

# Non-intrusive Physiological Monitoring for Affective Sensing of Computer Users

Armando Barreto  
*Florida International University*  
USA

## 1. Introduction

The last two decades have, undoubtedly, brought about an amazing revolution in the relationship between computers and their users. This relationship has evolved from an initial state in which the full “burden” of the communication was placed on the shoulders of the user, when early computer models had to be programmed one instruction at a time, toggling individuals switches (which restricted computer usage to very few, highly trained individuals), to the current status, in which, thanks to highly intuitive graphic user interfaces (GUIs), even young children can have some meaningful interaction with the personal computers that are now present at many homes.

Further, it is now possible for users to employ alternative means, such as their speech, or even the direction of their eye gaze, to interact with computers. In cases such as these, it is clear that now the computer has taken over a larger portion of the interaction “burden”, as ancillary programs (speech recognition, eye image processing, speech synthesis) will be running in the computer system to match the actions (speaking, shifting the point of gaze on the screen, listening) that the user naturally and almost effortlessly performs during the interaction.

One may think that computers are fast approaching a level of development in which they may recognize our speech, perceive our gaze shifts and speak to us just as well as another human could, to the point of being able to substitute, under certain scenarios, a human counterpart in a dialog. However, it is very likely that, in spite of the efficiency of the recognition of our speech and the fidelity and cadence of the synthesized voice, we would soon realize we are interacting with a machine, as the subtle modulation and adjustments that occur in human-human interaction due to phenomena such as empathy and sympathy would be found missing, substituted instead by mechanistic and often inflexible templates that have been pre-designed for short interaction segments, which disregard what the affective state of the user might be or how it might be changing. In summary, the goal of a human-computer interaction that should be inherently natural and social, following the basics of human-human interaction, as proposed by Reeves & Nass [Reeves & Nass, 1996], has not yet been reached.

## 2. Affective Computing and its Requirements

In response to the challenge outlined above, a whole new sub-discipline of Human-Computer Interaction has emerged, under the name of Affective Computing. One of its pioneers, Rosalind Picard, describes Affective Computing as ‘Computing which relates to, arises from, or deliberately influences emotions’ [Picard, 1997]. In analyzing the specific capabilities that a computer would require to fulfil Picard’s description, Hudlicka [Hudlicka, 2003] proposed that the following are the key processes involved:

1. Affect Sensing and Recognition
2. User Affect Modelling / Machine Affect Modelling
3. Machine Affect Expression

The interaction of these key processes involved in an Affective Computing implementation is shown in Figure 1.

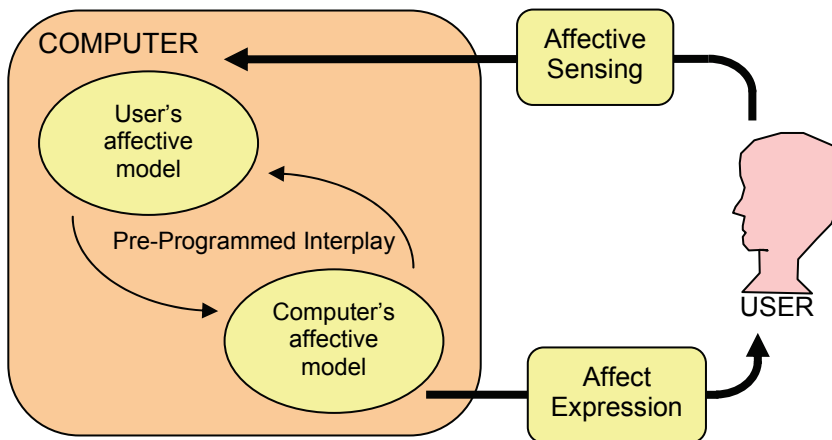


Fig. 1. Simplified diagram showing the interaction between the key processes in affective computing identified in [Hudlicka, 2003].

The interrelation between these basic building blocks for an affective computing system may be understood as follows: Just like humans are capable to “respond” to the affective state of another human by, for example, expressing empathy, an affective machine (affective computer) would first need to assess the affective state of its user, through affective sensing and recognition, in order to determine an appropriate reaction, based on an affective model of the user and its own affective model. The interplay between these models in the machine, for a given affective state identified in the user, and considerations derived from the functional purpose of the specific computing system, will determine what the reaction of the affective machine should be. Then, it is possible that the affective reaction of the machine may need to go beyond the modification of the state of its own (machine) affective model, and it may require that a resulting expression of affect be generated and directed to the user. For example, an affective avatar may be developed to provide support to the user during

computer-based tutoring activities. If this affective system were to detect sadness in the user (perhaps caused by being notified of a low score in a quiz), it would update the user's affective model and, if programmed to be supportive, it may also alter its internal affective model to a state of "sadness". Further, the affective machine may then express its empathy with the user by, perhaps, adopting a "sad" facial expression and changing to a soft and slow synthesized speech pattern.

### 3. Affective Sensing

It is clear, from the previous simplified description and example, that the initial task of affective sensing, i.e., the ability of a computer to remain aware of its user's affective states and transitions, is essential to the practical implementation of the complete affective computing paradigm. In fact, Picard has identified "Sensing and recognizing emotion" as one of the key challenges that must be conquered to bring the full promise of affective computing concepts to fruition [Picard, 2003]. However, it is just as clear that the implementation of a real-time, robust mechanism for the assessment of affective states and transitions in a computer user, as considered within a realistic interaction environment, remains an elusive goal.

