

Sensor-less Sensing for Affective Computing and Stress Management Technology

Pablo Paredes

Electrical Engineering and Computer
Science Department
UC Berkeley
paredes@eecs.berkeley.edu

David Sun

Electrical Engineering and Computer
Science Department
UC Berkeley
davidsun@cs.berkeley.edu

John Canny

Electrical Engineering and Computer
Science Department
UC Berkeley
jfc@cs.berkeley.edu

Abstract—This paper describes our vision on what should be the research around sensing and adaptive interventions to make affective computing and stress management technology pervasive and unobtrusive. With the use of common computer peripherals and mobile computing devices as affect sensors, personalized and adaptive intervention technologies can be developed. Furthermore, physiological sensing can be performed without the introduction of extraneous factors such as wearable devices or focused software. Different methods for sensing and complementary adaptable interventions and interactions are described and proposed. We show initial lab evidence of the use of a computer mouse in the detection of stress.

unintrusive sensing, affective computing, pervasive mental health, physiological measurements, mental states, mobile devices, psychological stress, bio sensors, activity sensors

I. INTRODUCTION

This paper presents the body of knowledge around the adoption, usability and design of non-invasive sensing techniques for affective computing and stress management, due to its big impact in mental health and productivity. We define sensor-less sensing as the umbrella term covering the opportunistic use of existing sensors embedded in daily use computing devices and peripherals (mouse, keyboards, cameras, etc.) and mobile devices to repurpose their signals to track different biometric states representative of mental or physiological states directly or indirectly. Sensor-less sensing offers a great opportunity to detect physiological metrics such as heart rate, heart rate variability (HRV), breathing rate, etc. Another exciting opportunity is to indirectly detect mental state changes. Preliminary research performed in our lab on this topic has been able to detect robust signals in voice [3], and by measuring the natural oscillation of the arm through a computer mouse [22]. In a first approach sensor-less sensing can be used to monitor affective states associated with the use of computer interfaces (such as frustration, anger, happiness, stress, etc.) There are a number of advantages in leveraging daily use computing devices as “sensors”:

Accessibility - peripherals such as mice and keyboard are ubiquitous and are indispensable to interaction with desktops graphic user interfaces.

Unobtrusiveness - There is no need for wiring sensors to the body or speak to a microphone (or cellphone) especially in semi-structured office spaces.

Long-term, in-situ monitoring - Many people spend a substantial amount of time using computers, which affords the opportunity to monitor and provide feedback while people are engaged in stressful tasks.

Application and content neutral monitoring - Mice motion and smartphone gestures is neutral to the application and the content with which the user is interacting. This implies better generalizability and alleviates privacy concerns compared with other more intrusive techniques, like monitoring keystrokes or camera usage.

II. BACKGROUND

A. Arousal and Stress

Arousal and stress affect almost all the body functions, including cognitive function and memory. Stress/arousal is not a single system but several related systems working in harmony (allostatic body function) [1]. Stress has a concrete function to trigger a reaction to a specific threat (stressor) [2]. Arousal generally improves memory, speed of work, association and pattern recognition; it also increases errors on unfamiliar tasks and causes cognitive and perceptual narrowing. The effects of these changes on performance depend on the task.

This general form has been observed in studies on the effects of emotional stress on recall [4]. However, a large body of research has given conflicting results under different stress/performance combinations, often with linear relationships between arousal and performance, or with performance that saturates at high levels of arousal. It is best to assume that this relationship needs to be learned for each class of tasks. Similarly, it has been shown that the optimal level of arousal varies from one individual to the next. This all suggests that in order to make use of stress/arousal feedback, a system needs to gather a large amount of data about each individual, and furthermore this data needs to be identified with the task the subject is doing.

B. Physiological Affect Sensing

Recent advancement in sensors and wireless technologies and related computational techniques have accelerated the push

towards wearable health-care devices capable of providing ambulatory monitoring of a variety of vital signs, such as electrocardiogram (ECG), electromyogram (EMG), pulse oximetry, bioimpedance, galvanic skin response (GSR) [17].

