Multi-channel Physiological Sensing of Human Emotion: Insights into Emotion-Aware Computing using Affective Protocols, Avatars and Emotion Specifications

Panagiotis D. Bamidis, Andrej Luneski, Ana Vivas, Christos Papadelis, Nicos Maglaveras, Costas Pappas

Lab of Medical Informatics, The Medical School, Aristotle University of Thessaloniki, Greece

Abstract

This paper introduces a methodology for combining multi-channel psycho-physiological recordings of affective paradigms into a framework where the scientific results of such experiments are utilized in the human computer interaction context to model the computer's response based on the emotional context of the user and the situation. An affective protocol is described the results of which are expected to be combined with anthropomorphic avatars that enhance the man-machine interaction. The technological infrastructure of the later component is provided by means of XML specifications of signal descriptions and emotion recognition, as well as avatar behavior generator descriptions.

Keywords:

affective computing, emotion identification, specification, avatar technology.

Introduction

The interaction between humans and computers (HCI) has been a subject of research and discussion for a long time. The goal in the enhancement of HCI is getting it closer to the interaction between humans [1]. The essential human ability, and the one that nearly all the research is focused on, is human intelligence and its incorporation into computers. The HCI has experienced the introduction of facial and speech recognition, natural language recognition, as well as, software intelligent agents than can learn and reason on their actions [2]. However, until recently, emotions were rarely the topic mentioned in the discussions of human intelligence and their application in computers was not even considered. The reason was that emotions were considered as extraordinary human ability. Nowadays, however, there are many proofs of the importance of the emotions in the expression of intelligence [3]. The arguments put alongside the significance of emotions gave birth to a new area of "emotional intelligence", defined as "the capacity to understand emotional information and to reason with emotions" [3]. All the work that is done with computing and is related to, arises from or deliberately influences emotions on computers is called "affective computing" [4]. Affective computing has been a hot topic in recent years mostly because it introduced a new area in computing and therefore an increasing number of applica-

tions in that area. Until recently research has been mostly focused on monitoring user emotional reactions and trying to distinguish between different emotion categories such as fear, anger, sadness, happiness etc. The ability of recognizing human emotions requires the computer to monitor the user and based on certain parameters or conditions classify his/her emotional state. Since we are intending to copy the essence of human-human interaction, it is mandatory to identify the natural way of accomplishing this task by humans. Namely, human beings use several ways of communication and emotion recognition using their biological sensors of sight, touch and sound. Facial communication has proven to be of unique importance since facial expressions help in a great deal in identifying people's emotions [5]. Moreover, vocal communication as well as gesture based can be successful, to certain extend, in sensing and distinguishing between several classes of emotions.

Apart from using human natural senses, emotions can be recognized by monitoring psycho-physiological changes in the user [6, 7]. Naturally, for this to take place certain sensors have to be used recording heart pulse/activity (e.g ECG), galvanic skin response (GSR or SC), respiration rate, electroencephalogram (EEG) etc. This type of emotional reaction recording has been most attractive recently and it is the main focus in this project as well. While in the field of medicine the influence of emotions on the human health is of major concern, in human-computer interaction the focus is on understanding the user's emotions and improving the quality of software programs accordingly.

Understanding emotions is far from simple due to the complexity of human physiology. It would be so if each emotion category was characterized by a specific physiological pattern. However, certain emotions, even as different as love and fear, can cause similar effects on human physiology and, therefore, may be misrecognized. The emotion recognition process follows a sequence of steps even for the preparation of data before the recognition process. The initial step is to acquire the data from the user physiological signals. Following is the extraction of the key features of the signals [8]. The features form the basis of the comparison method. The feature extraction procedure can vary and depends on the goals of each specific project and therefore cannot be discussed in general. The final step in emotion recognition is performing

the classification of the emotional data into categories based on specific classification techniques [9]. Mostly used pattern classification methods are the Hidden Markov Model [10], Fisher linear Projection [11], Support Vector Machines [9] etc. However, these techniques are appropriate for classification of large amounts of data into categories. Moreover, there is a clear need for collections of representative emotional signals in short samples of 20-50 seconds containing digital signal data, in a flat file format such as that used in the MIT-BIH file library [12]. There exists also a need for a more representative way of specifying an emotion into one data record. An XML file containing the required signal data was successfully used in [9], thereby providing evidence that an XML based representation of the emotion elements is a suitable format of data specification. Furthermore, the contents of the XML data record are of great importance for interconnection among different research results, platform independency and reusability purposes, which can be used in telemedicine, decision-making etc. [13]. Thus, it is essential to follow certain standards or widely accepted guidelines for the structure and names of the elements in the record. One related standard was introduced in [14]. Basically, it is an introduction of a markup language – ecgML for modeling and storing ECG data of patients based on XML representations.