The difficulty implicit in resolving the challenge of machine-based affective sensing may prompt researchers to actually question the feasibility of such task. Nonetheless, the very observation that has prompted researchers to pursue machine-based recognition of emotions provides evidence of the existence of clues that signal the emergence of affective states in humans, and the transitions between them. That is, we want to endow computers with affective sensing abilities because we wish they did what humans do when we interact with each other. Thus, the fact that humans are capable (according to their individual levels of perception) to "read" affective clues from people that engage them in interaction, indicates that such objective clues exist and that they are readable by an entity that is external to the subject being observed. Therefore, the actual challenge is to select the most appropriate set of clues to watch for, and then to develop systems with enough sensitivity and specificity to detect (and interpret) the occurrence of those clues, with a moderate level of "false detections".

It should also be noted that the goal of providing computers with an awareness of the affective states of their users for the purpose of enhancing the interaction between the user and the computer implies, indirectly, that the means utilized to achieve that assessment should not, simultaneously, be detrimental to the quality of the interaction. Specifically, this consideration limits the monitoring mechanisms that a realistic affective sensing system may involve to achieve its objective, restricting them to those that would not be considered hindrances for an ordinary computer user.

#### 3.1 Affective Sensing Approaches based on Audio, Video or Text

Perhaps due to the preference of avoiding extraordinary devices or components that could be considered to hinder the ordinary activities of the computer user in an abnormal manner, some research groups have placed a strong emphasis on attempting the assessment of user affective states using streams of data that are (more or less) commonly available to many

contemporary computing systems. In this sense, it could be considered that many personal computers these days are equipped with a camera that captures video of the face of the user and it can also be considered that a microphone could be continuously recording the speech of the user. Certainly, since much of our interaction with computers is still through text typed on a keyboard, this could also be considered a pre-existing stream of information from the user that could, potentially, be used to attempt the assessment of user affective status.

Zeng et al. [Zeng et al., 2007], provide an interesting survey of relevant systems that use video (typically of the user's face) or audio, or both, to attempt the assessment of the user's affective state. Most vision-driven approaches are based on the known changes that occur in the geometrical features (shapes of eye, mouth, etc.) [Chang et al., 2006] or appearance features (wrinkles, bulges, etc.) [Guo and Dyer, 2005] of the faces of the subjects, according to different affective states. The approaches based on speech monitoring search for affective clues in the explicit components of speech, i.e., in its linguistic aspects and also in the implicit (or paralinguistic) aspects. In this area, the work of Cowie et al. is significant, as it associated acoustic elements to prototypical emotions [Cowie et al., 2001]. Further, some groups have recently begun to explore the coordinated exploitation of audio-visual clues for affective sensing [Fragopanagos & Taylor, 2005].

Liu et al. argue that the utilization of text typed by the user is particularly important, since "the bulk of computer user interfaces today are textually based" [Liu et al., 2003]. They also provide a basic taxonomy for most common textual affect sensing approaches, consisting of four groups. The first group is "Keyword Spotting", in which the presence of words considered affective indicators is detected [Elliot, 1992]. The second group is formed by approaches based on "Lexical Affinity", where more than just obvious affect words are assigned a probabilistic "affinity" for a particular emotion [Valitutti et al., 2004]. The third category is for approaches based on "Statistical Natural Language Processing", in which a machine learning algorithm is trained using a large corpus of affectively annotated texts [Goertzel et al., 2000]. The final category is reserved for highly customized or "Hand-Crafted Models", such as Dyer's "Daydreamer" project [Dyer, 1987].

### **3.2 Affective Sensing Approaches based on Data Collected Directly From the User**

In contrast with the audio, video and text approaches outlined above, other research groups have considered that the changes in the subject's facial expression, or in the words the user speaks or types, are all external manifestations of much deeper changes that the user undergoes as he or she modifies his/her affective state. These research groups have set out to identify the intrinsic physiological modifications that are directly associated with the affective states and transitions that occur in human beings, and have proposed methods for sensing those physiological changes in ways that are non-invasive and non-intrusive for a computer user. The following sections of this chapter explain the rationale for this school of thought, present some of the most relevant implementations of physiological monitoring systems for affective sensing in computer users and preview some of the most innovative approaches in this area of work.

#### 4. Rationale for Physiological Monitoring Towards Affective Sensing

In trying to devise mechanisms that would enable a computer to gain awareness of the affective state of its user through monitoring of his/her physiological signals, one could also ask: How does a human become aware of the emotional state of another? How does one change when affected by an emotional stimulus? Most of us can attest to some clear, involuntary and unmaskable changes in our bodies as reactions to strong emotional stimuli: our hearts may change their pace during climatic moments in a sports event we witness; our hands may turn cold and sweaty when we are scared; we may feel “a rush of blood to the head”, when we get into a strong argument. These are not imaginary changes, but instead reflect the perception of an actual reconfiguration of our organism that takes place as a reaction to the psychological stimuli listed.

Just like we are capable of identifying an affective shift in another human being by sensing his/her physiological reconfiguration (e.g., seeing the redness in the face of an angry colleague, feeling the wetness and cold of a fearful person’s hand) computers could, potentially, measure these physical quantities from their users and utilize those measurements to assess their affective states. This approach to affective sensing follows the lead of studies on the “detection of deception” (lie detectors), in that it attempts to capitalize on the physiological reconfiguration associated with transitions between affective states. The reconfiguration experimented by a human subject as a reaction to psychological stimuli is controlled by the Autonomic Nervous System (ANS), which innervates many organs and structures all over the body. It is known that the ANS has the ability to promote a state of restoration in the organism, or, if necessary, cause it to leave such a state, favouring physiologic modifications that are useful in responding to the external demands. These changes in physiological variables as a response to manipulations of the psychological or behavioural conditions of the subject are the object of study of Psychophysiology [Hugdhal, 1995].

