Eustressed or Distressed? Combining Physiology with Observation in User Studies

Avinash Wesley
Computational Physiol. Lab
University of Houston
Houston, TX 77204 USA
awesley@uh.edu

Peggy Lindner
Computuational Physiol. Lab
University of Houston
Houston, TX 77204 USA
plindner@uh.edu

Ioannis Pavlidis
Computational Physiol. Lab
University of Houston
Houston, TX 77204 USA
ipavlidis@uh.edu

CHI 2012, May 5–10, 2012, Austin, TX, USA.
ACM 978-1-4503-1016-1/12/05.

Abstract
In this article the authors describe a novel way to conduct user studies via the combination of a physiological and an observational information channel. The method enables not only the quantification of arousing emotional states but also their disambiguation into positive or negative instances. The physiological channel targets sympathetic responses and is materialized as a perspiratory signal extracted from thermal imagery of the perinasal area. The observational channel is materialized via decoding of facial expressions. However, while such decoding is usually performed in the visible spectrum, the authors have developed an algorithm to carry this out in thermal imagery instead. Thus, thermal imaging is used for both physiological and observational analysis. The potential of this dual unobtrusive methodology is demonstrated with some examples from a stress study, where users (surgeons in this case) interact with laparoscopic training boxes.

Author Keywords
User studies; emotions; eustress; distress; sympathetic signals; facial expressions

ACM Classification Keywords
H.5.m [Information Interfaces and Presentation]: Miscellaneous;

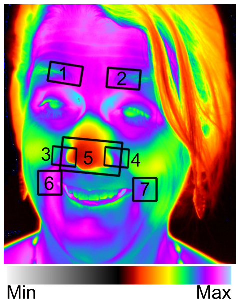

Figure 1: The seven ROIs used to capture facial muscle movement and deformation

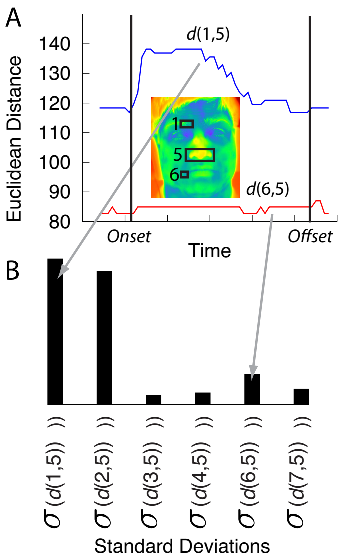

Figure 2: Feature vector formation for expression AU 1+2 (Inner+ Outer Eyebrow Raise)

General Terms

Design, Experimentation, Measurement

Introduction

One of the goals of human-centered computing is to unobtrusively monitor and understand human behavior, in order to assess interactions. Researchers have paid significant attention to the role of emotions on human behavior [2]. Emotions are not directly measurable, but can be inferred from expressive cues, self-reporting, physiological indicators, and context. Previous work demonstrated that during emotional arousal physiological signs materialize on the face, such as increased blood flow in the peri-orbital area [4] and transient perspiration on the perinasal area [5]. These signs have thermo-physiological footprints and quantification methods have been proposed based on thermal imaging in [6] and [5]. Between the periorbital and perinasal signal, the latter is of particular interest to this work, because it is part of a cluster of sympathetic responses on sensory organs (tactile and olfactory) that are closely related to emotions [5]. Because the perinasal response is sympathetic in nature, it is non-specific to negative or positive arousal. However, with the aid of an observational cue, such as facial expressions, it would be possible to disambiguate instances of negative (unpleasant) versus positive (pleasant) arousal.

Facial expressions are formed through coordinated muscle actions and can be classified using the Facial Action Coding System (FACS) [3]. FACS breaks down the development of expressions into sets of basic units. It was designed to measure visible facial behavior in any context, not just in emotions, and has become the gold standard for facial measurement systems. Automatic FACS decoding in visual imaging has proved challenging due to the effect of light variability [1].

In this work, the authors used the thermal imaging method described in [5] to derive the perinasal signals. In addition, they have developed an algorithmic method to decode facial expressions in thermal rather than visual imagery. Hence, both physiological and observational analysis can be carried out computationally under a single imaging modality.

