms with fast design and test procedures. Taking this into account, in this paper we select the LSLD, which is one of the fastest classification algorithm from the weka classification methods [23]. The objective of the linear discriminant analysis is to combine features of the set of data, to differentiate effectively between classes. The LSLD is optimized by minimizing the mean square error. Below, we show the mathematical formulation.

First of all, we define the desired output, as which we want the classifying give, and the coefficients of the linear combination are expressed in equation (15).

$$\mathbf{T} = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1N} \\ t_{21} & t_{22} & \dots & t_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ t_{C1} & t_{C2} & \dots & t_{CN} \end{bmatrix} \quad (15)$$

Where $t_{ij} = 1$ if the j-th pattern belongs to the i-th class, and -1 in other cases.
We define the design patterns as follows:

$$\mathbf{Q} = \begin{bmatrix} \mathbf{P} \\ ones(1, N) \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1N} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{L1} & x_{L2} & x_{L3} & \dots & x_{LN} \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \quad (16)$$

The obtained output is $\mathbf{Y} = \mathbf{V} \cdot \mathbf{Q}$, defined as the outputs of the linear combinations. The output is a matrix of C rows and N columns.
The error is defined as the difference between the desired output and the achieved output:

$$\mathbf{E} = \mathbf{Y} - \mathbf{T} = \mathbf{V} \cdot \mathbf{Q} - \mathbf{T} \quad (17)$$

This error is a matrix with C rows and N columns. The mean quadratic error (MSE) is derived from the coefficients and equaled to zero to minimize the error.

$$\mathbf{E} \cdot \mathbf{Q}^{\mathbf{T}} = zeros(C, L+1) \quad (18)$$

$$\mathbf{V} \cdot \mathbf{Q} \cdot \mathbf{Q}^{\mathbf{T}} = \mathbf{T} \cdot \mathbf{Q}^{\mathbf{T}} \quad (19)$$

The matrix $\mathbf{Q} \cdot \mathbf{Q}^{\mathbf{T}}$ is a quadratic matrix of $(L+1) \times (L+1)$, which could have a determinant different from zero only if $N > L$.

Finally, $\mathbf{V}$ is obtained as equation (20) indicates:

$$\mathbf{V} = \mathbf{T} \cdot \mathbf{Q}^{\mathbf{T}} \cdot \left(\mathbf{Q} \cdot \mathbf{Q}^{\mathbf{T}}\right)^{-1} \quad (20)$$

In this research, there are 3 classes to distinguish: stress on rest, highway and city. We train the classifier in a supervised mode. This is possible with a vector labels with the numbers 1 and 0, to indicate the class of each data. By this way, rest state is labelled as $[100]$, highway as $[010]$ and city as $[001]$. Like that, initially we adjust with well-known data, the vector of weights, which will classify unknown data subsequently.

4.2 Genetic algorithms

Genetic algorithms are inspired by evolution laws, described by Darwin's theory. They use the same combination of mutation, selection and crossover to evolve a solution to a problem.

In recent years, evolutionary computation and genetic algorithms have been successfully applied to the design and training of classifiers in many different pattern recognition problems. The basis of evolutionary computation techniques is the modeling of mathematical problems in an evolving way, so the same rules that are applied in natural evolution (mutation, selection and crossover) can be applied to find the mathematical solution of the problem. Each possible solution of the problem is considered as an individual, and a population of individuals is generated.

Taking into consideration the large number of features exposed in this work, brute force based feature selection methods are impractical in order to determine the best combination of features for the problem at hand. So, in this paper we use genetic algorithms, that allow to reduce the search space in the feature selection procedure. The objective of the GA is to determine the optimal set of features, which gives the minimum mean quadratic error of the least squares linear discriminant (LSLD) used over the set of data. We establish a size of population for the GA of 1000, and a number of generations of 100. We limit the maximum number of features chosen by the GA, making the experiments with different number of features. For example, the experiment with the whole set of biological signals is proven with a maximum number of features of 5, 10, 20, 30 and 40.

4.3 Bootstrap method

This technique is used to evaluate the error of the predictive model when we have not much information. The method consist of 3 stages:

- Create a train set.
- Use the train set to select the features with the genetic algorithm, and to design the LSLD classifier.
- Test through the test set.