The Autonomic Nervous System coordinates the cardiovascular, respiratory, digestive, urinary and reproductive functions according to the interaction between a human being and his/her environment, without instructions or interference from the conscious mind [Martini et al., 2001]. According to its structure and functionality, the ANS is studied as composed of two divisions: The Sympathetic Division and the Parasympathetic Division. The Parasympathetic Division stimulates visceral activity and promotes a state of “rest and repose” in the organism, conserving energy and fostering sedentary “housekeeping” activities, such as digestion [Martini et al., 2001]. In contrast, the Sympathetic Division prepares the body for heightened levels of somatic activity that may be necessary to implement a reaction to stimuli that disrupt the “rest and repose” of the organism. When fully activated, this division produces a “flight or fight” response, which readies the body for a crisis that may require sudden, intense physical activity. An increase in sympathetic activity generally stimulates tissue metabolism, increases alertness, and, from a global point of view, helps the body transform into a new status, which will be better able to cope with a state of crisis. Parts of that re-design or transformation may become apparent to the subject and may be associated with measurable changes in physiological variables. The alternated increases in sympathetic and parasympathetic activation result in a dynamic equilibrium achieved by the ANS, and produce physiological changes that can be monitored through

corresponding variables, providing, in principle, a way to assess the affective shifts and states experienced by the subject.

However, the physiological changes caused by sympathetic or parasympathetic activations are not well-focused, and do not impact just a few organs at a time. Instead, parasympathetic and sympathetic activations have effects that tend to be distributed over numerous organs or subsystems, appearing with a subtle character in each of them. So, for example, sympathetic activation (in general terms) promotes the secretion of adrenaline and noradrenaline, inhibits bladder contraction, promotes the conversion of glycogen to glucose, inhibits peristalsis and secretion, dilates the bronchi in the lungs, accelerates the heartbeat, inhibits the flow of saliva, dilates the pupils of the eyes and reduces the peripheral resistance of the circulatory system. In contrast, parasympathetic activation (in general terms) stimulates the release of bile, contracts the bladder, stimulates peristalsis and secretion, constricts the bronchi in the lungs, slows the heartbeat and stimulates the flow of saliva.

The distributed effects of the sympathetic-parasympathetic tug-of-war set up an interesting paradox for the assessment of affective states: There are (potentially) many points where the effects of ANS changes might be observed, yet none of those variables displays strong effects that could unequivocally reveal an affective state or transition. In some instances this ambiguity is further compounded by the fact that the observable physiological variables may be changed by ANS reactions to non-affective stimuli. That is, for example, the case of the pupil diameter, which is known to respond strongly to the amount of light impinging on the retina, through the Pupillary Light Reflex (PLR).

## **5. Selection of physiological signals that can be monitored non-intrusively**

In spite of the fact that the effects of sympathetic and parasympathetic activation, as physiological expressions of affective states and transitions, surface in numerous locations around the body, only a subset of those changes can be monitored by currently available means in ways that can still be considered “non-intrusive” in the context of human-computer interaction. According to this consideration, the following physiological signals, which are likely to be influenced by the ANS, are nonetheless impractical for affective sensing in the context of ordinary computer use:

Electrocardiogram (ECG) – The activity of the heart, directly reflected by the ECG is clearly affected by ANS changes. However, measurement of the ECG would require the application of electrodes to the chest of the computer user, which is an unrealistic assumption, even if the signals could then be transmitted wirelessly, to avoid having the user tethered to the computer.

Electroencephalogram (EEG) – The electrical signals produced by the activity of the brain may be influenced by ANS changes. Similar to the case of ECG, the measurement of the EEG would require the application of multiple electrodes to the scalp of the computer user, which is an impractical pre-condition for most computer users.

Pneumograph - The breathing pattern of a subject is likely to reveal ANS shifts. However, the collection of breathing data would ordinarily require the placement of a respiration transducer (pneumograph) fitted tightly around the chest of the computer user and the transmission (wired or wireless) of the signals to the computer, which is not practical by today's computer usage standards.

From the above remarks, it is clear that the collection of data from a computer user in ways that will not interfere strongly with the activities that are needed to operate the computer is an important limitation in the selection of physiological signals for affective sensing. An interesting alternative that emerged in the 1990's is the collection of physiological signals that can be retrieved by sensors that touch the skin of the user, particularly the skin of the hand. In 1999 Ark, Dryer and Lu [Ark et al., 1999] noticed that

"One obvious place to put the sensors is on the mouse. Through observing normal computer usage (creating and editing documents and surfing the web), people spend approximately 1/3 of their total computer time touching their input device. Because of the incredible amount of time spent touching an input device, we will explore the possibility of detecting emotion through touch"

Although the title of the paper in which these researchers included the above key reflection is "The Emotion Mouse" [Ark et al., 1999], it must be noted that the experiment described in the paper did not actually use a mouse-like device with the sensors. Instead they asked their subjects to hold two contact sensors (galvanic skin resistance and temperature) in their left hands while using a (normal) mouse with their right hands: "Participants were asked to sit in front of the computer and hold the temperature and GSR sensors in their left hand, hold the mouse with their right hand and wore the chest sensor". Through the means described in the previous excerpt, these researchers measured the heart rate (from a chest sensor), the temperature (from a contact sensor), the galvanic skin resistance (from a contact sensor) and assessed the General Somatic Activity (from the movement of the mouse), while the subjects attempted to emulate Ekman's six basic emotions: anger, fear, sadness, disgust, joy and surprise [Ekman and Rosenberg, 1997].

