

# A Stress-Detection System Based on Physiological Signals and Fuzzy Logic

Alberto de Santos Sierra, Carmen Sánchez Ávila, Javier Guerra Casanova, and Gonzalo Bailador del Pozo

**Abstract**—A stress-detection system is proposed based on physiological signals. Concretely, galvanic skin response (GSR) and heart rate (HR) are proposed to provide information on the state of mind of an individual, due to their nonintrusiveness and noninvasiveness. Furthermore, specific psychological experiments were designed to induce properly stress on individuals in order to acquire a database for training, validating, and testing the proposed system. Such system is based on fuzzy logic, and it described the behavior of an individual under stressing stimuli in terms of HR and GSR. The stress-detection accuracy obtained is 99.5% by acquiring HR and GSR during a period of 10 s, and what is more, rates over 90% of success are achieved by decreasing that acquisition period to 3–5 s. Finally, this paper comes up with a proposal that an accurate stress detection only requires two physiological signals, namely, HR and GSR, and the fact that the proposed stress-detection system is suitable for real-time applications.

**Index Terms**—Biometrics, fuzzy logic, galvanic skin response (GSR), heart rate (HR), physiological signals, stress detection, stress template.

## I. INTRODUCTION

CURRENT biometric systems attempt to identify/authenticate an individual in a unique and precise manner. Biometrics focus on such specific aim mainly under the assumption that registered or identified user will not act maliciously, and therefore, overall security will be ensured. However, a biometric system (whatever its complexity is) fails in one single aspect: What would happen if a registered user is forced to be utilized as a key to enter the system?

Nowadays, biometrics cannot provide any solution to that scenario. Nonetheless, those previous scenarios where a person is forced to access a system or entrance have a well-known common denominator: stress. Human body will react by increasing the blood volume pressure and the quantity of hormones. Furthermore, some basic functions will be avoided like hunger, sleepiness, and the like, focusing on the scenario and the malicious agent. In other words, the human body prepares to fight [1], [2].

This stress response is almost impossible to disguise and, therefore, is an accurate indicator about the security compromise.

Manuscript received June 11, 2010; revised September 28, 2010; accepted October 7, 2010. Date of publication January 6, 2011; date of current version August 30, 2011.

The authors are with the Group of Biometrics, Biosignals, and Security, Centro de Domótica Integral, Polytechnical University of Madrid, Campus de Montegancedo, 28223 Madrid, Spain (e-mail: alberto@cedint.upm.es; csa@cedint.upm.es; jguerra@cedint.upm.es; gbailador@cedint.upm.es).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIE.2010.2103538

A stress-detection system offers a solution for the previous situation where biometrics failed. This system will come up with an understanding of the state of mind of the user who attempts to access the system, so that if such a person is abnormally out of her or his normal values for stress/relaxation, then something strange is happening, and therefore, security could be compromised.

This paper proposes a stress-detection system based on two physiological parameters, namely, heart rate (HR) and galvanic skin response (GSR), together with fuzzy expert systems to elucidate to what extent an individual is under stress. This system is able to provide a very fast decision on the state of mind of the user, which is very suitable for real applications. Furthermore, its simplicity and noninvasiveness make this approach a possible system to be easily embedded not only in current biometric systems but also in general accessing systems.

Results highlight the fact that stress detection can be achieved with an accuracy of 99.5% by measuring HR and GSR during 10 s. Furthermore, rates higher than 90% can be achieved by decreasing that acquisition time.

The layout of this paper consists of a state-of-the-art in relation to stress detection (Section II), a description of the physiological signals involved in this system (Section III), and which experiments were carried out to validate the system (Section IV). The system is described in detail in Section V, and its implementation is presented in Section VI. An evaluation of the whole system is provided within Section VII. Finally, conclusions and future work (Section VIII) end this paper.

## II. LITERATURE REVIEW

Stress detection has been considered from different points of view and approaches. The work presented by Andren and Funk [3] provides a system that is able to compute the stress level of an individual by the manner and rhythm in which a person types characters on a keyboard or keypad. Furthermore, Dinges *et al.* [4] provide a study of stress detection based on facial recognition.

However, the method proposed in this paper focuses on stress detection based only on physiological signals. Considering this, there exist many previous works related to this topic. The essay presented by Begum *et al.* [5] presents a study of stress detection based only on finger temperature (FT), together with fuzzy logic [6] and case-based reasoning [3].

HR variability has been also considered as a stress marker in human body. Due to this reason, HR has been widely studied and analyzed. Several authors consider this signal in

their reports: Jovanov *et al.* [7] presented a stress monitoring system based on a distributed wireless architecture implemented on intelligent sensors. HR was recorded along different locations in individual body by means of sensors located beneath clothes.

Nonetheless, it is not common to focus only on one certain physiological feature but to focus on many of them in order to obtain further and more precise information about the state of mind. Considering this multimodal approach, there are several articles which study a variety of parameters and signals, as well as the combination of them.

Continuing on this research line, the research provided in [8], [9] proposes a system gathering FT, GSR, and blood volume pulse (BVP). The main characteristic of this system lies on the fact that signals are acquired in a nonintrusive manner, and furthermore, these previous physiological signals provide a predictable relation with stress variation.

However, not all physiological signals involve an electrical component. Pupil dilation (PD) and eyetracking (ET) provide very precise information about frame stress. When an individual is under stress, PD is wider, and the eye movement is faster. The article presented in [10] not only considers PD and ET but also GSR, BVP, and FT. The main purpose of this approach is to recognize emotions, interest, and attention from emotion recognition. Moreover, it is possible to deduce the intention of the individual from these results, a very remarkable conclusion for future computer applications and for the sake of a better human–computer interaction (HCI) [11], [12].

The work presented by Sarkar [11] proposes fuzzy logic (as Jiang and Wang [13]) to elucidate up to what extent a user is under stress. Furthermore, it introduces an approach oriented to improve HCI. Moreover, the work presented by Healey and Picard [2] deserves special mention, since they are considered to be pioneers on stress-detection field.

The research by Lisetti and Nasoz [14] provides a complete study on emotion recognition, including a deep literature review on the experiments carried out to provoke emotions considering populations, algorithms, approaches, and so forth.

### III. PHYSIOLOGICAL SIGNALS

Several indicators have been considered within the literature to detect stress (Section II). However, this paper proposes the use of only two signals: GSR, also known as skin conductance (SC), and HR. These two signals were selected based on their properties regarding noninvasivity when being acquired and because their variation is strongly related to stress stimuli [2], [10], [15].

