

An application to enrich the study of Auditory Emotion Recognition

Renato Rodrigues[†], Augusto J. Mendes[‡], Davide Carneiro^{†*}, Maria Amorim^{‡§}, Ana P. Pinheiro^{‡§} and Paulo Novais[†]

*CIICESI, ESTG, Polytechnic Institute of Porto Felgueiras, Portugal

Email: dcarneiro@estg.ipp.pt

[†]Algoritmi Centre/Department of Informatics University of Minho, Braga, Portugal

Email: pg27740@alunos.uminho.pt, dcarneiro@di.uminho.pt, pjon@di.uminho.pt

[‡]School of Psychology, University of Minho, Braga, Portugal

Email: a65563@alunos.uminho.pt, tinhamorim@gmail.com, ana.pinheiro@psi.uminho.pt

[§]Faculty of Psychology, University of Lisbon, Lisbon, Portugal

Abstract—The ability to recognize emotions in spoken words is central in human communication and social relationships. When studying one's ability to perceive emotions, the standard paradigm is to have listeners choose which one of several emotion words best characterizes linguistically neutral utterances made by actors attempting to portray various emotional states. Usually, generic experiment control software are used, which may present several limitations. In this paper we present a novel approach to the problem, based on a mobile application that can be easily configured by the researcher to set up the desired protocol. This approach not only facilitates and improves study design and data collection, but also provides a plethora of new variables about the participants that, to the best of our knowledge, have never been considered before in this domain, including behavioural research.

I. INTRODUCTION

Face-to-face conversations are very rich contexts, particularly in the exchange of non-verbal content [1], that is, information that is not conveyed by words but by our gestures, postures, intonation or speech rhythm. These aspects are essential for an efficient process of communication between speaker and listener, and also allow one to perceive the state and intention of the other.

The ability to convey, as well as to accurately and rapidly decode, emotions is fundamental for the success of communication and interactions [2].

Auditory emotion recognition refers precisely to the ability of a listener to infer emotion from sounds in the environment, including the voice. When studying auditory emotion recognition, the standard perception paradigm is to have listeners choose which one of several emotion words best characterizes linguistically neutral utterances or nonverbal vocalizations made by actors attempting to portray various emotions [3], [4]. In addition, listeners may be asked to classify the sound in several dimensions, such as its valence (a continuum ranging from 'unpleasant' to 'pleasant'), arousal (from 'calm' to 'arousing'), and dominance (from 'controlled' to 'in control')[5]. Other important dimensions include the intensity of the emotion that was communicated, as well as its authenticity (genuine vs. acted emotions are processed differently[6].

Traditional approaches involves setting up the experimental trials, as well as controlling for stimulus presentation and timing through software such as Presentation¹ (Neurobehavioral Systems, Inc., Albany, CA, USA) or Superlab² (Cedrus, San Pedro, CA). The few measures that are often the focus of those studies (e.g. accuracy rates, reaction time) are usually obtained by recording the participants' responses directly via the software, or by using a paper-and-pencil approach.

This typical approach is often time-consuming, prone to errors (e.g. when the results are transferred from the paper to the computer), and dependent on the availability of the software and equipment in the context in which the behavioural data are required.

This paper introduces an innovative instrument to assess auditory emotional recognition that can be used both in research and clinical settings, focused on a Tablet. The user interacts with a mobile application to provide feedback about the auditory stimuli. To do so, the participant selects which one of several emotion words (arranged in buttons and set by the expert when defining the study) best characterizes the emotion conveyed by the voice. The participant also classifies the valence, authenticity and intensity of the emotion that was expressed.

While developed specifically for the field of auditory emotion recognition, the system can be easily adapted to other areas. Compared with more traditional assessments, this application aims to provide a faster and more dynamic way of assessing vocal emotional recognition in healthy subjects as well as in clinical populations. Moreover, this application incorporates concepts from Context-aware Computing[7], Ambient Intelligence[8] and Behavioural Biometrics[9], providing an innovative and interesting plethora of new variables that will significantly enrich these studies.

¹Presentation is a stimulus delivery and experimental control program typically used in neuroscience and behavioural research. <https://www.neurobs.com/>

²Superlab is an environment for setting-up and running experimental studies, providing accuracy and reaction time measures. <http://www.superlab.com/>

II. ARCHITECTURE

In Auditory Emotion Recognition studies, there are usually two people involved: the participant, whose capacity to evaluate subjective dimensions of an emotional sound is being assessed, and the researcher, who is playing the stimuli or monitoring the software that does so, as well as registering the participant's answers.

In this new approach, the same two roles exist. However, there is now a looser coupling between them. In fact, the participant may be alone in a room, listening to the stimuli and interacting with the tablet to provide the answers, while the researcher is in an adjacent control room, eventually looking at the participant's answers and behaviours in real time (Figure 1).