At about the same point in time, it was proposed that the variations observed in the "Blood Volume Pulse" (BVP) signal, which can be recorded from the subject's finger using an infrared photoplethysmograph (PPG), may also be appropriate to evaluate the ANS function [Nitzan et al., 1998]. More recently, it has been confirmed that the BVP signal from a photoplethysmograph may provide basic information about the heart rate and its variability [Giardano et al., 2002] in a non-intrusive form. In a sense, the BVP signal may offer additional information about the ANS function, as it is also affected by the peripheral cardiovascular resistance changes associated with increased sympathetic or parasympathetic activation.

Therefore, three contact-based physiological measurements seem to be viable candidates for the assessment of affective states in computer users, since their corresponding sensors may be incorporated in a customized mouse-type device: The galvanic skin response (GSR), the blood volume pulse (BVP) and the skin temperature (ST).

Additionally, as “webcams” become more and more common in computer systems, it is feasible that, in a near future, the analysis of a fourth physiological signal: the pupil diameter (PD) could be also used for the purpose of affective sensing.

The next section provides additional information about these four physiologic variables and their expected behavior under sympathetic activation. Typically, increased parasympathetic activation would have the opposite effect in each of the variables.

## 6. Effects of ANS changes in GSR, BVP, ST and PD signals

When a subject experiences stress and nervous tension, associated with increased sympathetic activation, the palms of his/her hands become moist. Increased activity in the sympathetic nervous system will cause increased hydration in the sweat ducts and on the surface of the skin. The resulting drop in skin resistance (increase in conductance) is recorded as a change in electrodermal activity (EDA), also called galvanic skin response, or galvanic skin resistance (GSR). So, in everyday language, electrodermal responses can indicate ‘emotional sweating’ [Hansen et al., 2003]. The GSR is measured by passing a small current through a pair of electrodes placed on the surface of the skin and measuring the conductivity level. In spite of its simplicity, GSR measurement is considered one of the most sensitive physiological indicators of psychological phenomena. GSR is also one of the signals used in the polygraph or ‘lie detector’ test. Figure 2 shows a typical increase in the GSR signal, known as a “Skin Conductance Response”

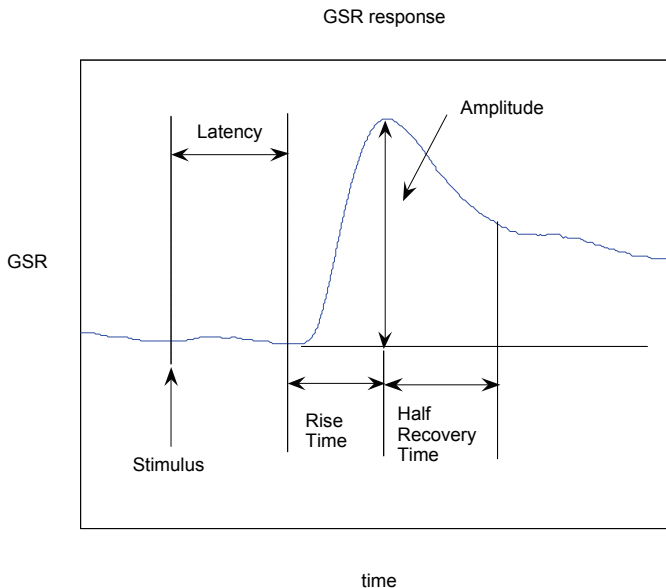


Fig. 2. Example of recorded GSR signal, showing a single Skin Conductance Response (SCR)

The measurements of blood volume pulse (BVP) may be obtained using the technique called photoplethysmography (PPG), to measure the blood volume in skin capillary beds, in the finger. PPG is a non-invasive monitoring technique that relies on the light absorption characteristics of blood, so it does not require costly equipment or specialized personnel. Traditionally, the Blood Volume Pulse has been used to determine the heart rate only. However, if measured precisely enough, it can be used to extract estimates of the heart-rate variability, which is another indicator of user affective state to be considered for human-computer interaction [Dishman et al., 2000; Picard & Klein, 2002]. Figure 3 shows a short segment (3 cardiac cycles) of a BVP signal recorded with a finger photoplethysmograph.

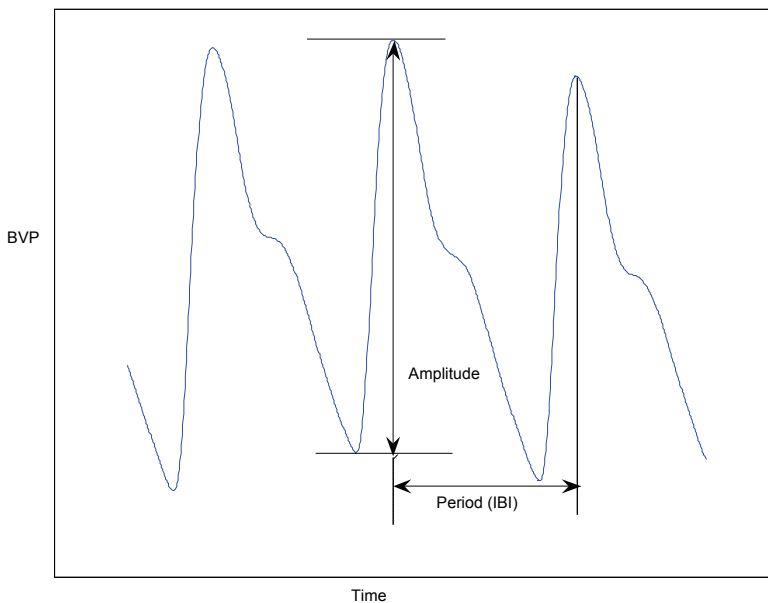


Fig. 3. Example of recorded BVP signal, showing three cardiac cycles (beats)

Changes of acral skin blood flow are also a commonly used indicator for sympathetic reflex response to various stimuli. In response to stimuli that produce sympathetic activation, the blood volume in the finger vessels is expected to decrease due to the vasoconstriction in the hairless areas of the hand but not in the hairy skin of the hand [Krogstad et al., 1995]. If this assumption is true, the finger temperature should transiently decrease according to this effect. A thermistor can be put in contact with the subject's finger to sense the temperature changes. Figure 4 shows an example of the temperature variations that may be observed as a manifestation of affective changes in a subject.

