Affect-Gaze-Gesture. A three layer model of post-WIMP interaction.

Krzysztof Krejtz
Interactive Technologies Laboratory,
Information Processing Institute
Al. Niepodleglosci 188B,
00-608 Warsaw, Poland
kkrejtz@opi.org.pl

Cezary Biele
Interactive Technologies Laboratory,
Information Processing Institute
Al. Niepodleglosci 188B,
00-608 Warsaw, Poland
cbiele@opi.org.pl

CHI 2013 Workshop on "Gaze Interaction in the Post-WIMP World", 1 April 27, 2013, Paris, France.

Abstract

The present position paper for the ACM CHI 2013 Workshop on "Gaze Interaction in the Post-WIMP World" is aimed at describing our vision and theoretical framework on how the interaction with computer devices and user interfaces may benefit from joint input of user's emotional state and visual attention. In natural settings, interactions with others (persons, physical objects and software) are mostly shaped by emotions which in turn guide human attention and then behavior. We postulate the three layer model of interaction between emotions, attention and behavior. Combining both user's gaze and their emotional state as a supporting input channel, may facilitate a more immersed and natural user experience with post-WIMP interfaces.

Author Keywords

visual attention, affect, multimodal, gaze-contingent interaction

ACM Classification Keywords

H.5.2 [Information interfaces and presentation]: User Interfaces input devices and strategies..

General Terms

Human factors, Theory

Introduction

During interaction with computers users are flooded with tons of visual information. Some of them are more salient and some are just unnecessary interruptions. The later often induce confusion and lower user's satisfaction and productivity, e.g. [7]. Modern trends in designing user interface (UI) try to address this issue, however there is still a space or even a necessity for a conceptual framework of more natural and smooth interaction designs.

A promising direction derives from the concepts of gaze-contingent interfaces and affective computing. Those try to employ user's visual attention and affective processes as an input sources for adaptive interfaces. However, there is a strong need for a more general conceptual model which will help to apply the relation between emotions, cognition and behavior into human-computer interaction (HCI). That need is especially salient for post-WIMP devices and UIs that rely on gestures as a more convenient way of interacting with computers. Our position attempts to fill in the gap by showing how emotions and attention jointly direct behavior during human interaction.

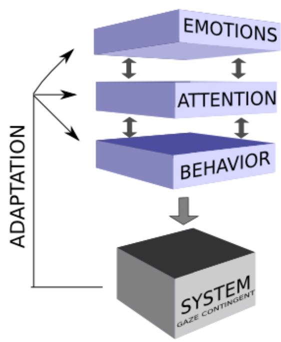

Figure 1: Three layer model of gaze- and emotion-contingent HCI.

We claim that the interplay of emotions and attention is crucial for the effective behavior regulation during interaction with an interface. Emotions shape formal and content aspects of human visual attention. Whereas attention as the mechanism of selection reduces the number of stimuli that can be potentially perceived. Thus, by measuring gaze and tracking user's emotional state we may with a greater certainty inform the system which elements presented on the screen should be featured. In our model, user's attention serves as a context for behavior (gestures) during interaction with computer system, yet the context for attention is formed by user's emotional state (see Figure 1). The present model consists of three interacting layers where affect and attention jointly build the regulatory context for acting with computer devices. We believe that UI implementing this model will produce an environment closely reflecting human-human communication when interpretation of other non-verbal signals is of the highest importance for the effectiveness of the interaction process.

Background

Affective biases of visual attention.

The concept of emotional computing which tries to bring emotions of the user as an important input for computer devices is not new, e.g. [4]. The novelty of our approach is the model of incorporating relations between user's emotions and attention in his/her interaction with post-WIMP interfaces. We claim that important for understanding of this relation is the concept of information reduction. Emotions select the information that comes to peoples perceptual system and thus limit the options among which attention will choose.

Consequently, one may say that affect directs attention toward some stimuli and suppress others. Affect directs attention basing on both valence and arousal. For example, people in positive mood show selective inattention to the negative images [8]. But what is more important in the context of HCI is the ability of mood and affect to shape formal characteristics of attention. 'Broaden and Build' theory [3] claims that positive emotions enhance creative, flexible, efficient thinking, and work productivity [14]. Moreover positive affect broadens visual field resulting in better recognition of stimuli presented in peripheral regions [15]. We may expect that during interaction with an interface, positive mood will enhance exploratory eye movements: longer saccades and shorter fixations while stress and negative emotions will effect in more focal mode of visual system. In the later situation any disrupting information from UI will produce user's dissatisfaction.

Regulatory role of affect and attention in multimodal HCI. In most of the current research in the field of human-computer interaction each of the channel of information (gaze, gesture, emotions) is processed independently, e.g. [5, 13] and the attempts to create multimodal interfaces are sparse, e.g. [11, 1]. In out position we propose to treat both emotions and attention during HCI with post-WIMP interfaces as a one system of information reduction. In this approach emotions serve as the mechanism of selection as well as regulation of the attention span in order to limit or broaden information stream. The interface may gather, combine and interpret such data from its user and try to continuously adapt to enhance user's ability to take more productive actions and rise it's satisfaction.

