5th Workshop on

Emotion in Human-Computer Interaction

- Real World Challenges -

held at the 23rd BCS HCI Group conference

1st September 2009

Cambridge University, Cambridge, UK

Organized by

Christian Peter, Fraunhofer IGD Rostock, Germany
Elizabeth Crane, University of Michigan, USA
Lesley Axelrod, Interact Lab, UK
Harry Agius, Brunel University, UK
Shazia Afzal, University of Cambridge, UK
Madeline Balaam, University of Sussex, UK

Kontaktadresse:
Fraunhofer-Institut für Graphische Datenverarbeitung
Joachim-Jungius-Straße 11
18059 Rostock
Germany

Telefon +49 381 4024-122
Telefax +49 381 4024-199
E-Mail christian.peter@igd-r.fraunhofer.de
URL www.igd-r.fraunhofer.de

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.
ISBN: 978-3-8396-0088-7

Fraunhofer-Informationszentrum Raum und Bau IRB
Postfach 800469, 70504 Stuttgart
Nobelstraße 12, 70569 Stuttgart
Telefon 0711 970-2500
Telefax 0711 970-2508
E-Mail verlag@fraunhofer.de
URL http://verlag.fraunhofer.de

Table of Contents

Emotion in Human-Computer Interaction – Real World Challenges

Christian Peter
Fraunhofer IGD
Joachim Jungius Str.11
18059 Rostock, Germany
+49 381 4024 122
cpeter@igd-r.fraunhofer.de

Lesley Axelrod
Interact Lab
University of Sussex,
Brighton, UK
l.axelrod@sussex.ac.uk

1. INTRODUCTION

The 2009 BCS conference's emotion workshop included a select group of researchers discussing 'Real World Challenges' for emotion in HCI. The day began with an icebreaker that required us to introduce ourselves in the style of various complex emotions which proved astonishingly difficult both to act out and for the rest of the group to recognise. A salient reminder of the challenges involved in this area and the fallibility of humans. This experience gave us plenty of insights: firstly, the development of emotion detection technologies are doing as well as can be expected (considering our own failings!) and secondly, they still have a long way to go...

The group became the subjects of affective research as we agreed to one of our numbers recording the day with a 'SenseCam' [1] as part of an ongoing project looking at affect in relation to recorded images. We drifted in and out of awareness of the continuous monitoring and the implications of our words of wisdom and affective states being recorded for posterity and later scrutiny!

We identified a number of topics we would like to work on: sense and nonsense of affective applications; problems with tools, particularly for data annotation and analysis; problems with studies, particularly the data aquisition; privacy and ethical issues; social aspects of our technologies and technology use; and what effect ubiquitous technologies and applications will have on the way we experience and show emotions.

A poster session that included some stunning visuals of robotic agents and applications gave an opportunity to present ongoing work, learn about each other's interests and to go deeper into project-related discussions.

We adjourned to the very pleasant grounds of the venue for our afternoon discussion where issues ranged around various topics including the respect due to people with autism who must work very hard every day to decipher and display emotional signals.

We were all pleased to note a growing emphasis in current work streams on the importance of context of affect – and a move away from basic emotions to representation of multi-dimensional complex emotional states. We had a short discussion why that might be and agreed finally that basic emotions had played an important role to get the field started, but that they are extremely limited when it comes to more sophisticated studies or implementing affective systems. Models and emotion representations, such as OCC [2] and the EmotionML standard draft [3, 4], were also discussed.

Context plays out in research in two ways: we need wider contextual information to make proper sense of emotion-related data, and we need emotion-related data as part of wider contextual information in order to design systems to support human behaviours. We concluded that emotion can't be sensibly separated from other 'meta' mental processes and should be considered in the context of information processing, decision making, cognition, motivation and so on. In fact, these processes influence each other and may result in physiological reactions, that can't easily be distinguished with sensors. There is a need for much more work to unpick all these areas in collaboration with other disciplines including psychology and sociology and in line with government and policy agendas and targets for well being and empowerment that have implications for work on affect. We felt the demographic shift to an ageing population will bring particular challenges for affective applications in health care and for collaborative learning and that HCI educators might need to increasingly address training students about affective issues.

Involving users in affective research poses particular problems for us as the language of emotion is so ambigous and it often proves difficult to communicate about our work with the public. Media reports tend to consist of sensationalist reports and we need to learn to make easily understandable 'cocktail party' explanations.

We progressed to a quite heated debate as to how suitable facial expressions really are for inferring a human's emotional state in the real world. Arguments in favour were that so much research has been done on this that it provides strong evidence that emotions result in unique patterns of changing facial expressions. Contrary opinions proposed that most facial activity is primarily for social communication, which might reflect underlying emotional states, might mirror the expression of the conversation partner, or might be a communication signal e.g. to show sympathy, interest, scepticism, dis-/agreement, etc.... In short, there is much more to read from a face than just emotion. We ended this day with a wrap-up session, reflecting what we had learned, which issues have been raised, what ideas we had, and what we ought to do back home again.

So what can we infer as to what that raised eyebrow means right now? We feel this debate needs to be continued and ideally underpinned with a thorough study - any volunteers?

2. ACKNOWLEDGMENTS

We would like to thank all participants for their lively discussions and valuable input they provided.

3. REFERENCES

[1] The SenseCam Project. http://research.microsoft.com/en-us/um/cambridge/projects/sensecam/

[2] Ortony, A., Clore, G. L., Collins, A., 1988. The Cognitive Structure of Emotions. Cambridge, England: Cambridge University Press.

[3] Schröder, M., Zovato, E., Pirker, H., Peter, C. & Burkhardt, F. (2007). W3C Emotion Incubator Group Final Report. Published online: http://www.w3.org/2005/Incubator/emotion/XGR-emotion-20070710

[4] Emotion Markup Language (EmotionML) 1.0 First Public Working Draft (2009). http://www.w3.org/TR/emotionml/

Putting the 'feel' into coding: FlowCoder, a prototype haptic coding tool

Lesley Axelrod Geraldine Fitzpatrick Eric Harris Madeline Balaam
University of Sussex Interact Lab, Informatics, Arundel 223, Falmer, Brighton BN1 9QH UK
44 (0)1273 872778
l.axelrod@ g.a.fitzpatrick@ erich@ madela@ sussex.ac.uk

ABSTRACT
Manual coding of emotion related (and other) behaviours is often necessary but is notoriously time consuming and tedious. Existing video coding tools are useful but all have limitations. We describe FlowCoder: a tool that is under development aimed to aid manual coding using multidimensional emotional spaces with haptic and auditory feedback and an alternative to the usual timeline bar display found in coding analysis tools.

Categories and Subject Descriptors
H.5.2. Evaluation/methodology

General Terms
Measurement, Design.

Keywords
Coding behaviour, FlowCoder; moodmat; measurement; haptic.

1. INTRODUCTION
Emotional complexity, multimodality and context are easily perceived and interpreted in real time by humans. Despite considerable advances in machine recognition of emotional states [8], manual coding of emotion related behaviour in research data such as videos is likely to be useful for some time to come. Manual coding is problematic because it is so costly of time and effort. Multiple run-throughs of video to mark discrete events or the onset and offset of various events or to validate coding have been estimated to need up to ten hours for 15 seconds of video [2].

We are interested in coding complex behaviours and emotional states, occurring in real life social situations. From our experience using different existing machine and manual coding tools, we collate here the features that we have found to be useful; while each tool has some features, none have all the features in the one tool. In addition a haptic element is introduced that may add value to the tool.

Existing coding tools and concepts have many interesting features that include the following:

Timeline bar displays are used in Transana, Anvil and Observer [9,12,16]. Mouse or keyboard entries add time stamps to transcripts or timelines displayed with video. Multiple behaviours can be defined, separately coded and displayed as differently coloured bars.
Likert scales such as the Self Assessment Manikin (SAM) [1] allow coding of the scale of a behavior (e.g positive to negative, valence or arousal). Scales are completed on paper or screen retrospectively. Resulting codes are not linked to video or timeline.
Two dimensions, colour and real-time representation are used in Feeltrace [3]. Input is via a mouse. Time line display shows activation on the y axis with colour shading under the curve to show emotion type.
Six simultaneous curves in real time show probability of different complex states in Facereader [4] using machine recognition to infer mental states and display them as colour coded probabilistic curves on a timeline.
Texture and colour conceptualise emotional states in the Sensory Evaluation Instrument and eMoto [5].
Continual measurement system (CMS) [10] is designed to facilitate continual measurement of affect by expert or non-expert coders using a joy stick or a mouse.

2. FLOWCODER
Our contribution is a prototype coding tool provisionally called 'FlowCoder' that combines the features we find most useful in other tools along with a haptic and sound enhanced interface (see Figure 1). We hope this tool will improve intuitive ease, quality and speed of manual coding and be useful with various behavioural and emotional scales and models. In this section we describe how FlowCoder is set up and delivers on the above features.

Our prototype involves using two screens and a haptic input device such as mouse or joystick. One screen displays a representation of behavior (a 'moodmat') with which the input device (mouse or joystick) interacts and the other screen displays the video to be coded with a timeline beneath it.

The representation of behavior might be a two dimensional representation such as a likert scale or affect circumplex e.g. SAM or Feeltrace or a more complex model e.g. Plutchik's 3D cone [1,3,14]. This graphical behavior representation or 'moodmat' is enhanced with sensory effects such as colour and/or texture (idea taken from eMoto and sensory evaluation instrument [5]).