Many of these vital signs are strongly linked with physiological changes induced by emotional arousal [19]. Healey and Picard used a combination of these physiological sensors to determine stress-levels for drivers in real-life driving tasks [8]. However, many have also reported that a variety of other conditions such as physical activity, attention or fatigue can introduce physiological responses very similar to stress/arousal [9].

Affect sensing using voice and facial features has also been explored. Specifically, voice analysis has demonstrated modest accuracy at estimating emotion and high accuracy of stress/arousal [3]. Computer vision has demonstrated good accuracy at emotion detection from facial images [13]. However, collecting sufficient data to train robust models for voice analysis is challenging, and in both cases deployment is further complicated because of privacy concerns [14].

In spite of their growing availability, friction continues to exist in public adoption due to the intrusive nature of body periphery sensing. By enabling daily use computing devices capable of converting some body signals into mental health metrics, new affective technology adoption could be improved.

III. PLANNED RESEARCH

A. Emotion and Arousal from Computer Peripherals

Mouse - A promising exploration of the use of mouse movements to detect affect was Wolfgang Maehr’s diploma thesis [15]. Maehr used several metrics on mouse movement, and emotions induced in subjects by watching short videos. Specific “motion breaks” that were discontinuities in mouse movement were significantly related to both arousal and disgust and close to significant for anger. One common correlate of arousal/stress is muscle tension. Tension in arm and wrist muscles would change the dynamics of the movement, e.g. its resonant frequency and damping ratio.

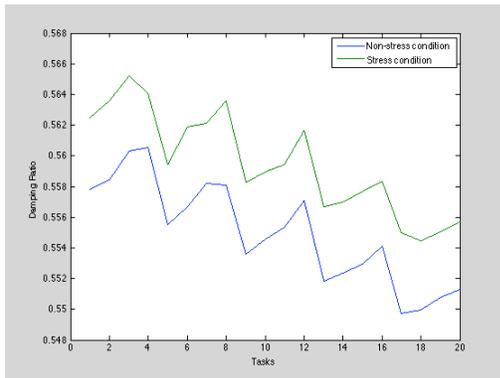


Figure 1. Mouse displacement damping ratio for 20 different motion tasks for mental Non-stress and Stress conditions [22]

In our work in progress (Sun, et. al. [22]) we have seen that it is possible to train simple controller models, such as Linear Predictive Coding (LPC), to capture the basic dynamic

parameters of the movement. From such a model, approximate muscle tension is easily computed. The exploration used “game grade” mice, which have update rates of 500 Hz and resolutions of several thousand dots per inch (DPI). We replicated the study analysis. Figure 1 shows the damping ratio aggregated over N=30 subjects performing 20 different motion tasks with 5 randomized repetitions each.

As it can be observed, the difference is larger than standard deviation of any condition. Under a t-test, we observe statistical significance ($p > 0.05$) for all the tasks. The accuracy averaged over all the 100 experiment tasks (20 tasks x 5 reps) was 90%. We plan to extend this work in a number of directions. We are examining the sensitivity of the model for normal mice with lower spatial resolution.

Keyboard - Affective inference from keyboard activity was recently described by Epp, et. al. [7]. The developers used timings from individual key and short key-sequence (bi- and tri-gram) features. Data were collected naturalistically, i.e. the system monitored subjects’ everyday computer use, and ESM (Experience Sampling) was used to gather self-assessments of emotional state. The system showed promising accuracy (70%-88%) for most emotion labels. While these results are encouraging, as the authors acknowledged, these raw classification rates for skewed categories masked a very modest gain over baseline classification (e.g. for excitement). Still the approach shows the power of this feature set.

We will explore the use of other classifiers, Support Vector Machines (SVM) and logistic models, and dimension reduction on the feature set (without emotion labels) using Latent Dirichlet Allocation (LDA) and Gamma-Poisson (GaP). Finally, keyboard and mouse models will be combined with a view to more robust classification.

