Multimodal behavioral analysis for non-invasive stress detection

Davide Carneiro [a,*], José Carlos Castillo [b], Paulo Novais [a], Antonio Fernández-Caballero [b], José Neves [a]

[a] *Universidade do Minho, Departamento de Informática, Campus of Gualtar, 4710-057 Braga, Portugal*
[b] *Universidad de Castilla-La Mancha, Instituto de Investigación en Informática de Albacete, 02071 Albacete, Spain*

ARTICLE INFO

ABSTRACT

This study presents and approach to measure the levels of acute stress in humans by analysing their behavioral patterns when interacting with technological devices. We study the effects of stress on eight behavioral, physical and cognitive features. The data was collected with the participation of 19 users in different phases, with different levels of stress induced. A non-parametric statistical hypothesis test is used to determine which features show statistically significant differences, for each user, when under stress. It is shown that the features more related to stress are the acceleration and the mean and maximum intensity of the touch. It is also shown that each user is affected by stress in a specific way. Moreover, all the process of estimating stress is undertaken in a non-invasive way. This work constitutes the foundation of a context layer for a virtual environment for conflict resolution. The main objective is to overcome some of the main drawbacks of communicating online, namely the lack of contextual information such as body language or gestures.

1. Introduction

In a pursuit of better remote communication frameworks, our society relies more and more in Virtual Environments (VEs). Blascovich and Loomis (2002) define VEs as synthetic sensory information that lead to perceptions of environments and their contents as if they were not synthetic. In other words, VEs may be seen as simulated environments that, in some way, try to look like the real environments being simulated, with the aim of implementing some kind of interaction scenario. Typical fields of application of VEs include teaching in classrooms, informal learning, distance learning, business, e-commerce, gaming, real-life simulation or conflict resolution. However VEs, as they are seen nowadays, are still not a suitable replacement for traditional face-to-face communication.

In fact, a VE is frequently regarded as "cold", with emotions and other traces of our complex interaction modalities playing little to no role at all. One of the most important aspects here is that of body language. In their day-to-day interactions people (unconsciously) rely on body language to express themselves in a richer way. Mehrabian (1980) concludes that in a face-to-face communication there are three key elements: the words, the tone of voice and the non-verbal behavior. The author also concludes that the non-verbal elements are particularly important for communicating feelings and attitudes, stating that they account for the majority of

the information transmitted. That is to say: the way that words are said is more important than the words themselves.

In a related line of research, Dodds, Mohler, and Bülthoff (2011) conclude that the lack of gestural information from both speaker and listener limits successful communication in VEs. The authors experimentally prove not only that body language is very important for transmitting information but it is also important to perceive feedback from that transmission, i.e., to perceive if the communication is being successful or a different approach should be followed. Both the lack of feedback from the environment and meaningful content are also pointed out as a drawback by other researchers (Campbell, 1997; Marucci, Mori, Paterno, & Costalli, 2001).

To deal with this issue, several approaches can be found. Recently, Alsina-Jurnet and Gutiérrez-Maldonado (2010) investigate the influence of five user characteristics – test anxiety, spatial intelligence, verbal intelligence, personality and computer experience – on the sense of presence. Also, Rehm, Bee, and André (2008) deal with the idea of the analysis of the user's behavior and interpretations regarding the cultural background, utilizing accelerometers to uncover the user's cultural background by analyzing his/her patterns of gestural expressivity in a model based on cultural dimensions. Jaimes and Sebe (2007) describe the concept of multimodal interaction as a way to communicate between humans and computers using more than one modality or communication channel (e.g. speech, gesture or writing).

Also important is the affective aspect of communication (Beale & Creed, 2009; Hudlicka, 2003). Emotions appear in almost all models of human communication: facial expression, gestures,

* Corresponding author. Tel.: +351 934335822.
  E-mail address: dcarneiro@di.uminho.pt (D. Carneiro).

voice tone, respiration, skin temperature, and so on. Moreover, depending on the emotions, the message changes: once again, the most important is not what is said, but how it is said. As noted by Picard (1997) affect recognition is most likely to be accurate when it combines multiple modalities, information about the user's context, situation, goal, and preferences.

The importance of stress must also be considered. Evidently, stress is a very important factor in interpersonal communication, as it is in virtually any other aspect in our lives. However, current approaches on VEs lack stress models that can support it. This constitutes an obstacle to effective communication between the participants. In fact, research on stress applied to VEs does not exist.

The loss of all this context information in virtual settings makes it hard for the intervening parties to understand the state of each other. This constitutes the motivation for this research. In fact, when communicating online, people tend to forget that there is another person behind the screen on the other side. In that sense, there is a disinhibition effect and people tend to forget about the other's feelings and simply do not worry that much about the consequences of the words they utter and the actions they commit. Thus, it is often easier to offend people online. We aim to develop techniques that allow such context information to be acquired from the real environment and be passed onto the VE, enabling a more accurate description of the real individual in the virtual setting.

More specifically, this paper focuses on non-invasive acquisition of data to be used for estimating the level of stress of the users of a VE. We collected empirical behavioral data from 19 human users, including the way in which they interact with the interfaces and their movement in the environment, and related that information with different levels of stress, induced by specific stressors. We then assess the importance of each of the features considered according to the existence or not of statistically significant differences between the measurements taken on calm versus stressed states. The result is an assessment of the different features and their relevance and relation with the visible effects of stress on the user. To acquire the empirical data, hardware such as video cameras, accelerometers or pressure-sensitive touchscreens (Park, Lee, & Kim, 2011) is used. Moreover, this is done in a non-invasive way as the awareness of the monitoring may interfere with the results.

The remaining of the paper is structured as follows. In Section 2 we provide a definition and high level model of stress as well as a description of some of its effects. In Section 3 the case study in On-line Dispute Resolution is described. Section 4 contains an overview of the developed system and its main functionalities. Section 5 contains a description of the tests performed and the results achieved. Finally, Section 6 discusses the results and in the last section some concluding remarks and future lines of research are presented.

2. About stress and its effects

One of the first definitions of stress was proposed by Selye (1956). According do Selye, stress can be seen as a non-specific response of the body to external demands. These demands (the load or stimulus that triggered a response) are denominated *stressors* while the internal body changes that they produce constitute the actual *stress response*. Selye was also the first to document the chemical and hormonal changes that occur in the body due to stress.

Nevertheless, the definition of stress is still not consensual in the scientific community, remaining as an open topic of discussion. In fact, stress involves a multiplicity of factors, many of them subjective, leading to multiple interpretations that make it difficult to be objectively defined. Thus being, some researchers argue that such a concept is elusive because it is poorly defined (Cox, 1985) while others prefer not to provide an actual definition of the concept until a more accurate and consensual view of the phenomenon is achieved.

In an attempt to address this issue, researchers started dealing with stress from an empirical point of view. In this sense, a strong focus was put on its cognitive and behavioral effects and it started to be viewed as a mind–body, *psychosomatic* or psycho-physiologic phenomenon. A more up-to-date view of stress can thus be provided that looks at it as a physico-physiologic arousal response occurring in the body as result of stimuli. It should also be added that these stimuli only become stressors by virtue of the cognitive interpretation of the individual, i.e., the effects of stressors depend on the individual. This is the interpretation of stress considered in this paper and is the starting point for the definition of a stress model.

Given the multiplicity of factors that influence stress and the different modalities of the behavior and cognition that are affected, a single-modality approach for measuring the effects of stress would not be suited, as some experimental results demonstrate (Liao, Zhang, Zhu, Ji, & Gray, 2006). In fact, for a sufficiently precise and accurate measurement of stress, a multi-modal approach should be considered. The diagram depicted in (Fig. 1) represents a multi-modal approach to the stress recognition problem. This diagram has two main parts. The part to the left of the "stress" node depicts the elements that can influence human stress. These elements are included in the "context" node and represent the generalization of two main categories (sources of contextual information): the "user-centric" context and "environmental" context. On the other hand, the rightmost part of the diagram depicts the observable features that reveal stress. These features include quantifiable measurements on the user's physical appearance, physiology, behaviors and performance.

The "context" node is divided in two types according to the source of contextual information: the "user-centric" and the "environmental" context. User-centric information is composed of two categories: the background and the dynamic behavior. The background is composed by several attributes that can be extracted from the user's profile. These attributes include age, gender, working area, social status, personality traits, among others. The dynamic behavior reflects the contextual attributes related to the activity being performed by the user.

The "environmental" information fuses the characteristics of the physical environment, social environment and computational environment. Physical environment includes attributes such as the time, temperature, location, noise level, and luminance. High levels of noise, extreme temperatures and low levels of luminance
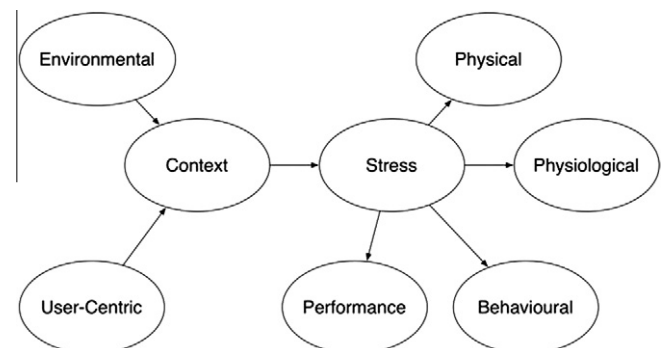

Fig. 1. A generic diagram for representing the multi-modality space in the recognition stress model depicting only the main features considered.

are well known stressors. The social environment concerns issues such as population density around the user or role/affinity/hierarchical position of the surrounding people. The computational environment can be characterized by the measurement of the electromagnetic field or the number of surrounding electronic devices.