GSR, known also as electrodermal activity, is an indicator of SC [15], [16]. In details, glands in the skin produce ionic sweat, provoking alterations of electric conductivity. First experiment on this nature dates back to 1907, when Jung first described some relations between emotions and the response of this parameter [8], [9].

GSR can be obtained by different methods, but the device proposed to acquire signals (Section IV-A) is based on an exosomatic acquisition. In other words, extracting skin conductivity requires a small current passing through the skin. GSR is

typically acquired in hand fingers, and its units of measure are microsiemens (micromhos) [8].

The main parameters of GSR, such as basis threshold, peaks, or frequency variation, vary enormously among different individuals, and thus, no general features can be extracted from GSR signals. Therefore, the parameters extracted from GSR signals are strongly related to each individual. However, despite the differences among individuals, GSR signal is not distinctive enough to identify an individual in terms of biometrics.

On the other hand, HR measures the number of heartbeats per unit of time. HR can be obtained at any place on the human body, being an accessible parameter that can be acquired easily [7], [17].

HR describes the heart activity when the autonomic nervous system (ANS) attempts to tackle with the human-body demands depending on the stimuli received [18]. ANS orders to increase the blood volume within the veins so that the rest of the body can react properly, if required. Concretely, a stressing stimulus provokes a similar reaction, increasing the number of heartbeats when facing these situations.

Among the wide number of methods to extract HR, the most common methods consider to measure the frequency of the well-known QRS complex in an electrocardiogram (ECG) signal [19], [20]. In contrast to ECG biometric properties [21], HR is not distinctive enough to identify an individual. Nevertheless, HR signal behaves in a different manner depending on the stimulus and the individual.

Summarizing, both HR and GSR behave differently for each individual, and therefore, posterior stress template must gather properly these unique responses in order to obtain an accurate result in stress detection.

### IV. DATABASE ACQUISITION

Every biometric system requires a set of different biometric acquisitions to train, validate, and assess the whole system [22]. This section provides an overview of how the data set was built considering the experimental setup and the characteristics of the database.

#### A. Overview

The experiments were carried out in the Faraday room in the Human Psychology Laboratory, Psychology Faculty, Complutense University of Madrid (UCM), endowed with electromagnetic, thermal, and acoustic insulation.

The main aim of this step is to collect HR and GSR signals from each participant. The device proposed to carry out these experiments is I-330-C2 PHYSIOLAB (J & J Engineering), which is able to process and store six channels, including electromyography, ECG, respiration rate, HR, and GSR. Sensors were attached to the hand's right (or left, but not both) fingers [23], wrist, and ankle in order to acquire both HR and GSR, avoiding sensor detachments, unplugged connectors to the analog-to-digital converter, and/or software acquisition errors. Moreover, the sample acquisition rate is made with one sample per second for both HR and GSR.

## B. Participants

The participants were mainly students from the Psychology Faculty (UCM) and Social Work (UCM), with a total of 80 female individuals of ages from 19 to 32 years old, with an average of 21.8 years old and a standard deviation of 2.15. The lack of male individuals is due to the Faculty where the experiment took place, since the percentage of male to female students is almost negligible. Therefore, no male candidates were included during collection of the database.

## C. Task Justification

Provoking stress on an individual requires a specific experimental design in order to obtain a proper arousal in terms of physiological signal [2], [4]. Concretely, this paper proposes to induce stress by using hyperventilation (HV) and talk preparation (TP) [24].

HV is defined as a certain kind of breath, which exceeds standard metabolic demands as a result of excess in respiratory rhythm.

As a consequence, several physiological changes emerge, such as arterial-pressure diminution in blood until a certain level, so-called hypocapnea [24], [25], and blood pH increment, known as alkalosis.

However, voluntary HV does not always produce an actual anxiety reaction [24], and therefore, an additional anxiogenic task is required to ensure that a positive valence in terms of stress response is provoked. Such a task is TP.

Results provided by Cano-Vindel *et al.* [24] and Zvolensky and Eifert [25] highlight that HV produces a physiological reaction (in terms of physiological registration, HR, and GSR) similar to the reaction induced by a threatening task of preparing a talk.

As a conclusion, TP and HV provoke both an alteration in physiological parameters together with different emotional experiences. These previous tasks have been widely studied and evaluated with positive results, and they are very suitable to induce stressing stimuli on individuals.

## D. Procedure

Two groups (namely, Group 1 and Group 2) were created, ensuring that the distribution of their respective anxiety levels, measured by psychological tests [26], [27], were similar. In other words, this selection seeks to avoid one group containing people which barely react against stress and another group with people which overreact under stressing conditions. Therefore, both groups must be well balanced in terms of anxiety levels in order to validate the experiments.

Participants from Group 1 underwent an experimental session using physiological and subjective signals under the following conditions: calm state (baseline, namely, BL1), stimulating task (HV), threatening task (TP), and baseline poststress (BL2). On the other hand, the order of tasks was swapped for participants from Group 2: calm state (baseline), threatening task (TP), stimulating task (HV) and baseline poststress. The main reason for altering the order is to make the task order independent from the results obtained [24], [26].

Obviously, BL1 implies no stressing stimuli on the individual in contrast to HV and TP. However, nothing can be assured in relation to BL2, since it cannot be considered either as a stressing or as a relaxing state [24], [26], [27].

## E. Database Discussion

Two questions arise from this database acquisition: Do these experiments assure that the final system is able to detect stress/relax states when performing real applications by being trained with this database? To what extent could the results obtained with these experiments be generalized to a wider population, including males and females with a larger age range?

The former question has an affirmative answer. The physiological response to a stressing agent is strongly related to each individual, and such a response is similar, independent of the time during the stressing stimulus provoked the response [28]. Moreover, stress mechanism extracts some information from the stimuli so that if such stressing agent appears again, the human body is able to react faster and better compared with the first time [1], [29]. This characteristic makes useless the repetition of the same tasks after a certain period of time and furthermore makes unnecessary a third session with different tasks, since the response will not be the same as the stimuli provided by different task provoke different responses.

Finally, the latter question is hard to answer since it is difficult, even for expert psychologists, to state whether the response among female and male individuals differs, inasmuch as the previous response varies within female individuals [30]. Several researchers support the idea that male and female individuals suffer different responses when stress agent endures through time, (e.g., a great amount of work at job, a bad economical situation, and so forth), but they have similar responses when the stress stimuli consist of specific actions in a very short period of time, e.g., an accident, an armed robbery, and the like [29].