Camera - The face is a rich source of affective information through facial expression. Systems such as AFA, based on the Facial-Action Coding System, have demonstrated greater than 90% accuracy on standard test datasets [12]. Research on facial expression recognition is quite mature and is now finding its way into commercial systems, for instance Affectiva’s FaceSense. Facial expression recognition is a complex computer vision task, and we do not plan to explore it. Instead there have been some exciting developments recently in measurement of vital signs data from facial images, which we wish to build on and extend. The first of these is pulse detection from webcam images of the face [20]. Blood flow through the vessels in the face causes tiny changes in the color of the face. This approach has been realized in MIT’s CardioCam, and in Philips “Vital Signs Camera” app for the iPad. The Philips app also measures breathing through small movements of the chest. Pulse rate and breathing rate are both a strong correlates of arousal. So is movement generally. So head and body movement captured from the tracker provide a further cue to arousal.

We will develop a camera-based arousal monitor as outlined above. We believe adaptive comb filters are a good compromise for isolating and enhancing the breathing signal. The core signals will be pulse from face color, and breathing from correlated body motion and face color. We will also explore the value of statistics on general body motion (velocity,

acceleration, distance) as cues to arousal level, and correlate these with other measures.

B. Emotion and Arousal from Smartphone Activity

Gestures - Following the methodology used in mice, we plan to extend this work to a three dimensional space to model common gestures used in smartphones for the detection of arousal and emotions. As described in the mouse section, we want to make sure gestures are not application or user specific and are commonly used, so that emotional detection can happen in the wild with common smartphone usage. By dynamically measuring accelerometer and gyrosopic data in three dimensions we want to be able to model the oscillations of the arm during normal smartphone usage, i.e. during normal tasks such as dialing, searching, browsing, etc. We expect that the already large level of hand holding of the device will help generate sufficient data without the need of additional tasks.

Voice - We expect to further enhance accuracy and complexity of emotions via a multi-modal approach. We plan to integrate our mechanical system with a voice sensing model based on Chang, K., et. al. toolkit, AMMON [3], which has already produced high accuracy emotion detection through the detection of voice changes both at an intonation level (via glottal timings) and at a prosodic changes (pauses, speech speed, etc.)

C. Adaptive Interactions and Interventions

Feedback and Interface Adaptation – The first and most basic interaction/intervention will be to present direct feedback of different emotional states to users. Direct feedback should raise awareness and drive behavior change. Another option is the automatic alteration of the interface to better adapt to the emotional state of the user. The interface could be adapted either at the background or foreground level [11] to help the user either maintain attention in the task at hand during a stressful, but productive state, or change and relax during a frustrating or overwhelming episode. Interface adaptation to human affect has been used in airplanes, using previous knowledge, self-reports, diagnostic tasks and physiological sensing and changing the interface at the content and format levels [10]. Work has also been done to adapt computing interfaces to cognitive and affective changes (the latter captured through changes in facial expressions) [6]; however little work has been done to unobtrusively sense affect and actively adapt computer interfaces to affective changes.

We plan to explore the use of sensor-less sensing to monitor different emotions commonly present during interactions with computing devices to help people use them more towards the accomplishment of their goals and to optimize the emotional engagement associated with it. In terms of mental health, the ability to maintain an “appropriate” level of stress, needed to fulfill the task will be beneficial especially in productive settings like the office or school.

Emotional Regulation and Psycho-education Interventions– Emotional regulation interventions such as calming technologies present a good opportunity to help deliver users that present initial (moderate) symptoms of stress or other emotional changes with a brief and effective intervention that would help people avoid unnecessary emotional alteration.

Paredes, et. al. [18] present an example of some mobile individual and social interventions that are brief and usable without the need of additional hardware and leveraging intrinsic and social aspects. Coyle, D. et. al. [5] has explored gaming interventions focused mental health based on Cognitive Behavioral Therapy (CBT) We expect to extend the work done with these interventions to further explore interventions based on social and gaming concepts. One specific technology already being developed in our lab is the use of machinima (machine + cinema) technology for the creation of short movies that can deliver some life skill or CBT concept teaching that could be beneficial to regulate emotional change or stress. Figure 2 shows snapshots of a simple movie created on the Sims 3 platform that teaches element of social interaction based on a CBT manual [16]. The manual content is converted into a movie script and later transformed into interactive version of it using an adaptable decision tree and simple game mechanics.



