Statistical Analysis of EKG, GSR and Respiration Features for Stress Recognition in Automobile Drivers

Erica Di Marco[1], Filippo Ventura[1], Elena Vitti[1]

[1]Università Politecnica delle Marche, Master course of Biomedical Engineering, Department of Information Engineering

Abstract

This paper presents a method for analyze physiological signals, Electrocardiogram (EKG), Skin Conductance (foot GSR) and Respiration, to extract features able to identify drivers' stress level while driving. The data used in this study come from an already existing database in which physiological signals were acquired while the drivers followed a prescribed path specifically designed to make the drivers experience different levels of stress. Features were extracted from 5 minutes intervals during the rest, highway and city driving conditions and the most significant ones were selected through statistical analysis. The results indicate that the three best indicators for stress recognition are features coming from foot GSR and Respiration signals for which an accuracy up 94% is achieved. These findings could be managed for in-vehicle information systems to improve drivers' safety and comfort.

1. Introduction

The progressive increase in the use of cars as a means of transport has led to an increase in car-induced accidents as consequence of driver stress or lack of attention which could be affected by emotional events. Driving in a stressful condition, for example in the city or highway, is associated with life-threatening situations and compromises the decision-making skills. The physiological response of the human body to stress causes an increase in heart rate, respiration rate, muscles contraction, sweating, etc. [1] In some previous studies, different metrices and signals have been employed to develop stress recognition algorithms. Some of these studies argued that considering a single feature coming from only one signal, is not sufficiently informative. In contrast, some other research, employ only one signal to detect stress levels. For example, some works focused their attention on skin conductance signal [2], some others on Electromyography (EMG) [3], others on Respiration signal [4-5].

This paper presents a method to measure stress using physiological signals such as Respiration, foot GSR and Heart Rate Variability (HRV), that are useful metrics to provide feedback from the drivers' state. Furthermore, from previous studies, we have noticed that features extracted from these signals provide the best indicators of driver stress.

The goal of this experiment is to select which features, among those we extracted from the three previously mentioned signals, singularly provide the best ability to distinguish between low, medium, and high stress during driving.

2. Methods

The data used in this study are taken from the "Stress Recognition in Automobile Drivers" database. This database contains a collection of multiparameter recordings from healthy volunteers, taken while they were driving on a prescribed route in the greater Boston area. The path was specifically designed to make the drivers experience different level of stress; the drive included periods of rest, highway and city driving that were assumed to produce low, medium, and high levels of stress [1]. For each driver of the database, six signals were acquired using five physiological sensors: an electrocardiogram (EKG) on the chest, an electromyogram (EMG) on the left shoulder, a chest cavity expansion respiration sensor (Resp.), around the diaphragm, two skin conductivity (SC) sensors, one on

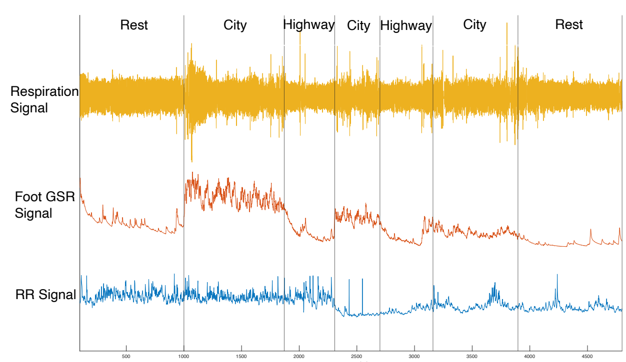

Figure 1. *Illustration of the Respiration, foot GSR and hearth rate signals.*

the left hand (hand GSR) and one on the left foot (foot GSR) and a marker signal, used to divide the driving path in the three different periods.

2.1. Feature extraction

In our study we analysed three signals: EKG, foot GSR and Resp. using the marker signal to distinguish the three levels of stress. Since some drivers' data sources lack some information, we only used ten drivers' data among the seventeen available. For each signal, 5 minutes nonoverlapping segments of data are taken from each of the rest, city, and highway driving periods, to represent a period of low, medium, or high stress.

The feature extraction was performed through MATLAB_R2021a®.

The EKG signal was used to extract the R peaks, using Pan-Tompkins peak detection algorithm. The Heart Rate Variability (HRV) signal, which reflects the variations of the period between consecutive heartbeats over time, was extracted from RR intervals. From the HRV signal six different features were extracted, see *Table 1.*

To analyse the foot GSR we followed two approaches: in the first approach we simply normalized the signal and in the second approach, instead, the signal was first smoothed by convolution with a 500 points Hanning window and then normalized. The normalization was performed in the following way:

$$G = \frac{g - \min(g)}{\max(g) - \min(g)} \quad (1)$$

where g is the raw signal and G is the normalized one.