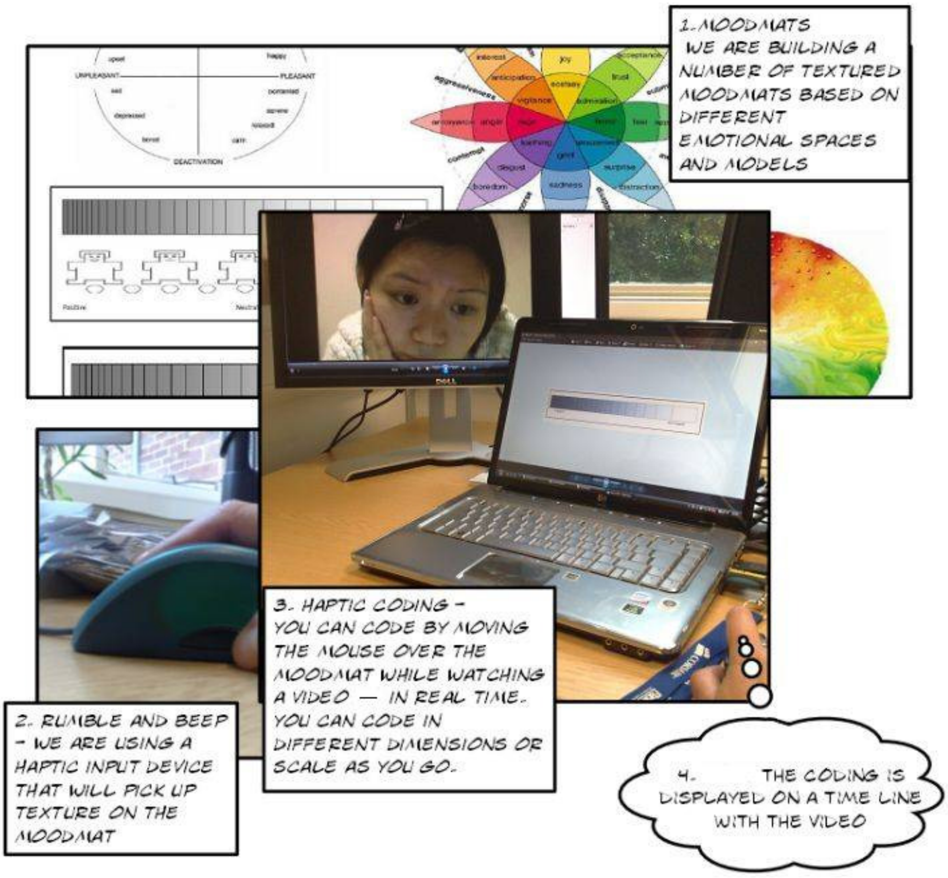

Figure1. Storyboard

As the input device (mouse or joystick) is moved across the behavior representation (moodmat) the user gets haptic and auditory feedback that tells them where they are in the space. For example if using a mouse and a likert scale the rate of 'beeps' and the sensation of texture increases as the mouse moves to higher point on the scale. Haptic and/or audio feedback reduces the need to look at the input device (mouse or joystick) and facilitates real time coding of emotional states with limited attentional overhead to the coding process itself.

The video to be coded is displayed on the second screen with an x/y axis timeline displayed below it (e.g. as in Facereader [4]). The haptic device is used to code the selected observed behavior and this is translated into a plotted line on the x/y axis graphical display. The timeline (as used by many analysis systems) is synchronized with the video [9,12,16]. Our timeline allows graphical display on an x/y axis (as in Facereader), which enables display of the degree of the coded behavior. For example (if coding emotional states) the degree of intensity. Our timeline can display multiple plotted lines (as in Facereader), so it is possible to code multiple different behaviours and see at a glance how they relate eg how much facial activity is going on as well as the degree of gestural activity. Alternatively the timeline can show simultaneous displays of coding of the same features by different coders allowing comparison of coding efforts.

Figure one shows a storyboard of how FlowCoder works.

1. Selecting a likert scale textured moodmat
2. mouse 'rumble' as it slides over the moodmat
3. coder watching video of emotional behavior on one screen and coding by moving mouse over the haptic moodmat displayed on the other screen
4. planned timeline for coding display using FlowCoder

The 'FlowCoder' prototype uses a Logitech haptic mouse, with a multimodal interface ('moodmat') developed using Immersion Software Development Kit [7] and IFeelPixel [6]. A video is displayed on one screen with a timeline displayed below it (see Fig2). The haptically enhanced likert scale (moodmat) is displayed on the other screen. The coder watches the video and moves the mouse over the moodmat to indicate coding of the emotional state portrayed in the video. Haptic and/or auditory feedback guides the movements, without the need to stop and look. The timeline display beneath the video has an x/y axis and allows display of simultaneous plotted lines. A vertical bar indicates the current point in the video. Each plot represents a coding input stream with the degree of perceived activity (y axis) and the point in time (x axis). Several coding efforts (made on different run-throughs) can be shown simultaneously. These might represent coded observations of different behaviours (e.g. degree of intensity and degree of valence) or coding decisions by different coders (e.g coder 1 and coder 2 for comparison or verification of coding decisions.)

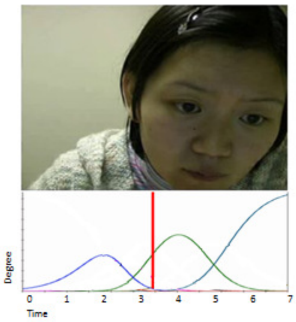

Figure2. Video and timeline display

Red vertical line shows point in time. 2 (or more) plotted lines on the x/y show:

- coding streams e.g. line 1 is amount of postural movement observed and line 2 is degree of boredom perceived.
- or comparisons of decisions made by different coders e.g. line 1 is boredom as coded by observer 1 and line 2 is boredom as coded by observer 2.

3. DISCUSSION

A growing appreciating of the complexity of emotion demands tools to reflect this [13]. In everyday life, people often rely on their 'gut' judgment of emotional behavior and there is some evidence that such emotion driven decisions are more reliable and valid. It may be, paradoxically, that the 'unnatural' degree of attention required for manual coding (constant pauses and replays) distorts interpretation. FlowCoder facilitates real-time manual coding, by relying on haptic and/or auditory feedback, so that the expert coder knows where they are in the emotional space represented, without needing to stop and look at the input device and consequently visual attention is not distracted.

Most existing manual coding tools rely on sequential coding of emotional dimensions, with levels carried out one at a time. For example using Observer to code degree of valence you must mark onset and offset of an intensity period and then indicate the degree of valence for that period. FlowCoder allows coding of 2D space with potential to explore 3D. Our timeline display shows the degree of any coded category such as perceived intensity or facial muscle activity.

Our timeline display allows us to compare several categories at once, offering fresh visualization of multiple activity patterns. Input from various coders can be compared, giving affordance for training coders or verification of coding decisions.

Many paths for future work lead from our early prototype. We plan to do comparative tests to validate use of FlowCoder and evaluate its cost effectiveness, usability and the quality of coding resulting from it. We will try a haptic joystick as an alternative input device and experiment with different emotional representations, including 2 and 3D models such as eMoto, Russell's circumplex and Plutchick's cone [14,15] and using a Falcon 3D input device [10] or gesture recognition. We will compare the haptic version against a tangible version of the FlowCoder. We will experiment with different visualizations of data on the timeline. There may be potential to add this type of timeline to existing and next generation computer assisted qualitative data analysis (CAQDAS) tools.

Our contribution is a prototype coding tool with multimodal feedback (auditory and haptic) to enhance usability. It can be used to code any observed behavior (not only for affect). The display on an x/y axis allows visualization and exploration of multidimensional emotion representations. It allows coding of degree of a category in a first pass on data. It may be easier to use for novice or non-expert coders. It has potential to enhance coding practice of experienced coders. It is quick to use and so will facilitate a 'quick and dirty' first pass on data. It has potential to assist in training inexperienced coders by allowing easy comparison of coding. FlowCoder may be more usable, more cost and time effective and provide higher quality coded data in a more detailed analysis of emotion.

4. REFERENCES

[1] Bradley MM; Lang PJ. Measuring emotion: the Self-Assessment Manikin and the Semantic Differential., *J Behav Ther Exp Psychiatry*, Vol. 25, No. 1. (1994)

[2] Cohn J; Ekman P Measuring Facial Action The new handbook of methods in nonverbal behavior research Jinni A. Harrigan, Robert Rosenthal, Klaus R. Scherer (eds) Oxford University Press, (2005) P35

[3] Cowie, C; Douglas-Cowie, E; Savvidou, AS; McMahon, E; Sawy, M; Schroder, M. 'FEELTRACE': An Instrument for Recording Perceived Emotion in Real Time. In Douglas-Cowie, E; Cowie, R &Schroder, M. eds, *Proceedings of the ISCA Workshop on Speech and Emotion: A Conceptual Framework for Research*, Pages 19-24, Textflow, Belfast, (2000).

[4] el Kaliouby R; Robinson P. Generalization of a Computational Model of Mind-Reading. *Proc 1st ACII*, pp 582-589, (2005)

[5] Fagerberg, Petra, Ståhl, Anna and Höök, Kristina (2004): eMoto: emotionally engaging interaction. In *Personal and Ubiquitous Computing*, 8 (5) pp. 377-381

[6] IFeelPixel http://www.ifeelpixel.com/download/

[7] Immersion software http://www.immersion.com/

[8] Jaimes;A; Sebe,N. Multimodal human-computer interaction: A survey *Computer Vision & Image Understanding* 108, 1-2 (2007)

[9] Kipp, M., 2001. "Anvil-a generic annotation tool for multimodal dialogue". In *7th European Conference on Speech Communication and Technology* (Eurospeech).

[10] Messenger D. S., Mahoor, M., Chow, S., & Cohn, J. F. (2009). Infant Smiling Dynamics and Perceived Positive Emotion. Infancy, 14.

[11] Novint Falcon 3d grip http://home.novint.com/

[12] Observer^{TM} Noldus http://www.noldus.com/

[13] Picard R. *Future affective technology for autism and emotion communication* Royal Society London (2009) available http://royalsociety.org/page.asp?tip=1&id=8454

[14] Plutchick R; *The nature of emotions* (1980)

[15] Russell, J. A. A circumplex model of affect. *Journal of Personality and Social Psychology,* 39, (1980).

[16] Woods, D, and Fassnacht, C. (2007). Transana v2.2x. http://www.transana.org. Madison, WI: The Board of Regents of the University of Wisconsin System.

Emotions in Technoscience: the performance of velocity

Simone Belli
Universitat Autonoma de Barcelona
B5-032 Social Psychology Dept.
08193 Bellaterra (Barcelona) SPAIN
+34 935811704
simone.belli@uab.cat

Rom Harré
Rom Harré
Georgetown University
303 White Gravenor,Psychology Dept.
20057 Washington, DC, USA
+01 2026872117
harre@georgetown.edu

Lupicinio Íñiguez
Lupicinio Íñiguez
Universitat Autonoma de Barcelona
B5-034 Social Psychology Dept.
08193 Bellaterra (Barcelona) SPAIN
+34 935812874
lupicinio.iniguez@uab.cat

ABSTRACT
In recent years, the topic of emotions has been influenced by postconstructionist research, particularly by the use of performativity as a key concept. According to Judith Butler (1993, 1997) the construction of emotions is a process open to constant changes and redefinitions. The final result of the emotion-language "natural" development is what is known as Technoscience. In this realm new ways of naming emotions have emerged. In our research on the use of Information and Communication Technologies (ICT) by Cyber-Cafés and Call Shops users, we came to understand how these technologies are significant in those users' daily life. The emphasis has been on analyzing emotions related to the use of ICT in the aforementioned settings. Using the concept of performance (Butler, 1990), we explore how discourse creates a need for particular emotions, which did not exist before their performance. To understand this performance, we use an ad hoc tool called Membership Categorization Analysis (MCA) as it is used by the Manchester School. Analysis has revealed a membership category in which velocity is salient as well as performance and primary emotional content constructed through language by users of ICT. This 'velocity', produced by discourse, seems to follow the evolution of Technoscience in Social Sciences, placed in the context created by two concepts, Donna Haraway's (1990) cyborg and Alessandro Baricco's (2007) mutant.