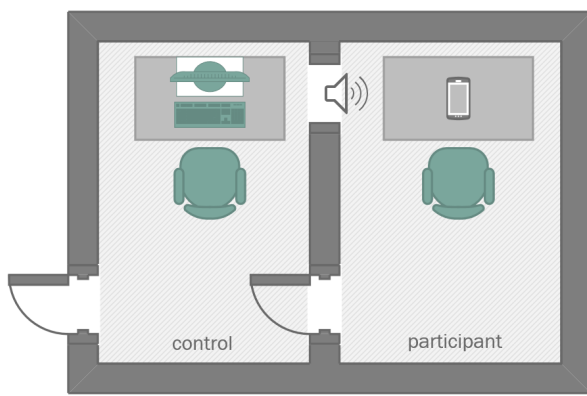


Fig. 1. Layout of the environment: the participant sits alone in a room, listening to vocal sounds through a sound speaker and interacting with the tablet, which is wirelessly connected to the computer in the control room.

There are three main components in the architecture, as depicted in Figure 2: a mobile application and a controller exist in the so-called room area (the location where the study is taking place) and the server exists in a potentially different location.

The mobile application, implemented in Android, exists in the user area, in which the user interaction takes place. The application receives the details of each study and dynamically generates the corresponding user interfaces. It also acts as a data-generation device, collecting not only the participants' answers but also data describing their behaviour during the study. These data are collected in a transparent way and sent to the server in real time.

The computer in the controller area communicates directly with the mobile application during the studies. The researcher uses this computer to configure new studies or manage existing ones as well as to start a new data collection procedure in a specific device. The computer is connected to a Logitech 5.1 Surround Sound System, which is used to play the auditory stimuli, as requested by the mobile application. The data collected can be visualized in real time in this computer by the researcher.

Finally, the server is the computer where data is stored after the conclusion of the data collection procedure. It includes tools for data processing, analysis and visualization.

Communication between these devices is implemented by means of a sockets API. Messages are exchanged in JavaScript Object Notation (JSON) format. This format is a lightweight, text-based, language-independent data interchange format. It was derived from the ECMAScript Programming Language Standard. JSON defines a small set of formatting rules for the portable representation of structured data. JSON can represent four primitive types (strings, numbers, booleans and null) and two structured types (objects and arrays).

These messages are used, for instance, by the mobile application to request the controller computer to play a given auditory stimulus, or by the controller computer to send the configuration of a study to the mobile application.

The following two sub-sections describe, in more detail, the functionalities of the control computer and the mobile application.

A. Life-cycle

From a high-level perspective, the life-cycle of the developed system is as follows:

- 1) The researcher creates or selects an existing study and uploads it to the intended mobile device;
- 2) The participant begins by providing the necessary personal information;
- 3) The participant goes through a training phase, similar to the real one, but in which instructions on how to proceed are provided. The aim is that the participant gets familiarized with the process and with the type of stimuli that will be presented;
- 4) Upon finishing training, the participant may choose to repeat the previous step;
- 5) The participant starts the actual experimental task. Each study is composed of one or more iterations. In each iteration:
 - a) The participant hears a stimulus (a nonverbal vocalization - e.g., laughs, growls). She/he may repeat it any number of times by clicking a button in the graphical interface;
 - b) The participant classifies the stimulus according to the perceived emotion: relief, fun, fear, neutral, disgust, pleasure, achievement, anger, surprise or sadness;
 - c) The participant classifies the stimulus according to valence, using a 9-point likert scale;
 - d) The participant classifies the stimulus according to intensity, using a 9-point likert scale;
 - e) The participant classifies the stimulus according to authenticity, using a 9-point likert scale;

B. Control

In what concerns the control computer, there are three main functionalities worth highlighting (Figure 3).