The diameter of the pupil is determined by the relative contraction of two opposing sets of muscles within the iris, the sphincter and dilator pupillae, and is influenced primarily by the amount of light and accommodation reflexes [Beatty & Lucero-Wagoner, 2000]. The pupil of

the human eye can constrict and dilate such that its diameter can range from 1.5 mm to more than 9 mm. The pupil dilations and constrictions are governed by the Autonomic Nervous System (ANS) in humans. Several researchers have established that pupil diameter increases due to many factors. Anticipation of solving difficult problems, or even thinking of performing muscular exertion will cause slight increases in pupil size. Hess [Hess 1975] indicated that other kinds of anticipation may also produce considerable pupil dilation. Previous studies have also suggested that pupil size variation is related to cognitive information processing. This, in turn, relates to emotional states (such as frustration or stress) since the cognitive factors play an important role in emotions [Grings & Dawson, 1978]. Partala and Surakka have found that using auditory emotional stimulation, the pupil size variation can be seen as an indication of affective processing [Partala & Surakka, 2003]. There are several techniques available to quantify pupil size variations [Grings & Dawson, 1978]. Currently, automatic instruments, such as infrared eye-tracking systems, can be used to record the eye information including pupil diameter and point of gaze. It is foreseeable that, in the near future, the resolution and quality of “webcams” and personal communication cameras may evolve to a point in which they will be able to assess the pupil diameter of a computer user in a continuous fashion.

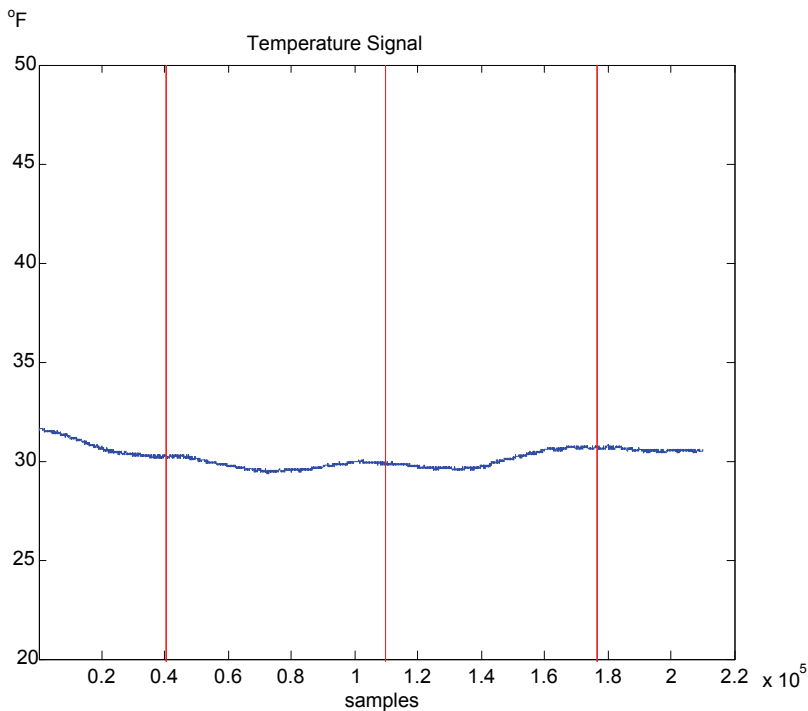


Fig. 4. Example of recorded skin temperature signal. Each vertical line indicates application of a stress stimulus. Total duration of segment shown is approximately 9.72 min (sampling rate was 360 Hz)

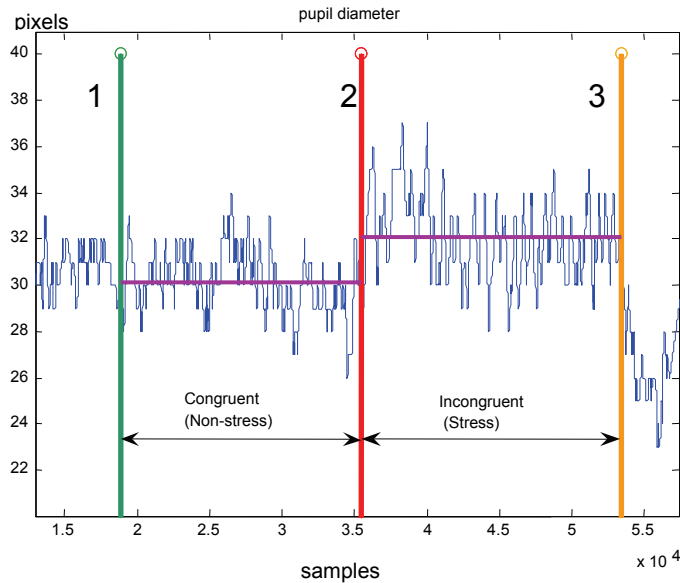


Fig. 5. Example of recorded pupil diameter signal (in camera pixels). A stress stimulus is applied in between vertical lines "2" and "3". Total duration of segment shown is approximately 2.08 min (sampling rate was 360 Hz)

## 7. Affective Sensing Systems Based on Physiological Monitoring

Several research groups have attempted the development of emotion recognition systems based on the analysis of physiological signals. The Affective Computing group of Dr. Rosalind Picard at the Media Laboratory of The Massachusetts Institute of Technology (MIT) has explored different approaches for affective sensing, which include the monitoring of physiological signals, since the mid-1990's [Picard, 1997; Picard et al., 2001; Picard & Klein, 2002].