Significant progress has been made in all the above fields and therefore new ideas arose which basically provide answers on the question "What can the computer do after recognising human emotions?". Lisetti and colleagues [15] have recently investigated the use of agent-centered modalities or modes (avatars), and multimodal feedback given to a system user. For example, an interface agent for an e-health system session can display empathy via an anthropomorphic avatar who adjusts its facial expressions and vocal intonation according to the user's emotional state, as the latter is depicted by the set of measurements.

In the light of the previous developments, the scope of this paper is twofold. First, to present a step-wise approach to the design of an experimental protocol that aims to enable multi-channel physiological sensing of a subject's emotion. The second goal is to prepare the theoretical and technological grounds for later direct adaptation of computer user interfaces guided by the elicitation and identification of emotions. The former goal is attained through the use of physiological sensors like EEG, ECG, and skin conductance, while the latter one is mainly driven by the notion of avatar technology and the preparation of the relevant technological platform that will exploit the results of the experimental analysis.

Material and methods

The AFFECTION project context

This piece of work is part of a collaborative project, called AFFECTION, between the Lab of Medical Informatics at the Medical School of the Aristotle University of Thessaloniki, Greece, and the Brain Science Institute of RIKEN in Japan. The project aims at creating a scientific foundation for the robust identification of human emotional states through fusion and correlation of data from a multi-sensor

research framework. The emotion-related research findings will be subsequently incorporated within usability evaluation methodological frameworks to enable new objective/direct evaluation methodologies, as well as, new interface adaptation strategies. The project is envisaged to contribute to dynamic characterization and recognition of the subjects' emotional state upon interaction with computer systems. To achieve this, the project will employ multimodal recordings such as vocal expressions and physiological signals of the autonomic (heart rate, blood pressure, skin conductance etc) and central nervous system (EEG, MEG). Research will be carried along the following main stages: in the first, procedures, examples and baselines of multiple-channel emotional responding approaches are utilised to obtain fundamental inter-correlations of the various physiological measures in the light of behavioral data. The last stage involves the use of specific paradigms in order to incorporate the findings into a unified framework that ascertains the changing emotional state of the user and allows for adaptation of computer user interfaces based on the user and application context. A block diagram of the project is given in Figure 1.
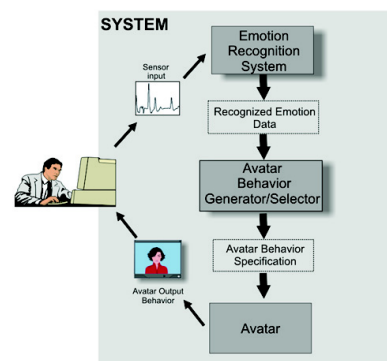

Figure 1 - A block diagram of the AFFECTION system

Affective protocol design

As already mentioned, emotions are complex phenomena, which include a wide range of observable behaviours, expressed feelings that are private and subjective, and changes in body states (distinctive somatic and autonomic responses). To achieve a first insight and start building on the scientific foundation of the AFFECTION project, the initial study aims at investigating the relationships between the different patterns of brain activation, autonomic responses, subjective experience and cognitive behavior related to emotionally-evocative photographs from the International Affective Picture System (IAPS) [16]. 50 healthy adults ranging in age from 18 to 40 years old form the participants' sample. Before the experiment, participants are updated and asked to sign a consent form. A full self evaluation questionnaire is given to them right afterwards, in order to have a (subjective) definition of the emotional state in general.

In Phase 1, we study the psychological and neurophysiological correlates of passive exposure to emotionally-