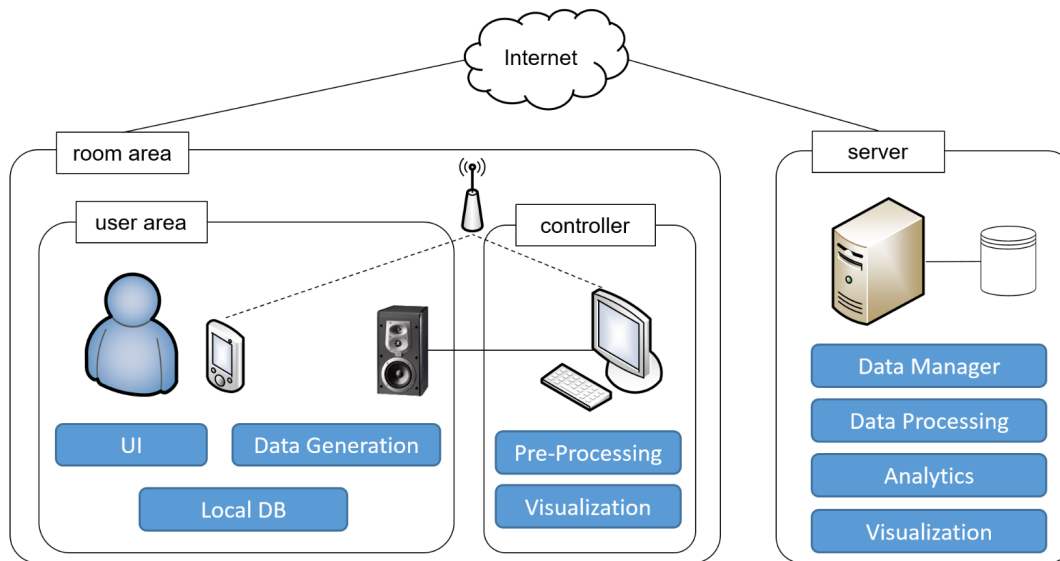


Fig. 2. Architecture of the data acquisition framework.

The first functionality is the management of existing studies or protocols and it is implemented by the first tab of the graphical interface. The first list in this tab shows the studies that currently exist in the server while the second shows the details (i.e. stimuli being used) of a study after it is selected. The date of the creation of the study as well as its description are also shown below. It is possible to delete the selected study or send it to a specific instance of the mobile application, running at a given IP address. The interface also allows to filter studies or stimuli inside studies by name and it reproduces the auditory stimuli locally whenever one is clicked. Finally, the interface also includes functionalities to check the communication with the client application.

The second functionality, implemented in the second tab, is the creation of new studies or protocols. It allows selecting from the vast list of available stimuli those that will constitute a new specific protocol or study, with a given name. By clicking each one the researcher may listen to the auditory stimulus. For convenience, stimuli can be searched by name. Once a new protocol is created it is stored in the server's database and becomes available to be used.

Finally, the third tab implements functionalities that allow the researcher to visualize the data. The researcher can visualize data that are being collected in real time or data that were collected in previous studies. Given the extent of the collected datasets, filters are also available to allow a more efficient visualization of the required data. Besides from being stored in the database, datasets can also be exported to a .csv file (Comma-Separated Values), which is a convenient method aimed at facilitating posterior data analysis by researchers. Section II-B1 provides a more detailed description of the structure of the generated datasets.

1) *Structure of the Dataset:* The dataset generated during each study describes a group of very different variables, both



Fig. 3. Graphical Interface of the control computer (in Portuguese) with its three main functionalities in three different tabs: study management, study creation, results.

personal, operational and behavioural.

Personal variables are included in the header of the dataset and are used to identify and characterize the participant including, among others, a unique id, gender, date of birth, level of schooling or job. Operational variables describe certain events such as the start of a study or advancing to the next stimulus. Finally, behavioural variables describe the behaviour of the user throughout the study.

Briefly, the dataset is composed of a sequence of lines, each line describing a specific event. Apart from the header, lines can have one of four types (touch, event, stimulus, emotion, likert). Depending on its type, each line holds different data, as follows:

- touch, timestamp, source, X, Y, P, S - it describes the event of a touch on the screen of the mobile device, at a given time. It identifies the source of the touch (e.g. a specific control, the background), its coordinates on the

screen and the average values of pressure and area during the touch;

- event, timestamp, description - it describes a specific event in a given time stamp, such as starting the instruction activities or advancing to the study;
- stimulus, timestamp, id - denotes the moment in which the participant in the study proceeds to a new stimulus;
- emotion, timestamp, type, #repetitions - it denotes that in a given moment, the participant classified the current stimulus as representing a particular emotion. It also describes the number of repetitions of the stimulus before registering the answer;
- likert, timestamp, type, value, #repetitions - it denotes the classification of the current stimulus in a likert scale of a specific value (i.e. valence, intensity or authenticity) with a certain value. It also describes the number of repetitions of the stimulus before registering the answer.

C. Mobile Application

The mobile application, developed for android, targets devices with large screens in order to facilitate interaction. It was designed to be simple to interact with (of particular relevance for special populations, such as children and older adults), and to minimize influence on results, namely by keeping text and colors to a minimum. Note that the graphical interfaces depicted in this section are in Portuguese since the mobile application was developed for the Portuguese context.

The application receives the protocol via socket in JSON format and dynamically generates all the corresponding graphical interfaces. Two examples are depicted in Figures 4 and 5.

Figure 4 depicts the graphical interface used by the participant to classify the emotion expressed while listening to the stimulus. The stimulus can be repeated any number of times by clicking on the sound speaker image. By clicking in one of the emotions the application advances to the following activity.



Fig. 4. Graphical interface for the participant to select the emotion portrayed by the stimulus heard: sadness, fear, disgust, anger, surprise, relief, pleasure, fun, achievement or neutral (in Portuguese).