One of the early efforts of this group, with respect to the monitoring of physiological signals for affective sensing, was reported in the paper by Healey and Picard [Healey & Picard, 1998]. In this effort, four physiological variables were collected from a single subject over an extensive period of time (32 days). In every session the subject would be asked to "experience and intentionally express eight affective states" when directed by a prompting system. The eight emotion states used were: no-emotion, anger, hate, grief, (platonic) love, romantic love, joy and reverence. The physiological signals monitored were the electromyogram (EMG) from the masseter muscle; the blood volume pulse measured with a finger photoplethysmograph; the skin conductance measured between the index and middle fingers of the left hand and the respiration pattern measured with a Hall-effect sensor strapped around the diaphragm. In this preliminary effort, Healey and Picard derived eleven features from these physiological variables and attempted the discrimination of the 8 emotions from the features using a Fisher linear discriminant projection. With this initial approach, the authors were not able to separate (classify) individual emotions, but were able

to distinguish six groups of three emotions each with correct classification levels ranging between 75% and 82%.

In 2001, Picard, Vyzas and Healey [Picard et al., 2001] published new results from more advanced analysis on the same type of data (also long-term recordings, from a single subject, of the same four physiological measurements, obtained while the subject expressed the same eight emotions). In this new report the researchers used an additional signal labeled "heart rate signal, H", said to be "derived from the blood volume pressure signal, B, by a nonlinear transformation performed automatically by the ProComp sensing system." The analysis involved the calculation of six features from each one of the five physiological signals (four measured directly and "H"): Mean of the raw signal; standard deviation of the raw signal; mean absolute value of the first differences of the raw signal; mean absolute value of the first differences of the normalized signal; mean absolute value of the second differences of the raw signal; and mean absolute value of the second differences of the normalized signal. In addition to these 30 features, these researchers also derived 10 "physiology-dependent" features (f1 - f10), for a total of 40 features. The classification approaches included the Sequential Floating Forward Search (SFFS) and Fisher Projection (FP), as well as a combination of both (SFFS-FP), which proved to be the most successful, achieving an overall classification accuracy of 81.25%, which consisted of only 30 misses in a test set of 20 instances of each of the 8 emotions (160 test instances in total).

In 2004 Kim and colleagues [Kim et al., 2004] reported the development of an emotion recognition system based on short-term monitoring of physiological signals from multiple (5) subjects. This system monitored four physiological signals: The electrocardiogram (ECG), measured between two electrodes ("from both upper arms"), the blood volume pulse obtained through a finger photoplethysmograph (PPG), the skin temperature measured from the ring finger of the left hand and the galvanic skin resistance, also known as electrodermal activity (EDA), measured between the index and middle fingers of the right hand. Although two different cardiovascular sensors were used, the ECG and PPG signals were used to study the same aspects: heart rate and heart rate variability (HRV). In this work, only a limited number of features were extracted from the original signals. The ECG signal was used to determine the basic heart rate by R-peak detection. The resulting "spike train" was transformed into a time series labeled by the authors "HRV time series". They studied the mean and standard deviation of the HRV time series, as well as its spectral composition in the low frequency (LF) band (0.03 Hz - 0.15 Hz) and high frequency (HF) band (0.15 Hz - 0.4 Hz). From the EDA signal, this group identified the occurrence of characteristic features called "Skin Conductance Responses" (SCRs) and used the frequency of their occurrence (in 50-second signal segments), the mean value of SCR amplitudes, their duration and the DC level of the EDA signal as features. Only two features were extracted from the skin temperature signal segments: its maximum and mean values. These features were then analyzed with a Support Vector Machine classifier, and this group reported a correct classification ratio of 78.4% in identifying instances of "sadness", "anger" and "stress", and 61.8% in identifying instances of "sadness", "anger", "stress" and "surprise".

In 2003, Barreto and Zhai [Barreto & Zhai, 2003] reported on the development of an instrumentation setup for the monitoring of four physiological signals towards the determination of the assessment of affective states in computer users. The four signals chosen for non-invasive and non-intrusive monitoring of subjects while they performed a specific computer task were: The galvanic skin resistance measured between two fingers of

the left hand; the blood volume pulse measured through a finger photoplethysmograph worn by the subjects in the ring finger of their left hands, the skin temperature measured with a integrated circuit temperature sensor attached to the thumb of the left hand of the users and the pupil diameter, obtained as a secondary measurement from a desk-mounted infrared eye gaze tracking system. This setup was capable of recording the GSR, BVP and ST signals, as well as additional time marker channels at 360 samples/second, for each signal. The pupil diameter was obtained as numerical values (expressed in pixels of the eye image captured by the eye gaze tracking system), every 1/60 of a second. These researchers used the monitoring setup to observe variations of BVP, GSR, ST and DP when the subject was presented with alternating neutral and stressing stimulation delivered as sequences of "congruent" and "incongruent" Stroop test trials. In a Stroop test trial the subject is presented with a word naming a color, written with a color font, and he /she is asked to identify verbally (or, in the case of a computerized version of the test, click on the screen button corresponding to) the font color [Stroop, 1935]. In a "congruent" trial the word presented spells the name of the color font used. In contrast, in an "incongruent" trial the color spelled by the word is different from the font color used, which elicits a mild level of mental stress in the subject [Renaud & Blondin, 1997]. Zhai and colleagues verified that the increased sympathetic activation during "incongruent" Stroop segments produced characteristic modifications on the four signals monitored. They derived a total of 11 features from the physiological signals monitored and used those features to attempt the differentiation of non-stress (Stroop congruent) and stress (Stroop incongruent) segments, by means of three different classifiers: A Naïve Bayes classifier; a Decision Tree classifier, and a Support Vector Machine classifier [Zhai et al., 2005; Zhai & Barreto, 2006]. These researchers found that the Support Vector Machine classifier performed best for the classification task, achieving a correct classification percentage of 90.10% [Barreto et al., 2007a].