Concerning the "behavioral" node, there are many features that can reveal stress. This may include the way an individual reacts before a conflict (Carneiro, Gomes, Novais, Andrade, & Neves, 2011) (a conflict style can be a coping strategy in response to stressful conflict (Tidd & Friedman, 2002)), the interactions patterns with the computer, the pressure of touches or clicks on touch screens or mouses, his/her agitation level, the input frequency and speed, among others. The "performance node" is in this work depicted in terms of accuracy and cognitive response. The accuracy feature is related to the precision of the touches/clicks on the controls of the interface. The cognitive response feature corresponds to the analysis (qualitative and temporal) of the user's responses to the conflict resolution demands. The physiological variables provide observable features about the user's stress state (Picard, 1997). These features include the Galvanic Skin Response of skin (GSR) (that assesses the electrical proprieties of the skin in response to different kinds of stimuli), the General Somatic Activity (GSA) (that assesses the movement of human body) and many others such as respiration or pupilographic activity. The physical appearance includes the visual features that characterize the user's eyelid movement such as pupil movement (e.g. eye gaze, papillary response), facial expression or head movement.

From an high level point of view, different types of stress can also be identified, namely acute and chronic stress. Acute stress comes from recent demands and pressures and from anticipated demands in the near future. On the other hand, chronic stress is a long-term one, due to social conditions, health conditions, dysfunctional families, among many other issues. This type of stress will have nefarious effects on the body and mind of the individual, slowly wearing him away day after day. On the other hand, acute stress, because it is short term, won't do the extensive damage associated with chronic stress. Nevertheless, it will instantaneously influence the performance of the actions being performed.

We are more interested in the analysis of acute stress and its effects given that they may be more determinant for the present outcome of a communication process. In that sense, in this work we focus on analysing the real-time effects of acute stress in order to evaluate it, rather on the analysis of the background information of each individual. Consequently, rather than evaluating the absolute level of stress of an individual (a utopian objective given the current state of research on stress), we will be evaluating changes in the level of stress throughout the communication process by analysing, in a non-invasive way, the effects of acute stress on the individual. This will allow to determine how an individual is being affected, in real-time.

3. A case-study on online dispute resolution

Given the well-known drawbacks associated to litigations in court, new methods appear aiming to solve disputes through alternative means. These are the so-called Alternative Dispute Resolution (ADR) processes (Moffitt & Bordone, 2005). They include negotiation, mediation, arbitration or conciliation, just to name a few. However, with globalization and the emergence of e-commerce, these models started to become outdated. Thus, a new trend of technology-supported conflict resolution gave birth to what is nowadays known as Online Dispute Resolution (ODR): conflict resolution that takes part over a telecommunication mean (Larson, 2006). One of the potential applications of telecommunication means in ODR (together with other supporting technologies) is actually to create VEs.

At first sight, using VEs to foster communication and information exchange between parties should represent an advantage. And, to a certain point, it does: it allows asynchronous meetings, it facilitates information exchange, and it eliminates the need for traveling, among others (Lodder & Zeleznikow, 2010). However, the fact that ODR takes place in a VE, without all the richness of face-to-face interaction, must also be seen as a serious drawback. Indeed, the lack of the context of the personal interactions seriously limits communication (Goodman, 2003). Context information is needed not only for parties to take better and more realistic decisions but also to interpret how the others are being affected by the issues. Moreover, at an unconscious level, context information is needed for parties to keep in mind that at the other end of the screen there are people with feelings, desires and fears and that each decision or word has to be thought carefully. Context is analyzed through a range of computer techniques. Smaragdis, Radhakrishnan, and Wilson (2009) describe some approaches for context extraction through audio signal analysis. Some other works use statistical approaches such as Hidden Markov Models or Gaussian mixture models to classify the sounds in the environment. In contrast, this paper deals with the use of computer vision techniques to analyze the activities carried out by the system users. Some related works deal with the recognition of body movements and posture analysis (Wu, Hua, & Yu, 2003; Jaeggli, Koller-Meier, & van Gool, 2009), gestures (O'Hara, Lui, & Draper, 2011) or gaze detection (Cuong & Huynh, 2010).

There are currently some ODR providers who already use VEs for conflict resolution. The Mediation Room[1] is a company that provides a virtual mediation space for parties trying to solve their disputes. When a client authenticates in the website, he/she has also access to all the information gathered so far, including previous conversations, current state of the process, information about the mediator, and about the outcome in a later phase. Another example is Smartsettle,[2] which is an online negotiation system described as a generic tool for decision-makers with conflicting objectives who wish to reach a formal agreement. This platform is used to solve problems relating to family, insurance, real estate, labor-management, contract negotiations, among others. Virtual Courthouse[3] is also an Internet-based service that enables parties to submit disputed claims, responses and supporting material in digital form for resolution by a neutral provider. It is a combination of multimedia technologies and business processes aimed at replicating the processes of dispute resolution in online environments, removing constraints of time, expense and distance. Finally, another example of existing tools in the field is MediateMe,[4] which builds on the concept of mediation for the masses, making it easily accessible to everyone. The MediateMe platform builds on virtual conference rooms that rely on messaging, voice feeds or video calls (Fig. 2).

However, all these tools disregard the issue of context information. Thus, these environments are highly objective and focused solely on the monetary values of the dispute, leaving aside aspects such as the evolution of the state of the parties. This may conduct mediators into taking risky decisions that they would easily spot and avoid in a face-to-face process.

We are developing a prototype of this stress-aware VE in the field of ODR. In fact, the UMCourt conflict resolution platform is being extended with a context-aware layer. This work is being

---

[1] The website of The Mediation Room is available at http://www.themediation-room1.com/ (accessed in April 2012).

[2] The website of Smartsettle is available at http://www.smartsettle.com (accessed in April 2012).

[3] The website of VirtualCourthouse is available at http://www.virtualcourt-house.com (accessed in April 2012).

[4] The website of MediateMe is available at http://mediateme.com (accessed in April 2012).
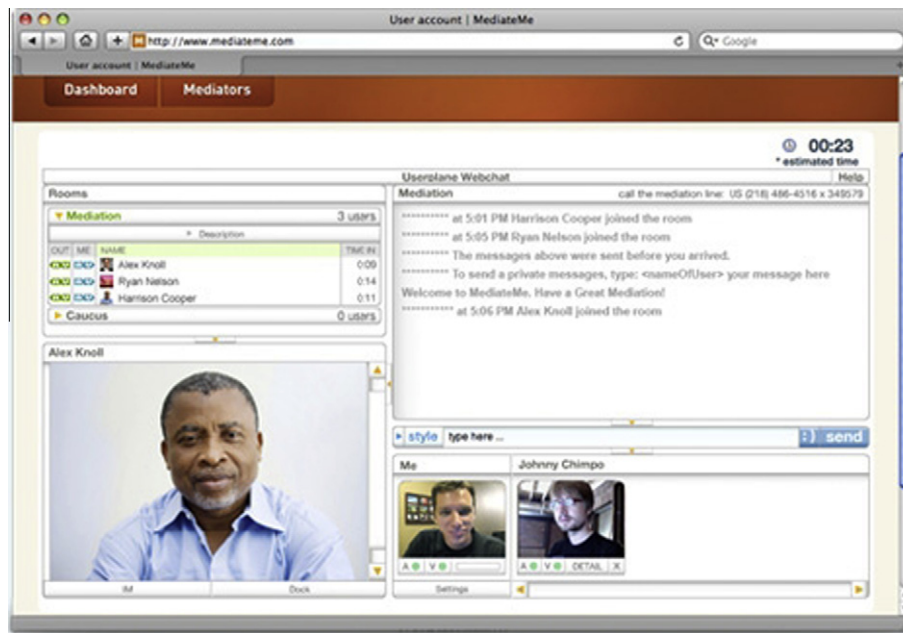
Fig. 2. Example of a virtual mediation room: MediateMe. Source: http://mediateme.com.

conducted under the frame of the TIARAC funded project – Telematics and Artificial Intelligence in Alternative Conflict Resolution. The main objective is to allow individuals involved in a technology-supported conflict resolution process to have access to context information about each other. From the point of view of a mediator, the access to this information allows him/her to better manage the process, taking breaks when emotions escalate or advising to better think a decision when the party is too stressed. From the point of view of a party, this permits him/her to assess the state of the other party, to perceive how the other party is being affected by each issue and, above of all, to gain the conscience that on the other side there is someone just like him/her, with equivalent fears, desires and expectations. Thus, parties also weight their decisions and words better, being less prone to attack or hurt each other. This work expects to achieve richer technology-supported conflict resolution processes, not solely focused on the objective issues, but taking into consideration the state of the parties. This will ultimately lead to more satisfactory outcomes. In previous work, and following the same line of thought, a module for estimating the conflict resolution style of the parties in a non-invasive and transparent way has already been developed (Carneiro, 2011, 2011). Moreover, the module has already been added to the UMCourt conflict resolution platform.