Thereby, it is justified to extend the results obtained with this database to a wider population.

## V. STRESS-DETECTION SYSTEM

A stress-detection system inherits several characteristics from biometric systems. First of all, a template extraction is required so that the system could create a profile in order to contrast, in future accesses, whether a user is actually under stress. This template is based on specific characteristics extracted from individual concerning parameters from the physiological signals HR and GSR.

On the other hand, once the user is associated to a template, the individual is able to access the system, and therefore, a template comparison is required. Both steps are described in following sections.

### A. Template Extraction

Mathematically, both HR and GSR are considered as stochastic signals. Therefore,  $\mathcal{H}$  represents the space of HR possible signals, and  $\mathcal{G}$  represents the space of GSR possible signals.

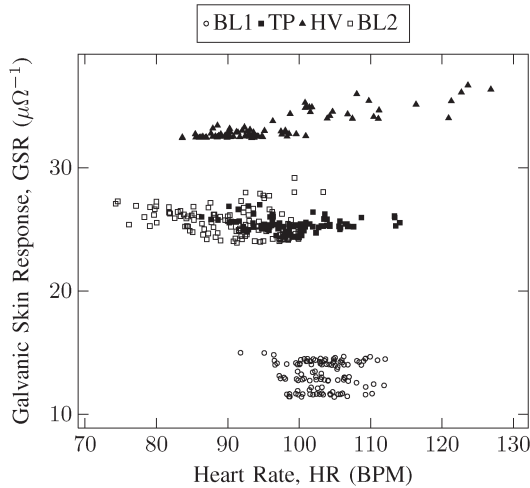


Fig. 1. Graphic representation of  $\gamma$ . Notice how the relation between HR and GSR varies depending on the stressing stimuli (BL1, TP, HV, and BL2).

Each stage will come up with a pair of signals  $h \in \mathcal{H}$  and  $g \in \mathcal{G}$  according to the experimental task conducted in each situation. Thus, a template extraction requires four pair of signals, namely,  $\gamma = [(h_1, g_1), (h_2, g_2), (h_3, g_3), (h_4, g_4)] \in \mathcal{H} \times \mathcal{G}$  corresponding to how the individual behaves under different states. Notice that signals  $h_i$  and  $g_i$  are not normalized, in contrast to previous approaches [2], [8]. The decision to avoid normalization was done based on experience, since data without normalization provide more accurate results in terms of stress detection.

Once  $\gamma$  is obtained, for each pair of signals  $(h_i, g_i)$ ,  $i = \{1, 2, 3, 4\}$ , a mean vector is obtained together with the deviation for each pair. In other words, four parameters are obtained:  $\zeta_{h_i} = \bar{h}_i$ , and  $\zeta_{g_i} = \bar{g}_i$ , which represent the mean of signals  $h_i$  and  $g_i$  in addition to  $\sigma_{h_i}$  and  $\sigma_{g_i}$ , related to the dispersion for each pair. Finally, the stress template, namely,  $\mathcal{T}$  is described by  $\mathcal{T} = (\zeta_{h_i}, \zeta_{g_i}, \sigma_{h_i}, \sigma_{g_i})$ ,  $i = \{1, 2, 3, 4\}$ .

Fig. 1 shows a visual example of a scattering representation of each pair of signals  $\gamma$ . Notice how nonstressing stimuli provokes a low excitation in GSR (Fig. 1,  $\circ$ ) and, on the contrary, the evidence of an arousal when undergoing an stressing tasks, such as TP (Fig. 1,  $\blacksquare$ ) and HV (Fig. 1,  $\blacktriangle$ ).

The aim of this action is to described the information in HR and GSR by four Gaussian distributions, centered in  $(\zeta_{h_i}, \zeta_{g_i})$  and with standard deviations  $\sigma_{h_i}$  and  $\sigma_{g_i}$ . This approach will facilitate the implementation of fuzzy antecedent membership functions by Gaussian distributions in a posterior fuzzy decision algorithm.

Let  $t_{\mathcal{T}}$  be the time used to acquire both signals in order to extract the stress template. Evidently, the performance of the system depends on this parameter since the longer  $t_{\mathcal{T}}$  is, the more information the system obtains, and therefore, the stress template may be more accurate. A study regarding this relation between  $t_{\mathcal{T}}$  and system performance is presented in Section VII-D.

Finally, after template extraction, the template must be stored. The template  $\mathcal{T}$  requires  $16 \times 32$  b since each template element (whatever  $\zeta_{h_i}$ ,  $\zeta_{g_i}$ ,  $\sigma_{h_i}$  or  $\sigma_{g_i}$  be), is represented by a float element or, in other words, 512 b, i.e., 64 B.

## B. Template Update

As a matter of fact,  $\mathcal{T}$  is different for each individual, and therefore, it must be stored as a whole template. However, these values do not remain unalterable but change along time. Consequently, this template must be refreshed every time the user accesses the system so that the template can adapt to the variation of the specific individual. Such template update must be performed each time the user uses the system. This template-update scheme remains as future work (Section VIII).

## VI. IMPLEMENTATION

This paper proposes two possible system implementations: a manual implementation, where the system parameters are set without any learning procedure, and an automatic implementation, where the system adapts its internal parameters to relate input with output during training stage.

The motivation for designing two systems is to compare an understandable system based on expert knowledge (manual implementation), containing very simple and intuitive rules, with an automatic system, which learns the rules according to the data, providing a more complex system.

In detail, the main difference between them relies on the fuzzy decision algorithm: First, manual implementation is based on a Mamdani [31] fuzzy decision algorithm using template  $\mathcal{T}$  to describe antecedent membership functions and implementing the output with triangular distributions. Second, the automatic implementation involves an adaptive-network-based fuzzy inference system [31], [32] fuzzy algorithm carried out to provide a fuzzy decision system adapted to the specific data (i.e., HR and GSR signals in different tasks—BL1, TP, HV, and BL2).

### A. Manual Implementation

As introduced previously, the manual implementation is designed with a Mamdani fuzzy system since it has been widely used in expert decision systems due to their comprehensible rules [31], [33]. This implementation provides a fuzzy output on the interval  $[0, 1]$  based on template  $\mathcal{T}$ .

