Toward Recognizing Individual's Subjective Emotion from Physiological Signals in Practical Application

Olivier Villon and Christine Lisetti*

Abstract

During last decade, an increasing interest for interpreting users' emotional subjective experience on the basis of physiological signals has led to various approaches. In this article, we focus on two different approaches toward emotion recognition : (1) the user-dependency of psycho-physiological data collected (i.e. do we choose to keep track of the specificity of individuals' responses or do we ignore such specificity), and (2) the degree of subjectivity of the stimuli used to elicit emotions (i.e. stimuli with high level of agreement in terms of what emotional experience they elicit among a population can be chosen versus stimuli without such an agreement). In order to assess the implications of adopting one of these methodologies on the personalization of emotion recognition from physiological signals we present our empirical results for emotion recognition from physiological signals based on an experiment involving 40 subjects. We conclude by proposing requirements for any chosen approach to achieve suitable online emotion recognition, in an out-of-the-lab context (e.g. interactive art, e-Health application).

1 Introduction

Research on emotion recognition from physiological signals has increased during last decade and is getting closer to achieve online recognition (i.e. most of the approaches tend to a 80% of recognition) either at the inter-individual level (using a common training database for different subjects) or at the intra-individual level (using one specific physiological training database for each subject). It is therefore becoming feasible to aim at building user-models including the user's emotions from that recognition process. In this article, we consider two main issues encountered in emotion recognition - user dependency and subjectivity of stimuli - and we study their implications for out of the lab scenarios (as home-health care, to bring back affective information in the communication channel between patients and health-care providers and/or artificial devices [7][6]). We especially formulate hypothesis and validate them with empirical results. We then propose requirements for out-of the lab applications that can guide in-the-lab research.

*Eurecom Institute, Multimedia Communications Department, Sophia Antipolis, FRANCE. Corresponding author : villon@eurecom.fr. This work was partially founded by ST Microelectronics in the framework of the Region Provence-Alpes-Côte-d'Azur (PACA) PACALab.

2 Related Work : User-Dependency and Subjectivity of Stimuli Approaches

Because we are particularly interested in the issues of the *user-dependency* and of the *subjectivity* of stimuli used for learning and testing emotion recognition techniques, we provide a brief review of existing approaches in terms of these two issues (summarized in Table 1[1]). Whatsoever the approach, learning and testing is made on what we refer as a set of a psycho-physiological pair. The psychological part is made of a discrete representation (e.g. 'joy','fear') or a dimensional representation (e.g. 'high valence','low arousal') of emotion. The physiological part is a set of feature values derived from the physiological measure performed on the subject.

User dependency : Considering the individual's psycho-physiological pair or not consider them (by mixing or averaging the pairs) as potentially different among users.

- In a *user-dependent approach*, a set of psycho-physiological pairs built by collecting data of one unique user (often recorded during several days) is used. A machine learning is then typically performed between the psychological evaluations in terms of emotion and physiological measures performed on the user, and recognition rate estimation is achieved on the database for that unique user.
- In a *user-independent approach*, a set of psycho-physiological pairs is built by collecting data from several users, and it is then averaged (or mixed, i.e. grouping pairs of different users) after normalization among the population studied. A machine learning algorithm is then performed between psychological and physiological representations of emotion of the averaged (or the mixed) studied population. Recognition rate is acheived on the database for any user, and still in the context of the experiment.

Subjectivity of stimuli : Considering or not consider the subjective evaluation of stimuli (by focusing or not on stimuli with a high level of agreement according to the emotion elicited across a population).

- In a *subjective rating of stimuli approach*, the user is requested to produce (via mental imagery) or to estimate subjectively the psychological emotion evaluations of stimuli, either with a discrete label (e.g. joy, fear) or on a dimensional spatial representation (e.g. in a 2D-space represent high arousal with a dot). This subjective estimation of the stimuli (self-report) is used for the training and testing.
- In a *social agreement of stimuli approach*, the user is not requested to estimate subjectively psychological emotion evaluations of stimuli, as a database of pre-validated stimuli chosen to have a high level of agreement among a representative set of a population is used.

We explain next the implications of adopting one of these methodologies on the personalization of emotion recognition from physiological signals.

3 Experiment and results

In order to assess the implications of adopting one of these methodologies on the personalization of emotion recognition from physiological signals, we tested empirically the psycho-physiological inter-individual differences in regards with the above mentionned approaches.