In this case-study a prototype of a context-aware VE in the legal domain for conflict resolution out of courts that considers additional sources of information generally disregarded by other approaches is developed. For this purpose, artificial intelligence (AI) techniques, namely machine learning (Alpaydin, 2010) and data mining (Witten, Frank, & Hall, 2011), are used. In fact, AI techniques are employed to leverage the work of legal practitioners and to improve the odds of the parties involved, potentially resulting in mutually more satisfactory outcomes (Bellucci, Lodder, & Zeleznikow, 2004).

4. System overview

In this paper a context-aware VE for conflict resolution is proposed, based on an Ambient Assisted Living approach (Novais, Costa, Carneiro, & Neves, 2010). We leverage on the UMCourt

Conflict Resolution Platform and extend it with an intermediate layer that receives sensory information and converts it into meaningful information about the context of interaction. This information is then used by the mediator to take better decisions, and by the parties to receive feedback from the environment, as described in the previous section.

Fig. 3 gives a general overview of the proposed approach. Each user participates in the conflict resolution from his/her own *real* environment. This environment is equipped with sensors and devices that acquire different kind of information from the user in a non-intrusive way. While the user conscientiously interacts with the system and takes his/her decisions and actions, a parallel and transparent process takes place in which this information is sent in a synchronized way to the conflict resolution platform. The platform, upon converting the sensory information into useful data, allows for a contextualized analysis of the operational data of the users. This contextualized analysis is performed by the platform itself (e.g. for performing decision-support related tasks) and by the mediator. Then, the parties receive feedback from the platform (e.g. a new proposal, information updates, notifications), which may also include some kind of feedback from the state of the parties (e.g. an avatar showing the state of the parties).

All this process is completely transparent to its users. That is, the data about the users' context is collected, processed and transparently added to the operational data. Thus, when parties participate in the technology-supported conflict resolution they sign in a usual virtual mediation or negotiation room, or they may even just be using traditional web forms. Anyway, the conflict resolution platform and the mediator have a way to assess their state when taking their decisions that allows for a better management of the whole process.

4.1. Functionalities

The system described builds on telecommunication technologies and on the UMCourt conflict resolution platform that provides a range of services previously developed (e.g. negotiation environment, information retrieval). Through this work, the platform is
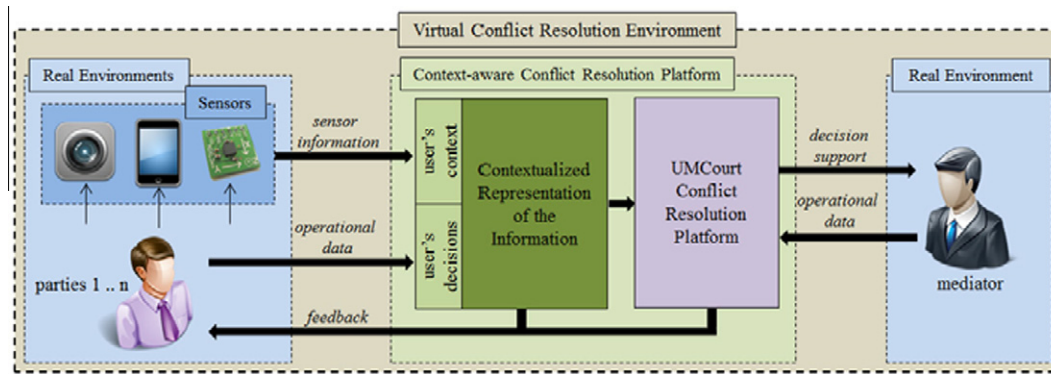
Fig. 3. General overview of the system.

extended with novel functionalities, provided by a group of devices being included in the user-area network (Fig. 4).

The main requirement when selecting the devices is that they have to provide as much information about the user environment as possible. Moreover, the user has to feel comfortable with them. In that sense, the selected devices are more or less common nowadays; so they do not represent a drawback for the user. Table 1 briefly describes each device and the main characteristics of interest.

Given that in this work the level of stress of the users is to be estimated, the focus is on devices capable of acquiring data related to stress. The following sources of information (from now on designated sensors), acquired from the respective devices, are considered:

- Touch pattern – the touch pattern represents the way in which a user touches the device and represents a variation of intensity over a period of time. This information is acquired from touchscreens with support for touch intensity.
- Touch accuracy – a comparison between touches in active controls versus touches in passive areas (e.g. without controls, empty areas) in which there is no sense in touching. This information is acquired from touchscreens.
- Touch intensity – the intensity of the touch represents the amount of force that the user is putting into the touch. It is analyzed in terms of the maximum, minimum and mean intensity of each touch. This information is acquired from touchscreens.
- Touch duration – this represents the time span between the beginning and the end of the touch event. This data is acquired from devices with touchscreens.
- Amount of movement – the amount of movement represents how and how much the user is moving inside the environment. An estimation of the amount of movement from the video camera is built. The image processing stack uses the principles

established by Castillo, Rivas-Casado, Fernández-Caballero, López, and Martínez-Tomás (2011) and uses image difference techniques to calculate the amount of movement between two consecutive frames (Fernández-Caballero, Castillo, Martínez-Cantos, & Martínez-Tomás, 2010).
- Acceleration – the acceleration is measured from accelerometers in mobile devices. It is useful for building an estimation of how much the user is moving and how he is doing it (e.g. is the user having sudden movements?). Moreover, information from the accelerometer is used to support the estimation of the intensity of touch.

The hypothesis to be tested is sensor values are influenced by stress in a significant way. Thus, changes in the level of stress result in changes in the readings from the sensors. That is, the following hypothesis is tested: when a user is stressed, he/she touches the interface in a different way, performing different movements, with less touch accuracy, and so forth. An environment is built with these devices and the described functionalities, in order to provide information to the application being used about the context and state of the user. In this case, the developed prototype is in the field of online dispute resolution.

5. Experiment and results

5.1. Methodology

As stated before, the main objective of this research line is to identify which factors of the human physiology influenced by stress can be pointed out using standard devices available on the market.

In this work we address the following research questions and hypothesis: (a) Does stress actually influences, in a significant manner, our interaction patterns while using common technologi-

Fig. 4. Devices used to implement the described functionalities.

Table 1
Description of the devices used.

Device	Brief description	Main features
HP Touchsmart	All-in-one PC	Touchscreen, web cam, large screen
Samsung Galaxy Tab	Tablet PC	Touchscreen, web cam, accelerometer, relatively large screen, mobile, Android OS
HTC PDAs	Smartphones	Touchscreen, camera, accelerometer, mobile, Android OS
Sony FCB-EX780BP	25x Super HAD PAL Color Block Camera with External Sync	25x Optical Zoom, Image stabilizer, Day/Night Mode, Privacy Zone Masking

cal devices? (b) Is it possible to accurately measure this influence of stress in a non-invasive and non-intrusive way?

With regard to the first question, we hypothesize that stress does influence our interaction patterns in a significant manner. We build on previous results in which significant differences in the parameters studied were found, with results varying from user to user (Novais, Carneiro, Gomes, & Neves, 2012). On the other hand, we are also supported by the work of other researchers who proved that stress does influence people's behaviors and physical responses (Healey & Picard, 2005; Rehm et al., 2008; Vizer, Zhou, & Sears, 2009).

Concerning the second question, we hypothesize that it is possible to measure, using non-invasive and non-intrusive methods, the level of stress of the users by analyzing key features in their interaction with technological devices. We build on previous work in which we were able to measure this without using any sensors expect for the ones available in the handheld devices (Carneiro et al., 2012).

There are also some sub-questions that, although not compulsory, will contribute to the achievement of better results. Namely: (1) Determine how the interaction patterns of each user are affected by stress and develop personalized stress models that will maximize the accuracy of the output. (2) Find groups of people that are affected in similar ways (e.g. same parameters show similar tendencies when subject to stress) in order to develop more accurate generic stress models, to be applied when a personalized model is not available.

To implement the research plan we set up a test environment in the Intelligent Systems Lab[5] of the University of Minho. In this environment, the user is isolated from external stimuli and is required to play a game that implies the interaction with the devices, as described in Subsection 5.2. The collection of the data was organized into two phases. In a first phase, test subjects were required to perform these tasks in a stress-free environment. In a second phase, the users performed the same tasks subject to stressors such as the vibration of the devices, loud and annoying sounds, unexpected behaviors of the devices, among others.

The empirical data gathered in both phases about the user interaction patterns and physical response is described in Table 2. This data was synchronized and transformed/normalized to allow its joint analysis. The participants of the proposed experiment were volunteer students and professors from our institution. 19 male and female individuals participated in the experiment aged between 20 and 57. All these individuals are familiar with the technological devices used and the interaction with them was not an obstacle.