This fuzzy system is described by Gaussian-based antecedent functions whose parameters coincide with centroids  $\zeta_{h_i}$  and  $\zeta_{g_i}$  and deviations  $\sigma_{h_i}$  and  $\sigma_{g_i}$ . In other words, the variables are represented by HR and GSR, considering linguistic labels for each task in the experiments (BL1, HV, TP, and BL2). Therefore, these functions describe the behavior of HR and GSR, respectively, under stressing stimuli provoked by experimental tasks BL1, TP, HV, and BL2. Gaussian-function selection is justified since they are very suitable for the data provided by HR and GSR signals.

Furthermore, two triangular functions describing each output possibility (stress or nonstress) characterize the consequent membership functions. The triangular functions were selected according to their adequate properties to provide an accurate output [33].

The proposed set of rules considers four rules, one for each stressing/nonstressing situation (BL1, HV, TP, and BL2). These

rules are very intuitive, and they are provided by an expert, involving the previously described parameters  $\zeta_{h_i}$ ,  $\zeta_{g_i}$ ,  $\sigma_{h_i}$ , and  $\sigma_{g_i}$ .

Before describing the rules in detail, some nomenclature is required. Let  $h_i^k$  and  $g_i^k$  be the physiological signals HR and GSR, respectively, for a given  $k$  subject ( $k \in \{1, \dots, 80\}$ ) in the state  $i$ ,  $i \in \{1, 2, 3, 4\}$ . This state  $i$  represents the four situations corresponding to the experiments to induce stress: BL1, HV, TP, and BL2. Let  $s_m^k$  be the output of the system, given a  $k$  subject and  $m$  representing the two possible aforementioned situations: stress ( $m = S$ ) or nonstress ( $m = NS$ ). Therefore, the rules are mathematically described as follows:

$$\begin{aligned} h_1^k \cap g_1^k &\rightarrow s_{m=NS}^k \\ h_2^k \cap g_2^k &\rightarrow s_{m=S}^k \\ h_3^k \cap g_3^k &\rightarrow s_{m=S}^k \\ h_4^k \cap g_4^k &\rightarrow s_{m=S \text{ or } NS}^k \end{aligned}$$

The latter rule makes reference to the poststress state, which is difficult to be classified as stress or nonstress, as stated previously. This duality will come up with two possible implementation schemes with each of them considering this output as stress and nonstress (Section VII-B).

Finally, the defuzzification method is carried out by a centroid method [6].

### B. Automatic Implementation

A different approach is used to improve the previous manual implementation since the algorithm learns which membership functions are to be used and which parameters must be selected in order to obtain a given output. This automatic implementation based on adaptive neuro-fuzzy inference system, which provides a Sugeno-type fuzzy inference system.

In other words, Sugeno implementation considers as initial conditions the membership functions from the previous Mamdani implementation. Consequently, this implementation needs a training stage to obtain the knowledge that is enough to properly detect stress [32].

### C. Differences Between Implementations

The main difference between the previous implementations (manual versus automatic) is related to the manner by which both systems are modeled. The Mamdani system models from the beginning both the output and the input and requires stress-template parameters (centroids and standard deviations). The output was modeled previously to the data, and that model was established manually based on a trial-and-error strategy.

On the other hand, the Sugeno implementation is able to come up with a more accurate and precise output starting from an initial condition based on  $\mathcal{T}$ . Furthermore, no previous knowledge about the output distribution is required since this system is able to provide an optimum output, by adapting its rules and membership function to the training data.

Regarding these rules, both implementations were eventually forced to have as many rules as the stages involved during

the training. For instance, if the training stages require only data from BL1 and TP, then there will be only two rules, one corresponding to each stage. Finally, there is a difference in terms of performance, since the Sugeno implementation offers a higher success rate in differentiating stress from relax states.

### D. Stress Measurement

Previous systems provide an output for each single pair of values  $(h_i, g_i) \in \mathcal{H} \times \mathcal{G}$ . In other words, these systems can provide an output each second.

However, the system provides an output each  $t_{\text{acq}}$  seconds in order to obtain a more representative result. Therefore, the output consists of a vector of  $t_{\text{acq}}$  points. The final value will result as a metric from this previous vector.

Two measurements are extracted from that output, namely, media  $\mu$  and median  $\mu_{1/2}$ . The result provided by these indicators is also verified to be in the interval  $[0, 1]$ .

These proposed measurements provide a result each  $t_{\text{acq}}$  seconds, and, as discussed posteriorly in Section VII, measurements based on median  $\mu_{1/2}$  usually produces a more accurate result when compared with  $\mu$  measurements.

Finally, the system must decide which output provided by previous metrics ( $\mu$  or  $\mu_{1/2}$ ) corresponds to a stress or a calm state. Although the output is provided within the interval  $[0, 1]$ , there must be a threshold on that interval indicating the boundary between values belonging to stress state or calm state. This threshold, namely,  $\rho_{th}$  will be obtained so that the performance of the system is maximized. Furthermore, threshold  $\rho_{th}$  is fixed and unique for the whole database. The idea of obtaining one threshold for each individual could improve the system performance. The study of this proposal remains as future work (Section VIII).

## VII. RESULTS

This section attempts to gather a study of the system performance in relation to those parameters proposed in the previous sections: threshold  $\rho_{th}$  and temporal parameters ( $t_{\mathcal{T}}$  and  $t_{\text{acq}}$ ), implementations (Mamdani or Sugeno), and stress measurements ( $\mu$  and  $\mu_{1/2}$ ).

### A. Database: Training, Validation, and Testing Data

In order to obtain valid results, the database must be divided into three groups.

- 1) Training data: Used to extract the template, i.e.,  $\mathcal{T} = (\zeta_{h_i}, \zeta_{g_i}, \sigma_{h_i}, \sigma_{g_i})$ .
- 2) Validation data: Used to fixed threshold  $\rho_{th}$  and temporal parameters ( $t_{\mathcal{T}}$  and  $t_{\text{acq}}$ ) in order to maximize the performance of the system.
- 3) Testing data: Used to obtain which implementation and metric are most suitable, and therefore, what is the performance of the whole system.

For each individual, a vector containing  $t_{\mathcal{T}}$  seconds of  $\gamma$  (for each task BL1, HV, TP, and BL2) was used for training data; a vector of  $t_{\text{acq}}$  seconds for each task was used to validate the system, and the rest of the data were used as testing data. These

latter testing data will be split in slots of  $t_{acq}$  seconds. Notice that 1 s corresponds to one sample in HR and GSR 1-D signals (Section IV-A). This assignment is done randomly. Notice that this validation scheme is similar to a  $K$ -fold cross validation.