Figure 2. Machinima movie snapshots filmed on the Sims 3 platform and decision tree for an intervention to teach social interaction skills

Additionally we want to explore foreground and background interventions that could be delivered via the graphical user interface (GUI) associated to the operating systems or the application being used. Some ideas in that could be developed are: a. the use of desktop themes, screen savers, menu bars, etc. could be used to deliver soothing messages or color combinations that can help reduce stress, b. the modification of fonts or illumination of the screen to help improve reading during moments of emotional arousal, c. help block or reduce the number of cues presented through automated messages, such as incoming email or chat, to help maintain focus during highly stressful situations.

IV. METHODOLOGY

A. Usability Studies

Given the new uses derived from extracting personal, and in many cases intimate, information, our research approach contemplates a fundamental focus on usability. In the case of sensing personal mental data, it is important to understand that many users may not have a clear conceptual model of this novel sensing technology. Additionally, their level of motivation and their level of engagement will be diverse and certainly should change over time.

We plan to do experiments to examine the interaction between the user and a device that “reads” your mind and how this can generate further confusion, anxiety, or behavioral changes, as well as desire or rejection of the system. We want to observe the differences between a) treating the device as

reader of one's (mental) self or b) as a companion ("pet") that is affected or altered by our mental states. We will perform these studies in all the previously mentioned technologies (mouse, keyboard, smartphone, camera) and will also add emotion recognition technology using the AMMON [3] toolkit, which provides emotion and stress voice features on a mobile phone. We plan to run single mode and multi-modal lab and field studies to observe adoption and engagement patterns.

B. Stress and Emotion from Sensor and Contextual Data

We will begin with lab studies with subject under induced stress and emotions. We have run some of these studies in our lab to date, using a variety of calibration methods [18] [22]. In reality it is difficult to train a stress or emotion sensor because there is no objective ground truth. While Heart Rate Variability (HRV) responds strongly to Autonomous Nervous System (ANS) tone, there are several confounds. Even chemical tests (e.g. cortisol samples from saliva) are confounded by body chemistry (esp. medications), and exhibit significant lag (minutes or tens of minutes). We make use also of induced stress, but different subjects respond very differently to particular stressors, so the best we can hope for is an increase in stress on average. This is still enough to train a model, and once trained we can study the accuracy of this model with or without the inclusion of particular sensors to measure their individual predictive value. Data for stress modeling will include continuous variables (estimates from the biometric sensors including voice, keyboard, mouse, HRV), a periodic time variable, and discrete variables (location, id of a nearby person). We will start with simple models, namely linear regression of sensor data on a consensus of "control" signals (cortisol, self-report and Tricorder HRV). Discrete changes will be modeled with additive linear coefficients. Going beyond pure supervised regression; we will experiment with latent factor models, which may expose useful patterns of thought and/or behavior, which predict stress. We have previous experience with rich factor models for activity classification from desktop activity data [21].

C. Integrated Model and Data Acquisition

Given the models as trained above we will move to live, in the field modeling. Using the models above, we will compute real-time estimates of emotional states and stress for users working at a desktop or using a mobile device. When significant changes occur in modeled emotional and/or stress level occur, the system will prompt the user to give a self-report. These self-reports will then be used for further training and model refinement. By issuing requests at times of change, we will be able to gather emotional and stress label data that should be as close as possible in time to triggers. Thus it should be of maximum value for building models of the complex dependencies on discrete emotional triggers and stressors.

D. Expected Results

Our expectation is to obtain representation of stress through PC peripherals and Smartphones correlated with traditional Psychometric and Biometrics, opening the door to present our system of tools to the research community to test them as potential new sensors of stress/arousal.

We thank our psychology, bioengineering, and physiology advisors such as Prof. Edmund Seto, Prof. Adrian Aguilera, Prof. Ricardo Muñoz, Prof. Stephen Schueller, Dr. Keng-hao Chang, Dr. Reza Naima and Jesper Jeppesen.

