Person-specific behavioural features for automatic stress detection

Jonathan Aigrain, Séverine Dubuisson, Marcin Detyniecki, Mohamed Chetouani

▶ **To cite this version:**

HAL Id: hal-01364098

https://hal.sorbonne-universite.fr/hal-01364098

Submitted on 12 Sep 2016

Person-specific behavioural features for automatic stress detection

Jonathan Aigrain[1], Séverine Dubuisson[1], Marcin Detyniecki[2,3] and Mohamed Chetouani[1]

[1]Sorbonne Universités, UPMC Univ Paris 06, UMR 7222, ISIR, Paris, France

[2] CNRS, UMR 7606, LIP6, Paris, France

[3]Polish Academy of Sciences, IBS, Warsaw, Poland

Abstract— This paper introduces behavioural features for automatic stress detection, and a person-specific normalization to enhance the performance of our system. The presented features are all visual cues automatically extracted using video processing and depth data. In order to collect the necessary data, we conducted a lab study for stress elicitation using a time constrained arithmetic mental test. Then, we propose a set of body language features for stress detection. Experimental results using a SVM show that our model can detect stress with high accuracy (77%). Moreover, person specific normalization significantly improves classification results (from 67% to 77%). Also, the performance of each of the presented features is discussed.

I. INTRODUCTION

One of the key elements in human-human interaction is to understand the nonverbal cues and the social signals of the person we are communicating with. It has been shown that nonverbal communication conveys an important part of the meaning of a message [8], [9], [32]. Designing machine sensing and understanding social signals is a difficult task, since the meaning of these signals is person and context dependent. In this paper, we focus on the specific context of stressful situations, and we aim at automatically detect the presence of stress by studying the body language.

Automatic detection of stress is an emerging domain. It faces issues related to the difficulty to obtain a natural database and to ethical problems. Current detection techniques are based mainly on speech signals [11], [23] and/or from physiological signals, using either traditional sensors (ECG, skin conductance, ...) [1], [12], [17], [29] or hyperspectral images [5].

In [11] and [17], stress is detected though voice and physiological signals during a real-world driving-task. However, in this context, voice and physiological signals are not necessarily the most suited sources in order to detect stress. Someone who is driving does not always talk enough, and physiological sensors are obstructive. In this case, cues extracted from the body language are likely to be the most convenient sources of information to assess stress since they can be captured by camera sensors. To our knowledge, there are few papers adressing stress detection using body language. Giakoumis et al. [13] showed that behavioural features enhanced the performance of automatic stress detection systems using physiological features. Soury [30] used postural features in a multimodal fusion model, but in this work, this modality turned out to be the least efficient.

However, the correlation between behavioural cues and affective states has been extensively studied, mainly for automatic emotion recognition [15], [18], [20] and expression synthesis [4], [18], [20], as well as in the psychology [8], [9], [32]. However, though most of the affective information studied can be used to effectively detect a large number of emotions, there likely are some visual cues that are specific to the expression of stress.

In this paper, we present the first results of a preliminary study on stress detection by extracting and analyzing human body features. This paper is organized as follows: Section 2 reports the stress concept by reviewing the different approaches in the psychology literature. Section 3 presents our lab study for stress elicitation, in particular the protocol, the setup and the annotation of the subjects' stress level. Section 4 describes our feature extraction framework. Section 5 presents the person-specific normalization and its motivation. Section 6 outlines and discusses the results obtained by our learning approach based on SVM. We present the impact of kernel functions, person-specific normalization and features used. Finally, Section 7 concludes by discussing future research directions.

II. THE STRESS CONCEPT

The term *stress* appeared in the medical literature in 1950s [25]. Since then, it has been widely used [21], [25]. Holmes and Rahe [19], when developing a tool called the Social Readjustment Rating Scale (SRRS), described stress as a set of events that require a fixed amount of adjustement. In this case, stress is used as the stimulus that triggers the adaptation, but it can also be used to describe the response to a stimulus. Hans Selye viewed stress as the *"nonspecific response of the body to noxious stimuli"* [28]. In his work, he described a *noxious stimulus* as a stimulus that threatens homeostasis, the body "preferred" set-point for physiological variables.