The justification for this division is based on the research carried out by Picard and Healey [18], where several physiological signals (not only HR and GSR) were recorded during a period of time of 32 days in the same person. Eight emotions were provoked during 30 min per day, and no substantial changes were appreciated during that period in each emotions. In other words, physiological signals behave similarly in each task through time, and therefore, the  $h$  and  $g$  signals can be divided into smaller parts, considering each segment as an independent acquisition.

### B. Evaluation Schemes

As a general idea and according to the procedure described in Section V-A, detecting stress involves the acquisition of physiological signals during four different stages: BL1, TP, HV, and BL2 (Section IV-D).

However, is it extremely necessary to make individuals come through all the previous tasks to detect their stress? In other words, would it be possible to detect stress being trained only with data extracted during a BL1 session?

In order to provide some light on this aspect, nine schemes are proposed with the sole purpose of answering the previous questions. These schemes consider all possible combinations among BL1, TP, HV, and BL2, as listed in the following.

- 1) BL1 + TP + HV + BL2 $\alpha$ : This scheme proposes the use of all stages, considering BL2 as a nonstressing stage.
- 2) BL1 + TP + HV + BL2 $\beta$ : This scheme proposes the use of all stages but considering BL2 as a stressing stage.
- 3) BL1 + TP + HV: Scheme proposed without BL2.
- 4) BL1 + TP: Scheme considering both stressing and non-stressing stages.
- 5) BL1 + HV: Scheme considering both stressing and non-stressing stages.
- 6) HV + TP: Scheme considering only stressing stages.
- 7) BL1: Scheme considering only a nonstressing stage.
- 8) HV: Scheme considering only a stressing stage.
- 9) TP: Scheme considering only a stressing stage.

In summary, these schemes attempt to study up to what extent all the previous stages (BL1, TP, HV, and BL2) are necessary for the sake of a precise stress detection. Notice that template  $\mathcal{T}$  is adapted to each scheme.

The subsequent experiments are based on this experimental setup.

### C. True Stress Detection versus True Nonstress Detection

A stress detection system must reach a compromise between detecting properly which individuals are under stress situations and which individuals are in a normal state of mind.

Thereby, two evaluation parameters are defined.

- 1) True stress detection rate (TSD): when the system properly detects stress when an individual is under stress

stimuli. This TSD factor corresponds to the sensitivity statistical measure since TSD can be described as follows:

$$\text{TSD} = \frac{\#\text{True Positives}}{\#\text{True Positives} + \#\text{False Negatives}} \quad (1)$$

where a True Positive means classifying as stressed an individual which is indeed under stress, and False Negative means classifying as relaxed an individual which is under stressing situations.

- 2) True nonstress detection rate (TNSD): when the system correctly detects no stress in an individual and the subject is indeed not under stressing situations. This TNSD factor corresponds to the specificity statistical measure since TNSD can be described by

$$\text{TNSD} = \frac{\#\text{True Negatives}}{\#\text{True Negatives} + \#\text{False Positives}} \quad (2)$$

where a True Negative means classifying as nonstressed an individual which is not under stress, and False Positive means classifying as stressed an individual which is calm and relaxed.

Obviously, TSD and TNSD depend strongly on threshold  $\rho_{th}$ . If  $\rho_{th} \rightarrow 0$ , then the system considers every output as a stress stimuli (TSD decreases and TNSD increases), and vice versa.

Therefore, a compromise must be reached by finding a certain threshold  $\rho_{th}$  where TSD is equal to TNSD. This threshold is defined as true equal stress detection rate (TESD).

This latter parameter is of outstanding importance since the performance of a stress-detection system can be given by TESD as it indicates that the precise point where the errors in detecting false positives (i.e., detecting a stressing situation when the user is calm) and false negatives (the opposite situation) is the same.

Notice that the higher the TESD is, the more accurate is the performance of the system.

At this point, one question arises: Which is the best indicator (TESD, TSD, or TNSD) to provide an evaluation on the performance of a stress-detection system? TESD is obtained with validation data, and therefore, threshold  $\rho_{th}$  and temporal parameters  $t_{\mathcal{T}}$  and  $t_{acq}$  are fixed to maximized TESD. These parameters are set *a posteriori* [22]. On the other hand, TSD and TNSD are obtained with testing data, i.e., TSD and TNSD give an understanding on how the system behaves with real data. Notice that the previous parameters ( $\rho_{th}$ ,  $t_{\mathcal{T}}$ , and  $t_{acq}$ ) have been already fixed and adapted with validation data, and therefore, the performance of the system might be barely unbalanced. In other words, TSD and TNSD will not be equal at TESD, but TSD could increase at the expense of TNSD, or vice versa.

This suggests that TESD is a fine system performance indicator since it provides an approximation based on validation data. However, TSD and TNSD provide a real rate of the performance. Obviously, TESD cannot be always calculated in the previous schemes (Section VII-B), as TESD requires both stressing and nonstressing data during the training and the validation of data. For instance, the HV + TP scheme has only

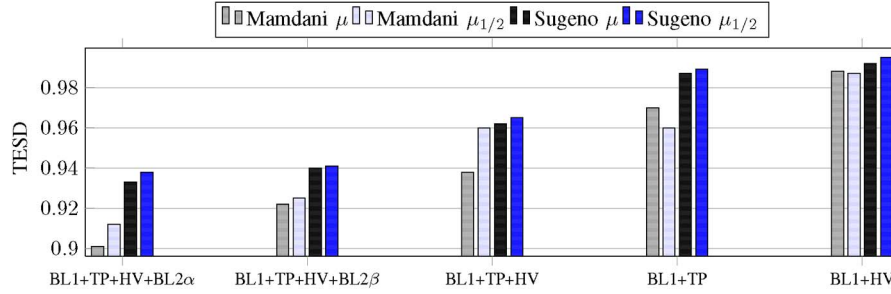


Fig. 2. System performance (TESD) for the first three schemes. Better performance is achieved when BL2 is not considered due to its confusing nature.