REFERENCES

- [1] Brindley, D., Rolland, Y., et al. Possible connections between stress, diabetes, obesity, hypertension and altered lipoprotein metabolism that may result in atherosclerosis. *Clinical science* (London, England: 1979) 77, 5 (1989), 453.
- [2] Burchfield, S., *The stress response: a new perspective*, American Psychosomatic Society, 1979
- [3] Chang, K., Fisher, D., Canny, J., and Hartmann, B. How's my mood and stress?: an efficient speech analysis library for unobtrusive monitoring on mobile phones. In *Proceedings of the 6th International Conference on Body Area Networks, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering)* (2011), 71–77.
- [4] Christianson, S., *Emotional stress and eyewitness memory: A critical review*. *Psychological Bulletin*, Vol 112(2), 284-309, Sep 1992
- [5] Coyle, D., McGlade, N., Doherty, G. & O'Reilly, G. (2011) *Exploratory Evaluations of a Computer Game Supporting Cognitive Behavioural Therapy for Adolescents*. ACM CHI 2011
- [6] Duric, Z., Gray, W., Heishman, R., Li, F., Rosenfeld, A., Schoelles, M., Schunn, C. and Wechsler, H., *Proceedings of the IEEE*, 2002
- [7] Epp, C., M. Lippold, and R.L. Mandryk, *Identifying Emotional States using Keystroke Dynamics*, in *ACM Conference on Human Factors in Computer Systems (CHI)2011: Vancouver, BC*.
- [8] Healey, J., and Picard, R. Hpl-20040229: Detecting stress during real-world driving tasks using physiological sensors.
- [9] Hong, J., Ramos, J., and Dey, A., *Understanding physiological responses to stressors during physical activity*. In *Proceedings of Ubicomp* (2012).
- [10] Hudlicka, E. and McNeese, M., *Assessment of User Affective and Belief States for Interface Adaptation: Application to an Air Force Pilot Task, User Modeling and User-Adapted Interaction*, 2002
- [11] Ishii, H., *Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms*, ACM - CHI '97
- [12] Tian, Y., Kanate, T. and Cohn, J. Recognizing action units for facial expression analysis, *IEEE Trans. Pattern Analysis Machin Intelligence*, 23 (2) (2001), pp. 97-115
- [13] Kurihara, K., ed. *Face Recognition Book*. 2007, Pro Literatur Verlag.
- [14] Lu, H., Rabbi, M., Chittaranjan, G., Frauendorfer, D., Mast, M., Campbell, A., Gatica-Perez, D., and Choudhury, T. *Stressense: Detecting stress in unconstrained acoustic environments using smartphones*. 14th Intl Conf. Ubiquitous Computing, 2012
- [15] Maehr, W., *eMotion: Estimation of the User's Emotional State by Mouse Motions*, in *Information and Communication Engineering 2005*, Fachhochschule Vorarlberg: Dornbirn, Austria.
- [16] Munoz, R. *Group Cognitive Behavioral Therapy of Major Depression*, Latino Mental Health Research Program, UCSF, 2000
- [17] Naima, R., and Canny, J. *The Berkeley tricorder: Ambulatory health monitoring*. In *Wearable and Implantable Body Sensor Networks*. Sixth International Workshop on, IEEE, (2009), 53–58.
- [18] Paredes, P., Chan, M. *CalmMeNow: exploratory research and design of stress mitigating mobile interventions*, ACM - CHI 2011
- [19] Picard, R. *Affective computing: challenges*. *International Journal of Human-Computer Studies* 59, 1 (2003), 55–64.
- [20] Poh, M.-Z., D.J. McDuff, and R.W. Picard, *Non-contact, automated cardiac pulse measurements using video imaging and blind source separation*. *Optics Express*, 2010. 18(10).
- [21] Rattenbury, T. and Canny, J. *Task Support using Activity-Based Context*. in *ACM Conference on Human Factors in Information Systems, CHI 2007*
- [22] [WORK IN PROGRESS PAPER] Sun, P., Paredes P. and Canny, J. *MouStress: Detecting Stress from Mouse Motion*. To be Submitted to ACM - CHI 2014

ACKNOWLEDGMENT

