M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 478
Appears in: Proceedings of the 1998 Workshop on Perceptual User Interfaces PUI'98

A New Affect-Perceiving Interface and Its Application to Personalized Music Selection

Jennifer Healey, Rosalind Picard and Frank Dabek
Rm E15-389, The Media Laboratory
Massachusetts Institute of Technology
20 Ames St., Cambridge, MA 02139
fenn@media.mit.edu, picard@media.mit.edu, fdabek@mit.edu

Abstract

A wearable computer that perceives and responds to the wearer's affective state offers a new kind of perceptual interface. Instead of asking the user to continuously select preferences from a menu, the affective wearable gets to know its wearer's preferences by recognizing and responding to signals that carry emotional information. This paper presents a new design for an affective wearable interface, aimed at applications that do not require extensive access to a keyboard and monitor. We replace the standard handheld keyboard and monitor worn over the eye with a Palm Pilot interface, and modify the Palm Pilot software to accept physiological input as well as traditional forms of input. The complete system can capture patterns from many kinds of user behavior, which can potentially be used to learn good predictors of user preferences. The new application presented here uses the new affective wearable system to aid in music selection by incorporating not only stated user preferences, but also physiological variables that might be indicative of the user's present mood.

1 Background and Motivation

Affective wearables are wearable computers equipped with physiological sensors and pattern recognition that can perceive and respond to the wearer's affective state. These computers are designed to be comfortably worn for long periods of time, and therefore will not require constant interaction with direct input devices such as a screen or a keyboard. Instead of asking us to repeatedly click a button to indicate our preferences through menu selections, the goal of the affective wearable is to "get to know" the wearer and their preferences through perception of the user's affective state [PH97]. The affect recognition algorithms and user preference profiles then act as a personal, private, perceptual user interface that can be controlled by the wearer to help manage task load or entertainment.

To learn the user's *natural* preferences and affective responses, it is necessary that the interface to the wearable computer be unobtrusive so that it does not cause unusual anxiety or social behavior. The natural interaction of the user and their social comfort is of primary concern for this kind of assessment. The loss of affective communication through eye movement, facial gesture, and the social suspicion aroused by an unseen virtual environment make heads up displays (HUDs) unattractive for affective wearables.

We have modified the existing interface to be less socially obtrusive by integrating an adapted form of the PalmPilot. With this pen based tablet interface, the wearer is still free to interact with people using normal facial expressions and


Figure 1: Direct access to the computer has been achieved through a modified PalmPilot interface, shown (left) as a monitor physiological sensors and (right) running the VT100 emulator.

Figure 2: Using the tablet interface it is clear when the user is attending to the computer and when they are attending to the other person.

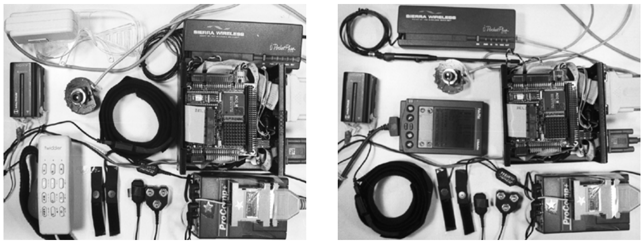

Figure 3: The original design of the wearable system (left) showing, clockwise from top left, the Private Eye HUD, CDPD modem, PC104 based wearable [Sta95], the AD converter, EMG, BVP GSR, a chordic keyboard and the battery, and, center the digital camera and the respiration sensor. In the new system (right), the PalmPilot replaces the HUD and chordic keyboard.

Figure 4: Shown left, the physiological sensors for the affective wearable, a respiration sensor, worn around the chest, an EMG, shown here on the bicep, a BVP attached to the wrist and a GSR sensor shown worn on two fingers. Shown right, the same sensors under clothing become unobtrusive. In this photo the respiration sensor is in the same location, EMG is on the trapezius muscle (shoulder), and BVP and GSR sensors are on the back. The sensors are connected to a wearable computer with the PalmPilot interface, worn in the satchel on the left shoulder. Photo courtesy of Justin Seger.

gestures as shown in Figure 2. This interface allows the user to have direct access to the computer for note taking, Internet communication, and control over peripheral devices such as digital cameras and physiological sensors (as shown in Figure 1). This direct input-output device is stored in the wearable's satchel most of the time, allowing the person to attend to other tasks and people in social situations, during which time the perceptual interface can continue to operate and try to improve its knowledge of the user's preferences.

2 The PalmPilot Interface

The programmable PalmPilot provides an easy-to-learn interface that can directly access the functionality of the 486 Linux based wearable computer, and be easily stored when not in use. The original design of the affective wearable [PH97] included a Private Eye HUD mounted on lab goggles, a Twiddler chordic keyboard, four physiological sensors: the electromyogram (EMG) , the photoplethysmograph (BVP), (for measuring heart rate and vasoconstriction), the skin conductance sensors (GSR), and a tuned analog to digital converter as shown in Figure 3. In the new design, the PalmPilot interface replaces both the Private Eye and the Twiddler. We developed software (available upon request to researchers) to enable the PalmPilot to communicate with the wearable computer through a serial port.