One of the main problems in both Holmes' and Selye's conceptions of stress is the absence of impact from individual perception of the stimulus. Ursin et al. [31] later demonstrated that coping behaviour reduced physiological activity. Thus, we think that transactional models of stress such as those presented in [22], [24] fits better our intuitive vision of stress. These models present stress as a process that includes the stimulus, the appraisal of the situation and the physiological, behavioural and psychological reactions.

In addition, Koolhaas et al. [21] stated that the stimulus had to be uncontrollable and/or unpredictable in order to define this process as *stress*.

The protocol of our stress-induction experiment is mainly based on the transactional definition of stress.

III. DATA COLLECTION

In this section, we present how we collected the data through a stress-induction experiment.

A. Stress induction procedure

According to Dickerson and Kemeny[7], the best way to increase a subject's cortisol level, which is largely considered as being correlated to psychological stress [6], [7], [25], [30] is to make her perform a task that combines public speaking and cognitive challenges. Indeed, tasks that contain uncontrollable traits and/or social evaluation are the most stressful ones. Thus, we designed an evaluated time-constrained mental arithmetic test.
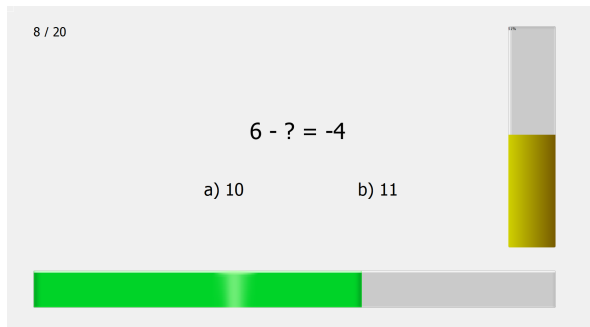

Fig. 1. Screenshot of the test software used for the study

Figure 1 shows a screenshot of the test. The question asked is shown in the middle of the screen. There are two possible answers below it. At the bottom, the progress bar represents the time limit for the current question. On the right is the score bar, which gives the subject a feedback of her performance. The color of the score bar has a meaning: green means "above average", yellow means "average" and red means "below average".

The test is composed of 6 steps of increasing difficulty. The three first steps contain 20 questions, the others contain 30 questions. The timing of each step is about 1 minute and there is a break period of 5 seconds between them. The subjects are told that both quickness and correctness of their answers are taken into account for their score. It gives the subject the impression of being evaluated, but in reality the values the score bar shows are set in advance. It displays an "above average" score at the beginning. Thus, the participant finds the test easy enough and feels like she should suceed. Then, the score drops to "average" and "below average" levels, giving the participant the impression she is failing the test.

B. Stress level annotation

At the end of the test, a questionnaire including stress self-assessment is conducted for each step of the test, using a Likert-scale (1-5). Since the quality of the annotation is greatly dependent of subjects' memory, they watch their own videos of the different steps before giving a stress level.

Self-assessment has been chosen because it is the most faithful annotation to the used definition of stress, which gives great importance to the personal perception. In addition, annotation based on judgement of human experts can induce a bias in the performance of some features, and there were some cases where annotators did not achieve a good agreement score [30]. Regarding the use of hormone levels for annotation, [33] has shown that the difference in dynamics between behaviour and hormone levels makes it complex, but it will be investigated in future works.

TABLE I
STRESS LEVEL DISTRIBUTION

Stress level	1	2	3	4	5
Number of videos	17	20	20	17	10

C. Experimental setup

Participants are standing at a distance of around 3 meters from the screen where the test is displayed. For each question, they have to clearly express to a staff member which answer they choose among the two propositions. Video and skeleton data were collected with a Microsoft Kinect placed under the screen. A high definition camera located above the screen provides an optic zoom of the particpant's face.

D. Participants

Fourteen subjects participated in the experiment (3 women, 11 men, mean age = 24.8 ± 2.8 years old). Each subject is aware of the purpose of the experiment, but do not know about the biased behaviour of the score bar, and gives written consent.

E. Acquired data

For each of the 14 partipants, for each of the 6 steps, the acquired data are:

- video of the whole body in 640×480 from the Kinect
- skeleton from the Kinect
- video of the face in 1440×1080 from the HD camera
- self-assessed stress level

The average duration of a video is 57.8 seconds ± 12.6.

IV. FEATURE EXTRACTION

In this section, extracted features and procedures to extract them are described. First, the features extracted from the Kinect video are presented, followed by those extracted from the video of the face. All features are computed over the whole video for each step of the protocol. They are summarized in Table II.