Categories and Subject Descriptors
J.4 [Computer Applications]: Social and Behavioral Science – Psychology, Sociology.

General Terms
Performance, Human Factors, Languages.

Keywords
Emotions, performance, velocity, Membership Categorization Analysis (MCA), Information and Communication Technology (ICT).

1. INTRODUCTION
In recent years, the topic of emotions has been influenced by postconstructionist research (Iñiguez, 2005), particularly by the use of performativity as a key concept. According to Judith Butler (1993) the construction of emotions is a process open to constant changes and redefinitions (Butler, 1997). In this vein, the ultimate effect of the "natural" evolution of emotion and language is the appearance of Technoscience. New ways of naming emotions have emerged from the viewpoint of Technoscience (Belli, S., Harré, R., Iñiguez, L., in press). Due to Information and Communication Technologies (ICT) there exist new emotional aspects in which philosophers, psychologists, and epistemologists have zeroed in their common interests, for instance, the affective machine (Rose, 1983, Brown 2005, Brown, Stenner, 2001, Michael, 2000, 2006), the notions of cyborg and techno-disembodiment (Haraway, 1989, 1995, James, Carkeek, 1997, Gibbs, 2006, Hollinger, 2000), or the notion of "disclosure", fuzzy phenomenon meaning the expression of emotions through a screen.

In our research on the use of Information and Communication Technologies (ICT) by Cyber-Cafés and Call Shops users, we came to understand how these technologies are significant in those users' daily life. The emphasis has been on analyzing emotions related to the use of these technologies in these specific settings. Using the concept of performance (Butler, 1990), we explore how discourse constructs a need for particular emotions, which did not exist before their performance. These events or "acts" are seen as natural through their repetition over time. They are a set of multiple everyday social interactions. To understand this performance in discourse, it is necessary to use a "tool". In our case the tool is called Membership Categorization Analysis [MCA], as it is used by the School of Manchester. The particularity of this approach is that recognizes emotions as a performance continuously changing in language, that is, they are iterative and progressive. This particular symbiosis between emotions and ICT opens up a new area of study in which emotion is understood as a symbolic process depicted discourse at its best. For this reason it is fundamental to see emotion as a process linked to the way people face and use ICT. Analysis has revealed velocity as a salient membership category. It is constructed by users through discourse as a performative and primarily emotional process. Thus, discursively produced 'velocity' seems to follow the "natural" evolution of Technoscience in Social Sciences, such as it is understood through by the concept of cyborg (Haraway, 1990) and the concept of mutant (Baricco, 2007).

Haraway's concept of cyborg helps us to understand why the relationship between the individual and the machine is like an extension of the person herself, created by discursive performance. This relationship constructed through discourse helps us to introduce what we will talking about in this article, that is, an analysis of ICT users' interviews in which a new emotional performance, i.e. velocity, emerges. As already asserted, we believe the use of ICT as an extension of one's own body, as stated by the concept of cyborg, is constructed through language. Hence, its use is always discursive. Regarding the relationship, or conflict, between individual and society, the discursively performative use of ICT constructs the former as well as the latter in a continual iteration of speech acts.

2. WHAT IS AN EMOTIONAL PERFORMANCE?

When it comes to the issue of performance in postconstructionist studies surely the name of Judith Butler is the most important. Butler (1993, 1997) considers emotion as a process constantly evolving in actual discourses. She offers a depiction of emotions in a completely new perspective through the concept of performance. Although in their texts the term "emotion" is not explicit, it spontaneously arises out of people discourses and positions. Using the concept of performance, Butler (1990) explores how discourse constructs a need for particular emotions: "Such acts, gestures, enactments, generally constructed, are performative in the sense that the essence or identity that are intended to express, manufacturing constituted and sustained through corporeal signs and other discursive means" (Butler, 1990: 136). In this way, emotion is a performance produced by these constructions. It is comprised by acts which are internally discontinuous; that is to say, emotions do not exist prior to their performance. There cannot exist a successful "copy" of an emotion of which one could expect to faithfully reproduce a previous event or a new emotion. These events or "fabrications" are seen as natural through their repetition over time. Eventually, they become standards, and they come to be seen as normal. They are taken as a set of multiple daily social interactions. Nevertheless, these performative acts are open to constant change and redefinitions.

Another very important characteristic of this concept is iteration (Butler, 1993). Butler uses Derrida's theory of iterative acts for deepening the understanding of performance: "The performativity cannot be understood outside of an iterative process, a controlled and limited repetition of norms. (...) This means that the iterative performance is not an act or event, singular, but a ritualized production (...)" (Butler, 1993: 95); in short, "a stylized repetition of acts" (Butler, 1990: 140).

3. FIND THE PERFORMANCE IN SPEECH

A key task for researchers is to find theory in practice, that is to say, finding performance in discourse. Thanks to MCA, as taken by the Manchester School, we have fulfilled this objective, for we have understood the role of performance in the study of emotions in ICT settings.

MCA is a formal analysis of the procedures used by people to make sense of other people and their activities. Sacks (1992) developed the MCA view from the writings of Lacan and Chomsky and his colleague in those years was Emanuel Schegloff. Harold Garfinkel also contributed to shape the MCA approach from the ethnomethodological perspective. He established as the main task of the method the depiction of how people categorize themselves and their worlds. Membership categories are comprised of "categories related to predicates, which may include dispositions, acts, tasks, beliefs, and values" (Watson, 1978). Most of MCA categories are relatively crystalline (e.g. "mother", "son") and we can discuss how these categories are used in discourse. But one of the main problems is that these categories are treated as if they were fixed. This problem especially emerges in social and political identity of the body researches, for new categories appear and disappear throughout the analysis. This particular characterization goes through and implies discourse. That is, we are interested in how participants manage the relationship between categories (Leudar, 1998). Their management is categorized into "collections", "classes", and "relationships". But, in our research, what interests us most is what Leudar (1998) calls "distributed discursive network", that is, how categories and relationships interactively emerge from discourse (Leudar, 1995, 1998; Nekvapil and Leudar, 1998).

Ivan Leudar's perspective primarily differs from Sacks' view because of his particular way of understanding "categories" of codes. According to Leudar, these "categories" of codes are constantly evolving inside discourse. They are never fixed. Instead they change continuously, and a member of a code can often change the category itself. It is an open and circular dialogic process. The perception of this categorial change is due to linguistic and lexical factors produced in social practices (Alvarado, Belli, Iñiguez, in press).

Another important aspect of Leudar's approach is that categories must be active. His conception of MCA is mainly defined departing from that key element. Leudar depicts categories as a circle, a bounded system which, at the same time, constantly evolves, coming and going from open to closed and vice versa.

Thus, discourse can be considered as an activity, since it is considered in an active fashion, and situations themselves are very active too, which is known as "talking in activity."

4. EMOTIONS IN TECHNO-SCIENCE

We saw that emotions have a strong relationship with language (Belli, Iñiguez, 2008), especially as regard the concept of performance (Butler, 1990). We will focus now on the "effects" of this approach in everyday language. Performance is better understood when we observe that some emotions appear and disappear in ordinary discourse. This is so because emotions are embedded in narratives. They are speech-acts (Oatley and Jenkins, 1992:75).

Performance is subject to the power of the discourse of emotions. The discourse of emotions is full of smaller discourses on new concepts and metaphors which serve to articulate and understand the lexicon of emotional words. Concepts such as techno-disembodiment and human-affective machine are the techno-scientific side of emotions. They represent new performances, new emotions appearing in the discursive arena.

The emotions keep on changing in the natural and spontaneous language of everyday life. New words enter in the discursive arena thanks to performance. Thus, new areas are produced in the technology sector (Belli, Harré, Iñiguez, in press). This is how emotions begin entering the technological discourse as another performance in everyday discourse. The concept of performance related to ICT, produces a new narrative in social sciences, such as techno-disembodiment and the affective machine. Only in recent years it has been understood that emotions and new technologies have a very close relationship. Especially in affective settings, we have had the opportunity to observe it several times; for instance, a mother speaking by phone with her children and family while she is crying; an immigrant that "goes out" to party with her friends in their country of origin through Internet; a chat conversation between young lovers who actually are separated by thousand miles.

The examples we have shown in the preceding paragraph are used to support the thesis that emotions have to be interpreted in the social context in which they occur. Thus, there is no wonder if emotions appear and disappear in the discursive arena. A recent field of research concerned with emotions is Technoscience. Mike Michael (1996, 2000, 2004, 2006), with a clear semiotic method, recognizes emotions as affective matter.

In Technoscience the topic of emotions is also related to the semantic conception of embodiment. It does not consider the cognitive treatment. It emphasizes on the communicative and linguistic side. The main authors addressing this issue are Haworth (1990), Niedenthal, Barsalou, Ric, and Krauth-Gruber (2005), Prinz (2005), Lyon (1999), Katz (1996), Haworth (1990), Malin and Peterson (2001), and Harré in his article "The necessity of personhood as embodied being "(1995). In a more extreme sense, sometimes embodiment is treated as techno-disembodiment, such as James and Carkeek (1997) do.

These different discourses on affect and emotions in Technoscience can be seen in the context offered by Nikolas Rose (1983), with the concept of affective matter and the figure of the machine that constructs the individual, i.e., the affective machine. This vision is embodied by the figure of the cyborg (Haraway, 1990, 1995; Hollinger, 2000). Also Gergen (1990) gives a reinterpretation of the affective matter in the postmodern society.

Steve Brown and Paul Stenner (2001; Brown, 2005) speak of collective emotions in Technoscience and consumer society, by using the writings of Spinoza as well as more recent authors such as Schaub (1933) who states that the affective matter becomes a human-machine being.