Figure 5 depicts the graphical interface in which the participant is able to classify the valence of a given vocalization using a 9-point likert scale. As in Figure 4, the stimulus can be

repeated by clicking on the sound speaker. After adjusting the desired value, the participant advances by clicking a button.

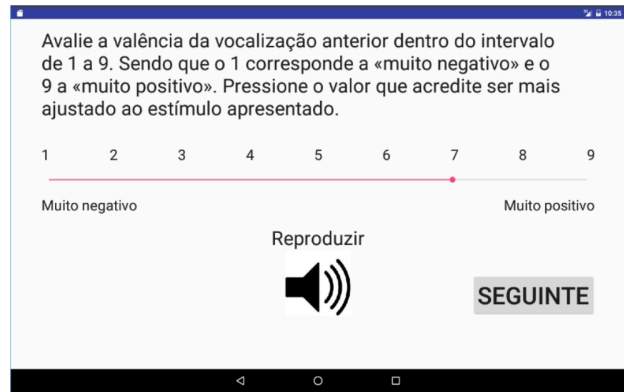


Fig. 5. Graphical interface with a 9-point likert scale for the assessment of the stimulus' valence by the participant (in Portuguese).

As the user interacts with the mobile application it registers the previously mentioned data and sends them, in real time, to the controller computer. Data are aggregated and sent via socket, in JSON format, at each new stimulus or at the end of the protocol in the case of the last stimulus.

III. VALIDATION

So far this paper has been devoted to the description of the developed system and its main functionalities. In this section we describe the process of its validation in the School of Psychology of the University of Minho, Portugal, where the application will from now on be used. For validation purposes ten individuals were selected to participate in a protocol with fifty stimuli, with five exemplars of each emotion (sadness, fear, disgust, anger, surprise, relief, pleasure, amusement, triumph, plus a neutral expression). A Samsung Galaxy Tab 3 with a screen of 10.1" was used for user interaction.

We should note that the main goal of this paper is not actually to study vocal emotion recognition in a given population but rather to validate the developed system. Therefore, this section does not focus on a comprehensive statistical analysis of the collected data, but instead exemplifies some of the critical information that can be obtained in a system such as the one described here.

1) *Behavioural Biometrics*: Behavioural Biometrics is the field of study related to measuring uniquely identifying patterns in human actions. Many of the things we do are done in a unique way: walking, talking, typing on a keyboard or using the mouse [10]. Given the uniqueness of our behaviours, Behavioural Biometrics are often used to implement continuous and post-login methods of security and authentication. Nonetheless, they can also be used to detect significant changes in the user's behaviour, which may indicate a significant change in the user's state. In previous work we have studied the effects of mental fatigue and stress on keystroke and mouse dynamics [11], [12].

The developed system provides an interesting number of new features related to Behavioural Biometrics that may enrich

TABLE I
GENERAL STATISTICS OF TOUCH INTENSITY

| | \bar{x} | \tilde{x} | S | #touches |
|-----|-----------|-------------|--------|----------|
| MP1 | 0.0992 | 0.1020 | 0.0294 | 256 |
| MP2 | 0.0905 | 0.0902 | 0.0199 | 486 |
| MP3 | 0.1130 | 0.1098 | 0.0366 | 468 |
| MP4 | 0.1129 | 0.1176 | 0.0250 | 627 |
| MP5 | 0.1053 | 0.1020 | 0.0358 | 376 |
| FP1 | 0.0705 | 0.0628 | 0.0315 | 403 |
| FP2 | 0.0884 | 0.0863 | 0.0274 | 406 |
| FP3 | 0.1180 | 0.1216 | 0.0284 | 504 |
| FP4 | 0.0928 | 0.0941 | 0.0233 | 252 |
| FP5 | 0.0823 | 0.0862 | 0.0303 | 485 |

the diagnosis and behavioral profile of the participant. These features include touch intensity, touch duration and touch area. In previous work, we have shown how the shape of the touch (intensity over time) is affected by the participant’s level of stress [13]. Researchers will now be able to understand how touch duration/intensity/area varies according to, among others, gender, type of stimuli, age or cognitive disabilities. This may, in the future, open the door to the development of mobile applications that are able to profile a user, namely in terms of cognitive disabilities or emotional deficits.

Table I shows the number of touches necessary for each of the ten participants to complete the protocol, as well as the values for the mean, median and standard deviation of touch intensity. In this table MP denotes Male Participants while FP denotes Female Participants.

Figure 6 depicts the distribution of touch intensity in each participant, showing the inter-individual differences.