A. Body activity features

There is few research of how stress can affect our body language. However, in order to recognize or to regenerate someone's affective state, there are several sets of features extracted from the body that are usually used: the body activity [4], [13], [15], [20], posture information such as symmetry [13], [15], center of gravity displacements [14], or the spatial extent [4], [15], kinematics information such as smoothness [4], [15] and detection of specific gestures [13], [15].

Using the skeleton and the video captured by the Kinect, 3 categories of features are extracted: the Quantity of Movement (QoM) and its derivatives, the detection of periods of high activity and posture changes and the detection of self-touching in the region of the head.

1) Quantity of Movement: The QoM is the activity of the body or of some parts of it. For a given frame, we compute the QoM in two different ways. The first one (SQoM) uses the skeleton extracted by the Kinect and is the sum of the displacements of each joint between two frames.

$$QoM_{skel}(i) = \sum_{j \in joints} \sqrt{(v_{j_i} - v_{j_{i-1}})^2}$$

with i the frame index and v_{j_i} the 3D position vector of the joint j in the ith frame. We also compute the QoM using RGB videos, and call it IQoM. This corresponds to the number of pixels that have changed between successive frames.

$$QoM_{video}(i) = Card(\{p_i | abs(p_i - p_{i-1}) > t\})$$

with p_i the RGB vector of pixel p for the ith frame and t a threshold. Once the QoM has been computed on all frames, the mean value is used as the feature. In order not to be biaised by the size of a person, it is also divided by the surface of the bounding box containing the skeleton.

The SQoM is computed separately for the whole body and the head (HeM). For the head only, we also compute the SQoM following Z-axis only (HeMZ). For the body only, the Fourier transform of the signal is calculated, divided into 10 bins (FFT1 - FFT10), which gives frequency information about the subject's activity.

2) Detection of high activity periods and posture changes: Using the peaks of the QoM computed for each frame, the periods of high activity are extracted (see Fig. 2) . The number of periods (HAPC), their mean duration (HAPMD) and their mean intensity (HAPMV) are used as features. For each peak, the frame at the beginning and the one at the end are compared to determine whether the person has changed her posture. The number of posture changes (PCC) is used as a feature. We make the hypothesis that posture changes may reflect an increasing uncomfortability.

3) Detection of self-touching: It has been shown that self-touching can be an indicator of negative affect [16]. We detect self-touching in two parts of the body: the head and the hands. To detect whether a person is self-touching in the region of the head, we compute the distance between hand, head and neck joints. If one of the distances is below a
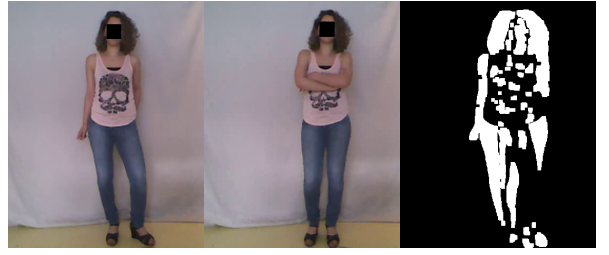

Fig. 2. Example of detection of a posture change. The left image is the frame at the start of the peak, the middle image is the frame at the end, and the right one is the difference between the two others

threshold, the person is considered as self-touching the head (see Fig. 3). The number of times we detect a single hand as self-touching the head (STHC) and the mean duration (STHMD) of it are used as features. We extract the sames features in cases of two hands face-touching (ST2HC and ST2HMD).

Fig. 3. Examples of detection of self-touching in the region of the head. On the left is an example of self-touching with one hand. On the right is an example with two hands

The detection of hand self-touching is more complex. Indeed, the skeleton given by the kinect cannot be used alone in order to detect movements such as fingers rubbing. Thus, we compute the extraction of this feature as follows:

- Detection of the hands: because the skeleton provided by the Kinect can be imprecise in terms of joint location, we also use skin detection to find a more reliable position of the hands. Thus, starting from the original hand joint position, we look for the closest pixel where skin has been detected. This becomes the new hand position.
- Computation of the QoM: using the position we detected, we extract the sub-image of the hands. Then, we compute the video QoM between following hands frames.

The QoM for each hand separately (LHM for the left hand and RHM for the right one) and the QoM of both hands (HM) are used as features.