When we are before a screen, it is generated a series of mechanisms from which emerge our most intimate thoughts and feelings. The scientific narrative fiction has labeled this process with the term disclosure (Aviram, Amichai-Hamburger, 2005; Qian, Scott, 2007). Disclosure is one of the most striking aspects that we have found throughout the analysis of the interviews we have done for our research. Disclosure helps us to explain how we love to talk with strangers, or someone that we already know, but only through a screen we dare telling them things we would never say vis a vis, specially things we only express in private settings like emotions themselves. This allows us to understand the success the use of these technologies have in different aspects of life. Phenomena such as Facebook ensure just that. We contact someone who we already know, but getting into a more intimate mode, where we can express our hidden emotions or take public relations to another level and all through language. This kind of reasoning permits us to rethink in other terms the affective relationships, the intimate aspect of new technologies, and the emotions we find when we are in front of a computer's screen.

Technology allows us to measure, quantify, and identify in real time people's emotional states, how people communicate each other emotionally, and consequences in the realm of machines. The concept of techno-disembodiment as defined by Paul James and Freya Carkeek (1997: 107) is "an increasing abstraction of the way we live our bodies and widespread technological mediation of social relations." Carkeek and James argue that the strength of this concept is related to an emotional charge attached to an erotic-romantic residue, such as techno-sexuality; i.e., sexual intercourse without the actual presence of the other or with the technological representation of a sexual organ, the wide range of phone-sex and sex-chat practices up to cosmetic surgery. These examples illustrate a general and widespread emerging process in postmodern times.

An example of an affective machine is the construction of technological systems that can interact with humans and transmit bio-psychological changes, e.g. the use of shoes, bracelets, t-shirts that receive and transmit emotional changes to the individual, so that changes in emotional states are registered in their social contexts. These gadgets are just a phase in man-machine interaction, i.e., the human-affective machine.

It strikes us very attractive to talk about file-selves (Harré, 1983). For Harré, the self may be present in many different places and may take part in various episodes, though the body is actually located in one space and one time. For example, a person may be playing tennis and, at the same time, may be being evaluated via her curriculum vitae for a job, or a bank may be making a money transfer on her behalf, or police may be investigating her as a potential terrorist, etc. This is possible because any individual can be distributed among a huge range of files, i.e., recordings of personal information. My body and my accompanying 'file-selves' get along together.

If we take this example, we can understand the way it has become a part of everyday life. A person is working on-line, while their friends watch her photos on Facebook, comment them, her boyfriend sends her an e-mail, her mother calls her using Skype, a co-worker sends her a SMS, and yet you can add what has been previously suggested by Harré. For these reasons, we must understand the use of new technology as an extension of self.

Haraway's cyborg concept helps us to understand why the relationship between the individual and the machine would be like an extension of the same individual, created through discursive performances. In turn, this relationship introduces what we will try to set out in the next section, i.e., an analysis of interviews with users of new technologies, in which a new emotional performance appears: 'velocity'.

5. THE EMOTIONAL PERFORMANCE OF VELOCITY

"Use Gmail on your mobile, it's super fast!"
(http://mail.google.com 10th October 2008)

In interviews an emotional performance continually emerges as a membership category. It is always present, and without any doubt, it is the main candidate for the study of emotions in the ICT realm. This emotional performance has spontaneously emerged and has always taken different connotations. It has been growing as an ever-changing category, i.e., as a performance. This phenomenon can be only understood through MCA, according to the Manchester School. As we saw before, this helps us to find the performance in emotional discourses. Thus, for the reasons already described, 'velocity' appears as a "new" emotional performance in the technoscience discursive arena.

Velocity is an emotion that continuously emerges in the interviews with Cyber-Cafés and Call Shops users, i.e., ordinary people who daily use ICT and who are not computer professionals. Through daily use of ICT, we have understood the importance of this emotional performance as a basic research priority.

This performance is an extension of our body and it is used to interact with others through ICT. It is a discursive production, since it is only through discourse that it can be constructed. It is like the concept of cyborg (Haraway), but entirely linguistic. In the next section, we will profile what an emotional performance is, and we will try to outline and see its course, finding a way to define 'velocity' without minimizing the importance of the constant and continuous evolution of its performance, its iteration. We will trace the path of this performance through discourse analysis and MCA.

This performance has been characterized by three aspects, which have emerged inductively: speed, efficacy, and ease. These three features have been useful throughout the analysis to find this performance in the text, and to understand their iteration.

This is a Membership Category [MC] related to the Bound Activity [BA] (i.e. activities, actions, verbs) associated with ICT comprising use, progress, time, and change. Following the cyborg concept, this use of one's own body as an extension is constructed through language; in other words, this activity is always discursively constructed. Thanks to language, it is possible to define this extension, this use of ICT, this BA, this performance. Regarding the relationship, or conflict, between individual and society, through the use of ICT and this discursive performance, is created a continuous iteration in order to relate with the others.

Thus, this performance is based on the use of technologies. This use is discursively constructed through the emotional performance of velocity. In the beginning of the analysis, 'velocity' appears as a cold emotion transmitting very little, but seducing because of its own performance. But in the long run, analysis leads us to see that velocity produces real emotions. Velocity goes from being cold and unemotional, or as Illouz (2007) says "a frozen emotion", and becomes a completely true emotion, not cold at all. We can see people laughing or crying in front of a screen (Belli, Iñiguez, 2008).

The use of the term "technology" refers to a whole list of terms related to a way of understanding emotions as cold, unemotional. This list includes design, operation, purchase, etc.. It is a list of terms that cannot be regarded as similar to the way we have specified emotion before. When it comes to talking about the emotional side of technology, 'unemotional' is the usual term to designate this realm. This term refers to the use of a "cold" device, like the computer, to communicate with other people. The cold light of the screen (Baudrillard, 1990:153) produces a cool seduction. This is the narcissistic charm of new technology, the cool charm of the Internet. But this absence of emotion, when related to the use of ICT to communicate with others, changes, evolves. It becomes a performance, an actual emotion.

<u>Extract 1</u>

```
210. -yeah ... a lot ... because they agreed to meet at this hour ... hour,  in the afternoon over there ... there ... morning over here …

211. The web-cam is wonderful ... you see your wife and your kids ... your family ... mother or your...

212. grandparents and two ... three hours ... there … crying. You'll see some crying ... mourning ... saying things

213. very very sweet ... then you hear and … it hits you ... doesn't it?
```

<u>Extract 2</u>

```
215. I... once I've got drunk with some friends watching the webcam

216. I was ... ah ... and cheers! ... (laughs) and drank too (laughs) and we've got drunk (laughs) Good

217. and "Bye" (laughs) ... It kills yearning... the longing ...
```

In extracts 1 and 2, we observe how this performance emerges and we can understand its "history". It is an iteration that changes from a unemotional condition, a cold emotion, until becoming a colorful real human emotion. But all this is possible to understand in a more detailed manner because of the extracts that we will be presenting.

We have considered appropriate treating this aspect of performance in this first part of analysis, highlighting the importance of the iteration process of this performance. Continually, we see technology as something cold, but suddenly, when we relate with it, we find out that it produces a huge amount of real emotions. Thus, "color" and "heat" change along with our analysis. We have just given the parameters, since it is not possible to define a priori what a cold emotion is or what a real emotion is. It is important to understand that at no time prior to analysis, can 'velocity' be defined as a MC, since it is during the analysis itself that this MC is defined. Only when a MC appears, it is the time for identifying and defining it as a MC. This MC is flexible, open, and dynamic, and in this way it can give meaning and evolve through discourse. Only treating this emotion as a performance, i.e. 'velocity', we can understand this iteration.

<u>Extract 3</u>

```
253. C: Interne: t (.) Maybe: ye: (0.5)  Internet and mo: bi: le

254. E: Yeah? (0.3) Is it more comfortable?

255. C:> It's the most comfortable <more fast: and (0.2) effective.
```

Velocity is a performance that has much to do with the linguistic metaphor of surfing, in the context of Internet. This term is taken directly from riding the waves on a beach. Baricco says (2008, p.111) “Don’t you see the lightness of that brain which is hooked on the foam of the waves?” To browse, i.e., to surf always refers to velocity, but not to a static velocity, but to a dynamic or performative one. Surfing is a very smooth motion, but we must be quick or we won’t have a chance to balance and we will fall into the deep sea. Such is the emotion of velocity in ICT. Nothing more than that. Baricco uses the example of Google, and how people “breathe” through it. They breathe, run, and surf in Google. These are new models and new skills (Baricco, 2008:114). In Google, there are paths across the surface which can be followed quickly, effectively, and easily. There are rapid sequences, drawn trajectories for surfing, movements linking different points in the space of the real. Velocity is a new emotional performance, because before the advent of ICT it was different. Before, we had to approach things one by one, to deepen relationships and intimacy in the long run. It was a patient task, a matter of study. For example, consider the review of scientific literature on a particular topic. We had to make a full and complete reading of a text for making sense of it, for seeing if it might be interesting for our research. Nowadays, this is not the case. You know at once, by looking for a keyword in the text, if the article or book might be useful for your research. Velocity pops up fast as a complete gesture, as a performance. Thus, experiencing requires to obtain a piece of information which is enough to lead to another piece of information. If the expectant and patient “mutant” stopped in each of them, as a man with lungs would do, the performance would be fragmented. The mutant breathes through gills to survive in Internet. He uses the search engine to see whether the text is useful for him or not. The man with lungs still has to read his text completely to know whether what he has in hands is helpful for him or not. Baricco’s metaphor is powerful and useful to explain why the emotion of velocity emerges from ICT. The performance is clear if we think of Baricco’s mutant as a discursive evolution of Haraway’s cyborg, which has learned what are the minimum and maximum time for doing things:

Extract 18

222-I use the computer in my leisure time ... for example ... when I have

223 nothing ... or I’m all alone ... and I’m not with Estefania… or my other friend is not at home... then I…

224 turn on the computer ... enter to Peru’s webpage ... I read the news ... my country ... what happened

225 I enter my email ... or listen to music … until

226 my body gets tired or I get bored and go home ... but ... I use it as a second...

227 option ... since the first one is phone. I use it when I’m idle

228 / if there is something / ... “Give me five minutes” ... or “give me an hour or so ...” “Oh, I’m going to check...