The data gathered was analyzed in order to determine statistically significant differences between phase 1 and phase 2 of the data collection. Measures of central tendency and variability were calculated for all variables of interest. The Mann–Whitney–Wilcoxon Statistical test was used to test whether there are actual differences in the distributions of the data. A 0.05 level of significance was considered. The data analysis was performed using Wolfram

Mathematica®, Version 8.0.

Based on this analysis of the data we were able to determine which parameters, for each individual, were effectively affected by stress. Using this knowledge, we are able to develop personalized models for stress estimation in real time. Moreover, a more generic model can also be developed taking into consideration the data of several or all the users. This generic model can be applied in the cases in which a personalized one is not available. The models developed are used to develop a real-time stress estimation software layer to be used in a VE or in other domains.

5.2. Data collection

In order to identify which factors vary with stress and in what magnitude, a game with the following features was developed: a mentally challenging objective and stressors. The main objective of the game is that the user performs mental calculations using the four basic arithmetic operations and a group of numbers given randomly in order to get as close as possible to a target number. The score is given in function of the distance to the result to the target number: the closer to a given target number, the higher the score, up to a maximum of 100 points when the result is equal to the target number. There are however some rules. In each round, it is only possible to use each number and each operator once. As an example, let us say that the target number is 198 and you have four random numbers such as 50, 45, 3 and 8. A fairly good solution would be $3*50+45$.

In each consecutive round of the game the user has to perform a calculation. There are however stressors that make the effects of stress visible. These are a time limit and vibration and sounds on the handheld devices. The time limit decreases as the number of rounds increases. Moreover, the longer the user takes to answer in a round, the smaller the time limit in the following round. Thus, there is a pressure on the user to answer quickly. At the same time, there is also the pressure to make the best possible score. Vibration and sounds are used to increase the level of stress when the time is almost over at each round. When only a few seconds remain, the handheld device starts vibrating and making a disturbing beeping sound whose frequencies are higher when the time limit is smaller. All this increases the physiological effects of stress on the users, as the results show. A screenshot of the game in the first round is shown in Fig. 5.

As stated before, the game is developed to induce stress into the user and to determine in which way each user is affected. To study this, 8 parameters are analyzed for each user: acceleration, maximum intensity of touch, mean intensity of touch, duration of touch, amount of movement, score, touch accuracy, and amount of touches classified as stressed. Fig. 6 highlights the stress analysis process regarding the aforesaid stressor elements.
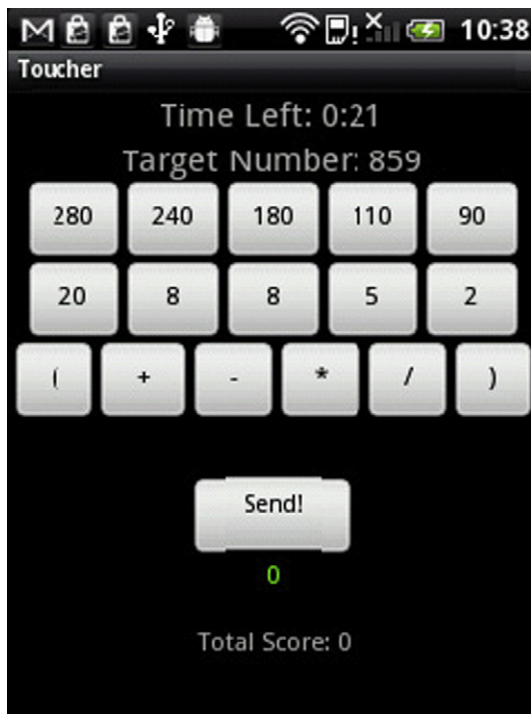
5.3. Preliminary data analysis

Each user was requested to play the game for some rounds without any source of stress. In this version, the game has no time limit and no vibration or annoying beeping sounds. In that sense, the user calmly plays the game, with enough time to think on

Table 2
Summary of the data generated during the tests.

Data	Brief description	Size of dataset
Acceleration	Data concerning the acceleration felt on the handheld device while playing the game	27291
Maximum intensity of touch	Data about the maximum intensity of each touch in a touchscreen	1825
Mean intensity of touch	This dataset contains data about the mean intensity of each touch event in a touchscreen	1825
Amount of movement	A dataset containing information about the amount of movement during tests	25416
Touches on target	This dataset contains information about the accuracy of the touches	1825
Stressed touches	A dataset containing information that allows to classify each touch as stressed or not stressed	1825
Score	A dataset describing the performance of the user playing the game, during the tests	321
Touch duration	A dataset containing the duration of each touch event	1825

Fig. 5. Screenshot from the game interface.

the different possibilities. This phase allows collecting some data about how the user normally behaves when he/she is not under stress. This enables establishing a baseline for comparison. Afterwards, the same data is collected when the user is playing the game with the stressors. The data gathered without stress is from now on designated as *training data*, whereas the data gathered with the influence of the stressors is from now on designated as *real data*.
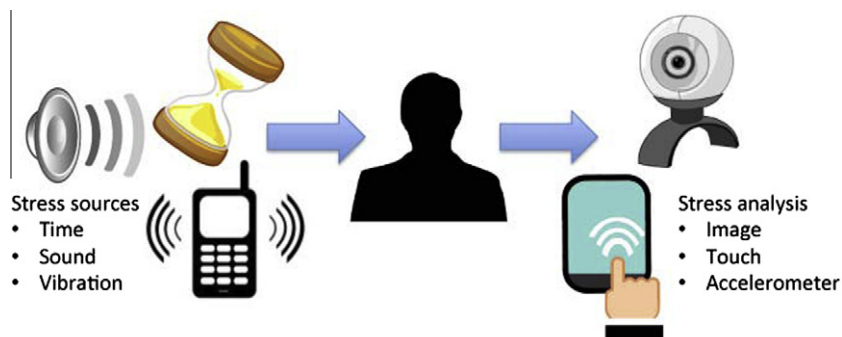
As the data about the state of the user comes from different sources (e.g. handheld device, video camera) and is synchronized by a timestamp, the clocks of the devices are previously synchronized. The data is organized into five groups. The first one consists of the training data. The second one contains all the real data. The other three are subgroups of the real data: sub-groups of the real data are created according to the time that is left for the end of the round at the instant of the event. Three groups were created: one for events that take place when there are less than 8 s left for the end of the round, another for events that take place when there are less than 5 s and the last one for events that take place when there are less than 3 s remaining. The notation depicted in Fig. 7 is used in graphics from now on in order to distinguish between the different groups of data.

5.4. Statistical data analysis

To determine to which extent each feature considered is or is not influenced by stress, the training data is compared with each of the remaining four groups. Provided that most of the distributions are not normal, the Mann–Whitney test is used to perform the analysis. This test is a non-parametric statistical hypothesis test for assessing whether one of two samples of independent observations tends to have larger values than the other. The null hypothesis is thus: H_0 = *The medians of the two distributions are equal*. For each two distributions compared, the test returns a p-value, with a small p-value suggesting that it is unlikely that H_0 is true. For each parameter, the training data is compared with the remaining four groups. In all the tests, a value of $\alpha = 0.05$ is used. Thus, for every Mann–Whitney test whose $p - value < \alpha$, the difference is considered to be statistically significant, i.e., H_0 is rejected.

A significant difference between the training data and the real data means that the parameter is effectively influenced by stress for this specific user. This is the most desirable result as it indicates a higher level of confidence. The analysis is from now on designated as *first order*. If this is not the case, the training data is compared with each one of the three subgroups of the real data, in order to look for increased levels of stress. If statistically significant
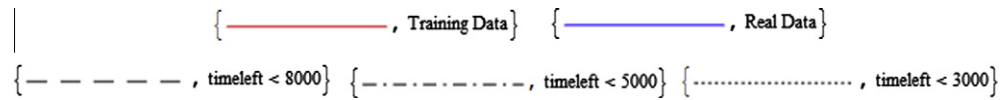

Fig. 6. Stress analysis process.

Fig. 7. Different lines used to depict the different representations of the data.

differences between one of these groups and the training data are found, it may still be concluded that the user is affected by stress in this parameter, although not in such an explicit manner. From now on, this kind of analysis is designated as *second order*.

In this section the results achieved in each of the analyzed parameters are described. Table 2 depicts the sizes and a brief description of each dataset generated for each considered parameter. Each of these datasets consists of a list of values describing the respective parameter as well as a timestamp and a username. In the following sub-sections the results for each of the studied parameters are provided. This is done by showing representative results from one or two specific users and then by showing the overall results considering all the users.

5.4.1. Acceleration

The information about the acceleration is provided by the handheld device. Our interest in this parameter lies in a potential relation between human movements and the level of stress. Specifically, we are interested in the way that the user moves his/her hands and interacts with the handheld device. In fact, a stressed user generally exhibits sudden hand gestures and movements and may also touch the device in a more brusque way. This may be measured by the accelerometer of the handheld device.