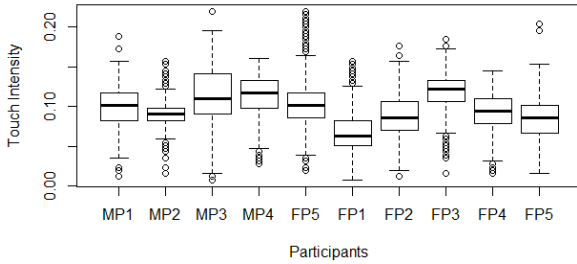


Fig. 6. Distribution of touch intensity by participant.

Table II analyses the same variable, but as a function of gender. According to the collected data, female participants performed lighter touches on the screen, showing a decreased mean and median values of intensity, despite a slightly larger standard deviation. Female participants also used, on average, less touches to complete the protocol. Figure 7 illustrates this important finding.

It is also possible to analyse the evolution of any of these features over time. Figures 8 and 9 show the evolution of touch intensity over time for female and male participants, respectively. This may reveal interesting individual information about each participant. For example: male participants 1 and 5 have two similar traits: a remarkable difference between the

TABLE II
TOUCH INTENSITY: COMPARISON BETWEEN MALE AND FEMALE PARTICIPANTS

| | \bar{x} | \tilde{x} | S | #touches |
|--------|-----------|-------------|--------|----------|
| Male | 0.1051 | 0.1020 | 0.0307 | 2213 |
| Female | 0.0913 | 0.0902 | 0.0332 | 2050 |

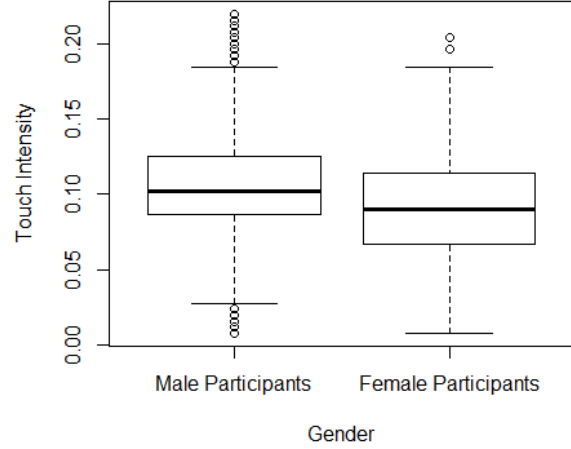


Fig. 7. Distribution of touch intensity by gender, differences are statistically significant (Kolmogorov-Smirnov test, p -value $< 2.2^{-16}$)

training phase and the actual protocol and a tendency for touch intensity to increase after time during the protocol.

Touch intensity should, in principle, be more or less constant over time. However, some users show significant differences at certain points (e.g. female participants 4 and 5). Were these differences caused by certain types of vocal emotions being experienced by the participants, or being more relevant for the participants? Does this mean that these participants were affected differently by these emotional expressions? Could this information pinpoint individuals with different emotional responses? Could this knowledge lead to the development of emotion-aware applications and devices?

Figure 10 shows another possible analysis: the intensity of the touch right after hearing a stimulus of a given type. In this participant, emotions classified as portraying fear were associated to a touch that is softer than average, while emotions classified as portraying relief caused stronger touches.

Finally, another feature that can be extracted from this dataset is the accuracy of touches, since the dataset distinguishes between touches on active controls or touches on the background.

We believe that this type of information may not only significantly enrich the diagnosis and study of auditory emotion recognition but also lead to the development of software and devices that are sensitive to emotions.

2) *Operational Features*: In this section we include features that are extracted from the actions of the participant during the protocol, and their timing.

One of the potentially interesting features is the number of repetitions of each stimulus. Table III highlights the differ-