This stages are included in a loop which is repeated so many times as bootstrap groups have been defined.

We have made a set of experiments applying bootstrap techniques for the feature selection. The defined loop

% ERROR OF EACH EXPERIMENT					
EXPERIMENT		**TIME SLOT (min.)**			
		0.5	**1**	**3**	**5**
All extracted features		17.65	14.04	13.43	19.67
ECG	All	20.59	38.47	41.35	29.51
	Statistical	45.19	44.67	47.6	46.72
	HR	48.13	46.58	50.96	45.9
	RR intervals	14.29	46.6	38.46	35.25
GSR	All	29.34	28.59	31.34	33.61
	Hand	35.88	34.28	38.61	31.97
	Foot	29.5	28.12	28.16	26.23
Respiration		17.48	16.16	17.79	14.75
EMG		35.73	34.6	36.06	33.61
All excepted GSR		19.12	15.79	18.75	17.21

Table 1. Error for each experiment, with its corresponding time slot.

No. OF SELECTED FEAT. OF THE GA					
EXPERIMENT		**TIME SLOT (min.)**			
		0.5	**1**	**3**	**5**
All extracted features		5	20	20	10
ECG	All	10	20	40	40
	Statistical	20	15	10	15
	HR	30	30	30	30
	RR intervals	15	5	8	8
GSR	All	10	5	5	5
	Hand	20	5	20	25
	Foot	10	5	15	5
Respiration		30	10	20	30
EMG		40	40	20	5
All excepted GSR		10	30	20	20

Table 2. Number of selected features by the GA, for each of the time slots under study.

has 8 repetitions, one per each driver. Each iteration of the bootstrap call the designed genetic algorithm.

For the feature extraction we have selected time slots of 5 minutes, 3 minutes, 1 minute and half a minute. Therefore, we make 4 experiments for each set of features.

5 Results

In order to study the performance of the algorithms, we carry out several experiments making different groups with the extracted features. The different groups were made with:

- All the features.
- All the extracted features of the ECG.
- Statistical features derivative directly from the ECG.
- Features corresponding to the HR.
- Derivative features from the RR intervals.
- All the features of the GSR.
- Features corresponding to the GSR in the hand.
- Features corresponding to the GSR in the foot.
- Features of the respiration.
- Features of the EMG.
- All the features except those belonging to the GSR.

For each experiment, we obtain an error for each of the time slots (5 minutes, 3 minutes, 1 minute and half minute), changing the maximum number of features used by the GA. Table 1 exposes the minimum error obtained for each experiment, with the corresponding time slot. Table 2 represents the number of selected features by the GA, for each of the minimum errors of the Table 1.

As Table 1 shows, features derived from the respiration and the RR intervals give the best results. In the study of the GSR, the foot offers better results than the hand. For the experiment with all the features, we can say that the time slot of 1 minute is better than 3 minutes, because with a similar error tax, it offers a 3 times bigger of temporal resolution. In general, the better time slots are 5 and 1 minute, being the worst selection 3 minutes.

It is possible to compare the errors of Table 1 with the number of features selected by the GA of Table 2. Focussing on the experiments which offer minimum errors (all extracted features, all the features except the GSR, and respiration), we conclude that a maximum number of features of 20, is a suitable choice.

Due to four of our experiments uses several signals, we can make an analysis of the percentage of each of the selected signals of these experiments, in order to have a global visualization of the contribution of each one. The set of features are: the statistical features over the ECG, the derived features from the HR, the features corresponding to the RR intervals, features of the GSR for the hand, features of the GSR for the foot, respiration and EMG. Table 3 shows those percentages.

	EXPERIMENT			
SIGNAL	**All feat.**	**ECG**	**GSR**	**All except GSR**
Statistical feat. of the ECG	0%	26%	-	14%
Feat. derived from the HR	9%	32%	-	15%
Feat. from the RR intervals	0%	41%	-	10%
GSR for the hand	7%	-	0%	-
GSR for the foot	15%	-	100%	-
Respiration	69%	-	-	56%
EMG	0%	-	-	5%

Table 3. % of use of each biological signal selected by the GA.