The primary benefit of the PalmPilot is that it increases the social comfort of the wearer. Additionally, many people are familiar with the handwriting recognition algorithm used by the PalmPilot and find it to be easier to learn to use than chordic keyboards. Although the physiological sensors can be hidden under clothes (as shown in Figure 4), a heads up display and chordic keyboard are more difficult to disguise and can interfere with social interaction.

The social advantages of a tablet-based interface over a heads up display are subtle but important. When the user's eyes are obscured, it is difficult for another person to determine whether the user is attending to them or to the computer (See Figure 2). Even if the HUD is on transparent glasses, it can seem that the computer always has precedence, since a screen change will often inadvertently grab the wearer's attention away from the other person. The PalmPilot interface allows the user's attention to be focused on other people rather than on the computer. By distancing the visual interface, the other person can be the primary focus of the user's attention. The tablet interface also facilitates the social acceptance of the wearable by allowing others to "glance over the shoulder" of the user to see what is being written[Bea97]. However, by keeping the tablet out of the view of others, some input can be kept private, even in a public setting, which is an advantage over speech interfaces. This interface provides a low power, low cost solution to wearable interface issues and is especially valuable when the primary task of the computer does not require frequent direct input from the user.

3 The Affective DJ: Music Selection with the New Affective Wearable

The perceptual interface for the affective wearable consists of sensors and algorithms that attempt to recognize the affective state of the user. The wearable can then try to learn which affective state variables are good predictors of user preferences in different physical and situational contexts [Pic97]. For example, the wearable could learn that when indoors studying, a high level of muscle tension is correlated with a preference for calming music, but while outside walking, a different pattern of physiological variables is correlated with a preference for more energizing music. The Affective DJ is presented here as an example of a computer application for the ambulatory environment that can be potentially improved by perceiving changes in the user's state, without requiring the user to directly input those changes through a keyboard or mouse.

The Affective DJ currently processes information about the user's affective state by detecting changes in the user's skin conductance on the palm of the hand. This response, one of two referred to as the galvanic skin response, has been shown to be an indicator of emotional responsiveness and to be only minimally involved with thermoregulation[SF90]. The measure correlates well with the emotional dimension of arousal and is most appropriate for making decisions about user preferences along this dimension. To avoid interrupting in the middle of a song, the Affective DJ only makes music selections at the end of every whole song.

Based on the whole song criteria, we chose a metric which captured the overall effect of a song. The current program calculates the average of the skin conductance for the last 30 seconds of the song and compares it to the average from the end of the previous song. As the skin conductivity signal is received by the wearable computer at 20 samples a second, through the A/D converter, a running average of the last 600 samples is calculated. At the end of each song this average is piped to the selection algorithm. This coarse metric of the difference in skin conductivity between the beginning and the end of the song seeks to sum up the overall effect of the song while ignoring the fast variations of the signal that are likely to be related to momentary qualities of the music.

It may be that there is much more information to be extracted from the dynamic nature of the skin conductance response. The number of sudden increases in the skin conductance (see Figure 5) can be tracked and recorded[HP98] and may be useful as another metric for quantifying the response of the music. In future versions of the system, this and other measures, such as heart rate variability, respiration rate, and muscle tension may be included in the decision algorithms.

Testing such a system is tricky and time-consuming, as there are a huge number of variables to control for, and the system really needs to have a lot of music and to be worn for a lengthy time before it can provide the planned advantages. Presently, we have some interesting initial findings from an informal experiment involving four subjects using our prototype system. The system trained on an individual during a period in which a random play list was presented, then the algorithm tried to create an "energizing" play list which would alternate songs to take the user's baseline to a level 1 micro-Siemen higher than the previous level and finally through a "relaxing" play list which would decrease the user's baseline by 1 micro-Siemen. Each session consisted of 10 songs and lasted approximately a half hour. The random play list consisted of selections from the pool of songs considered to be "high arousal," (modern alternative rock) and "low arousal," (mostly classical). During the first session of random song play, the change in skin conductivity for each song was recorded in a database for that user. The baseline was calculated as the average skin conductivity over the random song play session. This baseline was reassessed at the end of each type of play list, random, energizing, relaxing.

To simulate using the system in a natural environment, our users were allowed to work on their homework or use the computer during the experiment. Four subjects were run, but for one subject the skin conductivity sensor was plugged in to the wrong channel of the A/D converter and the signal was not recorded properly. During the hour-and-a-half session, users were asked to rate each song they heard based on how much they liked the song (on a scale of 1-hate to 7-love) and how exciting they found the song (on a scale of 1-very relaxing to 5-very energizing). When the songs were dummy-coded as being in the original categories of high arousal as "1" or low arousal "0" was found that the skin conductance was significantly correlated to this rating (p¡0.001). Skin conductance was also correlated to the users perception of the excitement level of the song (p¡0.005). In all cases the relaxing play list was able to lower skin conductivity; however, the arousing play list was not able to increase skin conductance in all cases. A possible explanation for this failure of the algorithm is that given the small number songs that were rated, ten, the best fit song was often a repeat which two of the users reported to be annoying.