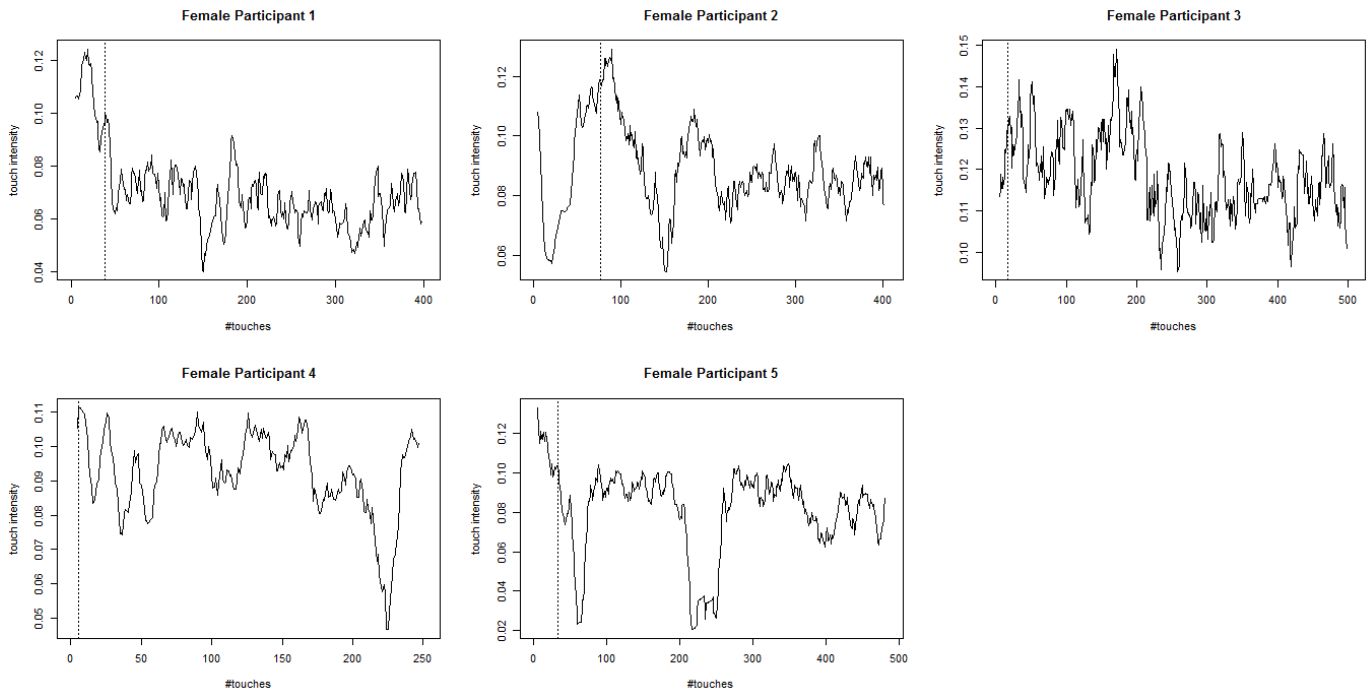


Fig. 8. Evolution of touch intensity during the protocol for each female participant. The vertical line denotes the moment in which the participant finishes the training phase.

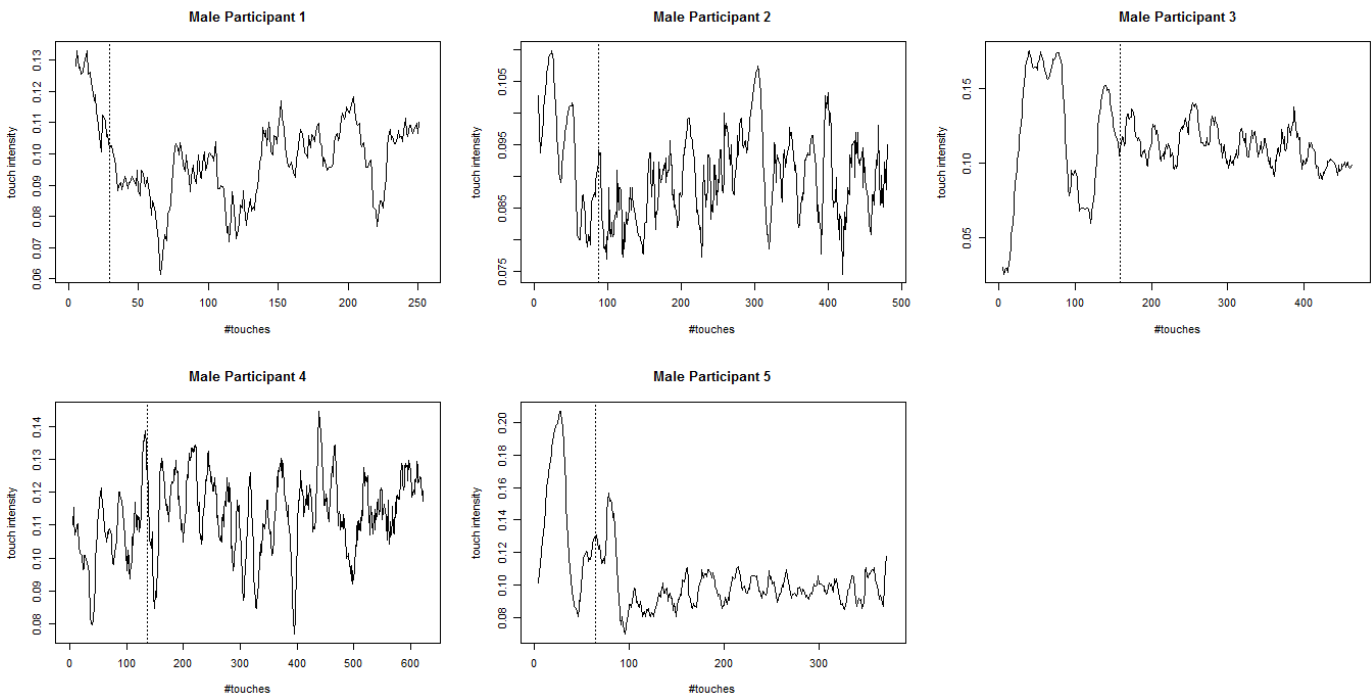


Fig. 9. Evolution of touch intensity during the protocol for each male participant. The vertical line denotes the moment in which the participant finishes the training phase.