[1] We are aware that there exist a wider number of noticeable differences among these approaches (e.g. choice of classifier, choice of emotion representation of stimuli) but they are less relevant for the personalisation of emotion recognition and therefore for this discussion

	User-dependency		Subjectivity of Stimuli	
Authors	User Independent	User dependent	Social Agreement	Subjective Rating
Picard et al.[8]		✔		✔
Haag et al.[3]		✔	✔	
Kim et al.[4]	✔		✔	
Lisetti and Nasoz[5]	✔		✔	
Anttonen and Surakka[1]	✔		✔	
Wagner et al.[12]		✔		✔
Changchun et al.[2]		✔	✔	
Villon and Lisetti[10]	✔	✔		✔

Table 1. Description of approachs of emotion recognition from physiological signals, from the user dependency and the subjectivity of emotion estimation.

3.1 Materials and methods

We performed an experiment involving 40 subjects (half men and women, average 32 years-old), to build a database of psycho-physiological measures based on subjective ratings. While measuring heart rate and skin conductance, we provided a set of 60 stimuli (31 images, 25 sounds, and 5 videos) and requested subjects to then rate each stimulus according to valence (pleasure/displeasure) and arousal (calm/exciting) emotional dimensions. Some stimuli were from previous studies using social agreement (e.g. the images which comes from the International Affective Picture System, see [11]). Because we are interested by inter-individual differences which could exist between the stimulus and the affective experience (which are considered by subjective ratings), we did not only choose stimuli with a previously tested social agreement. The way individuals evaluate stimuli could be considered to engage an *embodied affective relationship*, dependent of the personal history of each individual ([9]) : we thus chose arbitrary stimuli varied in contents and multimedia type, to potentially elicit different subjective emotion for each individual. We collected physiological measures using a Polar-based system for the heart rate, and using the Armband from Bodymedia (a wearable device which embed a set of physiological sensor) [11] for the skin conductance. We collected psychological measures (subjective ratings) by using a 2-Dimensional (valence and arousal dimensions, and not discrete emotion, e.g. Joy) computer interface [11] where users could place stimuli according to their subjective ratings.

3.2 Results : Empirically Comparing inter-individual differences of psychological and physiological responses to stimuli

When adopting a set of stimuli with a social agreement, it is a mean to control the certainty of the emotion elicited (as subjective expression of subject might be misleading), and usually meant to ensure a rather uniform physiological response among a population. We can then consider the following problematic : Is the fact that individuals agree or disagree about the subjective emotion elicited by a stimulus (i.e. several individuals disagree about the pleasure/displeasure and/or the arousal elicited by a stimulus (fig. 1 (1) σ_Ψ) is an indicator of the uniformity of the measured physiological responses accross a population (fig. 1 (2) σ_Φ) ? We investigate the influence of social agreement of stimuli on physiological response similarities by the following hypothesis 1:

Hypothesis 1 *The fact that individuals have a social agreement about the subjective emotion elicited by a stimulus (σ_Ψ) is not an indicator of the uniformity of the measured physiological responses across a population (σ_Φ)*

For each stimulus, and for each subjects, a set of 28 physiological features was computed from heart rate and skin conductance(see [11]) for a description of the features).

To estimate the hypothesis 1, we first computed the standard deviation for each stimulus of the set of psychological coordinates expressed by subjects and the standard deviation of the set of features associated to physiological responses measured on each subject. Figure 2 plots the physiological

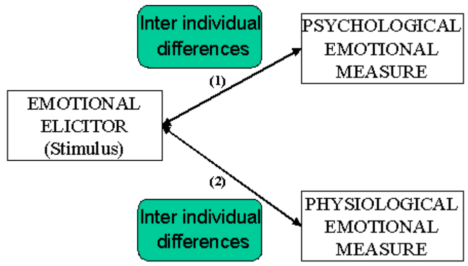

Figure 1. Comparison of inter-individual differences in the psychological and physiological evaluation of stimuli.

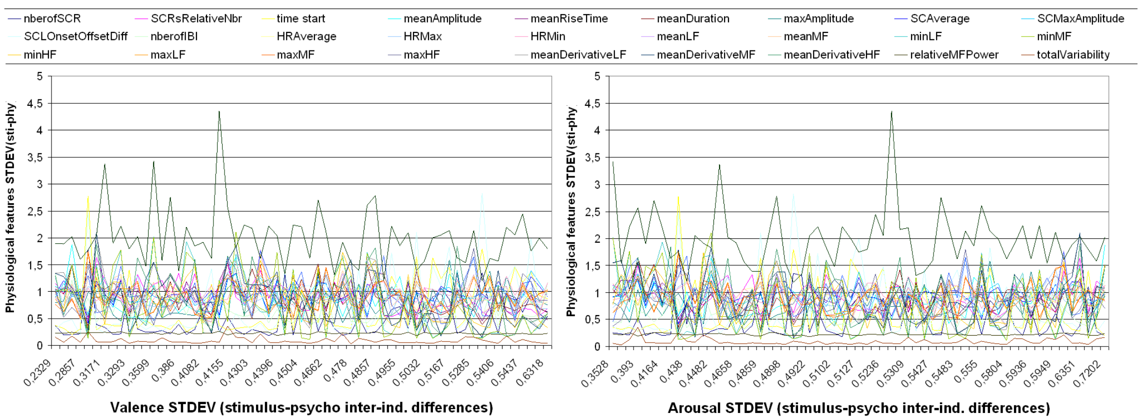

Figure 2. Relation between psychological and physiological standard deviations.