TABLE I

BEST TSD AND TNSD RATES OBTAINED DURING THE TESTING STAGE FOR EACH SCHEME INDICATING WHICH IMPLEMENTATION WAS INVOLVED, HOW MUCH TIME WAS REQUIRED FOR BOTH TEMPLATE EXTRACTION AND SAMPLE ACQUISITION, AND THE THRESHOLD

| Scheme                 | TSD (%) | TNSD (%) | Implementation      | $t_{acq}(s)$ | $t_{\mathcal{T}}(s)$ | $\rho$ |
|------------------------|---------|----------|---------------------|--------------|----------------------|--------|
| BL1+TP+HV+BL2 $\alpha$ | 94.1    | 92.2     | Sugeno $\mu_{1/2}$  | 10           | 7                    | 0.3    |
| BL1+TP+HV+BL2 $\beta$  | 93.6    | 95.3     | Sugeno $\mu_{1/2}$  | 12           | 7                    | 0.28   |
| BL1+TP+HV              | 97.3    | 96.7     | Sugeno $\mu_{1/2}$  | 12           | 10                   | 0.32   |
| BL1+TP                 | 96.3    | 98.9     | Sugeno $\mu_{1/2}$  | 12           | 10                   | 0.3    |
| BL1+HV                 | 99.5    | 97.4     | Sugeno $\mu_{1/2}$  | 10           | 7                    | 0.3    |
| HV+TP                  | 88.1    | 80       | Mamdani $\mu_{1/2}$ | 10           | 7                    | 0.3    |
| BL1                    | 92.1    | 71       | Mamdani $\mu$       | 15           | 10                   | 0.4    |
| HV                     | 74.3    | 93.1     | Mamdani $\mu_{1/2}$ | 12           | 10                   | 0.41   |
| TP                     | 69.2    | 85.1     | Mamdani $\mu_{1/2}$ | 12           | 7                    | 0.45   |

stressing data in the validation, and thus, TESD lacks meaning for this scheme.

#### D. Template Time ( $t_{\mathcal{T}}$ ) and Acquisition Time ( $t_{acq}$ )

The performance of the system (TESD) not only depends on the previous threshold  $\rho_{th}$  but also on two temporal parameters: template time ( $t_{\mathcal{T}}$ ) and acquisition time ( $t_{acq}$ ). The former time defines the required time to obtain the template, and the latter is related to the time demanded to acquire stress information from an individual.

Evidently, the longer  $t_{\mathcal{T}}$  and  $t_{acq}$  are, the more accurate will be the system. However, in real applications, time is the most valued asset, and therefore, a balance among  $t_{\mathcal{T}}$ ,  $t_{acq}$ , and TESD must be achieved.

These temporal parameters are fixed during the validation step and remain constant during the testing stage.

Finally, the most important result regarding these two parameters is the fact that high values of TESD (i.e., close to one) are achieved with bounded values for  $t_{\mathcal{T}}$  and  $t_{acq}$ .

In fact, only eight values were considered during this evaluation task for the sake of simplicity ( $t_{acq}, t_{\mathcal{T}} \in \{3, 5, 7, 10, 12, 15, 17, 20\}$ ) since TESD does not vary significantly between two values from the previous set and because more than 20 s is far beyond the sake of real applications. Furthermore, the time to elucidate about the degree of stress due to both implementations is negligible (Section VI-A and Section VI-B).

#### E. Evaluation Performance

Once the previous parameters, i.e., threshold  $\rho_{th}$ ,  $t_{\mathcal{T}}$ , and  $t_{acq}$  are obtained in order to maximize TESD, the system is finally implemented.

However, this paper proposes different schemes for possible implementation and considers the following: two measurements  $\mu$  and  $\mu_{1/2}$  (Section VII); two approaches, Mamdani (Section VI-A) and Sugeno (Section VI-B); and nine possible schemes (Section VII-B).

This current section attempts to compare all the possible combinations.

Fig. 2 shows visual information on the performance of the system under several schemes, highlighting the fact that the accuracy in detecting stress properly increases when BL2 is not considered in the training data. This conclusion is due to the fact that the BL2 state is difficult to be classified as stressing stimulus [24] and is therefore considered as a noisy signal during the decision step.

Better results are obtained when neglecting BL2, achieving a rate of TESD = 96.2% (Fig. 2, Sugeno  $\mu$ ) and TESD = 96.5% (Fig. 2, Sugeno  $\mu_{1/2}$ ).

Furthermore, Fig. 2 shows the gathered TESD results corresponding to schemes BL1 + TP (98.9%, Sugeno  $\mu_{1/2}$ ) and BL1 + HV (99.5%, Sugeno  $\mu_{1/2}$ ), where only two stages were considered to train and validate the implementations.

For clarity's sake, Table I presented the gathered performance of the system in relation to the parameters involved ( $\rho_{th}$ ,  $t_{\mathcal{T}}$ , and  $t_{acq}$ ), together with the scheme proposed and the TSD and TNS ratios obtained during the testing stage. Notice that schemes HV + TP, BL1, HV, and TP can only be implemented by means of the Mamdani approach due to the fact that the Sugeno approach requires a model of both states (relax and calm), and the previous schemes cannot meet that goal. Therefore, they only model a stressing distribution in case of HV + TP, HV, and TP or a nonstressing distribution (BL1).

The results provided in Table I are obtained based on the average of the random experiments using a cross-validation approach. Therefore, the results correspond to an average

TABLE II  
COMPARISON BETWEEN APPROACHES COMPARING STRESS-DETECTION RATES, PHYSIOLOGICAL SIGNALS, AND POPULATION INVOLVED. THE INITIALS ST STAND FOR SKIN TEMPERATURE

| Reference | Stress Detection Rate (%) | Physiological Signals | Population   |
|-----------|---------------------------|-----------------------|--------------|
| [2]       | 97.4%                     | ECG, EMG, RR, GSR     | Not provided |
| [34]      | 79.5-96.6%                | ECG, EMG, RR, GSR     | 1 subject    |
| [23]      | 85-96%                    | BVP, ST, RR, GSR      | Not provided |
| [35]      | 75-85%                    | ECG, EMG, RR, GSR     | 1 subject    |
| [36]      | 76%                       | ECG, EMG, GSR         | 8 subjects   |
| [19]      | 60-78%                    | ST, GSR               | 35 subjects  |
| Proposed  | 99.5%                     | HR, GSR               | 80 subjects  |

TABLE III  
CONFUSION MATRIX CORRESPONDING TO THE BEST SCHEME (BL1 + HV) AND ITS CONFIGURATION ACCORDING TO TABLE I

|           |            | Actual |            |
|-----------|------------|--------|------------|
|           |            | Stress | Non-Stress |
| Predicted | Stress     | 25473  | 228        |
|           | Non-Stress | 127    | 8572       |

accuracy obtained with different templates and testing data set, randomly selected, considering the whole database of 80 individuals.