B. Facial features

Using the method described in [27], we extract the activation level of 12 Action Units (AU). Action Units are presented by Ekman in the Facial Action Coding System (FACS) [10], which is a common standard to systematically categorize the physical expression of emotions. The 12 AU extracted are described in Table II. Since the activation level

is given by frame, we compute the mean over all frames and the standard deviation for each AU as features.

TABLE II

EXTRACTED FEATURES

Feature	Description
SQoM	QoM computed with the skeleton
IQoM	QoM computed with the RGB frames
FFT*i*	*i*-th bin of the Fourier transform of the IQoM (from 1 to 10)
HAPC	Number of periods of high activity
HAPMD	Mean duration of periods of high activity
HAPMV	Mean highest value of periods of high activity
PCC	Number of posture changes
STHC	Number of times self-touching with one hand in the region of the head occured
STHMD	Mean duration of self-touching with one hand in the region of the head
ST2HC	Number of times self-touching with two hands in the region of the head occured
ST2HMD	Mean duration of self-touching with two hands in the region of the head
LHM	QoM for the left hand
RHM	QoM for the right hand
HM	QoM for both hands
HeM	QoM for the head
HeMZ	QoM for the head only along Z-axis
AU1	Inner Brow Raiser
AU2	Outer Brow Raiser
AU4	Brow Lowerer
AU5	Upper Lid Raiser
AU6	Cheek Raiser
AU9	Nose Wrinkler
AU12	Lip Corner Puller
AU15	Lip Corner Depressor
AU17	Chin Raiser
AU20	Lip Stretcher
AU25	Lips Part
AU26	Jaw Drop

V. PERSON-SPECIFIC NORMALIZATION

In addition to testing the features, the impact of a person-specific normalization is evaluated. As said in [29] regarding physiological signals or in [2] regarding body motions, the evolution of the signal compared to a personal baseline is likely to be more meaningful than its absolute value. Thus, for each person, we compare the value of each feature to its value during the first step of the experiment, considered as non-stressful.

$$\tilde{f_{pj}} = \frac{f_{pj} - f_{p1}}{f_{p1}}$$

where p a person, j the step the video has been extracted from, f_{pj} the vector of features for the person p on step j and $\tilde{f_{pj}}$ the normalized vector. Thus, the features are expressed as an evolution percentage rather than a raw value.

VI. STRESS DETECTION

This section presents the results obtained by our model. First, the evaluation process is described. Then, the impact of several parameters - the SVM kernel function, the person-specific normalization and the different features - are discussed.

A. Evaluation process

In order to test whether our features are predictive or not, we use Support Vector Machines (SVMs) to build a classification model. Each video is associated with a label: stress (S) for video associated with a stress level above or equal to 4, non-stress (NS) for the remaining ones. For each evaluation, we use a leave-one-subject-out cross validation: videos from all people but one person are used as the training set, and the videos of the remaining person is used as the testing set. This cross validation is repeated 10 times in order to reduce the impact of randomness in the SVM margin initialization. The mean accuracy and its standard deviation are presented.

We discard videos that are labelled as non-stress if they followed a stress labelled video. These videos represent people acknowledging that they psychologically gave up because of the difficulty. Thus, the accuracy is computed on 76 out of the 84 videos for raw features, 49 labelled as non-stress and 27 labelled as stress. Since the person-specific normalization uses the features extracted from the first step, the accuracy is computed on 62 videos for normalized features, 35 labelled as non-stress and 27 labelled as stress.

B. Results

TABLE III

AVERAGE AND STANDARD DEVIATION OF THE ACCURACY OBTAINED WITH RAW AND PERSONNALIZED FEATURES AND WITH THREE KERNEL FUNCTIONS.

kernel type	raw	normalized
Poly 5	0.64 ± 0.04	0.77 ± 0.02
RBF	0.65 ± 0.03	0.76 ± 0.02
Linear	0.67 ± 0.01	0.77 ± 0.01

TABLE IV

IMPACT OF SEVERAL SET OF FEATURES WITH AND WITHOUT NORMALIZATION ON ACCURACY WITH A LINEAR SVM

features set	raw	normalized
All	0.67 ± 0.01	0.77 ± 0.01
Face	0.68 ± 0.01	0.65 ± 0.03
Body	0.63 ± 0.03	0.80 ± 0.01

1) Impact of the kernel function: Table III shows the results obtained by 3 different kernel functions: the Radial Basis Function (RBF), the polynomial of degree 5 (Poly 5) and the linear kernel. There is no significant difference between the 3 kernels concerning the mean accuracy. Since the linear kernel is known to be faster to train, we have chosen to keep it for the following experiments.