229 who has written” ... “I’ll check my mail” ... So I read my email ... and then I go

230 to the news from my country ... then I go to a tarot webpage…

231 any chafardeo [gossiping]... or for killing time ... ... but those are the moments when …

232. /-PAGES OF PERU ... THE TAROT ... …

233. / all ... for instance I enter the site Peru dot com ... which is a page that

234 gives you ... news ... all the things happening in Lima ... everything ... everything on my country ... right?

235 ... then, from there, when I get bored I visit a women’s site on tarot…

236 it’s a piece of nonsense to waste time ... to have fun ... whatever it comes

237 If I get bored... I visit a music and video clips site ... and if I continue bored,

238 well ... that’s it …

239 NO LONGER THERE …

240 the machine is over ... and I go home ...

In this extract, we can understand what velocity means in performance: a quick surf which does not permit to get bored by using a technological device. This velocity is an emotional performance allowing the individual to keep himself far from the “deep”, which, at this point, for him is an unjustified waste of time. It is a futile impasse (Baricco, 2008:115) which destroys the flow of performance. Emotion is found in performance, and such a performance must be fast or nonexistent. The deep is boredom, and the surface is entertainment (Extract 18, lines 236-237).

In order to move quickly on surface, everything is simplified. It is a rate of experiences which allows us to put them in sequence with other things. This is the emotional performance of velocity. Everything has to be on the surface, not in the depth. To facilitate the flow of communication, there must be a common language and a universal grammar based on movies or television (Baricco, 2008:115) .

This performance of velocity looks for gestures to enter easily the cyberworld and to get out in a equally easy way. It privileges which creates a movement, a performance, and any space generating an acceleration, an emotion. The goal of velocity is movement, performance itself. It does not seek experiences, it is experience itself. As seen in the last extract, everything becomes fast, effective, and easy systems until the individual gets bored (Extract 18, lines 223-226).

Boredom is slowness. It means going back to the queue (Extract 19, line 193), a natural component of past time. The past refers to boredom and slowness. It has nothing to do with velocity, i.e., a new emotional performance characterized by speed, efficacy, and ease. This is why today we try to get rid of boredom and slowness. To a contemporary child, Baricco explains (2008:116), boredom is almost unknown. He or she is continually doing things. He or she is engaged in various activities in different levels and contexts. If one decreases speed, one falls from the bicycle. The metaphor of the bicycle helps us to understand the performance of velocity, the need for fast and constant movement for not getting bored; for example, surfing the net, and its superior level: multitasking. Baricco (2008:116) defines the phenomenon of multitasking with the example of a child who can play Nintendo, eat a hot-dog, phone his grandmother, watch cartoons on TV, pat the dog with one foot and whistle the T-mobile melody. Or a teenager who does his homework while chatting on Messenger, using his iPod, sending sms, searching Google for the address of a pizzeria and playing with a rubber ball. Multitasking, of course, must be done quickly. If it were otherwise, it would cause negative emotions. Not replying an email after a reasonable interval of time, not being available on Messenger or not answering the cell phone are disastrous effects in everyday life. Velocity, therefore, must be a continuous and constant performance. Whether it is a matter of genius or idiocy, the brain is on fire for more velocity. This velocity allows us to inhabit whatever areas we wish with a rather low amount of attention, and this is what this emotion is all about. It is a way to do many things in a single gesture, in a single moment. In order to do so we have to use the three feature of the velocity performance: speed, efficacy, and ease. These three features ensure that we do not isolate any gesture in the course of multitasking. This can be understood due to the concept of performance, since we have to observe at the very moment it occurs. It is understood as a contemporary movement only. It is an evolution of the ICT practices, i.e., it is a continuous performance.

Surfing must always be fast. It is like riding a bicycle. We never stop to see the sea, the beach or the boats, because we see them while we are pedaling. And if we stop, it is only for a short time. Then we continue with our movement.

Using new technology helps us to get closer to others, and this we do through language. Thanks to language we can produce an emotional performance of velocity, in an instantaneous and intuitive way. It is something we have as individuals. It has evolved in recent years. In the past things did not have such an aspect. Nowadays, we have a constant evolution, a continuous approach to ICT. We have come to understand ICT as an extension of our own body, as on the cyborg model of Haraway. It is a comprehensive symbiosis as the mutant concept of Baricco asserts. Thus, we no longer see new technologies as cold and impersonal machines. Instead, we endorse them, emotionally speaking, as part of us. Before going to bed or when we wake up in the morning, often the first things we do is to check our inbox for new email, and such an act is already part of our everyday life.

In this sense, velocity is a basic emotional performance for the study of ICT, a primary emotion. It must be taken into account when approaching the study of new technologies.

Now then, if we go back to the first part of this analysis (Extract 1), we can now understand in a more comprehensive way velocity, the use ICT, the progress, the relationship with others, the mastery of time, and finally, the features of speed, efficacy, and ease. These are the constituent parts of what we call emotional performance.

Have we now understood this transition? Individuals see new technologies as parts of themselves, allowing them to interact with others, laugh or cry in front of a flat screen, but it is as if we cry or laugh with the person who is on the other side of the screen, our son or our friends. We do this through language, offering an infinite number of emotions, in what we say is an emotional performance. Velocity is always present in this relationship. Velocity is a constant when we use new technologies. Velocity is what allows us to move from understanding technologies as cold and unemotional, to something that is part of us and produces real emotions. Velocity allows us to understand that machines are extensions of our body, an integral part of our body, and we no longer believe they are just machines. The transition from the model of the cyborg, (body and machine together), to the mutant in which a distinction between the body and the machine no longer exists, is part of an advanced biology, a natural evolution of human beings.

6. CONCLUSIONS

Emotions have a strong relationship with language. We can express emotions through language. Express emotions means putting something in common with others. It is possible to express emotions according to the concept of emotional performance developed by Judith Butler. Emotions are not fixed, defined, and static. They are constantly evolving, continuously implied in an iteration process, and they do it through language, which is natural and subjective. This constant iteration makes them appear and disappear from the discursive arena. We can see how we abandon some forgotten emotions and discover new ones in Technoscience (Belli, S., Harré, R., Iñiguez, L. in press). Emotions undergo a constant evolution in everyday discourse and it has its highest expression in Technoscience. In fact, the term emotion can be associated with very specific areas in Technoscience, thus emotional expressions come into play in the discursive arena. The emergence of concepts such as techno-disembodiment, or the emotional connection between people and new technologies, namely the concept of machine-affectivity of Nikolas Rose, are some examples of what the current scope of Technoscience embraces. They allow us to use the concept of disclosure to think of love in the age of Technoscience.

In this article, we treated the concept of emotional performance, which we used to argue for the thesis that emotions are changing in discourse. Surely there is an infinite variety of emotions occurring in the individual when using new technologies to communicate with others. They can range from boredom to happiness, from anger to love. We do not know them independently, since only through language, irrational and subjective, can we know them. When addressing the issue of Technoscience, language produces what we call an emotional performance. Performance which emerges in the course of the analysis is velocity. Velocity has three features: speed, efficacy, and ease. Velocity is constructed through language, first as a cyborg, since technologies are an extension of our own body. They are still seen as cold and unemotional machines. But throughout the analysis we have shown how, due to the performance of velocity, the individual becomes a mutant. The technologies are part of our own body and they are no longer cold and alien machines, since we have been metabolized into them. Their use has become part of our daily practices. They are not understood as machines, but as a means to interact with others and their discourses, which reveals the concept of emotional performance. This performance must be fast, and without it, it is nothing or a return to slowness. It has to be multitasking, to do everything at once, and it must be well done. This allows us to feel excited, and to express it through language, using the emotional performance of velocity to indicate what is happening to us.

People can feel happy, bored, etc.. using new technology, but the relationship, or conflict, between individual, society, and ICT is always discursive, always constructed as a discursive performance, and in this performance the issue of velocity emerges. What is expressed is velocity, and not happiness or boredom. We talk to others using this category of velocity, which changes continuously. It can never be defined.

The extension of our own body constructed discursively provides the emotional performance of velocity. Through this characteristic it is argued that our discursive performance is defined. It may happen that the interviewee feels happy because of the purchase of technology, but that is something we do not know. Happiness on purchasing a new blackberry, boredom by having a previous version, these are reasons for wanting to change things. But, in terms of discourse, it means technological advance, which falls under the emotional performance of velocity. It may be that the individual experience certain amount of emotion, but we do not know. We can only know what he mentions when he uses new words for explaining emotions that already exist. This coincides with a change in the discursive arena. We use new terms when discussing the relationship between individual and new technologies. This extension of our own body, the emotional performance, adopts new terms. And velocity, though it may seem extravagant to take such a performance as emotional, is one of them.

7. REFERENCES

[1] Aviram, I., & Amichai-Hamburger, Y. (2005). Online infidelity: Aspects of dyadic satisfaction, self-disclosure, and narcissism. Journal of Computer-Mediated Communication, 10(3)

[2] Baricco, A. (2008). Los bárbaros: Ensayo sobre la mutación. Barcelona: Anagrama.

[3] Baudrillard, J. (1990). Seduction. New York: St. Martin’s.

[4] Bauman, Z. (2008). Happiness in a society of individuals. Soundings London, 38, 19.

[5] Belli, S., & Íñiguez-Rueda, L. (2008). El estudio psicosocial de las emociones: Una revisión y discusión de la investigación actual. PSICO, 39(2), 139-151.

[6] Belli, Simone; López, Cristian; Romano Javier (2009). “Producción, distribución y consumo de conocimiento: ¿un virus fuera de control?” Revista Iberoamericana de ciencia, tecnología y sociedad.

[7] Brown, S. D. (2005). Collective emotions: Artaud’s nerves. Culture and Organization, 11(4), 235-247.

[8] Brown, S. D., & Stenner, P. (2001). Being affected: Spinoza and the psychology of emotion. International Journal of Group Tensions, 30(1), 81-105.

[9] Butler, J. (1990). Gender trouble. New York: Routledge.

[10] Butler, J. (1997). Excitable speech: A politics of the performative. New York: Routledge.

[11] Butler, J. P. (1993). Bodies that matter: On the discursive limits of sex. New York: Routledge.

[12] Gergen, K. J. (1990). Affect and organization in postmodern society. In S. Srivastva, & D. L. Cooperrider (Eds.), Appreciative management and leadership: The power of positive thought and action in organizations. (pp. 153-174) Jossey-Bass.

[13] Gibbs, R. W. J. (2006). Embodiment and cognitive science. Cambridge: University Press.