inter-individual differences as a function of the psychological inter-individual differences (valence on left part and arousal on right one). The x axis values correspond to standard deviation of the valence (left), or arousal (right) computed on the 40 subjects, for each of the 61 stimuli rated. On the y axis are the values of standard deviation of each physiological features values for the associated stimuli. Because no correlation was found to be significant between thoses variables, the hypothesis 1 is confirmed.

This means that adopting a social agreement approach for the stimuli *does not guarantee* a more uniform physiological response across the population nor a more robust emotion recognition *outside the context of the experiment*.

3.3 Results: Empirically estimating the effect of the Social Agreement versus Subjective Rating approaches on emotion recognition

To analyse the modeling posibilities of psychophysiological representations, we first averaged psychological and physiological measures of affective states elicited by the stimuli. Then, we selected psychophysiological representations of stimuli with a psychological index of dispersion around the mean of different amounts. For each amount, we selected psychophysiological representations as belonging to the data series on which we test correlation if : $2\sigma_i <$ amount $* (x_{max} - x_{min})$, with $x_{max} - x_{min} = 2$ in the valence arousal space made of $[-1, 1]$. In previous studies (see section 2) based on social agreement, dispersion is necessarily low because pilot studies select stimuli with high level of *psychological* agreement among subject. We propose the following hypothesis, related to the possibility of using the results of the social agreement stimuli in other contexts :

Hypothesis 2 *The level of agreement of individuals about the subjective emotion elicited by a stimulus (* σ_Ψ *) is related to the possibilities of modeling using a user-independent approach*

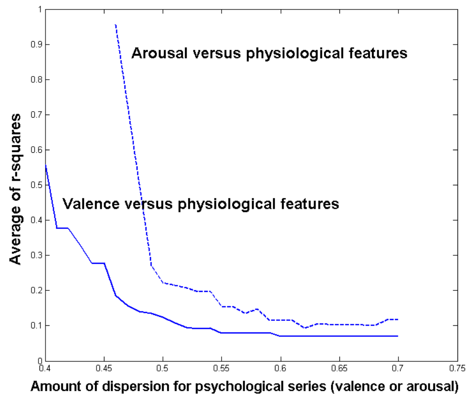

Figure 3. Effect of Psychological evaluation dispersion on significant correlation averages between

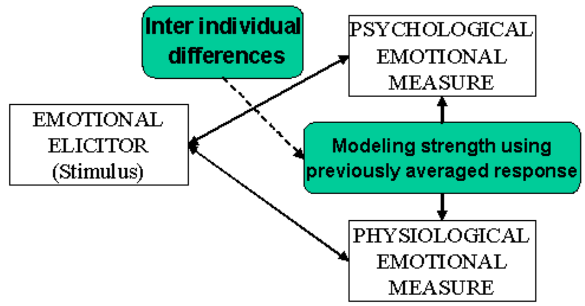

Figure 4. Effect of psychological dispersion on psychophysiological modeling strength.

Our result (figure 3)shows that the average of linear statistical test (r-square) is related to the dispersion (psychophysiological correlation for the population approach to 1 for stimuli with low dispersion, i.e. strong agreement). This means that for a low agreement (i.e. stimuli which elicit different subjective experiences for different subjects) the nature of the data will lead to a lesser performance in emotion recognition by using an averaged and normative approach. Thus, selecting subsets of psychophysiological representations associated to stimuli according to the agreement (inter-individual psychological differences) has effect on the modeling possibilities (see figure 4). The hypothesis 2 is thus confirmed.

4 From In-the-lab to out-of-the lab : Methodological Questions for Inferring Subjective Emotion from Physiological Signal

The field of emotion recognition from physiological signals has been mainly investigated in what we can call *in-the- lab* scenario. We consider in-the-lab the approaches which apply constraints on the subject (even if realized outside of the lab, as experimental multiple day data collection)In an experiment, we provide 'validated' stimuli (i.e. with a social agreement), or ask subjects regarding their subjective emotion regarding the set of stimuli. Then a division in learning and test sets is performed (random or stratified holdouts, or LOOCV) to evaluate the methodology and classifier in accuracy of recognition. We consider the problem of connecting user emotion estimation from physiology to real-life scenarii as a set of constraints regarding user dependency and subjectivity of stimuli, as shown in table (2). The left part of the table corresponds to the "in-the-lab" approach which goal is to demonstrate the possibility to predict emotion from physiology. At the opposite, the right part presents the "out-of-the-lab" approach focused on the personalization of the prediction to the user. Each line is related to (1) the amount of learning stimuli, (2) the psychological agreement of stimuli, (3) the time to train the system and (4) the possibility of repetition of stimuli.