2) Impact of the person-specific normalization: It clearly appears that using the person-specific normalization significantly improves classification results, by 13%, 11% and 10% for the polynomial, RBF and linear kernels, respectively as shown in Table III. However, as we can see with Table IV, the normalization is not effective on facial features. This might

be explained by the much smaller range of values the facial features can take compared to body features. Indeed, facial features values vary from 0 to 5, and there is little difference in scale among the participants. In comparaison, the IQoM can be 30 times bigger from a subject to another.

	classified as Stress	classified as Non-Stress
Stress	19.8 ± 0.42	7.2 ± 0.42
Non-Stress	7.1 ± 0.31	27.9 ± 0.31

	classified as Stress	classified as Non-Stress
Stress	14 ± 0	13 ± 0
Non-Stress	12 ± 0.66	37 ± 0.66

3) Performances of facial versus body features: Table IV shows the impact of the choice of features on the classification accuracy. When using raw features, facial features give better results than body features. This is probably due to smaller inter-individual differences with facial features. However, information on the evolution of body features are much more discriminative than information on the evolution of facial expressions. This confirms the idea that body language can bring valuable information to determine someone's emotion state.

feature	accuracy
PCC	0.73 ± 0.02
AU9 - std	0.72 ± 0.01
FFT2	0.72 ± 0.01
AU4 - std	0.72 ± 0.01
AU4 - mean	0.72 ± 0.01
AU25 - mean	0.60 ± 0.03
RHM	0.58 ± 0.01
AU5 - std	0.58 ± 0.02
AU9 - mean	0.56 ± 0.01
HAPMD	0.55 ± 0.04

4) Performances of individual features: In order to better understand the relevance of each feature, we calculated their classification accuracy if used alone. We use the RBF kernel to compute these results, since a linear kernel would allow only a single "split value" along the feature axis. The results for raw features are presented in Table VII, those for person-specific features in Table VIII. Concerning raw features, we can see that several features achieve better average classification accuracy when used alone than when combined with others. FFT1 for person-specific features also achieve alone a similar score to what we obtain with all normalized

feature	accuracy
FFT1	0.76 ± 0.02
HAPC	0.74 ± 0.01
FFT7	0.73 ± 0.02
HAPMV	0.73 ± 0.02
PCC	0.73 ± 0.02
AU1 - std	0.49 ± 0.03
AU25 - mean	0.46 ± 0.02
AU26 - mean	0.45 ± 0.02
AU15 - mean	0.45 ± 0.03
AU17 - std	0.43 ± 0.04

features. Thus, we can see that a single feature can be enough to detect stress with good accuracy.

Among these features, it appears that brows activity as raw features is a valuable information for our study. Indeed, AU4 and AU9 are correlated since AU4 is often activated when AU9 is. The number of times a subject changes its posture also achieve good classification results, both for raw and person-specific features. Features on body activity level such as FFT1 or information on periods of high ativity also achieve good accuracy as person-specific features. It is however difficult to interpret what FFT7 really means in an intuitive way.
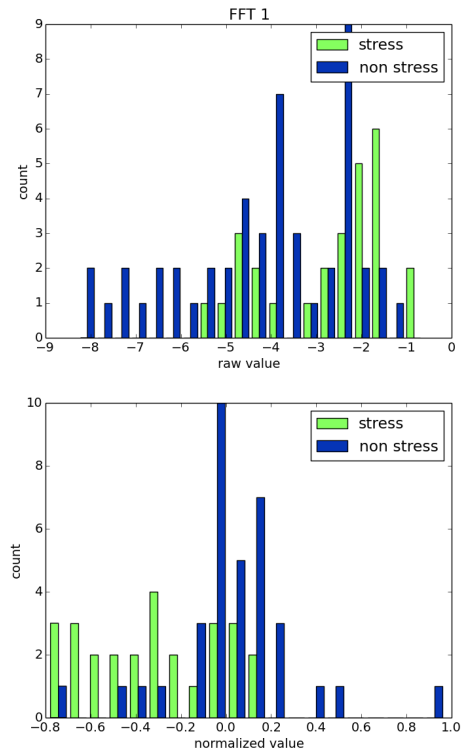

Fig. 4.   Distribution of FFT 1 raw and normalized values

Finally, we can see that person-specific normalization works well for body features, but not for facial ones. Figure 4

shows how normalization on FFT 1 improved the seperation between the two classes. The 5 best person-specific features are body features, and the 5 worst are all facial features. This observation can not be made for raw features, where body and facial information are both represented in the best and worst features.