## 8. Future Research Direction

An additional finding of Zhai and colleagues was that if the pupil diameter signal was removed from the ensemble of physiological signals monitored in their experiments, the performance of the classifiers, even the Support Vector Machine, would decrease significantly (to 58.85%), while the classification performance would remain essentially unaltered if, for example, the skin temperature signal were to be removed from consideration [Barreto et al., 2007a]. This observation has prompted further study of the potential of the pupil diameter signal, specifically, to determine the affective states or transitions of a computer user. Barreto et al. [Barreto et al., 2007b] were able to verify that the populations of PD values measured before and after a transition from a non-stress (congruent Stroop) experimental segment to a stress (incongruent Stroop) experimental segment were statistically different. Furthermore, these researchers also compared the Receiver Operator Characteristic (ROC) curves of individual features derived from the (mean) PD signal, the ST (mean slope) signal, the mean value of the BVP signal and the mean period from the BVP signal, considered as single-variable detectors, and found that the detector derived from the PD signal exhibited clearly superior characteristics (the area under the ROC curve was 0.96, versus 0.65 for the second-best detector) [Barreto et al., 2007c]. The indications of potential use of the pupil diameter measurement as a strong contributor to the identification of affective states and shifts in computer users is exciting

because this signal is not currently considered in many instrumental setups developed for affective sensing purposes. As such, it may very well represent an additional source of information that could be very useful in future studies of affective sensing. It should be noted, however, that all the studies described above which measured the variations of pupil diameter were performed in controlled environments, in which the ambient illumination and the light intensity emanating from the computer display were kept reasonably constant, by design, to minimize the unwanted influence of the pupillary light reflex (PLR) on the measured pupil diameter values. The practical application of the pupil diameter measurement for affective sensing purposes depends on the emergence of signal processing techniques that would be capable to differentiate pupil diameter changes caused by PLR from those derived from affective responses in the subject. The definition of such signal processing techniques is currently an open research topic.

## 9. Conclusion

It is clear, from the considerations briefly outlined in this chapter, that the definition of robust, non-intrusive methods for affective sensing in human-computer interactions is still an open challenge, which must be conquered as an essential pre-requisite to the fulfillment of the promise of Affective Computing concepts and their widespread application in everyday computing. While the goal of robust affective sensing may seem distant, it is also evident that a tremendous amount of progress has been made in the past two decades in many of the aspects that will necessarily be involved in a viable solution. Our understanding of affective states and their correlates to physiological changes has evolved, the sensing mechanisms used to monitor physiological variables have improved, the signal and image processing techniques used to analyze the physiological signals from the computer user continue to be enhanced, and the computing power that can be utilized to implement them (potentially in real-time) increases continuously. Additionally, research groups are now contemplating the use of multi-modal collaborative approaches in which the strengths of physiological monitoring can be combined with other sources of information about the user's affect available to the computer, such as face expression recognition and textual assessment of affect. When all of this is brought into consideration, it is foreseeable that practical solutions to the affective sensing problem might be found in a near future.

## 10. References

- Ark, W.; Dryer, D. and Lu, D. (1999). The Emotion Mouse. *Proceedings of the 8th International Conference on Human-Computer Interaction*. Vol. I, Ergonomics and User Interfaces, pp. 818-823.
- Barreto, A., & Zhai, J.(2003) Physiologic Instrumentation for Real-time Monitoring of Affective State of Computer Users, *WSEAS Transactions on Circuits and Systems*, Vol. 3, Issue 3., pp. 496-501.
- Barreto, A.; Zhai, J. & Adjouadi M. (2007a) Non-intrusive Physiological Monitoring for Automated Stress Detection in Human-Computer Interaction, *Lecture Notes in Computer Science* Vol. 4796, pp. 29-38, Springer-Verlag.