The remainder of the paper features a description of the thermal imaging methodology for the extraction of perinasal signals and the recognition of facial expressions. This is followed by a description of the human-machine interaction study where the dual analysis method was applied. The paper concludes with accuracy results of the facial expression recognition method and examples of its disambiguating role on the physiological signals.

Methods

Facial Expression Recognition

The authors tracked 7 regions of interest (ROIs) on the thermal imagery of the face (Fig. 1). The tracking algorithm used is described in [7]. The ROIs were carefully chosen to align with facial muscles heavily involved in emotional action units. Each ROI was abstracted by its centroid that was tracked over time forming a trajectory. The centroid of ROI-**5** (nose) was used as a reference, because the nose is the most stable part of the face and is largely invariant under expressions.

Evolving Euclidean distances between centroid trajectories were used as indicators of muscle actions. Specifically, the algorithm computed the Euclidean distances $d(\mathbf{x}, \mathbf{5})$ between each ROI-**x** ($\mathbf{x} \neq 5$) and ROI-**5** from the onset till the offset of every expression (Fig. 2A). A feature vector for each expression was then formed by computing the standard deviations of these Euclidean distances (Fig. 2B).

These feature vectors capture the characteristic inter-muscle deformations over the course of expressions, and hence can be used to train a classifier. The authors chose a feed-forward multilayer perceptron for classification. The multilayer perceptron featured 13 input nodes, 12 sigmoid nodes in the hidden layer, and 5 output nodes to classify expressions.

Study Design

The authors imaged thermally[1] and visually the faces of surgeons while they engaged in training in an inanimate laparoscopic skills lab at the Methodist Hospital, per a protocol approved by the local institutional review board. The surgeons ($n = 17$) were performing manipulation, precise cutting, and suturing drills on a laparoscopic training box over the course of several weeks. In 977 training trials (1-4 min each), the authors found 244 expressions made out of 5 action unit combinations (AU 1+2, 4, 9, 10, and 12). The subject population included both novice and experienced surgeons. Hence, the challenge level of the training was differing among subjects, producing interesting cases of distress and eustress, in a human-machine interaction paradigm that can generalize across many domains. Indeed, due to increased mechanization and computerization, interactions in areas as different as laparoscopic training and unmanned aerial vehicle (UAV) piloting, start looking increasingly similar.

Actual AUs	Predicted AUs				
	1+2	**4**	**9**	**10**	**12**
1+2	**6**	5	3	3	1
4	3	**39**	0	3	0
9	2	1	**5**	7	5
10	0	2	1	**69**	3
12	1	1	0	4	**80**

Table 1: Confusion matrix for thermal FACS classification

Results and Discussion

The perinasal signals quantified arousals during the training trials. Novice surgeons exhibited a preponderance of arousals with respect to experienced surgeons. We informed the type of arousal events in the physiological signals via the observation channel. Specifically, a certified FACS expert decoded facial expressions on the visual stream into three categories: positive, negative, and neutral. ations on the visual stream served as ground-truth. Then, the authors used 10-fold cross validation and percentage split to test the accuracy of the algorithmic FACS method on the thermal stream, using the expert's FACS annotation on the visual stream as ground-truth. The overall recognition rate found to be 81.55% (Table 1). This performance is likely to hold across different application scenarios, as thermal imaging is relatively insensitive to lighting conditions [1].

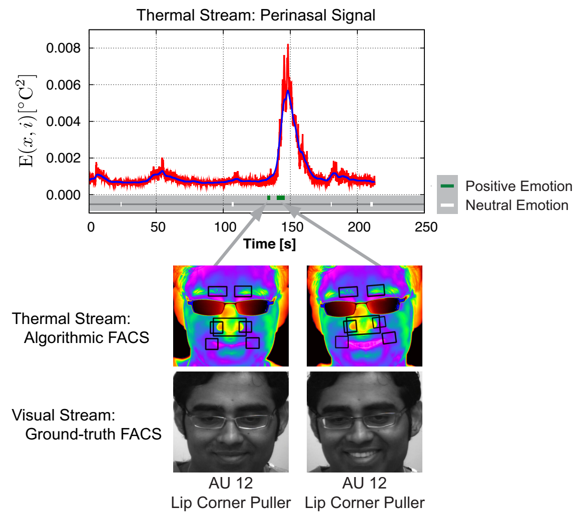

Figure 3: A bout of eustress for an experienced surgeon (D005), as indicated by the locally elevated perinasal signal and the positive facial expressions.