Notice that the best result is achieved with scheme BL1 + HV, which means that, for an accurate stress detection, only two tasks are required: a relaxing situation and a stressing situation. This is an outstanding result since it allows decreasing (in terms of time) the template-extraction step among the other aspects discussed in Section VIII.

The results shown in Fig. 2 and Table I yield to the conclusion that stress can be detected with an accuracy of 99.5%, recording the signal of the user during 10 s to create the template and 7 s for stress detection. This result highlights the improvement achieved in comparison with other approaches. This comparison is gathered in Table II, providing the following parameters to be compared: stress detection rate (TSD), the physiological signals involved, and the population used to evaluate the proposed approach. This improvement is achieved not only in terms of accuracy in stress detection but also in relation to the number of physiological signals (only HR and GSR) and the population.

Moreover, the confusion matrix for the most accurate scheme is provided in Table III. Each element on this confusion matrix represents each segment of  $t_{acq}$  seconds that have been compared with the corresponding template  $t_\tau$  through the different  $K$ -fold cross validations.

### VIII. CONCLUSION AND FUTURE WORK

The proposed stress-detection system is able to detect stress by using only two physiological signals (HR and GSR), providing an accurate output that indicates up to what extent a user is under a stressing stimulus. The manner that HR and GSR are extracted makes a future integration of this system on a current biometric system possible, increasing overall security.

The main characteristics of this system is its noninvasiveness, fast-oriented implementation, and outstanding accuracy in detecting stress when compared with the previous approaches. In other words, the system can detect stress almost instantly,

allowing possible integration in real-time systems. Notice that only two physiological signals are involved in contrast to the amount of features required to elucidate on the stress degree provided by the previous approaches.

Furthermore, the proposed approach allows adapting the behavior of an individual as time goes by. In other words, the template created can be modified (updated) in order to achieve a more accurate decision on the stress degree. This update will allow keeping the performance accuracy even if the individual is suffering a stressing situation that endures through time. Therefore, the system could detect both long- and short-term stress. This template update will be considered in a future research.

On the other hand, not only template  $\mathcal{T}$  but also  $\rho_{th}$ ,  $t_{\mathcal{T}}$ , and  $t_{acq}$  must be adapted for each individual so that the overall performance can be increased. These parameters ( $\rho_{th}$ ,  $t_{\mathcal{T}}$ , and  $t_{acq}$ ) have been fixed for the whole database, and therefore, if each individual possesses a different version of these parameters, then the accuracy of the system could be increased. This implementation remains as future work.

The database acquisition was based on psychological experiments carried out by expert psychologists. These experiments ensure that stressing situations are provoked on an individual, validating posterior HR and GSR acquisitions.

This paper has provided a decision system that is able to detect stress with an accuracy of 99.5% using 10 s to extract the stress template and 7 s to detect stress on an individual, using two physiological signals HR and GSR measured only during two tasks: a stressing task and a relaxing stage. Furthermore, both measurements  $\mu$  and  $\mu_{1/2}$  provide outstanding results as applied to both implementations, achieving, in most cases, an accuracy greater than 90%.

Finally, this system may be applicable in scenarios related to aliveness detection (e.g., detecting if an individual is accessing a biometric system with an amputated finger), civil applications (e.g., driver control), withdrawing money from a cash dispenser, electronic voting (e.g., someone is forced to emit a certain vote), and so forth. In other words, a wide variety of scenarios can benefit from this approach due to its noninvasiveness, the likelihood to be embedded on current security systems, and its possibility in detecting stress in real time [20], together with the capability of being combined to other stress-detection methods based on computer-vision algorithms. Moreover, future research entails integration with mobile devices.

### REFERENCES

- [1] P. J. Rosch, *Handbook of Stress, Medicine, and Health*. Boca Raton, FL: CRC, 1996, pp. 27–60.
- [2] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 156–166, Jun. 2005.
- [3] J. Andren and P. Funk, "A case-based approach using behavioral biometrics to determine a user's stress level," in *Proc. ICCBR Workshops*, 2005, pp. 9–17.
- [4] D. F. Dinges, S. Venkataraman, E. L. McGlinchey, and D. N. Metaxas, "Monitoring of facial stress during space flight: Optical computer recognition combining discriminative and generative methods," *Acta Astronaut.*, vol. 60, no. 4–7, pp. 341–350, Feb.–Apr. 2007, 2005.
- [5] S. Begum, M. U. Ahmed, P. Funk, N. Xiong, and B. von Scheele, "Using calibration and fuzzification of cases for improved diagnosis and treatment of stress," in *Proc. 8th Eur. Conf. Case-Based Reason. Workshop*, M. Minor, Ed., Sep. 2006, pp. 113–122.