VII. CONCLUSIONS AND FUTURE WORKS

In this paper, we have presented behavioural features for automatic stress detection, and a person-specific normalization to enhance the performance of our system. We have discussed the stress concept and presented a set of 49 automatically extracted behavioural features. We also observed that a person-specific normalization significantly improved the performance of our model, and that, for our study, the body language of the subjects provided more valuable information than their facial expressions. Finally, we saw that it is possible to have an unobtrusive stress detection system using only video and depth data.

However, the results obtained can still be improved. The results presented in Table VII and VIII clearly show that some features do not bring valuable information and thus should not be included training the SVM. Preliminary tests on feature selection showed promising results for classification accuracy. Moreover, we saw that person-specific normalization is effective only for body features. Trying to include both raw facial and person-specific body features in our model may give interesting results.

Finally, more subjects will participate in our experiment. However, they will not be aware of the purpose of the study. Thus, we will be able to observe the impact of individual perception of a situation on stress level.

REFERENCES

[1] A. Barreto, J. Zhai, and M. Adjouadi. Non-intrusive Physiological Monitoring for Automated Stress Detection in Human-Computer Interaction. In *Human-Computer Interaction*, pages 29–38, 2007.

[2] D. Bernhardt and P. Robinson. Detecting affect from non-stylised body motions. In *Affective Computing and Intelligent Interaction*, pages 59–70, 2007.

[3] N. Bianchi-Berthouze, P. Cairns, A. Cox, C. Jennett, and W. Kim. On Posture as a Modality for Expressing and Recognizing Emotions. In *Emotion in HCI workshop at BCS HCI*, pages 74–80, 2006.

[4] G. Caridakis, A. Raouzaiou, K. Karpouzis, and S. Kollias. Synthesizing Gesture Expressivity Based on Real Sequences. *"Multimodal Corpora. From Multimodal Behaviour Theories to Usable Models" In : International Conference on Language Resources and Evaluation*, pages 19–23, 2006.

[5] T. Chen, P. Yuen, M. Richardson, G. Liu, Z. She, and S. Member. Detection of Psychological Stress Using a Hyperspectral Imaging Technique. *IEEE Transactions on Affective Computing*, 5(4):391–405, 2014.

[6] S. Cohen, D. Janicki-deverts, and G. E. Miller. Psychological Stress and Disease. *Journal of American Medical Association*, 298(14):1685–1687, 2007.

[7] S. S. Dickerson and M. E. Kemeny. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological bulletin*, 130(3):355–91, May 2004.

[8] P. Ekman and W. V. Friesen. Hand Movements. *The journal of communication*, 22(4):353–374, 1973.

[9] P. Ekman and W. V. Friesen. Detecting Deception From The Body Or Face. *Journal of personality and Social Psychology*, 29(3):288–298, 1974.

[10] P. Ekman and W. V. Friesen. Facial action coding system. 1977.

[11] R. Fernandez and R. W. Picard. Modeling drivers' speech under stress. *Speech communication*, 40(1):145–159, 2003.

[12] A. Gaggioli, G. Pioggia, G. Tartarisco, G. Baldus, M. Ferro, P. Cipresso, S. Serino, A. Popleteev, S. Gabrielli, R. Maimone, and G. Riva. A system for automatic detection of momentary stress in naturalistic settings. *Studies in health technology and informatics*, 181:182–6, Jan. 2012.

[13] D. Giakoumis, A. Drosou, P. Cipresso, D. Tzovaras, G. Hassapis, A. Gaggioli, and G. Riva. Using activity-related behavioural features towards more effective automatic stress detection. *PloS one*, 7(9):e43571, Jan. 2012.

[14] T. Giraud, D. A. G. Jáuregui, J. Hua, B. Isableu, E. Filaire, C. L. Scanff, and J. C. Martin. Assessing postural control for affect recognition using video and force plates. In *Humaine Association Conference on Affective Computing and Intelligent Interaction*, pages 109–115, 2013.