With the smoothed signal we have extracted onsets and peaks using the Healey et al. algorithm [6]. The foot features are reported in *Table 1.*

The respiration signal was first filtered with a Butterworth pass band filter of the sixth order with a cut off frequency range between [0.16 0.5] Hz, to remove the noise. To extract the features, we have identified the peaks using the function 'findpeaks'. The respiration features we have considered are four, as shown in *Table 1*.

2.2. Statistical Analysis

The statistical analysis was performed on MATLAB_R2021a®. For all the features obtained for each driver we performed the one-way analysis of variance (ANOVA) test to evaluate whether the means of several groups are equal. The p-value produced by the ANOVA test was compared with the significance level that we assumed equal to 0.05. If the p-value was smaller than 0.05, the means feature differences were considered statistically significant, otherwise not. Multiple comparison of means was performed to obtain the p-value and the corresponding confidence interval for each pair of stress levels to statistically discriminate or not each group. Finally, we computed the Receiver Operating Characteristics (ROC) curve to evaluate the accuracy of each feature to recognize low, medium, or high stress. The ROC curve was obtained using the 'perfcurve' function in MATLAB, that returns the Area Under the ROC Curve (AUC). Also, we have calculated the Optimal Operating Point (opt) of the ROC curve as the point in which the Sensitivity is equal to Specificity.

Table 1. *Description of the extracted features*

Symbol	Unit	Feature description
RR features		
mRR	s	Mean of RR intervals
sRR	s	Standard deviation of RR intervals
rMSSD	s	Root mean square of successive differences of RR intervals
pNN50	%	Corresponding percentage of number of interval differences of successive RR intervals greater than 50ms
HRV	n.u.	Ratio of the low (LF) and high (HF) frequency heart rate spectral energies (LF/HF) [1]
HRV2	n.u.	Ratio of the low (LF) and high (HF) frequency heart rate spectral energies using the midfrequency (MF) range ((LF+MF) / HF) [1]
Foot GSR features		
mfoot	mV	Mean of normalized foot GSR signal
sfoot	mV	Standard deviation of normalized foot GSR signal
area_foot	mV*samples	Area under the normalized foot GSR signal
mfoot_smooth	mV	Mean of normalized and smoothed foot GSR signal
sfoot_smooth	mV	Standard deviation of normalized and smoothed foot GSR signal
area_foot_smooth	mV*samples	Area under the normalized and smoothed foot GSR signal
npeak	n.u.	Number of peaks of normalized and smoothed foot GSR signal
rise_time	s	Mean rise times between onsets and peaks
amp	mV	Mean of amplitudes between onsets and peaks
Respiration		
mlocs	s	Mean distance between peaks
mpks	mV	Mean of peaks amplitude
spks	mV	Standard deviation of peaks amplitude
Rrate	cycles/min	Respiration Rate

3. Results

In *Table 2* are reported the p-values obtained with the ANOVA test for each of the nineteen features, their AUCs, and the corresponding thresholds.

Note that the AUCs and the opt-thresholds have been calculated only for those features that have a p-values lower than 0.05, except for the mfoot_smooth and area_foot_smooth.

The results of the Multiple comparison of means are reported in *Figure 2,* in which are represented the differences between each couple of stress classes (Low-Medium, Low-High, and Medium-High). The multiple comparison has been performed on all the nineteen features while the AUC and the opt-threshold has been calculated