evocative stimuli. Four blocks of emotionally-evocative stimuli are selected from the IAPS, taking into account mainly two major dimensions, pleasure and arousal. Pictures are selected from the "Affective Space" as defined by the mean rating on the Valence (pleasure) and Arousal dimension. Each condition of "Affective Space" is manipulated between blocks, and the order of the Affective Space-block is counterbalanced across participants to avoid the order effect. In each trial, each stimulus is presented during 1 second. No inter-stimulus interval is scheduled for these blocks. There are 40 trials for each block, randomly selected from a larger set of pictures from each affective space condition. After each block, subjects fill in a questionnaire indicating how they feel right after this sequence block. In the second phase, the same pictures are used in a visuospatial attention paradigm, in order to investigate the effect of emotional processing on cognitive behavior. The independent variables are Affective Space block condition, and the visual field (left or right). On each trial, a central fixation cross appears for 500 ms followed by two pictures (one with high or low mean rating on valence and arousal and the other one with middle mean rating on both dimensions) for 500 ms. Then a target (a small asterisk) appears on the location of the emotional picture or in the location of the neutral picture (see Figure 2). The target remains until a response is made or until 2000 ms elapse. Participants are required to detect the appearance of the target; this can be done by pressing a key. During the experiment, recordings of the 10-20 EEG are conducted. In addition, simultaneous recordings of GSR, EOG, ECG are also taken.
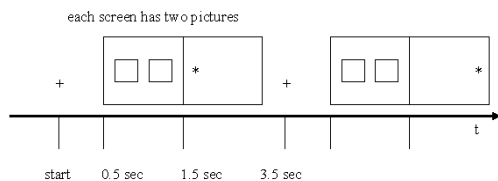

Figure 2 - Phase 2; block sequence of the affective protocol

The emotion recognition subsystem

Figure 3 shows a Logical View of the system Architecture.
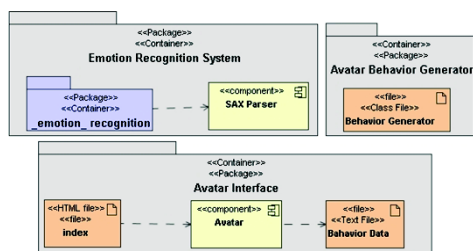

Figure 3 - Design model – logical view of the architecture

The emotion recognition system component is the most important element of the system since effective task exe-cution of all other system components depends on it. It is the first component of the execution sequence and the component that interacts with the environment outside the system i.e. the component that accepts the input from the user physiological signals. As mentioned before, that input comes in the form of XML data specifications. There are two input requirements that this system needs. The first is the filename of the XML representation for the input emotion to be recognized and the other is actually the invocation of the recognition procedure. The operation of this subsystem can be divided into three main tasks, namely, reading user data; recognizing the emotion by comparison to its knowledge base; outputting the recognizing emotion into a proper format.

The emotion recognition system currently requires different types of physiological data in order to perform a successful recognition on the user emotion: Electromyogram (EMG), Blood volume pressure (BVP), Skin conductivity (SC), Respiration rate (RR), Electrocardiogram (ECG). It is envisaged to include EEG in the near future to fully accommodate the needs of the affective protocol analysed above.

Apart from the signal data on which the actual recognition method is applied, additional information is required by the system. This is: (i) User identification (e.g. personal information, possible medical information etc), (ii) Signal sample data (e.g. time length, measuring unit etc.), and (iii) Measurement information (recording time, date, place, temperature etc).

Since the input to the emotion recognition system is an XML file having all the above required data, an internal part of this system is a parsing functionality that extracts only the data needed for the recognition process.

Emotion recognition is the key feature of this system and the "reasoning" part of it. The reasoning is actually a comparison between the emotion data obtained from the user and the emotions that the system can recognize. The number of these emotions can be variable and depends merely on how many XML data files are stored, each representing one "classical" emotion such as fear, happiness, sadness etc (Figure 4). The comparison is done between each of those emotions and the input emotion data by comparing the four signals independently. A simple method of comparison is chosen at the time being, as the scope is to merely test the feasibility of the overall methodological approach. To be more precise, following an extensive literature review on emotion recognition using pattern recognition techniques, sample signals were extracted for several emotions (anger, fear etc.). These samples can be considered as characteristic signals of the appropriate emotions and further matching can be done by simply comparing any new signal to each of them. The comparison is done using a weighted version of the dot product (cosine matching) technique on data samples of the signal. For each signal type we chose a number of characteristic features like the mean value, the number of peaks, Average amplitude of the peaks etc.