In fact, considering the first order analysis, according to our Results 80% of the users show a significant difference between training and real data. Moreover, when comparing the training data with the data from higher levels of stress (second order analysis), the results are even more expressive: 93.3(3)% of users show a statistically important difference between the histograms. In that sense, the acceleration felt in the handheld devices is effectively different between calm and stressed users.

Moreover, it should also be concluded that for most users the amount of acceleration measured tends to increase. Fig. 8 shows two examples in which the training values of the acceleration are more centered in a given value (less deviation) while the real data is slightly shifted to the right. The three subgroups of the real data are even more shifted, although the standard deviation also increases. For these two examples, the *p*-value returned by the Mann–Whitney test for the first order analysis is $5.9975*10^{-7}$ (for the data in Fig. 8(a)) and $2.75591*10^{-14}$ (for the data in Fig. 8(b)).

5.4.2. Maximum intensity

From the touchscreen of the handheld device it is possible to acquire data about how the user touches it. In this case, the maximum intensity of the touch is analyzed. The general assumption is that a more stressed user touches the screen with more intensity. The results obtained prove the assumption (see Fig. 9(a) and (b) that provide two example histograms). In both cases there is a clear shift in the values of the intensity towards higher values. Moreover, the differences observed between the distributions of the training data and the real data are statistically significant: *p*-value = $1.94289*10^{-11}$ for Fig. 9 (a) and *p*-value = 0.00169036 for Fig. 9 (b). And if taking into consideration the second order analysis, we notice that the higher the level of stress, the more the distributions are shifted to the right (black lines in Fig. 9).

And, this is in fact the general tendency: for all users who show a statistically significant difference, the tendency is that higher levels of stress are associated to an increased touch intensity. From a global point of view, considering a first order analysis, 53.3(3)% of the users under stress exhibit important differences in their touch intensities. If considering the second order analysis, this value rises to 60% of the users. Thus being, it is concluded that for approximately half of the users, the maximum intensity of the touch is significantly related to stress.

5.4.3. Mean intensity

A similar conclusion is achieved when the mean value of the intensity is analyzed during the touches. As depicted in Fig. 10, the mean value of the intensity tends to increase as greater levels of stress are considered. These two particular examples have a *p*-value = 0.00265927 for Fig. 10(a) and *p*-value = $6.5901*10^{-11}$ for Fig. 10(b), which means that the differences observed are statistically significant. Moreover, when analyzing data from all the users, the results seem to be slightly better than the ones of the maximum intensity. In fact, in a first order analysis 60% of the users show a considerable difference, while in a second order analysis this number is around 73.3%. In this sense, the mean intensity is also used to detect the effects of stress.

5.4.4. Amount of movement

The amount of movement represents a measure of the movement in front of the camera. The process is carried out from the information captured by cameras placed in the users' environment and using computer vision techniques to extract features regarding the users' states. Our initial hypothesis was that a stressed user moves more and in more sudden ways in an increased level of stress. However, the obtained results point the other way around: when users are under increased levels of stress they tend to move less. In fact, users become more tense and rigid and highly focused on what they are doing, decreasing the normal movements that people have when they are calm. Qualitative examples are offered in Figs. 11 and 12. Notice the differences of movement detected between the two figures. The first figure belongs to a non-stressed user while the second one belongs to the same user in a stressed state. Furthermore, a quantitative example is offered in Fig. 13. Given the nature of the data on the amount of movement, histograms are not the best way to depict them. Instead, a Box-and-Whisker plot that shows how the values are distributed for each of the five different analysis is shown. The results show how the user moves more during the training phase (without stress). On the contrary, when the level of stress rises the user moves less (Fig. 13 (a)). More accurate values for this specific example are shown in Fig. 13(b). It is possible to see that not only the mean value of the amount of movement decreases but also the values of the standard deviation do.

From a general point of view, the amount of movement shows statistically significant differences for around 47% of the users in a first order analysis. In a second order one, this value increases to 60%. Although the results achieved with this parameter point out to a different conclusion than the expected one, the parameter is considered for the estimation of the levels of stress.

5.4.5. Duration of touch

The duration of the touch stands for the time between the beginning and the end of each touch. Our initial assumption is that stressed users have longer touches (also backed up by the assumption that stressed users have more intense touches). In fact,
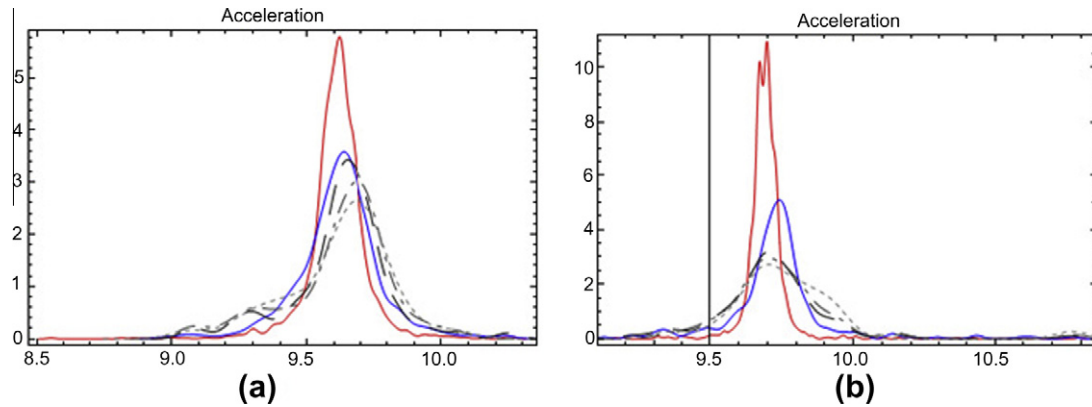
Fig. 8. Histograms of data from the module of the acceleration concerning two different users. The difference between the training data and the real data (and its subgroups) is clearly visible: the data from stressed users has more variability, i.e., stressed users move more or move in more sudden ways.
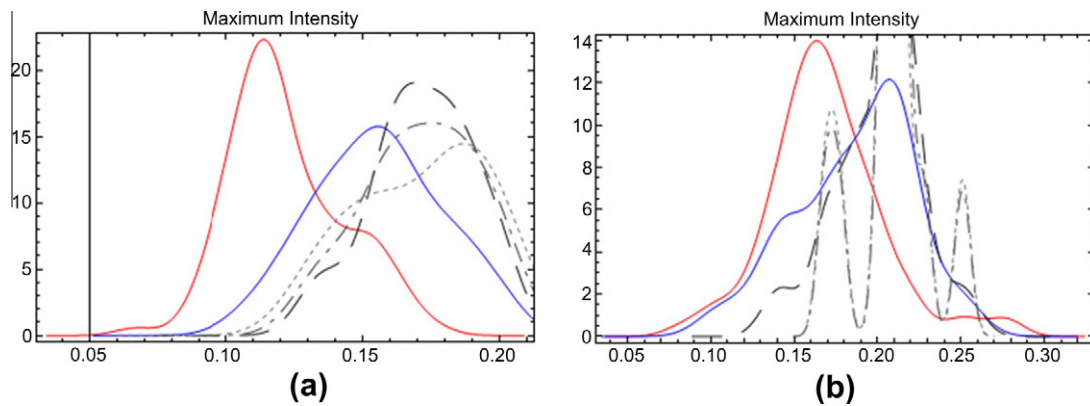

Fig. 9. Histograms of two different individuals concerning the maximum intensity of the touch. These two histograms show the tendency observed in most of the data: stressed individuals touch the screen with more intensity.
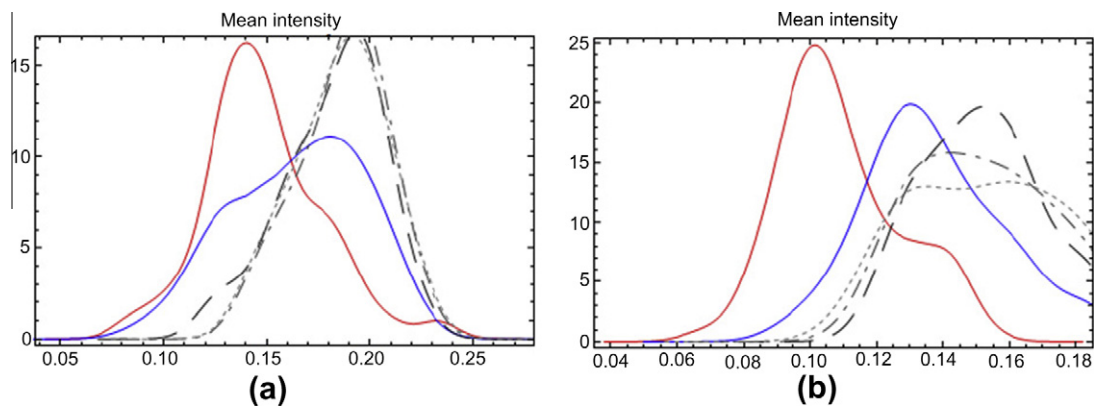

Fig. 10. Histograms of two different individuals concerning the mean intensity of the touch. As in almost every individual, the mean intensity of the touch increases with increased levels of stress.

according to the obtained results, it is not possible to state that there is a marked tendency towards this: there are test subjects that have longer touches when stressed while others have shorter ones. Fig. 14 depicts two examples of these opposing behaviors: (a) shows data from a user whose touch duration increases with stress, while (b) shows data from a user whose touch duration decreases. We must also state that the fact that there is a time limit in

each round has an influence on it, as users must finish the calculation rapidly, making use of shorter touches.