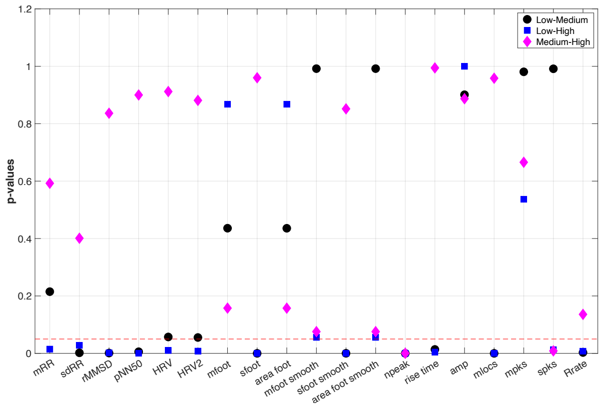

Figure 2. *p-values in multiple-comparison analysis*

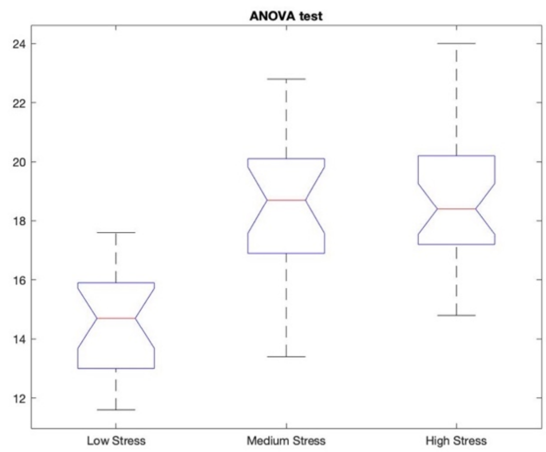

Figure 3. *ANOVA test for the Respiration Rate*

Table 2. *p-values, AUC values and opt-threshold values*

Features	p-value	AUC	opt-threshold
mRR	0.0201	0.6806	0.8850
sRR	0.0022	0.6949	0.0683
rMSSD	$4.5040*10^{-4}$	0.7684	0.0497
pNN50	$4.0945*10^{-4}$	0.7510	27.8195
HRV	0.0104	0.7571	2.4468
HRV2	0.0085	0.7592	3.1465
mfoot	0.1785		
sfoot	$5.8167*10^{-6}$	0.8190	0.2167
area_foot	0.1784		
mfoot_smooth	0.0291		
sfoot_smooth	$4.4127*10^{-6}$	0.8320	0.2444
area_foot_smooth	0.0291		
npeak	$8.0826*10^{-16}$	0.9470	8
rise_time	0.0032	0.7130	11.7276
amp	0.8765		
mlocs	$2.0903*10^{-10}$	0.9290	3.6738
mpks	0.5083		
spks	0.0028	0.6290	0.0519
Rrate	0.0031	0.9280	16.6

only for the thirteen selected features. It is important to recall that p-values higher than 0.05 are considered as errors that reduce the precision of classification.

In *Figure 3* are plotted the results of the ANOVA test done for the Respiration Rate, through three box plots in which are represented the median value (in red) and the 25^{th} and 75^{th} percentiles for each class of stress (upper and lower blue lines).

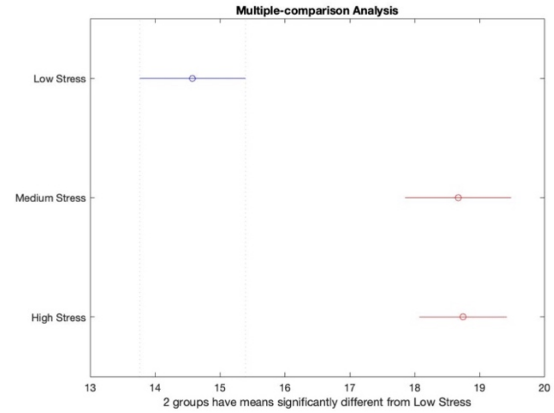

Figure 4. *Multiple-comparison analysis for the Respiration Rate*

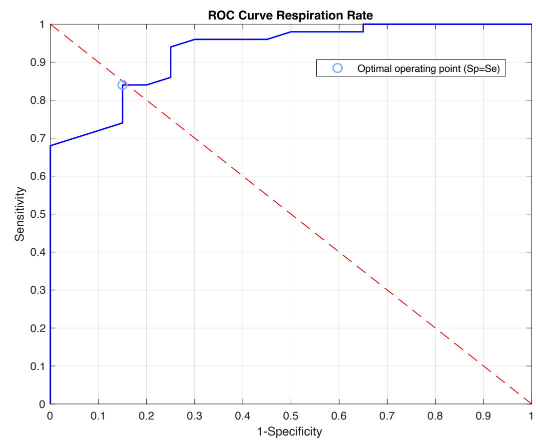

Figure 5. *ROC curve of the Respiration Rate*

In *Figure 4* is reported the result of the multiple comparison done for the Respiration Rate, in which are represented the 95% confidence intervals of that feature among the classes of low, medium, or high stress.

Figure 5 shows the ROC curve computed for the Respiration Rate feature, where it's highlighted with a circle the optimal operating point, i.e., the point in which the ROC curve intersects the bisector of the plane, which is also the point where the Specificity equals the Sensitivity of the test.

4. Discussion

In our study we wanted to analyze biological signals acquired during the drive to understand if there are some features that can be used to recognize the level of stress of the driver. The results we have obtained may be implemented in a stress recognition algorithm in the automobile applications to monitor the real time stress level of the driver and to consequently modify the driving environment.