5 Discussion

We pointed out main difficulties in emotion prediction from physiological signals regarding differences between individuals. We first analyzed different approaches by taking into consideration the user dependency criterion (user dependent versus user independent data) and the Subjectivity of stimuli criterion (social agreement versus subjective rating). Then, empirically showed that (1) physiological inter-individual differences are not related to psychological inter-individual differences

In the lab *: Demonstrating possibility of prediction*	**Out of the lab** *: tune and personalize the predictive system to users*
Large set Stimuli	None (pretrained classifier) or small set of learning stimuli(related to time spent on tuning the system before re-using it)
Psychological evaluation of stimuli based either on Subjective Rating or Social Agreement	Psychological evaluation of stimuli should allow Subjective Rating (see results section)
Long time to generate cases and to train the system	Short time (can't ask user to tune completely again the system each time (s)he use it.)
Could provide a standard set of stimuli as one-time experiment	Could not ask user to rate every period (e.g. week) the same set of stimuli to tune the system (habituation).

Table 2. Sets of constraints to apply emotion recognition from in-the-lab to out-of-the-lab

and that (2) choice of stimuli based on psychological agreement does not involve the fact results are generalizable for subjective stimuli. Finally, we considered such issues along with the requirements for an out of the lab use of emotional prediction. We shown the need to focus on the user dependency and subjectivity of stimuli notions to enhance the affective computing approaches on psychophysiology. The PsychoPhysiological Emotion Map (PPEM) model in its parametric form PPEM'i [10], may fulfill out-of-the-lab requirements by combining user-dependency and subjectivity of stimuli, and by modeling individual's physiological responses associated with stimuli with high inter-individual differences.

References

[1] J. Anttonen and V. Surakka. Emotions and heart rate while sitting on a chair. In *CHI 2005, Portland, Oregon, USA*, pages 491–499, 2005.

[2] L. Changchun, P. Rani, and N. Sarkar. An empirical study of machine learning techniques for affect recognition in human-robot interaction. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2-6 Aug.*, pages 2662– 2667, 2005.

[3] A. Haag, S. Goronzy, P. Schaich, and J. Williams. Emotion recognition using bio-sensors: First steps towards an automatic system. *LNCS*, 3068:36–48, January 2004.

[4] K. H. Kim, S. W. Bang, and S. R. Kim. Emotion recognition system using short-term monitoring of physiological signals. *Medical & Biological Engineering & Computing*, 42, 2004.

[5] C. Lisetti and F. Nasoz. Using noninvasive wearable computers to recognize human emotions from physiological signals. *EURASIP Journal on Applied Signal Processing*, 11:1672–1687, Sept 2004.

[6] C. Lerouge and C. Lisetti. Triangulating Design Science, Behavioral Science, and Practice for Technological Advancement in Tele-home Health. *International Journal of Health Care Technology and Management*, 7:pp-pp, 2006.

[7] C. Lisetti and F. Nasoz, C. Lerouge, O. Ozyer, and K. Alvarez. Developing Multimodal Intelligent Affective Interfaces for Tele-Home Health Care. *International Journal of Human-Computer Studies*, 59:245-255, 2003.

[8] R. Picard, J. Healey, and E. Vyzas. Toward machine emotional intelligence, analysis of affective physiological signals. *IEEE transactions on pattern analysis and machine intelligence*, 23(10), Oct 2001.

[9] O. Villon. Conscious dynamic of affective state and embodied relationship to music. Master's thesis, Université de Bordeaux 2, Bordeaux, June 2002. Supervisor : Antonio Camurri.

[10] O. Villon and C. Lisetti. A user-modeling approach to build user's psycho-physiological maps of emotions using bio-sensors. In *IEEE RO-MAN 2006, The 15th IEEE International Symposium on Robot and Human Interactive Communication, pages 269–276,* Hatfield.

[11] O. Villon and C. Lisetti. A user model of psycho-physiological measure of emotion. In *UM2007 User Modeling: Proceedings of the Eleventh International Conference*, 2007. to appear.

[12] J. Wagner, Jonghwa Kim, and E. Andre. From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification. In *Multimedia and Expo, ICME 2005. IEEE International Conference on*, pages 940– 943, 2005.