Psychophysiological measures of emotional valence and arousal. Psychophysiology provides a number of well established measurements that can serve as an input for a such adaptive interface, e.g. indicators of emotions' valence, arousal, and mental effort, e.g. [10, 6]. One of the key measures used in emotional assessment during interaction with computer include cardiovascular measures, electromyography of facial muscles, electrodermal reaction and respiratory measures [2]. Such physiological data can be combined and transformed, e.g. by fuzzy logic models [9] into emotional dimensions of arousal (low/high) and valence (negative/positive) based on circumplex model [12] and send on-line to computer system. Joining these information with gaze data will result in a unique opportunity for interaction designers to produce interfaces which are not only gaze-contingent but also 'emotion-contingent' - make more salient parts of a screen on which user is looking and adjust the attention span to users emotional state.

Position statement

The number of computers and computer-like devices i.e. smart-phones, tablets, game consoles or smart TVs that we use in our every day situations has recently increased significantly. Also the technology became available to a broader audience, and as a result the level of technological fluency of the users became very differentiated. This diversity stresses the need for adaptive and customizable interfaces that could gather the information about the user's needs in order to help them accomplish their task sand achieve their goals. With the implementation of the proposed three layer model, we see the possibility for the improvement of human-computer interaction in different areas e.g. digital entertainment (gaming, TV interfaces), education (adaptive tutors), digital art (adaptive design) or in workplaces of high cognitive demands, e.g. flight control.

We posit that in order to effectively improve human–computer interaction it is necessary to take into account the characteristics of attentional and emotional processes of the user. System should adapt to users action in a way coherent with the features of attention, e.g. focal/ambient, that stem from the current user emotions. Thanks to gaze information it would be possible to establish the currently attended spot on the screen and change the parameters of the surrounding interface according to attention parameters. One of the sample implementation of this process could look as follows: emotions are constantly monitor during interaction, which gives us information about focality of the attention, then knowing the fixation point on the screen it is possible to highlight it and make the rest of the screen less obtrusive, e.g. by dimming or by disabling or notifications.

Advances in the area of multimodal adaptive interfaces can change the way of interaction with the technology both for the professional and naive users but requires still both an extensive theoretical investigation and well-designed empirical research.

References

[1] Duric, Z., Gray, W. D., Heishman, R., Li, F., Rosenfeld, A., Schoelles, M. J., Schunn, C., and Wechsler, H. Integrating perceptual and cognitive modeling for adaptive and intelligent human-computer interaction. In *PROC. OF THE IEEE* (2002), 1272–1289.

[2] El-Nasr, M., Morie, J., and Drachen, A. A scientific look at the design of aesthetically and emotionally engaging interactive entertainment experiences. In *GAffective Computing and Interaction: Psychological, Cognitive and Neuroscientific Perspectives* (New York, 2011), 281–303.

[3] Frederickson, B., and Joiner, T. Positive emotions trigger upward spirals toward emotional well-being. *Psychological Science 13* (2002), 172–175.

[4] Gockay, D., and Yildirim, G. *Affective Computing and Interaction: Psychological, Cognitive and Neuroscientific Perspectives*. Information Science Reference, Hershey, New York, 2011.

[5] Grauman, K., Betke, M., Lombardi, J., Gips, J., and Bradski, G. Communication via eye blinks and eyebrow raises: Video-based human-computer interfaces. In *Universal Access in the Information Society* (2003), 2–4.

[6] Hazlett, R., and Benedek, J. Measuring emotional valence to understand the user's experience of software. *Int. J. Human-Computer Studies 65* (2007), 306–314.

[7] Iqbal, S., and Horvitz, E. Disruption and recovery of computing tasks: Field study, analysis, and directions. In *ACM CHI 2007* (2007).

[8] Isaacowitz, D. The gaze of the optimist. *Personality and social psychology bulletin 31*, 3 (2005), 407–415.

[9] Mandryk, R. A physiological approach for continuously modeling user emotion in interactive play environments. *Measuring Behavior 2008* (2008), 93–94.

[10] Mandryk, R., and Atkins, M. A fuzzy physiological approach for continuously modeling emotion during interaction with play environments. *International Journal of Human-Computer Studies 6*, 4 (2007), 329–347.

[11] Oviatt, S., and Cohen, P. Perceptual user interfaces: multimodal interfaces that process what comes naturally. *Commun. ACM 43*, 3 (Mar. 2000), 45–53.

[12] Russell, J. A circumplex model of affect. *Journal of personality and social psychology 39*, 6 (1980), 1161–1178.

[13] Sebe, N. Multimodal interfaces: Challenges and perspectives. *J. Ambient Intell. Smart Environ. 1*, 1 (Jan. 2009), 23–30.

[14] Seligman, M. *Authentic Happiness*. Free Press, New York, 2002.

[15] Wadlinger, H., and Isaacowitz, D. Looking happy: The experimental manipulation of a positive visual attention bias. *Emotion 8* (2008), 121–126.