[14] Gil, A. (2003). El consum com a emoci. Psicologia econmica i del comportament del consumidor.Barcelona: UOC.

[15] Gregson, N., & Rose, G. (2000). Taking butler elsewhere: Performativities, spatialities and subjectivities. Environment and Planning D, 18(4), 433-452.

[16] Haraway, D. J. (1989). Primate visions: Gender, race, and nature in the world of modern science. New York: Routledge.

[17] Haraway, D. J. (1995). Ciencia, cyborgs y mujeres: La reinvención de la naturaleza Valencia: Universitat de Valencia.

[18] Harré, R. (1995). The necessity of personhood as embodied being. Theory & Psychology, 5, 369-373.

[19] Haworth, J. T. (1990). The embodyment theory of pre-reflexive thought and creativity. In K. J. Gilhooly, M. T. G. Keane, R. H. Logie & G. Erdos (Eds.), Lines of thinking: Reflections on the psychology of thought, vol. 2: Skills, emotion, creative processes, individual differences and teaching thinking. (pp. 203-215) John Wiley & Sons.

[20] Hollinger, V. (2000). Cyborgs and Citadels/Between monsters, goddess, and cyborgs (book review). Signs: Journal of Women in Culture & Society, 25, 577.

[21] Illouz, E. (2007). Cold intimacies: The making of emotional capitalism London: Polity.

[22] Iñiguez Rueda, L. (2003). Análisis del discurso Barcelona: UOC.

[23] Iñiguez Rueda, L. (2005). Nuevos debates, nuevas ideas y nuevas prácticas en la psicología social de la era “post-construccionista”. Athenea Digital, 8

[24] James, P., & Carkeek, F. (1997). This abstract body: From embodied symbolism to techno-disembodiment. In D. Holmes (Ed.), Virtual politics: Identity and community in cyberspace. (pp. 107-124) London: Sage Publications, Inc.

[25] Katz, J. (1996). Families and funny mirrors: A study of the social construction and personal embodiment of humor. American Journal of Sociology, 101, 1194.

[26] Lazzarato, M. (2006): Transformacions de la subjectivitat en el capitalisme cognitiu, Barcelona: MACBA.

[27] Lefebvre, H. (1991). The production of space. London: Blackwell publishers.

[28] Leudar, I. (1995). Reporting political arguments. Reconstruction and Application: Proceedings of the Third Conference on Argumentation, , 3 42-59.

[29] Leudar, I. (1998). Who is martin McGuiness 1: On contextualizing reported political talk. Dialogue Analysis, 6, 217-224.

[30] Leudar, I., & Nekvapil, J. (1998). On the emergence of political identity in the czech mass media: The case of the democratic party of sudetenland. Czech Sociological Review, 6(1), 43-58.

[31] Lyon, M. L. (1999). Emotion and embodiment: The respiratory mediation of somatic and social processes. In A. L. Hinton (Ed.), Biocultural approaches to the emotions. (pp. 182-212) Cambridge: University Press.

[32] Malin, B., & Peterson, T. R. (2001). Communication with feeling: Emotion, publicness, and embodiment. Quarterly Journal of Speech, 87, 216.

[33] Michael, M. (1996). Constructing identities: The social, the nonhuman and change, London: Sage.

[34] Michael, M. (2004). Reconnecting culture, technology, and nature: From society to heterogeneity, London: Ebrary.

[35] Michael, M. (2006). Technoscience and everyday life: The complex simplicities of the mundane, London: Open University Press.

[36] Michael, M. (Ed.). (2000). Reconnecting culture, technology and nature : From society to heterogeneity. London: Routledge.

[37] Niedenthal, P. M., Barsalou, L. W., Ric, F., & Krauth-Gruber, S. (2005). Embodiment in the acquisition and use of emotion knowledge. In L. F. Barrett, P. M. Niedenthal & P. Winkielman (Eds.), Emotion and consciousness. (pp. 21-50) London: Guilford Press.

[38] Oatley, K., & Jenkins, J. M. (1992). Human emotions: Function and dysfunction. Annual Review of Psychology, 43, 55-85.

[39] Prinz, J. J. (2005). Emotions, embodiment, and awareness. In L. F. Barrett, P. M. Niedenthal & P. Winkielman (Eds.), Emotion and consciousness. (pp. 363-383). London: Guilford Press.

[40] Qian, H., & Scott, C. R. (2007). Anonymity and self-disclosure on weblogs. Journal of Computer-Mediated Communication, 12(4), 1428-1451.

[41] Rose, N. (1983). Effects of rational emotive education and rational emotive education plus rational emotive imagery on the adjustment of disturbed and normal elementary school children. New York: ProQuest Information & Learning.

[42] Sacks, H. (1992). Lectures on conversation: Vol 1 & 2. Oxford: Basil Blackwell,

[43] Schaub, E. L. (1933). Spinoza: His personality and his doctrine of perfection. Monist, 43, 1-22.

[44] Watson, D. R. (1987). Interdisciplinary considerations in the analysis of pro-terms. Talk and Social Organisation, , 261–289.

[45] Wittgenstein, L. (1958). Philosophical investigations, New York: Macmillan ed.

Case study on the assessment of the affective dimension in virtual learning environments

Eva de Lera
Universitat Oberta de Catalunya
Avda. Tibidabo, 47
08035 Barcelona (Spain)
(+34) 932535700
edelera@uoc.edu

Carles Fernàndez
Universitat Oberta de Catalunya
Avda. Tibidabo, 47
08035 Barcelona (Spain)
(+34) 932535700
cfernandezba@uoc.edu

Magí Almirall
Universitat Oberta de Catalunya
Avda. Tibidabo, 47
08035 Barcelona (Spain)
(+34) 932535700
malmirall@uoc.edu

ABSTRACT

This paper presents a proposal for fully online or blended learning universities to assess the affective dimension in online learning design.

The paper focuses especially on affective assessment by presenting a study of emotional inference through the triangulation of facial expression interpretations, pupil size and students as self-evaluators.

Keywords

E-learning, affective computing, affective learning, pupil size, FACS, affective assessment

1. INTRODUCTION

Most of the universities that are currently offering online courses usually have some kind of system for course design assessment. Traditionally these evaluation tools have been based on a set of recognized indicators provided by experts in graphic design, instructional designers, technologists… Up to now, the evaluation of affect in online learning usually means the use of questionnaires where students respond to items like '*do you think the course is motivating enough?*' or '*has the course met your expectations?*". Although questionnaires are a short and cost-effective way to gather data, we already know that users may answer just to offer an 'appropriate' opinion or are mediated by other variables. As Picard states, "self-report is colored by awareness of internal state, reflections on how such a report will be perceived, ability to articulate what one feels, etc." [1]. Also, the use of questionnaires requires the learning experience to be interrupted (or in any case, they cannot be used at the same time).

In fact, designers tend to implicitly propose a particular emotional scenario when they design virtual learning environments, but these decisions are generally not based on affective learning theories or findings.

In conclusion, at the present time the Universitat Oberta aims to introduce a new methododology that provide us with reliable data on the affective states of students in order to design engaging educational environments based on research outcomes.

2. INTRODUCING AFFECT IN ASSESSMENT

2.1 A reliable, feasible and cost-effective methodology

A method that requires adding more effort to any current assessment system needs to be reliable, feasible and cost- or time-effective. Affective evaluation must be integrated within the existing evaluation system and should not imply a considerable consumption of time so that institutions do not show resistance to it.

Our proposal consists of the use of three techniques: pupil size analysis, interpretation of facial and body expressions and students' self-assessment. All the data gathered through these techniques is compared to the real-time interaction of students.

In order to validate the combination (henceforth triangulation) of these techniques as a specific methodology, we planned a test with seven students interacting with the new mainpage of the virtual learning environment. Until now, after learners logged into the virtual campus they just found a mainpage with some news. The new mainpage is a personalized space based on widgets where students can choose the kind of information they would like to have when accessing the campus. It is very similar to the recognized and awarded *Igoogle* personal area and students are allowed to import external modules (gmail, google calendar, news and many other applications) as well as information from the university (forums, groups, news…), and personalize the design and position of these modules.

The sample was selected following heterogenic criteria: age, gender, studies and ICT skills were representative according to our target students at the university.

Although one of the objectives is to determine how affective the mainpage is, the essential focus of this research is to analyze the possibilities for inferring emotional data from pupil size, facial expressions and student self-assessment as a triangulation method.

2.2 Brief state of the art of the techniques used

The three techniques we have used have been tested as shown in the state of the art.

Pupil size has been widely studied by Picard and Hess (who coined the term in 1975), and accurately summarized by Timo

Partala [2] The latest research on pupilometry has confirmed that the increase of pupil size has a linear relationship with affective arousal. Arousal is significantly increased except with neutral stimuli. Nevertheless, other factors such as cognitive load and light reflex may have an effect and must be appropriately corrected in any research study.

One of the main advantages of this technique is that it does not require special sensors and measures involuntary responses from individuals, so that they cannot control their reactions.

Secondly, while pupilometry relies on arousal, FACS (Facial Action Code System) and other methods to analyze non-verbal information are related to valence. To date, there is some consensus on the existence of six basic emotions that can be codified according to FACS (anger, disgust, fear, joy, sorrow and surprise). The technique created by Ekman codes expressions as a combination of 44 facial movements called Action Units. [3]

Finally, concerning self-assessment; this technique has been partially criticized due to the fact that adults' self-report is colored by awareness of internal state, reflections on how such a report will be perceived, ability to articulate what one feels, and other factors [1] On the other hand, using self-assessment as a complementary method can bring interesting results. As a example of self-assessment interest, Sang-Hoon Jeong has developed a tool called VideoTAME [4] where subjects can view the recorded video of them performing tasks. This data can be useful as complementary information for emotion inference or confirmation of affective states.

2.3 The methodology in the test design

Students have to carry out five tasks on the mainpage: read an email in their webmail, add a new module (widget), change the position of another module, access the forum of a subject and access the virtual library. All these tasks can be done directly from the new mainpage.

More important than the tasks themselves are the *'key events'* that students need to perform in order to accomplish the task. For example, a key event in the task 'add a new module' is to find the right button that allows widgets to be added. Another key event for any of the tasks would be to know how to go back to the mainpage after the student uses webmail or the forums.

The key events are new stimuli within a 'steady/stable interaction', and represent opportunities for increase of arousal, especially when individuals cannot accomplish the event easily.