[15] D. Glowinski, N. Dael, a. Camurri, G. Volpe, M. Mortillaro, and K. Scherer. Toward a Minimal Representation of Affective Gestures. *IEEE Transactions on Affective Computing*, 2(2):106–118, Apr. 2011.

[16] J. Harrigan. Self-touching as an indicator of underlying affect and language processes. *Social Science and medicine*, 20(11):1161–1168, 1985.

[17] J. Healey and R. Picard. Detecting Stress During Real-World Driving Tasks Using Physiological Sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6(2):156–166, June 2005.

[18] J. Hoey, D. Kuli, M. Karg, A.-a. Samadani, R. Gorbet, and K. Kolja. Body Movements for Affective Expression : A Survey of Automatic Recognition and Generation. *IEEE Transactions on Affective Computing*, 4(4):341–359, 2013.

[19] T. H. Holmes and R. H. Rahe. The social readjustment rating scale. *Journal of psychosomatic research*, 11(5):213–218, 1967.

[20] A. Kleinsmith and N. Bianchi-Berthouze. Affective Body Expression Perception and Recognition: A Survey. *IEEE Transactions on Affective Computing*, 4(1):15–33, Jan. 2013.

[21] J. M. Koolhaas, a. Bartolomucci, B. Buwalda, S. F. de Boer, G. Flügge, S. M. Korte, P. Meerlo, R. Murison, B. Olivier, P. Palanza, G. Richter-Levin, a. Sgoifo, T. Steimer, O. Stiedl, G. van Dijk, M. Wöhr, and E. Fuchs. Stress revisited: a critical evaluation of the stress concept. *Neuroscience and biobehavioral reviews*, 35(5):1291–301, Apr. 2011.

[22] R. S. Lazarus. From psychological stress to the emotions: a history of changing outlooks. *Annual review of psychology*, 44:1–21, Jan. 1993.

[23] I. Lefter, L. J. Rothkrantz, D. A. van Leeuwen, and P. Wiggers. automatic stress dectection in emergency calls. *International Journal of Intelligent Defence Support Systems*, 4(2):148–168, 2011.

[24] S. Levine. Developmental determinants of sensitivity and resistance to stress. *Psychoneuroendocrinology*, 30(10):939–46, Nov. 2005.

[25] B. L. Lyon. Stress, Coping, and Health. In *Handbook of Stress, Coping, and Health*, pages 2–20. 2012.

[26] P. Mundy, M. Sigman, J. Ungerer, and T. Sherman. Defining the social deficits of autism: The contribution of non-verbal communication measures. *Journal of child psychology and psychiatry*, 27(5):657–669, 1986.

[27] J. Nicolle, K. Bailly, and M. Chetouani. Facial Action Unit Intensity Predicition via Hard Multi-task Metric Learning for Kernel Regression. *"Facial Expression Recognition and Analysis Challenge" In : IEEE International Conference on Automatic Face and Festure Recognition*, 2015.

[28] H. Selye. *The stress of life*. McGraw-Hill, New York, NY, US, 1956.

[29] Y. Shi, M. H. Nguyen, P. Blitz, B. French, S. Fisk, F. D. Torre, A. Smailagic, D. P. Siewiorek, M. Absi, E. Ertin, T. Kamarck, and S. Kumar. Personalized Stress Detection from Physiological Measurements. In *International Symposium on Quality of Life Technology*, 2010.

[30] M. Soury. *Détection multimodale du stress pour la conception de logiciels de remédiation*. PhD thesis, Université Paris-sud, 2014.

[31] H. Ursin, E. Baade, and S. Levine. *Psychobiology of stress - a study of coping men*. Academic Press, 1978.

[32] H. G. Wallbott. Bodily expression of emotion. *European journal of social psychology*, 28(1):879–896, 1998.

[33] O. Weisman, E. Delaherche, M. Rondeau, M. Chetouani, D. Cohen, and R. Feldman. Oxytocin shapes parental motion during father–infant interaction. *Biology letters*, 9(6):20130828, 2013.

[34] M. Zuckerman, B. M. DePaulo, and R. Rosenthal. Verbal and nonverbal communication of deception. *Advances in experimental social psychology*, 14:1–59, 1981.