Nevertheless, both examples show statistically important differences: p-value = $2.70933*10^{-8}$ for (a) and p-value = $9.54313*10^{-6}$ for (b). In fact, when analyzing all the users, nearly 47% show significant differences in a first order analysis and around 60% show considerable differences in a second order analysis. Thus being,

Fig. 11. Results of movement detection of a non-stressed user. Row (a) shows input images, row (b) shows binarized and filtered movement and row (c) shows the amount of movement detected.
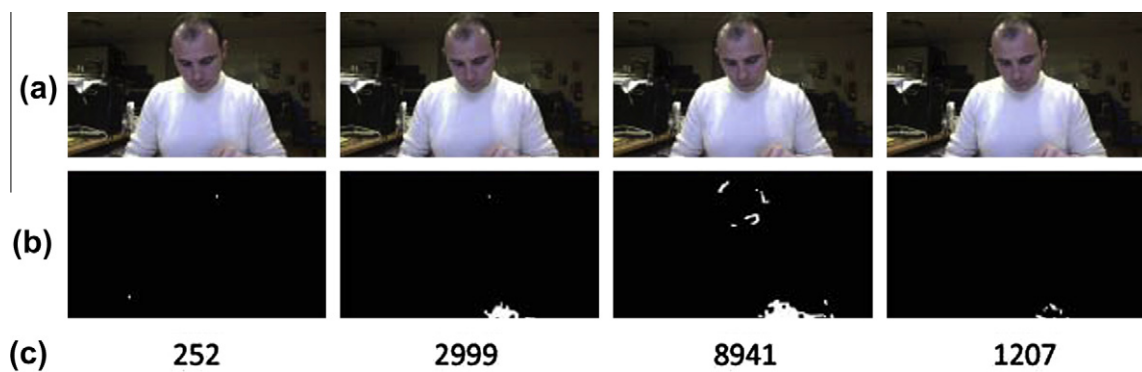

Fig. 12. Results of movement detection of a stressed user. Row (a) shows input images, row (b) shows binarized and filtered movement and row (c) shows the amount of movement detected.
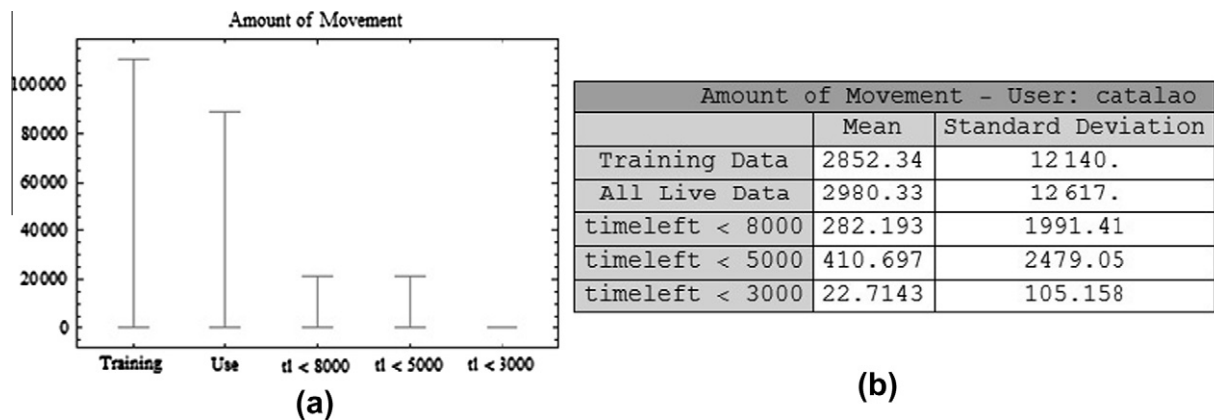

Amount of Movement – User: catalao		
	Mean	Standard Deviation
Training Data	2852.34	12140.
All Live Data	2980.33	12617.
timeleft < 8000	282.193	1991.41
timeleft < 5000	410.697	2479.05
timeleft < 3000	22.7143	105.158

Fig. 13. An example of how the amount of movement generally varies with the amount of stress: Higher stress is related to a smaller amount of movement. (a) Range of the values of movement for each dataset. (b) Mean and standard deviation for the different datasets for a specific user.

although different users react differently for this parameter, it still provides valuable data about how the individual user is affected.

5.4.6. Scores

In analysing the scores of the users, our main objective is to determine to which extent the cognition of the users is affected by stress. As it would be expected, the scores achieved by the users tend to decrease as the level of stress increases. Fig. 15(a) shows this tendency for a specific user: while training it is fairly easy and common to achieve good scores, with the mean of the scores decreasing under stress. The histogram in Fig. 15(b) depicts the distribution of the same data. For this specific user, although a visual analysis may point otherwise, the differences are not statistically significant: the Mann–Whitney test returns a p-value = 0.398694.

In fact, when analyzing all the users, the score does not appear as a parameter that can be said to be related to stress. In a first order analysis, only around 7% of the users show statistically significant differences between the scores in the training data and in the real data. Moreover, in a second order analysis this number only increases to 13%. In that sense, the score does not seem to be very preponderant in analysing the effects of stress in an individual. Nevertheless, score is solely used in these experiments as a way to induce an objective "worth fighting for" and making users
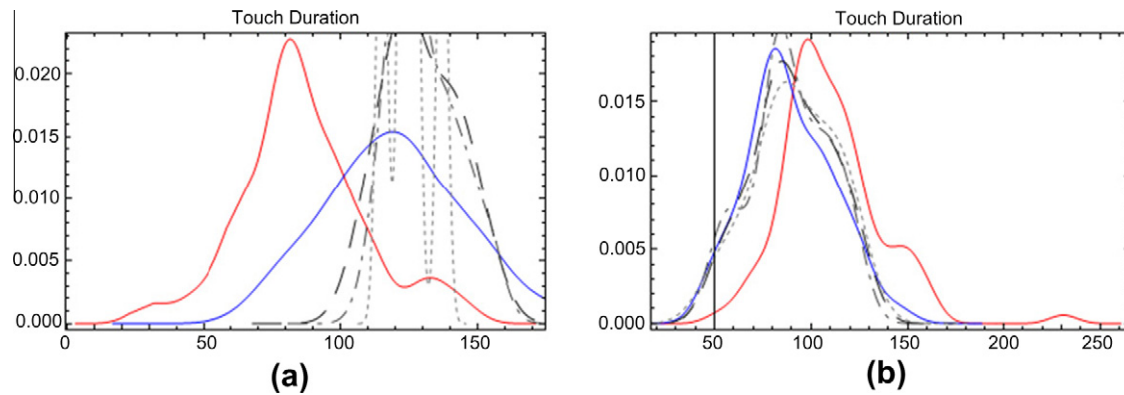
Fig. 14. Histograms of two different individuals concerning the duration of their touches. This feature does not have a homogeneous behavior: some stressed individuals have shorter touches while others have longer ones.
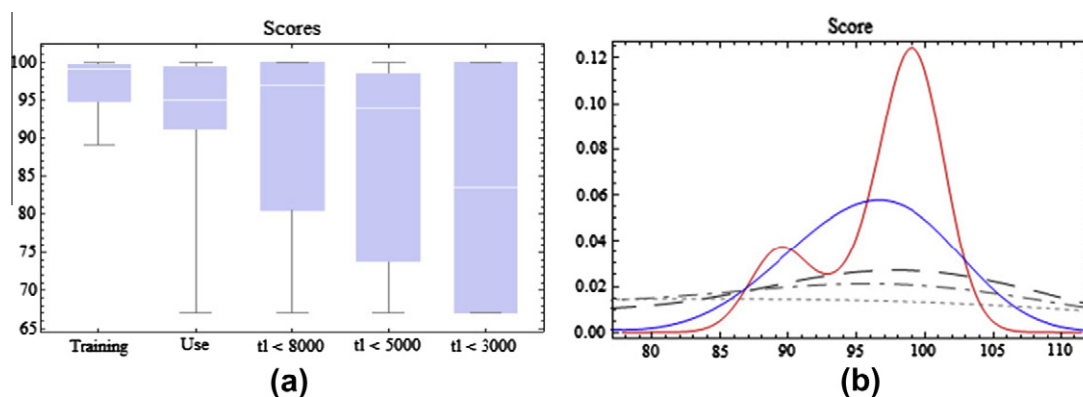

Fig. 15. Box-and-Whisker plot denoting the distribution of the scores for a given individual, in different levels of stress (a). The same data is shown in (b). Although the several distributions are visually different, these differences are not statistically significant. However, the general tendency is that users under stress tend to have worse performances.

commit with the test. In that sense, it is most likely not present in a real conflict resolution scenario or if it is, it is certainly estimated differently. Thus being this constitutes, in our opinion, no drawback.