We have identified the following key events:

- To find the key buttons such as: 'add a module', 'personalize your mainpage', 'virtual library', 'webmail' and 'forums' (a total of 5 events)
- To be able to go back to the main page after a task
- To understand the concept of module as a box that can be dragged and dropped

Apart from key events, we also identified a group of events that require special attention by the user such as typing the password, waiting for the pages to load, reading instructions, accessing wrong pages… We call them '*secondary events*'.

As stated before, our methodology is based on a combination of pupilometry, gesture expression heuristics and student self-assessment.

Each technique has its own objective. We used pupilometry to detect arousal increase, whether related to positive or negative affective states. Increased pupil size seems to indicate that 'something is happening internally' and may be a clear sign that impels us to also analyze the valence of the affective state through facial and gestural expressions. Moreover, triangulating pupil size with gestural heuristics and self-assessment should allow us to assess how precise pupil size is by comparing the average of increased pupil size.

In this sense, it is very interesting to analyze the correlation between increased pupil size and a body expression that confirms this physical response in arousal. We consider that pupil size is increased when size is 10% bigger than the average pupil size in the test for a particular subject.

Secondly, our gesture coding system is used to determine the valence of emotions, since pupilometry itself cannot do anything with valence. Our own technique, called *ten heuristics*, is based on observation and does not require extra implementation effort since most interface evaluations are conducted observing and recording the user as he or she interacts with the interface. In such a scenario, facial and body expressions are often observed and recorded, but generally not measured in a structured manner. The ten heuristic tool has been developed and adapted thanks to previous studies in which we tried to identify the most common expressions of students in virtual learning environments.

Finally, as for the third technique, the objective of self-assessment in our methodology is to confirm that both pupilometry and gestural interpretation can really infer affective states.

The triangulation between these three techniques has two objectives:

- To analyze the correlations between the data gathered with the techniques in order to assess the specific effectivity of each one
- To conclude whether the overall methodology is appropriate for affect measurement

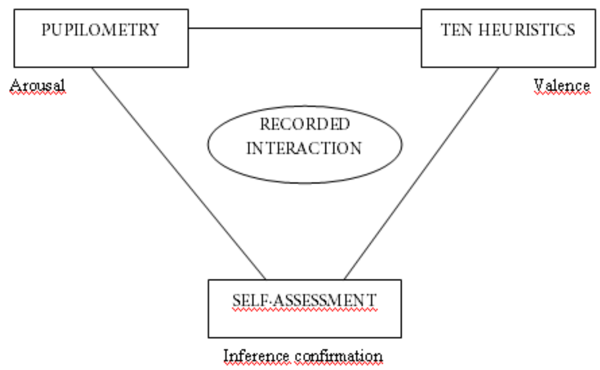

Figure 1. The methodology based on triangulation

2.4 Results and conclusions

As we stated before, due to the fact that we only tested a group of seven people, we cannot generalize conclusions. Nevertheless, there are a few results for reflection on the techniques we used:

Both pupilometry and 'ten heuristics' together were able to measure most of the key events and secondary events. We have analyzed the events where a key event could not be detected, the reason being that the student solved it without any problem, which means no arousal. The higher the level of success in a task, the less probable it is for arousal or non-verbal cues to appear.

These techniques have also showed that pupil size was increased or students reacted non-verbally to some events we had not considered as key or secondary.

We have also compared whether pupil size or ten heuristics were more reactive to key and secondary events. Pupil size responded to 74% of these events while ten heuristics did so in 65%.

One of the most important questions for our proposed methodology is the correlation between positive responses in pupil size and positive responses in heuristics. For most of the students (five) only 40% of the events with increased pupil size also showed any gestural reaction. The two others showed over 80% of correlation. Our conclusion is that pupil size does not seem to be a very good predictor if we only want to use it to infer or to state that the student is 'feeling something'. The size of pupils does not seem to be enough to assess arousal in our particular environment or conditions. Also, almost 40% of the events with increased pupil size did not show any particular facial or body expression.

Very interestingly, the correlation between high pupil size and specific gesture (average of 47.8%) is higher than the correlation between normal pupil size-specific gesture (20.7%) and high pupil size-no gestures (31.1%). This data may show that affective states are more commonly accompanied by the presence of both high pupil size and non-verbal expression, though evidence is not great enough to focus on pupilometry alone.

The analysis of gestural expressions by itself is richer and allows a good determination of valence. Most of the key events showed a particular expression, and the emotions inferred through these were confirmed through students' self-assessment.

Through students' self-assessment we have particularly checked whether the expressive reactions in Ten Heuristics can be inferred as specific emotions. Some of the heuristics were frequently repeated in the pilot test, such as frowning, compressing the lips, expressing vocally or hands touching the face.

The amount of non-verbal expression depends on each individual, and is probably the most important factor for heuristics performance. There was a relationship between success in tasks and amount of expressions (more difficulties, more presence of body expressions), but there was an individual who had important problems in some of the tasks but did not show many expressions. In her self-assessment she let us know she is not very expressive.

Globally, students' self-assessment confirmed more than 96% of the events detected by both pupilometry and heuristics. Almost 100% of the heuristics were confirmed and more than 90% of the events detected through increased pupil size.

Through the qualitative analysis of each specific case study we have identified other variables that have a particular impact on affect and the techniques we have selected to measure:

1. Pupil size tends to be higher during approximately the first minute due to some kind of tension
2. Sustained concentration makes pupil size increase but is not comparable to pupil size inrease in presentation of new interfaces, errors produced or difficulties in performance.
3. Concerning patterns of timeline analysis, pupil size is reactive usually 500 ms. before the key event and it keeps increasing depending on how well the user finds a solution but generally maintains the maximum size during 2 seconds.
4. Pupil size is not especially high when users speak, probably because talking is a way to decrease arousal.
5. We have analyzed the highest moments of increased pupil size but could not find specific body expressions related to high arousal.
6. The two most successful users did not show high increases in pupil size although they also had to solve new problems. Overall confidence may be the explanation.

In conclusion, the methodology of triangulation we have proposed for data collection and affect inference is useful, but we cannot rely on one technique alone, especially on pupilometry, since it showed that some key events and some difficulties experienced by users did not increase pupil size (less than half of events). Nevertheless, it is convenient to collect pupil size data since it can make us focus on events where subjects did not express anything.

Concerning other considerations of implementation such as feasibility, usability or time or cost-effectiveness, we can argue that these techniques do not require a big investment and the way they gather and export data is very useful for research. Pupilometry only requires an eyetracker with infrared rays, facial expression monitoring is carried out with a simple webcam, and for students' self-assessment we only need learners. These technologies are quite affordable.

With the appropriate guidelines for integrating this methodology in course assessment, staff should not notice a significant increase in work, or they should at least feel a balance between time investment and higher quality in course assessment and design.

3. REFERENCES

[1] Picard RW. et al., 2004. Affective learning – a manifesto. BT Technology Journal (22), 255

[2] Partala, T. and Surakka, V., 2002. Pupil size variation as an indicator of affective processing. International Journal of Human-Computer Studies, 59 (1-2), 185-198

[3] Michel, P., and El Kaliouby, R., 2003. Real time facial expression recognition in video using support vector machines. In proceedings ICMI International Conference on Multimodal Interfaces 2003 (Vancouver, Canada, 2003) ACM Press, New York, NY 258-263. DOI= http://doi.acm.org/10.1145/958432.958479

[4] Sang-Hoon, J., 2007. Suggestion of Methods for Understanding User's Emotional Changes While Using a Product. In Human Interface, M.J. Smith, G. Salvendy, Eds. HCII 2007, Springer-Verlag Berlin Heidelberg, 59-67

Automatic Recognition of Spontaneous Pain Expression Intensities

Zakia Hammal
Brain Music and Sound Laboratory, Université de Montréal
C.P. 6128 - Station Centre ville, Montreal (QC) H3C 3J7
(001) 514-343-6111
zakia_hammal@yahoo.fr

ABSTRACT
Automatic recognition of Pain expression and measuring its intensities has potential medical significance. In the current paper we present initial results of automatic recognition of Pain expression intensity in a 2-altenative choice (weak vs. strong). The proposed model is based on the Transferable Belief Model (TBM) [1] fusion process of facial information resulting from the permanents facial feature behaviors (eyes, eyebrows and mouth), the nasal root wrinkles and the context information (medical application context). Pain expression is first recognized from the six basic facial expressions and Neutral. Its intensity is then measured based on the intensity of the facial feature behavior according to each subject. To do that the recognized Pain expression images was passed to a 2-alternative classification model (weak vs. strong). In order to evaluate the model performances Naïve human observers were tested on the same data for recognizing Pain expression and the corresponding intensity. The proposed model recognizes successfully Pain expression and was successfully able at recognizing weak from strong Pain expression intensity. Preliminary results based on 20 participants videotaped while undergoing thermal heat stimulation at non-painful and two painful intensities (week and strong Pain) compare favorably to human performances at recognizing weak vs. strong spontaneous Pain expression.

Categories and Subject Descriptors
I.2.10 **[Vision and scene understanding]:** [Video analysis]; I.5.4 [**Pattern Recognition Applications**]; H.1.2 **[User/Machine systems]**: [Human information processing, Human Factors].

General Terms
Algorithms, Experimentation, Human Factors.

Keywords
Pain expression, Intensities, Uncertainty, Human behavior, Computer vision.

1. INTRODUCTION
In the current paper we are interesting in the automatic identification of Pain expression intensities in a medical context application where videotaped people are not able to verbally express there own painful state. The proposed model aims at overcoming several challenges encountered by human-machine interaction (HCI) systems applied in real-life applications: context application (hospital), doubt between expressions and non-prototypic facial expressions (Pain expression recognition). We present new developments of a previously proposed model [2, 3] that is extended here to the dynamic recognition of spontaneous "Pain intensities" in videos sequences. We focus on Pain as HCI because of its potential medical significance (for example, as a pain assessment tool in individuals who are not able to communicate Pain verbally (e.g. newborns, individuals with pronounced cognitive impairments [4]). The proposed model is based on the Transferable Belief Model [1] fusion process of facial information resulting from the permanents facial feature behaviors (eyes, eyebrows and mouth), the nasal root wrinkles and the context information (medical application context). The first step of the proposed model is to identify if the videotaped person have Pain or not [3] and then to extend the recognition to the identification of a weak or a strong Pain. The proposed model has already proved its robustness to identify Pain from 8-alternative facial expressions [3]. In the current contribution we will present our current advances extending the model to measure the intensity of the detected Pain expression.