Avatar behavior generator

This subsystem is quite simple. Its one and only goal is to generate a file with specific parameters that will instruct the avatar behavior. The reason for considering it as a separate component is its functionality. The reaction of the entire system to the input emotion depends on this component, and for experimentation purposes in this project it is created with a simple classification function. The basic idea of the classification function is to discover the level of emotional reaction and to connect that to the output behavior. For example, if the emotion recognition system gave as output that it recognized anger at user side and with 80% correctness or the user was feeling very angry, this component needs to find out which parameters are the appropriate for such a strong emotional reaction on the user side. This component takes two inputs from the emotion recognition system: the recognized emotion and the comparison value. There is much freedom in the decision for the output behavior. In different studies or experiments, the implementation of this component varies. In this first project demonstration, it was decided only to mimic the emotional reactions of the user, in the facial area, since the avatar represents only facial human-like characters. The output is a file that has a predefined structure, since the avatar is deployed into this system and has specific requirements that cannot be changed. The Haptek platform was used for the avatar creation. The avatar has to be embedded into a web page and thus there are two kinds of inputs. First, is the input that the user gives to the web page for running the avatar; the second input is the file containing the behavior parameters. The web page contains the avatar and one button for starting the avatar behavior.
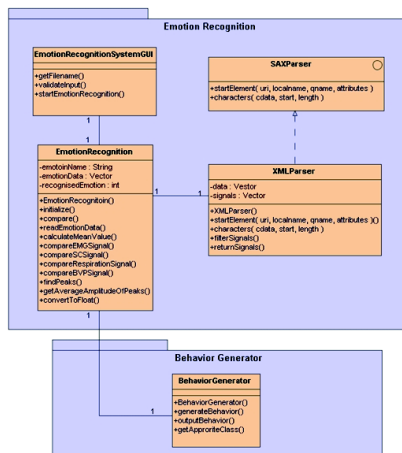
The main element in the file is the EmotionRecord. It contains four records, each corresponding to the physiological signals, taken from the user emotion for example, EMG, BVP, SC, and RR in the current version. Furthermore, the record element consists of the record data (the main element) and some additional information such as the recording device, the recording date etc. Figure 5 shows a portion of the XML specification. An example of the comparison based on the features of the signals is given in Figure 6 only for the skin conductance case, since the implementation of the rest of the features is quite similar. The avatar implementation is a combination between the HTML page representing the interface with the user, the scripting files with the avatar code and the external data file for the avatar behavior. The appearance of the avatar is possible by installing the Haptek Player plug-in, which can run the Haptek HyperText commands as the interface between the HTML and the JavaScript functions. The behavior of the avatar and its output to the user is triggered by pressing a kind of "Run button" at the moment, but it is envisaged to be automatically triggered in future versions.


```
...
<EmotionRecord studyID="ECG00001">

    <Record type="EMG">
    <!--Recording device can be here-->
    <RecordData>
        <Channel>MLII</Channel>
        <Waveforms>
        <XValues>
        <XOffset dataType="time">00:00:00.000</XOffset>
        <Duration dataType="time">00:05:06.000</Duration>
        <SampleRate unit="Hz">20</SampleRate>
        </XValues>
        <YValues>
        <RealValue>
        <From dataType="time">00:00:00.000</From>
        <To dataType="time">00:00:10.000</To>
        <Data>1.15,2.50,2.10,2.15,2.10,2.15 … </Data>
        </RealValue>
                </YValues>
        </Waveforms>
    </RecordData>
    </Record>
...
```

Figure 5 - XML specification of one record

Figure 4 - Design model (class diagram) – emotion recognition container and behavior generator


```
public void compareSCSignal()

    Vector storedSCSignal = (Vector)emotion.elementAt(2);
    float x=0;
    int numberOfSCR = 0;
    int numberOfPeaks = 0;

    for(int i=1; i<(storedSCSignal.size()-1); i++)
    {
        if( (Float)storedSCSignal.elementAt(i) >
        (Float)storedSCSignal.elementAt(i-1) &&
            (Float)storedSCSignal.elementAt(i) >
        (Float)storedSCSignal.elementAt(i+1))
        {
            numberOfPeaks+=1;
            if(checkForSCR(storedSCSignal,i) == true)
            {
                numberOfSCR+=1;
            }
        }
    }
}
```

Figure 6 - The compareSCSignal method for SC features

Results

The emotions stored in the emotion recognition system, as well as, the input emotions are represented as XML files. The file is a modified representation of the standard for ECG data representation described in the literature review.

Discussion

The intention of this paper was neither to introduce a new way of identifying an emotional signal, nor to put across a new affective protocol, but rather to carefully explain the individual methodological steps for the scientific exploitation of affective computing solutions. To this end, the

affective protocol design, currently under data collection, will provide the means to build up a knowledge base of emotional signals. One of the main strengths of this project was the idea for XML specifications of emotional data for the user. This can introduce new ways of emotion recognition by qualitative classification of emotions and not only by statistical quantitative experimentation. Furthermore, the contents of the XML data representation are represented by a modified standard for ECG patient data representation used elsewhere.