5.4.7. Accuracy

Here, the accuracy of the touches of the users, i.e., the relation between the touches on active versus passive areas and the levels of stress is analysed. The type of data from this parameter is different from the previous ones. In fact, instead of having a list of continuous values, there is a list of true or false values, indicating whether touches did occur or not in an active area. Thus being, this parameter is not analysed with the Mann–Whitney test but by looking at plots such as the one depicted in Fig. 16, that shows how the accuracy varies when considering the data for all the users.

It can be concluded that touch accuracy remains relatively high, even for increased levels of stress (above 95%). However it is possible to identify the expected decreasing tendency.

5.4.8. Intensity curve

The last parameter considered in our tests is the nature of the touches on the screen. For this purpose a classifier with data from several users was previously built and trained. The idea is to classify the variation of intensity and time during a touch as belonging or not to a stressed user. However, each touch has an arbitrary duration and this may generate more or less intensity points during the touch. Thus, given that the number of values is arbitrary for
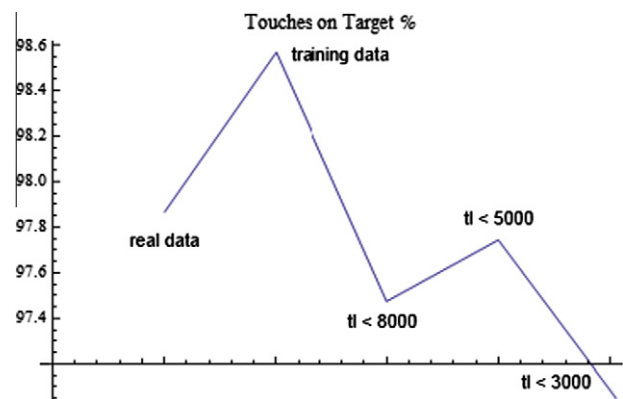

Fig. 16. This plot shows the percentage of touches on target in the five different analysis for all the users.

each touch, we can not simply make a classifier using this data as it is.

Provided that touch variations generally have a curve shape, a quadratic function is fit so that it minimizes the distances to the intensity curve. The quadratic function is of type $f(x) = ax^2 + bx + c$, $a \neq 0$. After having a curve of this type for each touch pattern, they can be compared in terms of parameters a, b, and c of the curve. Given this, a classifier that uses these three parameters a, b, and c to classify the touches was built. The tree of the J48 classifier is depicted in Fig. 17. In a total of 349 instances used, the classifier cor-
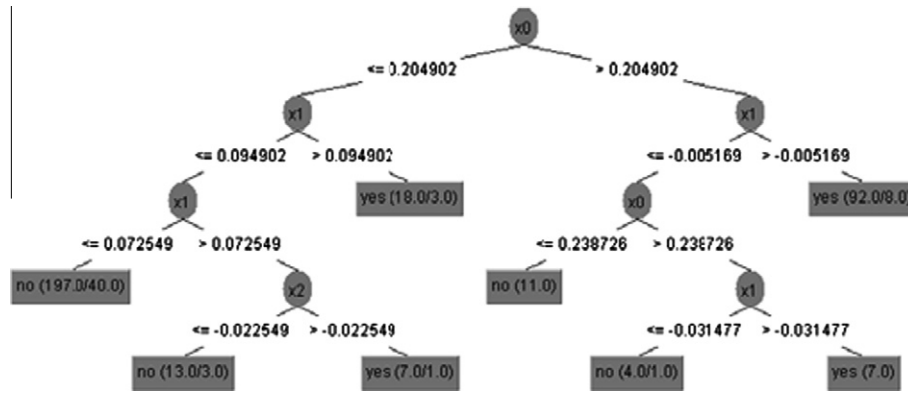
Fig. 17. Tree of the classifier used for classifying touches as stressed or not.

rectly classifies nearly 78%. Moreover, this is built using data from all the users. Training a classifier for each specific user would increase the statistic and generate more accurate classifiers. Nevertheless, considering that this is a generic model, the results are considered satisfactory.

Fig. 18(b) depicts the considered approach. The orange line is a plot from the variation of the intensity during the touch. It contains 6 points. The blue line represents the curve built to fit the variation of the intensity. As expected and depicted in Fig. 18(a), the number of touches classified as stressed tends to greatly increase when moving to higher levels of stress. This figure considers data from all the users.

6. Discussion

Let us now make a critical analysis of the results presented before by analyzing the data from all the users (Fig. 19). We conclude that the most significant feature is the acceleration, with 80% of the users exhibiting important differences between training and real data. On the other hand, the parameter that presents the worst results is the score. However, we would like to stress that this last parameter is merely indicative and is not considered in the real application. It is used here solely with the purpose of inducing an objective on the user: without the objective of maximizing the score, the user would not feel any stress.
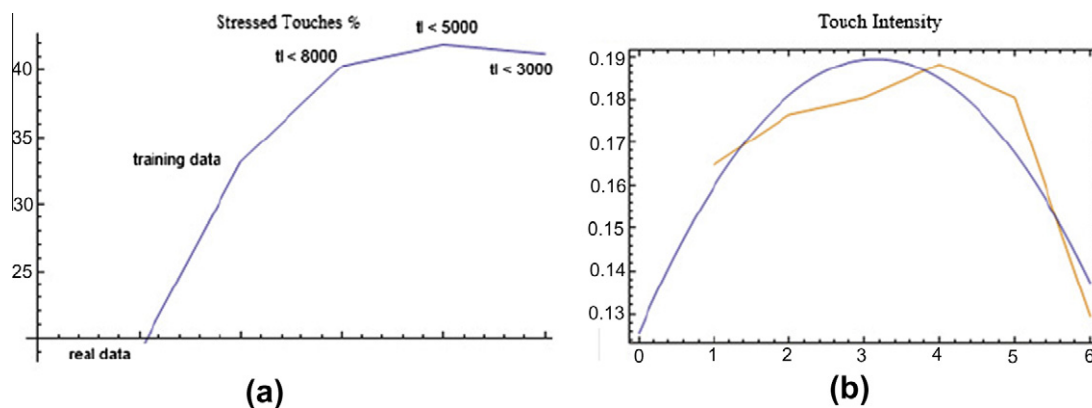

Fig. 18. The plot depicted in (a) shows an increase in the number of touches classified as stressed when considering datasets of increased level of stress. The data concerns all the users. The plot depicted in (b) shows the variation of the touch intensity with time (orange line) and the quadratic function that best fits it (blue line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
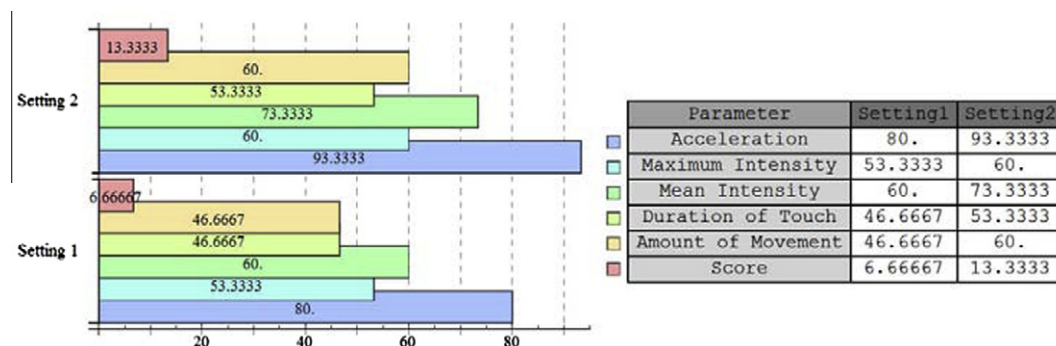

Parameter	Setting1	Setting2
Acceleration	80.	93.3333
Maximum Intensity	53.3333	60.
Mean Intensity	60.	73.3333
Duration of Touch	46.6667	53.3333
Amount of Movement	46.6667	60.
Score	6.66667	13.3333

Fig. 19. Percentage of users that reveal significant differences between training and real data considering a first order analysis (Setting 1) and second order analysis (Setting 2).

We also conclude that, as expected, all the parameters show a better performance in a second order analysis. However, as already mentioned before, the first order analysis is preferred as it considers all the data (and not only data from more stressed states), which means that a statistically significant difference (when it exists) is more solid. Analysing all the data, each test subject has in average 2.93(3) parameters with considerable differences (out of 6). Moreover, in the worst case there are 3 users with only 1 parameter with important differences, and in the best case there are 4 users with 5 parameters with significant differences. Given this, it can be stated that a generic model can be applied to users for whom no training data exists. However, personalized models are certainly more accurate and that is the aim.

Moreover, from the results it is also concluded that the most affected parameters are the ones that people do not conscientiously control. Consider as an example the acceleration of our hands (i.e., how much they move) or the intensity of the touch. On the other hand, the parameters that are more rational (in this case the score) are not influenced in such a significant manner. In our opinion, this is positive in the sense that unconscious behaviors and reactions are more difficult to forge and are usually true reactions of human's body. Thus, the results that stem from their analysis is more solid.