- [6] L. A. Zadeh, "Fuzzy logic = computing with words," *IEEE Trans. Fuzzy Syst.*, vol. 4, no. 2, pp. 103–111, May 1996.
- [7] E. Jovanov, A. O'Donnell Lords, D. Raskovic, P. G. Cox, R. Adhami, and F. Andrasik, "Stress monitoring using a distributed wireless intelligent sensor system," *IEEE Eng. Med. Biol. Mag.*, vol. 22, no. 3, pp. 49–55, May/June 2003.
- [8] F. Angus, J. Zhai, and A. Barreto, "Front-end analog pre-processing for real-time psychophysiological stress measurements," in *Proc. 9th WMSCI*, 2005, pp. 218–221.
- [9] J. Zhai, A. Barreto, C. Chin, and C. Li, "Realization of stress detection using psychophysiological signals for improvement of human-computer interactions," in *Proc. SoutheastCon*, Apr. 2005, pp. 415–420.
- [10] H. Prendinger and M. Ishizuka, "Symmetric multimodality revisited: Unveiling users' physiological activity," *IEEE Trans. Ind. Electron.*, vol. 54, no. 2, pp. 692–698, Apr. 2007.
- [11] N. Sarkar, "A novel interface system for seamlessly integrating human-robot cooperative activities in space," NASA Inst. Adv. Concepts, Atlanta, GA, Tech. Rep., 2002.
- [12] J. Kim and E. Ande, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008.
- [13] M. Jiang and Z. Wang, "A method for stress detection based on FCM algorithm," in *Proc. 2nd Int. CISP*, Tianjin, China, Oct. 2009, pp. 1–5.
- [14] C. L. Lisetti and F. Nasoz, "Using noninvasive wearable computers to recognize human emotions from physiological signals," *EURASIP J. Appl. Signal Process.*, vol. 11, pp. 1672–1687, 2004.
- [15] A. Barreto and J. Zhai, "Stress detection in computer users based on digital signal processing of noninvasive physiological variables," in *Proc. 28th Annu. Int. Conf. IEEE EMBS*, Aug. 2006, pp. 1355–1358.
- [16] Y. Shi, N. Ruiz, R. Taib, E. Choi, and F. Chen, "Galvanic skin response (GSR) as an index of cognitive load," in *Proc. CHI Ext. Abstr. Human Factors Comput. Syst.*, 2007, pp. 2651–2656.
- [17] J. Choi and R. Gutierrez-Osuna, "Using heart rate monitors to detect mental stress," in *Proc. 6th Int. Workshop Wearable Implantable BSN*, 2009, pp. 219–223.
- [18] W. Picard and J. A. Healey, "Wearable and automotive systems for affect recognition from physiology," MIT, Cambridge, MA, Tech. Rep., 2000.
- [19] M. S. Sharawi, M. Shibli, and M. I. Sharawi, "Design and implementation of a human stress detection system: A biomechanics approach," in *Proc. 5th ISMA*, Amman, Jordan, May 2008, pp. 1–5.
- [20] A. Bar-Or, J. Healey, L. Kontothanassis, and J. Van Thong, "Biostream: A system architecture for real-time processing of physiological signals," in *Proc. 26th Annu. Int. Conf. IEEE IEMBS*, Sep. 2004, vol. 2, pp. 3101–3104.
- [21] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. S. Wiederhold, "ECG to identify individuals," *Pattern Recognit.*, vol. 38, no. 1, pp. 133–142, Jan. 2005.
- [22] S. Yanushkevich, P. Wang, M. Gavrilova, M. Nixon, and S. Srihari, *Image Pattern Recognition: Synthesis and Analysis in Biometrics*. Singapore: World Scientific, 2007.
- [23] H. Cai and Y. Lin, *An Experiment to Non-Intrusively Collect Physiological Parameters Towards Driver State Detection*. New York: Academic, 2007.
- [24] A. Cano-Vindel, J. J. Miguel-Tobal, H. Gonzalez-Ordí, and I. Iruarrizaga-Diez, "Hyperventilation and anxiety experience," *Anxiety Stress*, vol. 13, no. 2/3, pp. 291–302, 2007.
- [25] M. J. Zvolensky and G. H. Eifert, "A review of psychological factors/processes affecting anxious responding during voluntary hyperventilation and inhalations of carbon dioxide-enriched," *Clin. Psychol. Rev.*, vol. 21, no. 3, pp. 375–400, Apr. 2001.
- [26] J. J. Miguel-Tobal and A. Cano-Vindel, *ISRA: Inventory of Situations and Responses of Anxiety*, 2002.
- [27] A. Cano-Vindel and J. J. Miguel-Tobal, *IACTA: Cognitive Activity Inventory on Anxiety Disorders*, 2001.
- [28] R. M. Yerkes and J. D. Dodson, "The relation of strength of stimulus to rapidity of habit-formation," *J. Comparative Neurol. Psychol.*, vol. 18, no. 5, pp. 459–482, Nov. 1908.
- [29] T. Lin, M. Omata, W. Hu, and A. Imamiya, "Do physiological data relate to traditional usability indexes?" in *Proc. 17th Australia Conf. Comput.-Human Interact.*, OZCHI, Narrabundah, Australia, 2005, pp. 1–10.
- [30] R. M. Sapolsky, "Individual differences and the stress response: Studies of a wild primate," *Adv. Exp. Med. Biol.*, vol. 245, pp. 399–411, 1988.
- [31] K. Guney and N. Sarikaya, "Comparison of Mamdani and Sugeno fuzzy inference system models for resonant frequency calculation of rectangular microstrip antennas," *Prog. Electromagn. Res. B*, vol. 12, pp. 81–104, 2009.
- [32] G. Bailador and S. Guadarrama, "Robust gesture recognition using a prediction-error-classification approach," in *Proc. FUZZ-IEEE*, 2007, pp. 1–7.
- [33] G. Bailador and G. Triviño, "Pattern recognition using temporal fuzzy automata," *Fuzzy Sets Syst.*, vol. 161, no. 1, pp. 37–55, Jan. 2010.
- [34] J. Wagner, N. J. Kim, and E. Andre, "From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification," in *Proc. ICME*, 2005, pp. 940–943.
- [35] L. Guang-yuan and H. Min, "Emotion recognition of physiological signals based on adaptive hierarchical genetic algorithm," in *Proc. World Congr. Comput. Sci. Inf. Eng.*, 2009, vol. 4, pp. 670–674.
- [36] D. Kulic and E. Croft, "Anxiety detection during human-robot interaction," in *Proc. IEEE/RSJ Int. Conf. IROS*, Aug. 2005, pp. 616–621.



**Alberto de Santos Sierra** was born in Madrid, Spain. He received the M.S. degree in telecommunication engineering from the Polytechnical University of Madrid, Madrid, in 2007, where he is currently working toward the Ph.D. degree, with doctoral thesis based on topics related to hand biometrics in mobile devices.

His research interests cover stress detection by means of physiological signals and iris biometrics.



**Carmen Sánchez Ávila** received the Ph.D. degree in mathematics science from the Complutense University of Madrid, Madrid, Spain, in 1993.

Since 1985, she has been with the Applied Mathematics Department, Polytechnical University of Madrid, Madrid, where she is currently a Professor and lectures on different subjects related to mathematics and numerical analysis, together with subjects covering topics in wavelets and biometric systems in Ph.D. courses. Her research lines cover topics related to signal processing, cryptography, and biometrics.



**Javier Guerra Casanova** was born in Madrid, Spain. He received the M.S. degree in telecommunication engineering from the Polytechnical University of Madrid, Madrid, in 2008, where he is currently working toward the Ph.D. degree, with doctoral thesis related to biometric identification by means of in-air signatures carried out with devices embedding an accelerometer.



**Gonzalo Bailador del Pozo** was born in Madrid, Spain. He received the M.S. degree in computer science from the Polytechnical University of Madrid, Madrid, in 2003, where he is currently working toward the Ph.D. degree in computer science, with the doctoral thesis entitled "Modelado de patrones temporales usando técnicas de Soft Computing. Aplicación al análisis de movimientos del cuerpo humano."