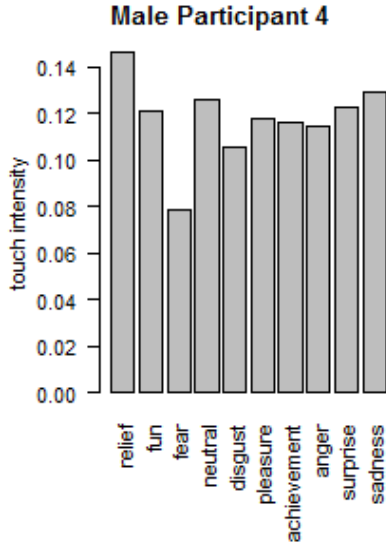


Fig. 10. Touch intensity immediately after hearing each type of stimulus for one of the participants.

TABLE III
NUMBER OF STIMULI PLAYED: COMPARISON BETWEEN MALE AND FEMALE PARTICIPANTS

| | \bar{x} | \tilde{x} | S | #stimuli played |
|--------|-----------|-------------|-------|-----------------|
| Male | 86.6 | 81.0 | 34.22 | 433 |
| Female | 70 | 68 | 16.60 | 350 |

ences between male and female participants in this respect, complemented by Figure 11. The data show that, on average, male participants played the stimuli nearly 87 times, whereas female participants played the sounds 70 times on average. Note that each of the stimuli used is automatically played once and then each user has the option to repeat it any number of times. This means that male participants repeated a stimulus nearly 37 times (on average), whereas female did so 20 times.

Although this difference is remarkable, it is not statistically significant (note, though, that the sample size is small, which precludes the generalization of these findings; perhaps, a larger sample size will show significant differences). Nonetheless, this type of information may lead to interesting questions. For example, are female individuals better at perceiving emotions since they apparently require less repetitions of the stimuli to classify them (and as demonstrated by many studies probing sex differences in vocal emotional processing - e.g., [14], [15], [16])?

More detailed features could also be explored, such as the number of repetitions per type of stimulus, which could point out, for each individual, emotions that are eventually harder to decode. A related measure that could also be extracted from the dataset is the time spent for each type of stimulus, with longer times spent to classify a stimulus of a given type potentially pointing out an increased difficulty in assessing the corresponding emotion.

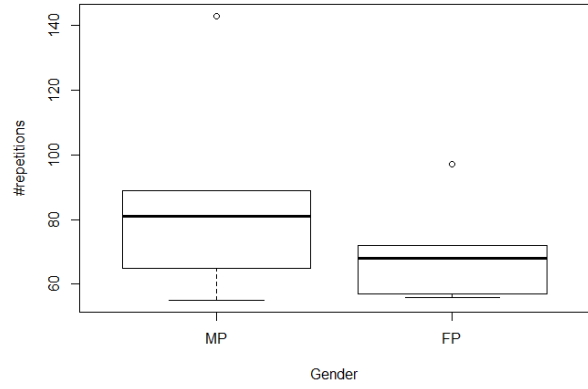


Fig. 11. Distribution of the number of stimuli played, by gender, differences are not statistically significant (Kolmogorov-Smirnov test, p -value = 0.873)

TABLE IV
PERCENTAGE OF CORRECTLY CLASSIFIED EMOTIONS

| | P1 | P2 | P3 | P4 | P5 | \bar{x} |
|--------|-----|-----|-----|-----|-----|-----------|
| Male | 70% | 74% | 64% | 70% | 70% | 69.6% |
| Female | 52% | 72% | 80% | 80% | 82% | 73.2% |

Another very interesting feature that can be extracted from the dataset is the percentage of correctly classified emotions. This is possible to calculate since in the database each stimulus is associated to the emotion it portrays. Table IV shows the results extracted from the dataset, with female participants showing a better accuracy at classifying emotions. This supports a previous claim that female individuals may be better in decoding emotions from vocal cues (e.g., [15]). Nonetheless, and as mentioned before, it is not the aim of this paper to put forward this type of conclusions but rather to show the richness of information that can be extracted from the developed system.

Other so-called operational features could be extracted, or more detailed versions of the described ones, such as the percentage of correctly classified emotions by type of emotion, potentially pointing out vocal emotions that, for each individual, are easier to identify than others.

IV. CONCLUSION

In this paper an innovative approach was presented to improve auditory emotion recognition studies. The system, composed of a mobile application and a server application, was developed and validated in cooperation with the School of Psychology of the University of Minho, where it will be used to assess auditory emotion recognition both in research and clinical settings.