As well as Table 1, Table 3 shows that respiration is the best signal in this research. The foot give better results than the hand, and the EMG seems to be the worst signal to

detect stress.

In the experiment with all the biological signals, there are two very similar solutions with low error. These are the cases with time frames of 1 minute (with an error of 14.04%) and 3 minutes (which has an error of 13.43%), respectively. We have analyzed the sets of selected features in these cases, in order to determine which features are more frequently selected by the GA. Table 4 shows these features, from the most frequently used to the least ones, discarding those features with an use below a 25%, since we consider they are not relevant.

FEATURE	SIGNAL	% USE
Entropy of the recurrence analysis	Respiration	100%
Entropy of the 2nd derivative	Respiration	87%
Sum of the amplitude	GSR of the foot	87%
Entropy of the 3rd derivative	Respiration	81%
Percent of determinism	Respiration	50%
Mean inspiratory flux Vt/Ti	Respiration	44%
Log. of $HF/(LF+MF+HF)$ ratio	HR	44%
Maximum of the ECG	ECG	37%
Entropy of the tidal volume	Respiration	31%
Frequency corresponding to the max. power of the Welch periodogram	Respiration	31%
$(LF+MF)/HF$ ratio	HR	31%
Sum of the total tidal volume	Respiration	25%
Range	Respiration	25%
Interquartile range	Respiration	25%
Interquartile range	GSR of the hand	25%
RMSSD parameter	RR intervals	25%
Geometric mean	ECG	25%

Table 4. More significant features between the two cases of minimum error, in the experiment with all the biological signals.

As we can see, those features based on the respiration signal are the most frequently selected. Mainly, we highlight the use of the entropy and those features obtained from the recurrence analysis. In the case of the GSR, the sum of the amplitude of the electrodermal response in the foot is the most important. The frequency ratios for the HR are also important.

6 Conclusions

The aim of this research based on achieving the classification of three levels of stress through a database with biological signals such as ECG, GSR, respiration and EMG for a group of drivers. For that purpose, we propose a genetic algorithm combined with LSLD and bootstrap techniques, to select the combination of features. The feature extraction is one of the most important stage on the design of the classifying. In this study, we contribute of a wide collection of features derived from each of the biological signals. We describe temporal and frequency domain features, and its corresponding ability to discriminate the three levels of stress under study. This large number of features has been obtained from a large bibliography review. Although the study of the stress is not new, most of the literature is based only on one of the signals (i.e. [1][2][3]), or on several signals but with a very limited number of features (i.e. [4][5]). Therefore, this research makes a bigger study, with a higher collection of features about other studies.

One of the best contributions of this study is based on the conclusions obtained from the feature combinations. Entropy is the most frequently selected feature. The use of respiration outperforms the use of any other biological signals included in the study. Inside the respiration, the entropy and the recurrence analysis are the principal methods in this research. Concerning the analysis of the GSR, the foot is the best place of detection of the stress, being the best feature the sum of the amplitude of the electrodermal response of the foot. In the case of the EMG, it is not obviously a good signal for detecting stress. Frequency domain features are less used than temporal ones, and time slots of 5 minutes are the more suitable, followed by the time slots of 1 minute.

As future work, it is possible to extend the database with other biological signals. That is because is well-known that exist other biological signals to detect stress, as the blood volume pulse, the corporal temperature, EEG, pupil diameter, etc. Bear in mind that the entropy and the recurrence analysis are the best features for the classification, we suggest the use of them over other biological signals. Furthermore, there are many studies to detect stress through the variability on speech. So a possible future study is to combine the biological signals of this work, with the study of the speech.

Acknowledgement

This work has been funded by the Spanish Ministry of Defense (DN8644-ATREC) and by the University of Alcala under project UAH2011/EXP-028.

References

[1] K. Dishman, Y. Nakamura, E. Garcia, W. Thompson, L. Dunn, and N. Blair, Heart rate variability, trait anxiety, and perceived stress among physically fit men and women, *Int. J. of Psychophisiology*, 37, 2000, 121-133.