7. Conclusions and future work

In this work some parameters and their relation with stress have been analyzed. One of our main objectives was to do it in a non-invasive way, so that the experiments performed would not influence the results. In that sense, the focus has been set on parameters capable of being transparently acquired from the environment of the application, without the user being explicitly aware of it. Specifically, it has been studied how stress influences the acceleration, the maximum and mean intensity of touch, the duration of the touch, the amount of movement and the cognitive performance. Two additional experiments have also allowed to build a classifier of touch patterns (stressed and not stressed) and to analyze the accuracy of the users' touches. This work defines a way to measure how each user is affected by stress. The approach opens the door to define personalized models to measure the influence of stress in the users.

In future work, we will focus on finding and including new parameters, although still following a non-invasive approach. This will allow the consideration of more features and modalities, thus improving the accuracy of the solution. One of the future additions is a mouse with incorporated luminosity and galvanic skin response. These sensors will activate when touched and will allow acquiring information about the heartbeat rate (through the variations in the luminosity in the finger due to the arrival of blood) and the amount of moisture of the skin (by measuring its electrical conductance). These two parameters are directly related to stress and are definitely very hard to deceive. Regarding the visual sensors, our aim is to go beyond the simple movement analysis to have a broader range of parameters detected. The use of pattern recognition techniques and classifiers will enable the detection of emotions and actions to analyse the users' mood. These techniques, in addition to eye-related features (gaze detection, blinking rate), will complement the stress detection to enhance the conflict resolution process. We also plan to use functional magnetic resonance imaging in order to validate the approach and to further calibrate it, in cooperation with the School of Medical Sciences of our institution.

Acknowledgments

References

Alpaydin, E. (2010). *Introduction to machine learning.* The MIT Press.

Alsina-Jurnet, I., & Gutiérrez-Maldonado, J. (2010). Influence of personality and individual abilities on the sense of presence experienced in anxiety triggering virtual environments. *International Journal of Human-Computer Studies, 68*(10), 788–801.

Beale, R., & Creed, C. (2009). Affective interaction: How emotional agents affect users. *International Journal of Human-Computer Studies, 67*(9), 755–776.

Bellucci, E., Lodder, A. R., & Zeleznikow, J. (2004). Integrating artificial intelligence, argumentation and game theory to develop an online dispute resolution environment. In Proceedings of the 16th IEEE international conference on tools with artificial intelligence (pp. 749–754).

Blascovich, J., & Loomis, J. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry, 13*(2), 103–124.

Campbell, D. A. (1997). Explorations into virtual architecture: A HIT lab gallery. *IEEE Multimedia, 1,* 74–76.

Carneiro, D., Gomes, M., Novais, P., & Neves, J. (2011). Developing dynamic conflict resolution models based on the interpretation of personal conflict styles. In Proceedings of the 15th Portuguese conference on artificial intelligence, progress in artificial intelligence, lecture notes in computer science (Vol. 7026/2011, pp. 44–58).

Carneiro, D., Gomes, M., Novais, P., Andrade, F., & Neves, J. (2011). Automatic classification of personal conflict styles in conflict resolution. In *Frontiers in artificial intelligence and applications. Proceedings of the 24th international conference on legal knowledge and information systems* (pp. 43–52). IOS Press. ISBN 978-1-60750-980-6.

Carneiro, D., Carlos Castillo, J., Novais, P., Fernández-Caballero, A., Neves, J., & López, M. (2012). Stress monitoring in conflict resolution situations. In Paulo Novais, Kasper Hallenborg, Dante I. Tapia, & Juan M. Corchado Rodríguez (Eds.), Ambient intelligence – software and applications – 3rd international symposium on ambient intelligence (ISAmI 2012), Springer – series advances in intelligent and soft computing (Vol. 153, pp. 137–144), ISBN 978-3-642-28785-5.

Castillo, J. C., Rivas-Casado, A., Fernández-Caballero, A., López, M. T., & Martínez-Tomás, R. (2011). A multisensory monitoring and interpretation framework based on the model-view-controller paradigm. In Proceedings of the 4th international workshop on the interplay between natural and artificial computation (Vol. 1, pp. 441–450).

Cox, R. (1985). Sport psychology: Concepts and applications. W.C. Brown.

Cuong, N. H., & Huynh, T.H. (2010). Eye-gaze detection with a single webcam based on geometry features extraction. In Proceedings of the 11th international conference on control, automation, robotics and vision (pp. 2507–2512).

Dodds, T. J., Mohler, B. J., & Bülthoff, H. H. (2011). Talk to the virtual hands: Self-animated avatars improve communication in head-mounted display virtual environments. *PLoS ONE, 6*(10), e25759.

Fernández-Caballero, A., Castillo, J. C., Martínez-Cantos, J., & Martínez-Tomás, R. (2010). Optical flow or image subtraction in human detection from infrared camera on mobile robot. *Robotics and Autonomous Systems, 58*(12), 1273–1281.

Goodman, J. (2003). The pros and cons of online dispute resolution: An assessment of cyber-mediation websites. *ReviewDuke Law and Technology, 4.*

Healey, J., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems, 6*(2), 156–166.

Hudlicka, E. (2003). To feel or not to feel: The role of affect in human-computer interaction. *International Journal of Human-Computer Studies, 59*(1), 1–32.

Jaeggli, T., Koller-Meier, E., & van Gool, L. (2009). Learning generative models for multi-activity body pose estimation. *International Journal of Computer Vision, 83*(2), 121–134.

Jaimes, A., & Sebe, N. (2007). Multimodal human-computer interaction: A survey. *Computer Vision and Image Understanding, 108*(1–2), 116–134.

Larson, D. A. (2006). Technology mediated dispute resolution (TMDR): A new paradigm for ADR. *Ohio State Journal on Dispute Resolution, 21*(3), 629.

Liao, W., Zhang, W., Zhu, Z., Ji, Q., & Gray, W. (2006). Toward a decision-theoretic framework for affect recognition and user assistance. *International Journal of Human-Computer Studies, 64*(9), 847–873.

Lodder, A. R., & Zeleznikow, J. (2010). *Enhanced dispute resolution through the use of information technology.* Cambridge University Press.

Marucci, L., Mori, G., Paterno, F., & Costalli, F. (2001). Design criteria for usable web-accessible virtual environments. <http://www.museumsandtheweb.com/node/9254>.

Mehrabian, A. (1980). *Silent messages: Implicit communication of emotions and attitudes.* Wadsworth Publishing Company.

Moffitt, M. L., & Bordone, R. C. (2005). *The handbook of dispute resolution.* Jossey-Bass.

Novais, P., Costa, R., Carneiro, D., & Neves, J. (2010). Inter-organization cooperation for ambient assisted living. *Journal of Ambient Intelligence and Smart Environments, 1*(2), 179–195.

Novais, P., Carneiro, D., Gomes, & M., Neves, J. (2012). Non-invasive estimation of stress in conflict resolution environments. In Yves Demazeau, Jörg P. Müller,

Juan M. Corchado RodrÃguez, & Javier Bajo Pérez (Eds.), Advances on practical applications of agents and multi-agent systems – 10th International conference on practical applications of agents and multi-agent systems (PAAMS 2012), Springer – series advances in intelligent and soft computing (Vol. 155, pp. 153–160), ISBN 978-3-642-28785-5.

O'Hara, S., Lui, Y. M., & Draper, B. A. (2011). Unsupervised learning of human expressions, gestures, and actions. In Proccedings of the 2011 IEEE international conference on automatic face & gesture recognition (pp. 1–8).

Park, D., Lee, J. H., & Kim, S. (2011). Investigating the affective quality of interactivity by motion feedback in mobile touchscreen user interfaces. *International Journal of Human-Computer Studies, 69*(12), 839–853.

Picard, R. W. (1997). *Affective computing*. Cambridge, MA, USA: MIT Press.

Rehm, M., Bee, N., & André, E. (2008). Wave like an Egyptian: Accelerometer based gesture recognition for culture specific interactions. In Proceedings of the 22nd British HCI group annual conference on people and computers: Culture, creativity, interaction (Vol. 1, pp. 13–22).

Selye, H. (1956). *The stress of life*. McGraw-Hill.

Smaragdis, P., Radhakrishnan, R., & Wilson, K. W. (2009). Context extraction through audio signal analysis. In Multimedia content analysis (pp. 1–34).

Tidd, T. S., & Friedman, R. (2002). Conflict style and coping with role conflict: An extension of the uncertainty model of work stress. *International Journal of Conflict Management, 13*(3), 236–257.

Vizer, L. M., Zhou, L., & Sears, A. (2009). Automated stress detection using keystroke and linguistic features: An exploratory study. *International Journal of Human-Computer Studies, 67*(10), 870–886.

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

Wu, Y., Hua, G., & Yu, T. (2003). Tracking articulated body by dynamic Markov network. In Proceedings of the ninth IEEE international conference on computer vision (Vol. 2, pp. 1094–1101).