According to previous studies we have noticed that the most meaningful signal for this purpose is the EKG signal and the derived heart rate signal [1]. Otherwise, some other studies have focused on the skin conductivity signals, finding that the foot GSR signal provides more precise information about the drivers' stress level with respect to the hand GSR signal [2]. Another study focused on the fact that using features coming from further signals increased the accuracy of the classification algorithm [3]. For this reason, we chose to also use the Respiration signal in our study.

The goal of our study was to understand which features are better discriminators for stress levels. To achieve this goal, we have first performed the ANOVA test for all our nineteen features to understand if numerical values assumed by the three stress levels are statistically different. We assumed the significance level of the ANOVA test at 5%, so that we could discard all the features with a p-value greater than 0.05, since are considered not significant. Looking at the results obtained, we decided to reject four features, because their p-value was greater than 0.05, see *Table 2*. Then we also performed the multiple comparison to verify if there are individual differences in the means of each pair of stress levels. The results, shown in *Figure 1*, confirmed what has been observed in the ANOVA test, except for the mfoot_smooth and area_foot_smooth features, for which the p-value obtained from ANOVA was lower than 0.05, however the confidence intervals obtained with the multiple comparison analysis included the 0 value, meaning that the differences are not significant. Moreover *Figure 2* suggests that, in many of the cases, the most significant differences are between Low-Medium and Low-High levels of stress, meaning that the Medium and High classes may be considered as a unique class of stress. Consequently, in the next steps of the Statistical Analysis, we have considered only two classes: the stress class (Medium and High level of stress) and the non-stress class (Low level of stress).

Thus, we considered only thirteen features as relevant among the nineteen available. For those thirteen features we have found the ROC curve and we have calculated the AUC and the optimal threshold values. The AUC value reflects the accuracy of the feature to recognize the level of stress: if AUC=0.5 the test is not informative, if $0.5<AUC\leq0.7$ the test has a low accuracy, if $0.7<AUC\leq0.9$ the test has a moderate accuracy, if $0.9<AUC<1$ the test is highly accurate and if AUC=1 the test is perfect. Referring to *Table 2,* the most accurate features are npeaks, mlocs and Rrate, which have an $AUC>0.9$. The optimal threshold has been calculated as the value for which the Sensitivity is equal to the Specificity in the ROC curve. This is the usual approach when you are performing a test for a non-critical clinical diagnosis, as in our case.

The results of this study illustrate that despite introducing the EKG (HRV) as the best stress indicator in previous studies, the properly selected foot GSR and Respiration signals are able to detect the stress level, as well as the EKG.

5. Conclusions

The present study proposed a method to identify the stress levels through EKG, foot GSR and Respiration signals. Our research has shown that these signals can successfully identify the presence of stress with an accuracy that ranges between 62% and 94%. Since the three features npeaks, mlocs and Rrate show high accuracy, we can conclude that, in our study, foot GSR and Respiration signals are the most reliable data source for stress recognition. One possible limitation of this study is the fact that our stress recognition algorithm is based on individual features, whereas it may be better to consider several features simultaneously to improve the accuracy and the precision of the classification. On the other hand, it is convenient to use less sensors, to have less interference with the driver during the driving task, endorsing the fact that a single signal measurement procedure would be an ideal perspective for future stress detection devices. Since this research is based on pre-acquired datasets on the Physionet website, utilizing more accurate and complete data would improve the accuracy of classification. Despite these limitations, our measures could be managed for in-vehicle information systems to improve drivers' safety and comfort. The perspective of future research could be implementing a single wearable sensor that automatically acquires signals, processes them, and communicates to the car the drivers' stress condition.

References

[1] Healey, Jennifer A., and Rosalind W. Picard. "Detecting stress during real-world driving tasks using physiological sensors." *IEEE Transactions on intelligent transportation systems* 6.2 (2005): 156-166.

[2] Memar, Maryam, and Amin Mokaribolhassan. "Stress level classification using statistical analysis of skin conductance signal while driving." *SN Applied Sciences* 3.1 (2021): 1-9.

[3] Pourmohammadi, Sara, and Ali Maleki. "Stress detection using ECG and EMG signals: A comprehensive study." *Computer methods and programs in biomedicine* 193 (2020): 105482.

[4] Masaoka, Yuri, and Ikuo Homma. "Anxiety and respiratory patterns: their relationship during mental stress and physical load." *International Journal of Psychophysiology* 27.2 (1997): 153-159.

[5] Vlemincx, Elke, et al. "Sigh rate and respiratory variability during mental load and sustained attention." *Psychophysiology* 48.1 (2011): 117-120.

[6] Healey, Jennifer Anne. *Wearable and automotive systems for affect recognition from physiology*. Diss. Massachusetts Institute of Technology, 2000.