Figure 7 - Avatar appearance after emotion identification

The standard representation was modified to satisfy the needs of the emotion XML specification for this project. Obviously, the current XML specification of emotion data can be significantly improved by adding elements that will more accurately describe physiological signal data rather than only having numerical signal representations (e.g. EEG feature descriptions). Furthermore, the emotion recognition method based on the dot product comparison of the features can become more consistent so that it represents a proven classification of the emotions based on the results of the psychophysiological (affective) experiments. The generator of the avatar behavior is the component that can be also easily extended. Improvements can be envisaged in the "reasoning" method for generation of behavioral parameters. In specific, if the component is used in relation to larger amounts of emotion categories, it will be obviously more efficient and it can produce more classes of behavior parameters appropriate for the specific emotions. Finally, the avatar itself might go through several enhancements in its appearance to the user. The ultimate goal of embedding emotional awareness into the computer is to produce a system that can recognize emotions, and respond intelligently and appropriately in real-time, just like humans do. This paper provides a methodology for enabling the construction of more intelligent systems based on scientific reasoning and experimental results and not mere technological artifacts. The envisaged incorporation of the findings into a unified framework that ascertains the changing emotional state of the user and allows for adaptation of computer user interfaces based on the user and application context seems an exciting future prospect of this project. The wide variety of application areas that can be associated with such a system e.g. interactive games, learning systems, e-health/home care systems, etc call for a careful continuation of the project development.

References

[1] Reeves B, Nass C. The Media Equation: How People Treat Computers, Cambridge Univ. Press, 1996.

[2] Ghauoi C. Human computer interaction encyclopedia, Idea Group Inc., London, UK, 2006.

[3] Mayer JD, DiPaolo M, Salovey P. Perceiving affective content in ambiguous visual stimuli: a component of emotional intelligence. J Pers Assess 1990 ;54(3-4):772-81.

[4] Picard RW. Affective Computing. MIT Press, Cambridge, USA, 1997.

[5] Ekman P. Emotion in the Human Face, Cambridge University Press, New York, 2nd edition, 1982.

[6] Hone K. Empathic agents to reduce user frustration: the effects of varying agent characteristics. Interacting with Computers 2006; 18(2):227-45.

[7] Bamidis PD, Papadelis C, Kourtidou-Papadeli C, Pappas C, Vivas AB. Affective computing in the era of contemporary neurophysiology and health informatics. Interacting with Computers 2004; 16(4):715-21.

[8] Kim K, Bang S, Kim S. Emotion Recognition System Using Short-term Monitoring of Physiological Signals. Med. Biol. Eng. Comp. 2004; 42:419-27.

[9] Katsis CD, Ganiatsas G, Katertsidis N, Fotiadis DI. Emotional Understanding Using A Wearable Multisensorial Platform. In Bamidis PD et al (Eds), Improving the Quality of Health Information: an international perspective, Proc. iSHIMR 2005:188-97.

[10] Scheirer J, Fernandez R, Klein J, Picard RW. Frustrating the user on purpose: a step toward building an affective computer. Interacting with Computers 2002; 14(2):93-118.

[11] Healey J and Picard RW. Digital Processing of Affective Signals; In Proceedings of ICASSP, Seattle, 1998.

[12] MIT PhysioNet. PhysioNet: the research resource for complex physiologic signals. http://www.physionet.org/

[13] Varni A, Kemp B, Penzel T and Schlogl A. Standards for biomedical signal databases. IEEE Eng Med Biol Mag. 2001;20(3):33-7.

[14] Wang H, Azuaje F, Jung B, Black N. A markup language for electrocardiogram data acquisition and analysis (ecgML). BMC Medical Informatics and Decision Support 2003; 3(4).

[15] Lisetti C, Nasoz F, LeRouge C, Ozyer O, Alvarez K. Developing multimodal intelligent affective interfaces for tele-home health care. Int J Hum-Comput Stud 2003; 59:245-55.

[16] Bradley MM, Cuthbert BN, Lang PJ. Picture media and emotion: effects of a sustained affective context. Psychophysiology 1996; 33(6):662-70.

Address for correspondence

Panagiotis D. Bamidis,
Lab of Medical Informatics,
Medical School,
Aristotle University of Thessaloniki,
P.O. Box 323, 54124,
Thessaloniki, Greece.