Figure 3 shows a characteristic example of eustress from the experimental set. About 130 sec into the drill, the perinasal signal of the surgeon exhibits elevation, which is

[1] The thermal imaging system included a MWIR camera from FLIR (model SC6000), outfitted with a MWIR 100 mm lens f /2.3.

characteristic of arousal. From the observational channel, either via manual FACS decoding in the visual stream or via algorithmic FACS decoding in the thermal stream, can be inferred that this is a bout of eustress. Indeed, the experimental context supports this conclusion, as this is an experienced surgeon who successfully addressed a technical challenge towards the end of the drill (sense of accomplishment).

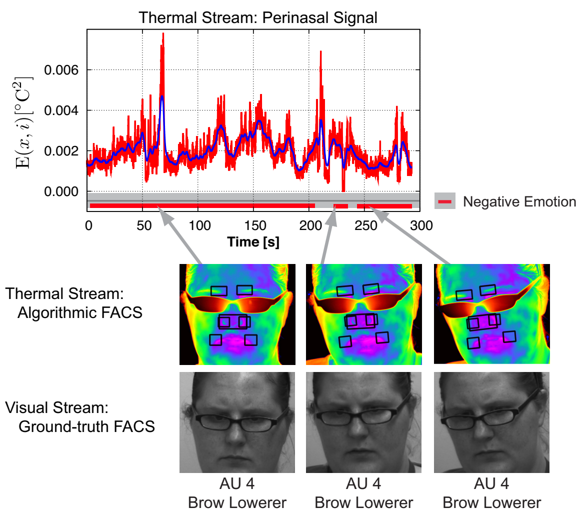

Figure 4: Bouts of distress for a novice surgeon (D025), as indicated by the fluctuating perinasal signal and the negative facial expressions.

Figure 4 shows a characteristic example of distress from the experimental set. The surgeon seems to be undergoing a roller-coaster of arousals. The distressing type of arousals is informed by the observational channel (both in the visual and thermal streams). Indeed, this is a novice surgeon who performed multiple errors during the execution of the drill (sense of foreboding).

Future studies may benefit from the proposed analysis method that is not only comprehensive (quantitative and qualitative), but also economical (single imaging modality with no labor).

Acknowledgements

Support provided by NSF award # IIS-0812526.

References

[1] Buddharaju, P., Pavlidis, I., Tsiamyrtzis, P., and Bazakos, M. Physiology-based face recognition in the thermal infrared spectrum. *IEEE Transactions on Pattern Analysis and Machine Intelligence 29*, 4 (2007), 613–26.

[2] Ekman, P. *Emotions revealed*. Times, New York, NY, USA, 2003.

[3] Ekman, P., Friesen, W., and Hager, J. C. *Facial Action Coding System Investigators Guide*. 2002.

[4] Levine, J., Pavlidis, I., and Cooper, M. The face of fear. *Lancet 357*, 9270 (2001), 1757.

[5] Shastri, D., Merla, A., Tsiamyrtzis, P., and Pavlidis, I. Imaging facial signs of neurophysiological responses. *IEEE Transactions on Biomedical Engineering 56*, 2 (2009), 477–484.

[6] Tsiamyrtzis, P., Dowdall, J., Shastri, D., Pavlidis, I., Frank, M., and Ekman, P. Imaging facial physiology for the detection of deceit. *International Journal of Computer Vision 71*, 2 (2007), 197–214.

[7] Zhou, Y., Tsiamyrtzis, P., and Pavlidis, I. Tissue tracking in thermo-physiological imagery through spatio-temporal smoothing. *Proc. of the 12th Int. Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2009) 5762* (2009), 1092–1099.