[2] V. Vuksanovic and V. Gal, Heart rate variability in mental stress aloud, *Medical Engineering and physics*, 29, 2006, 344-349.

[3] P. Melillo, M. Bracale, and L. Pecchia, Nonlinear heart rate variability features for real-life stress detection. Case study: students under stress due to university examination, *Biomedical Engineering Online*, 2011.

[4] Y. Shi, M. Hoai, and P. Blitz, Personalized stress detection from physiological measurements, *International Symposium on Quality of Life Technology*, 2010.

[5] J. Wijsman, B. Grundlehner, H. Liu, H. Hermens, and

J. Penders, Towards mental stress detection using wearable physiological sensors, *33rd Annual International Conference of the IEEE.*, 2011, 1798-1801.

[6] E. Bruce, Measures of respiratory pattern variability, *Bioengineering approaches to pulmonary physiology and medicine*, 1996, 149-159.

[7] J. A. Healey and R. W. Picard, Detecting stress during real-world driving tasks using physiological sensors, *IEEE Transactions on Intelligent Transportation Systems*, 6(2), 2005.

[8] F. T. Sun, C. Kuo, H. Cheng, S. Buthpitiya, and P. Collins, Activity-aware mental stress detection using physiological sensors, Silicon Valley Campus, 2010.

[9] M. Khezri and M. Jahed, Real-time intelligent pattern recognition algorithm for surface emg signals, *BioMedical Engineering OnLine*, 2007.

[10] S. Ahmad and P. Chappell, Surface emg classification using moving approximate entropy and fuzzy logic for prosthesis control, *Proc. of the 2008 MyoElectric Controls/Powered Prosthetics Symposium*, New Brunswick, Canada, 2008.

[11] G. Sundkvist, L. Almer, and B. Lilja, Respiratory influence on heart rate in diabetes mellitus, *British medical journal*, 1(6168), 1979, 924-925.

[12] D. Kim, J. Kim, and C.Youn, Poincare geometry-characterized arrhythmia identification scheme un grid, *International journal of engineering science and technology*, 1(3), 2009, 146-152.

[13] J. Piskorski and P. Guzik, Geometry of the poincaré plot of rr intervals and its asymmetry in healthy adults, *Physiological measurement*, 28, 2007, 287.

[14] J. A. Healey, *Wearable and automotive systems for affect recognition from physiology*, Ph.D. dissertation, Massachusetts Institute of Technology, 2000.

[15] C. M. de Muller, *Invasive and non-invasive assessment of upper airway obstruction and respiratory effort with nasal airflow and esophageal pressure analysis during sleep*, Ph.D. dissertation, Institut de Bioenginyeria de Catalunya, 2010.

[16] E. Bruce, Measures of respiratory pattern variability, *Bioengineering approaches to pulmonary physiology and medicine*, 1996, 149-159.

[17] C. J. Webber and J. Zbilut, Dynamical assessment of physiological systems and states using recurrence plot strategies, *J. Appl.Physiol*, 1994.

[18] R. Picard, E. Vyzas, and J. Healey, Toward machine emotional intelligence: Analysis of affective physiological state, *IEEE Transactions on pattern Analysis and Machine Intelligence*, 23(10), 2001, 1175-1191.

[19] M. Gruber, *Statistical digital signal processing and modeling* (New York: John Wiley and Sons Inc., 1996).

[20] Zhai, J., and A. Barreto, Stress detection in computer users through non-invasive monitoring of physiological signals, Florida International University, 2008.

[21] A. Ahmad, J. Ishak, H. Ali, and H. Chappell, Review of electromyography control systems based on pattern recognition for prosthesis control application, *Australian Journal of Basic and Applied Sciences*, 5, 2011, 1512-1518, 2011.

[22] W. Kang, J. Shiu, C. Cheng, J. Lai, H. Tsao, and T. Kuo., The application of cepstral coefficients and maximum likelihood method in emg pattern recognition, *IEEE Transactions on Biomedical Engineering*, 8, 1995.

[23] Ian H. Witten, Eibe Frank, Mark A. Hall, *Data Mining: Practical machine learning tools and techniques* (San Francisco: Morgan Kaufmann,2011).