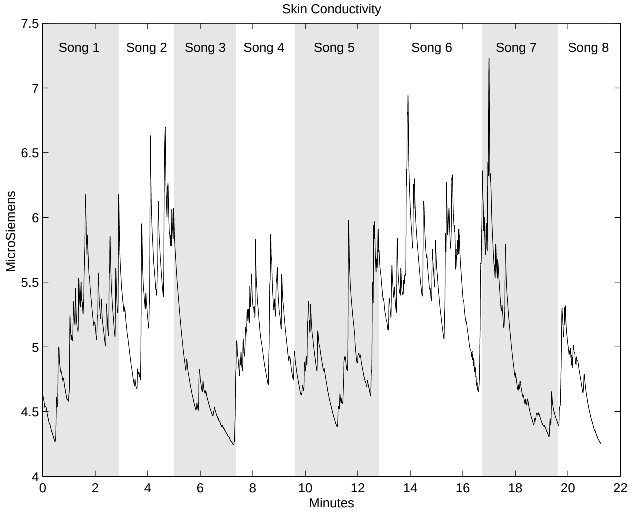

Figure 5: An example of the skin conductivity signal monitored by the wearable for a seated user. The signal is segmented to show how the Affective DJ algorithm would choose the next song based on each segment of this GSR signal.

This example illustrates how physiological signals could be perceived by an interface and used to adjust musical selection. In the not too distant future when large databases of music can be automatically downloaded by a wireless system, the interface could try to pre-load selections that it thinks are most likely to be preferred by the user, and then offer that user a choice of music that is most likely to agree with their momentary preferences. Instead of offering the user an unwieldy list of "everything out there" it might, for example, offer faster access to forty pieces of the type that best reflect the user's preferences, including their present mood. The user can then choose from this selection, or tell the system to choose; in either case, the system's knowledge of user preferences can help keep it from overloading the user with irrelevant selections. Affective responses, both physical and behavioral, are a key part of the context information that an intelligent interface should perceive in order to learn how to adjust system behavior on the user's behalf.

4 Summary

The ambulatory environment presents a new set of requirements for perceptual interfaces, especially a demand for sensitivity to social situations. When direct interaction through keyboards and screens is infrequent, a minimally obtrusive interface such as the PalmPilot may be preferred. We have presented a new design for an affective wearable computer that can perceive and respond to changes in its wearer's physiology without necessarily prompting the user for input.

The Affective DJ is a work in progress, that has only begun to be fully tested. The challenges of assessing multiple users' affective states and learning their preferred responses in different situations is a grand one. The present version of the Affective DJ uses only a portion of the data available; other data from the electromyogram, which indicates muscle tension, or from changes in heart rate with respect to respiration may prove more useful for predicting music preference. Ultimately, a learning algorithm should be combined with several features of these physiological signals together with features such as time of day and location of person (office, commuting, home, etc.) to help determine music preference.

The current system relies upon a very simple algorithm based on skin conductivity, which was confirmed in our (limited) tests to have a significant correlation with perceived excitement level of a song. We have also found certain skin conductivity changes (related to an orientation response) to be useful in controlling a wearable digital camera [HP98]. These two projects are examples of new efforts to fully understand affect in context, and to develop more sophisticated algorithms incorporating multi-modal sensing and pattern recognition for building interfaces that are more human-centered in design.

5 Acknowledgments

The authors would like to acknowledge the support of BT, HP, and the TTT consortium. We gratefully acknowledge the help of Justin Seger who wrote a new version of the code for the Affective DJ, Fernando Padilla for creating the digital images, Mary Obelnicki for demonstrating the use of the Private Eye system and Justin Seger, Deepa Iyengar and Erik Trimble for proofreading.

References

[Bea97] Ivan Bretan and et al. Leave the office, bring your colleagues. In *CHI-Extended Abstracts*, pages 335–336, 1997.

[HP98] J. Healey and R.W. Picard. Startlecam: A cybernetic wearable camera. *To appear : Proceedings of the Second International Symposium on Wearable Computers*, Oct. 1998.

[PH97] R. W. Picard and J. Healey. Affective wearables. *Personal Technologies*, (1):231–240, 1997.

[Pic97] R. W. Picard. *Affective Computing.* The MIT Press, Cambridge, MA, 1997.

[SF90] Michael E. Dawson Anne M. Schell and Diane L. Filion. The electrodermal system. In Cacioppo and Tassinary, editors, *Principles of Psychophysiology: Physical, social and inferential elements*, chapter 10, pages 295–324. Cambridge University Press, Cambridge, first edition, 1990.

[Sta95] T. E. Starner. Wearable computing. Perceptual Computing Group, Media Lab 318, MIT, Cambridge, MA, 1995.