- Barreto, A.; Zhai, J, Rische, N. & Gao, Y. (2007b) Measurement of Pupil Diameter Variations as a Physiological Indicator of the Affective State in a Computer User, *Biomedical Sciences Instrumentation*, Vol. 43, pp. 146-151.
- Barreto, A.; Zhai, J., Rische, N. & Gao, Y. (2007c), Significance of Pupil Diameter Measurements for the Assessment of Affective State in Computer Users, in *Advances and Innovations in Systems, Computing Sciences and Software Engineering*, K. Elleithy, ed., Springer, (ISBN: 978-1-4020-6263-6). Dordrecht, The Netherlands, pp. 59 – 64.
- Beatty, J. & Lucero-Wagoner, B. (2000) The Pupillary System. In: *Handbook of Psychophysiology*. J.T. Cacioppo, L.G. Tassinari, and G.G. Berntson, editors. pp. 142-162. Cambridge University Press
- Chang, Y.; Hu, C.; Feris, R. & Turk, M. (2006), Manifold Based Analysis of Facial Expression, *J. Image and Vision Computing*, Vol. 24, No. 6, pp. 605-614.
- Cowie, R.; Douglas-Cowie, E.; Tsapatsoulis, N.; Votsis, G.; Kollias, S.; Fellenz, W. & Taylor, J. (2001). Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine* 18(1): 32-80.
- Dishman, R.K.; Nakamura, Y.; Garcia, M.E.; Thompson, R.W.; Dunn, A.L. & Blair, S.N. (2000) Heart rate variability, trait anxiety, and perceived stress among physically fit men and women. *International Journal of Psychophysiology* 37:121-133
- Dyer, M. (1987). Emotions and their computations: Three Computer Models. *Cognition and Emotion* 1(3): 323-347.
- Ekman, P. & Rosenberg, E. (1997) *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression using the Facial Action Coding System (FACS)*, New York: Oxford University Press.
- Elliott, C. (1992). *The affective reasoner: a process model of emotions in a multi-agent system*, Doctoral Dissertation, Northwestern University.
- Fragopanagos, N. and Taylor, J. (2005). Emotion recognition in human-computer interaction. *Neural Networks* 18(4): 389-405.
- Giardano, N.; Lehrer, P & Edelberg, R. (2002) Comparison of finger plethysmograph to ECG in the measurement of heart rate variability, *Psychophysiology*, v.39, pp. 246–253.
- Goertzel, B.; Silverman, K.; Hartley, C.; Bugaj, S. & Ross, M. (2000). The Baby Webmind Project. *Proceedings of the AISB'00 Symposium*, The Society for the Study of Artificial Intelligence and the Simulation of Behaviour.
- Grings, W.W. & Dawson, M.E. (1978) *Emotions and Bodily Responses A psychophysiological Approach*. Academic Press.
- Guo, G. & Dyer, C. (2005). Learning from examples in the small sample case -- face expression recognition. *IEEE Transactions on Systems, Man and Cybernetics - Part B* 35(3): 477–488.
- Hansen, A.L.; Johnsen, B.H. & Thayer, J.F. (2003) Vagal influence on working memory and attention. *International Journal of Psychophysiology* 48:263-274
- Healey, J. & Picard, R. (1998) Digital processing of affective signals. *Proceedings of the 1998 IEEE International Conf. on Acoustics, Speech and Signal Processing*, pp. 3749-3752.
- Hess, E. H., (1975), *The Tell-Tale Eye: How your eyes reveal hidden thoughts and emotions*, Van Nostrand Reinhold Co.
- Hudlicka, E. (2003). To feel or not to feel: the role of affect in human-computer interaction. *International Journal of Human-Computer Studies* 59(1-2): 1-32.

- Hugdahl, K.(1995) *Psychophysiology: The Mind-Body Perspective*, Harvard University Press, Cambridge, MA.
- Kim, K.H.; Bang, S.W. & Kim, S.R. (2004) Emotion Recognition System using Short-term Monitoring of Physiological Signals. *Medical & Biological Engineering & Computing* 42:419-427
- Krogstad, A.L.; Elam, M.; Karlsson, T. & Wallin, B.G. (1995) Arteriovenous anastomoses and the thermoregulatory shift between cutaneous vasoconstrictor and vasodilator reflexes. *J. Auton. Nerv. Syst.* 53:215-222
- Liu, H.; Lieberman, H. & Selker, T. (2003) A model of textual affect sensing using real-world knowledge. *Proceedings of the 8th international conference on intelligent user interfaces*. Miami, Florida, USA, ACM, pp. 125-132.
- Martini, F.H.; Ober W.C.; Garrison C.W.; Welch, K.; & Hutchings, R. T.(2001) *Fundamentals of Anatomy & Physiology*, 5th Edition, Prentice-Hall, Upper Saddle River, NJ.
- Nitzan, M.; Babchenko, A.; Khanokh, B. & Landau, D. (1998) The variability of the photoplethysmographic signal – a potential method for the evaluation of the autonomic nervous system, *Physiological Measurement*, Volume 19, pp. 93 – 102.
- Partala, T. & Surakka, V. (2003) Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies* 59:185-198
- Picard, R. (1997). *Affective computing*, MIT Press.
- Picard, R.; Vyzas , E. & Healey, J. (2001) Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23(10): 1175-1191.
- Picard, R. & Klein, J. (2002) Computers that recognize and respond to user emotion: theoretical and practical implications. *Interacting with Computers* 14:141-169
- Picard, R. (2003) Affective computing: challenges. *Int. J. Hum.-Comput. Stud.* 59, pp. 55-64.
- Reeves, B. & Nass, C. (1996) *The Media Equation*. Cambridge Univ. Press
- Renaud, P.& Blondin, J.-P. (1997) The stress of Stroop performance: physiological and emotional responses to color-word interference, task pacing, and pacing speed. *International Journal of Psychophysiology* 27:87-97
- Stroop, J.R. (1935) Studies of the interference in serial verbal reactions. *Journal of Experimental Psychology* 18:643-662.
- Valitutti, A.; Strapparava, C. & Stock, O. (2004) Developing Affective Lexical Resources. *PsychNology Journal* 2(1): 61-83.
- Zeng, Z.; Pantic, M; Roisman, G. & Huang, T. (2007). A survey of affect recognition methods: audio, visual and spontaneous expressions. *Proceedings of the 9th international conference on Multimodal interfaces*. Nagoya, Aichi, Japan, ACM, pp. 126-133.
- Zhai J., & Barreto A. (2006) Stress Detection in Computer Users Through Noninvasive Monitoring of Physiological Signals, *Biomedical Science Instrumentation*, vol. 42, pp. 495-500.
- Zhai, J., Barreto, A., Chin, C., & Li, C. (2005) User Stress Detection in Human-Computer Interactions, *Biomedical Sciences Instrumentation*, Vol. 41, pp. 277-282.