In this paper the developed system was described, with a special emphasis on the functionalities implemented. Moreover, the main innovative aspects of the work were described, accompanied with examples of the information that can be extracted from each study. These key innovative aspects can thus be summarized:

- Studies are easy to design and share among researchers. The mobile application generates all the necessary graphical interfaces according to the study design, in a transparent way for the researcher;
- Data collection and storage in a structured manner is automatized and requires no Human intervention, improving the efficiency of the process and its validity by eliminating potential Human error;
- Many new variables are now considered that may provide important information about participants' behaviour. This may be very important to clarify how each participant is affected by different types of emotions;
- Several studies can be conducted simultaneously as now there is not a dependence on the researcher to play the stimuli and record the the participants' responses;
- Data are readily available during and immediately after the collection, facilitating and accelerating its analysis.

However, we believe that the advantages of the developed system go beyond the enrichment of auditory emotion recognition studies: we are convinced that this approach may lead to the collection of relevant data to create specific user profiles. Indeed, the School of Psychology frequently deals with populations with varying characteristics, including age, gender, disabilities, psychiatric, among others.

The collection of these data, properly contextualized with the participants' characteristics, may allow the characterization of user interaction profiles that can later be used to develop software and hardware sensitive to human emotions or characteristics.

ACKNOWLEDGMENT

This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT – Fundação para a Ciência e Tecnologia under projects UID/CEC/00319/2013 and PTDC/MHN-PCN/3606/2012.

REFERENCES

- [1] M. L. Knapp, J. A. Hall, and T. G. Horgan, *Nonverbal communication in human interaction*. Cengage Learning, 2013.
- [2] P. N. Juslin and P. Laukka, "Communication of emotions in vocal expression and music performance: Different channels, same code?" *Psychological bulletin*, vol. 129, no. 5, p. 770, 2003.
- [3] J.-A. Bachorowski, "Vocal expression and perception of emotion," *Current directions in psychological science*, vol. 8, no. 2, pp. 53–57, 1999.
- [4] C. F. Lima, S. L. Castro, and S. K. Scott, "When voices get emotional: a corpus of nonverbal vocalizations for research on emotion processing," *Behavior research methods*, vol. 45, no. 4, pp. 1234–1245, 2013.
- [5] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.
- [6] A. Anikin and C. F. Lima, "Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations," *The Quarterly Journal of Experimental Psychology*, no. just-accepted, pp. 1–53, 2016.
- [7] A. Schmidt, "Context-aware computing: context-awareness, context-aware user interfaces, and implicit interaction," *The Encyclopedia of Human-Computer Interaction, 2nd Ed.*, 2013.
- [8] P. Novais, R. Costa, D. Carneiro, and J. Neves, "Inter-organization cooperation for ambient assisted living," *Journal of Ambient Intelligence and Smart Environments*, vol. 2, no. 2, pp. 179–195, 2010.

- [9] M. Sultana, P. P. Paul, and M. Gavrilova, "A concept of social behavioral biometrics: motivation, current developments, and future trends," in *Cyberworlds (CW), 2014 International Conference on*. IEEE, 2014, pp. 271–278.
- [10] A. Pimenta, D. Carneiro, P. Novais, and J. Neves, "Monitoring mental fatigue through the analysis of keyboard and mouse interaction patterns," in *International Conference on Hybrid Artificial Intelligence Systems*. Springer, 2013, pp. 222–231.
- [11] M. Rodrigues, S. Gonçalves, D. Carneiro, P. Novais, and F. Fdez-Riverola, "Keystrokes and clicks: Measuring stress on e-learning students," in *Management Intelligent Systems*. Springer, 2013, pp. 119–126.
- [12] D. Carneiro and P. Novais, "Quantifying the effects of external factors on individual performance," *Future Generation Computer Systems*, vol. 66, pp. 171–186, 2017.
- [13] D. Carneiro, P. Novais, M. Gomes, P. M. Oliveira, and J. Neves, "A statistical classifier for assessing the level of stress from the analysis of interaction patterns in a touch screen," in *Soft Computing Models in Industrial and Environmental Applications*. Springer, 2013, pp. 257–266.
- [14] M. Besson, C. Magne, and D. Schön, "Emotional prosody: sex differences in sensitivity to speech melody," *Trends in cognitive sciences*, vol. 6, no. 10, pp. 405–407, 2002.
- [15] A. Schirmer, T. Striano, and A. D. Friederici, "Sex differences in the preattentive processing of vocal emotional expressions," *Neuroreport*, vol. 16, no. 6, pp. 635–639, 2005.
- [16] A. P. Soares, A. P. Pinheiro, A. Costa, C. S. Frade, M. Comesaña, and R. Pureza, "Affective auditory stimuli: Adaptation of the international affective digitized sounds (iads-2) for european portuguese," *Behavior research methods*, vol. 45, no. 4, pp. 1168–1181, 2013.