2. DATABASE
We used a spontaneous video database consisting of 20 participants videotaped while undergoing thermal heat stimulation: 1 non-painful (1°C below the individual pain threshold) and 2 painful intensities ((week and strong pain) 2–3°C above the individual pain threshold) [4]. The faces of subjects were videotaped at a distance of 4m (see Figure 1). The onset of the stimulation is marked on the videotape, by switching on a light signal concurrently (see Figure 1). Videos were captured at a resolution of 720x536 pixels, out of which the face area spanned an average of approximately 320x470 pixels. Facial expressions displayed on each frame of the obtained videos have been analyzed using the Facial Action Coding System (FACS) [5].

Figure 1. Examples of Neutral and Pain expression intensities (middle apex in weak condition and right apex in strong condition) for two subjects

3. RECOGNITION OF PAIN EXPRESSION INTENSITIES

The general purpose of the current modeling is the automatic recognition of Pain expression intensities (among the six basic facial expressions and neutral) of videotaped people in a medical environment (see Figure 2). The first step is then to identify if the videotaped person have Pain or not before identifying its intensity. In the following we will succinctly summarizes Pain expression recognition [3].

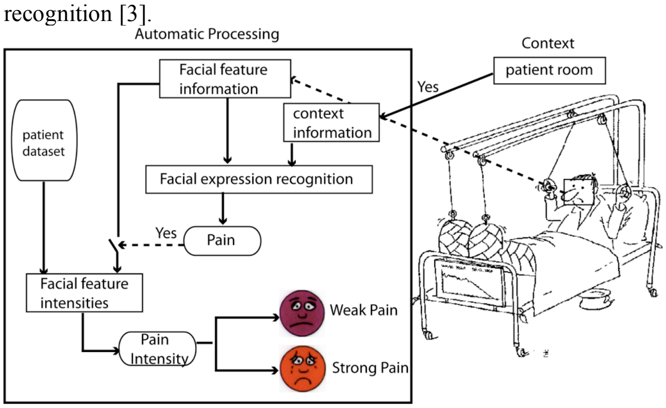

Figure 2. Overview of the proposed system

3.1 Pain Detection

3.1.1 Sensors information

The first step of Pain expression recognition model is the extraction of the contours of the permanent facial features (eyes, eyebrows and mouth see Figure 3) [6] where the corresponding deformations are measured by five characteristic distances D_1 to D_5 (see Figure 3). For each characteristic distance, five symbolic states reflecting the magnitude of the deformation according to the neutral state are defined: S_i if the current distance is roughly equal to its corresponding value in the Neutral expression, C_i^+ (vs. C_i^-) if the current distance is significantly higher (vs. lower) than its corresponding value in the Neutral expression, and $S_i \; Or \; C_i^+$ (vs. $S_i \; Or \; C_i^-$) if the current distance is neither sufficiently higher (vs. lower) to be in C_i^+ (vs. C_i^-), nor sufficiently stable to be in S_i. Secondly, based on the positions of the characteristic points of the eyes (inner eyes corners) the nasal root area is selected for wrinkles detection (see Figure 3). In the selected area the Nasal root wrinkles detection is based on the Canny edge detector. The presence or absence of wrinkles is decided by comparing the number of edge points in the nasal root in the current expressive image with the number of edge points in the nasal root of a Neutral facial image. If there are about twice more edge points in the current image than in the reference image, wrinkles are considered to be present "present" P otherwise they are considered as "absent" A. Finally, the context variable is coded with two states: MC if the videotaped sequence is in a medical context for example "Hospital" (and then the aim is to know if the current videotaped expresser has Pain) and NMC if not (and then the current expression is one of the 8 possible facial expressions). In a real-life application the medical context can be assessed by the medical doctor or the nurse using a simple pressure button which changes the system parameters towards the recognition of Pain expression [3].

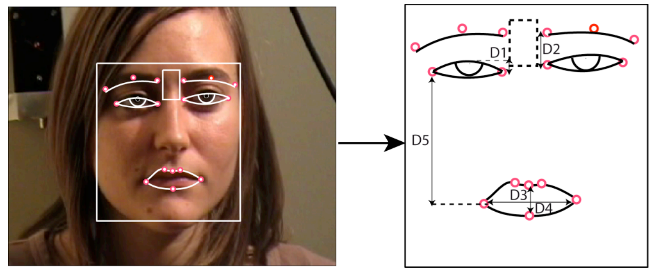

Figure 3. Example of facial features segmentation (left) and the corresponding characteristic distances (right) [14].

At this point Pain expression is modeled by a set of rules (Tables 1, 2 and 3) according the sensor states (permanent facial features, transient features and context). However Pain expression as all facial expressions may include a blend of expressions, which makes human observers often hesitating between several expressions [9]. Then an all-or-none system based only on the logical rules is not sufficient to reliably recognize Pain expression from the six basic facial expressions and Neutral. This issue is directly tackled by the TBM [1] modeling process of all the knowledge resulting from the characteristic distance states, the nasal root wrinkles and the context information (see section3.1.2).

Table 1. Rules table based on the characteristic distances states for Pain expression recognition

	D_1	D_2	D_3	D_4	D_5
Pain	C^-	C^-	C^+	C^+ Or C^-	C^-
Neutral	S	S	S	S	S

Table 2. Rules table based on the nasal root wrinkle states and context information for Pain expression recognition

	Pai	Ang	Dis	Hap	Sur	Fea	Sad	Neu
Nasal root Wrinkles	P	P	P	A	A	A	A	A
Context	MC	MC Or NMC	MC Or NMC	MC Or NMC	MC Or NMC	MC Or NMC	MC Or NMC	MC Or NMC

3.1.2 Sensors belief modeling

As reported above each sensor can takes a set of states (see Table 1,2,3). The belief $m_{Cue}^{\Omega_{Cue}}$ of each sensor state is defined as:

$$m_{Cue}^{\Omega_{Cue}} : 2^{\Omega_{Cue}} \rightarrow [0,\ 1]$$

$$A^{\Omega_{Cue}} \rightarrow m_{Cue}^{\Omega_{Cue}}(A), \quad \sum_{A \in 2^{\Omega_{Cue}}} m_{Cue}^{\Omega_{Cue}} = 1 \qquad (1)$$

Where $Cue = \{D_i, TR, CT\}$, $\Omega_{D_i} = \{C_i^+, C_i^-, S_i\}$, $\Omega_{TF} = \{P, A\}$, $\Omega_{CT} = \{MC, NMC\}$ correspond to the space of discernment associated to each sensor, $2^{\Omega_{D_i}} = \{C_i^+, C_i^-, S_i, S_i \cup C_i^+, S_i \cup C_i^-\}$, $2^{\Omega_{TF}} = \{P, A, P \cup A\}$, $2^{\Omega_{CT}} = \{MC, NMC, MC \cup NMC\}$, correspond to the sets of possible focal elements called frame of discernment. $m_{Di}^{\Omega_{Di}}$ is the piece of evidence (belief) of each characteristic distance state D_i, $m_{TF}^{\Omega_{TF}}$ the belief of each nasal root wrinkle state and $m_{CT}^{\Omega_{CT}}$ the belief of the each context variable state. P means that the nasal root wrinkles are present and A means that they are absent. $P \cup A$ corresponds to the doubt state between P and A ($P\ Or\ A$). From the frame of discernment only the states P (we are sure that the wrinkles are present) and the state $P \cup A$ (we don't know) are considered (this corresponds to the doubt on there state). MC means medical context (the expresser is in a medical context then it is more likely that the expected expression corresponds to Pain) and NMC means not medical context, $MC \cup NMC$ ($MC\ Or\ NMC$) means that the context of the expresser is unknown (then the expected expression can be one of the six basic facial expressions, Pain or Neutral). From the frame of discernment only the states MC (we know the context is medical) and the state $MC \cup NMC$ (we don't know the context) are taken into account.

3.1.3 Dynamic and progressive belief fusion

At each time *t* (frame) of the sequence once the beliefs on the sensors states defined (see section 3.1.2) the temporal fusion consist in combining their previous behaviors inside an increasing window Δt from the *beginning* until the *end* of the sequence. Then, at each time *t* inside the window Δt, the belief on the state of each sensor is selected based on the combination of the current state and of the whole set of the past states since the beginning of the expression. The dynamic fusion of the beliefs is made according to the number of appearance of each symbolic states noted $\underset{\Delta t}{Nb}(state)$ and their integral (sum) of plausibility noted $\underset{\Delta t}{Pl}(state)$ computed inside the temporal window Δt.

$$K_t(state) = \begin{cases} 1 \ \ if \quad m_{Sensor}(state) \neq 0 \\ 0 \ \ otherwise \quad 1 \leq t \leq \Delta t \end{cases} \tag{2}$$

$$\underset{\Delta t}{Nb}(State) = \sum_{t=1}^{\Delta t} K_t(State) \tag{3}$$

$$\underset{\Delta t}{Pl}(state) = \sum_{t=1}^{\Delta t} m_{Sensor}(state) + m_{Sensor}(state') \tag{4}$$

$$state \in state'$$

From the two parameters $\underset{\Delta t}{Nb}(state)$ and $\underset{\Delta t}{Pl}(state)$, the selected states of each visual cues at each time *t* inside the temporal window Δt are chosen as:

$$\underset{\Delta t}{State}(Cue) = \max(\underset{\Delta t}{Pl}(state_{Cue}) / \underset{\Delta t}{Nb}(state_{Cue})) \tag{5}$$

$$state_{D_i} \in 2^{\Omega_{Di}}, \ state_{TF} \in 2^{\Omega_{TF}}, \ state_{CT} \in 2^{\Omega_{CT}}$$

$$1 \leq i \leq 5$$

Once the beliefs on all the cue states are defined, the conjunctive combination rule (Eq. 6) [1] is used to combine all the resulted information and results in m^{Ω} the BBA of the corresponding expression or subset of expressions:

$$m^{\Omega} = \oplus m_{D,TF,CT}^{\Omega}(H) = (m_D^{\Omega} \oplus m_{TF}^{\Omega} \oplus m_{CT}^{\Omega})(H) \tag{6}$$

Where H denote expressions or subset of facial expressions.

After the fusion process of all the available information the resulted expression or subset of expression with the maximum of piece of evidence (belief) is chosen.

Once Pain expression is identified from the six basic facial expressions and Neutral, its intensity is measured using a 2-alternative classification model (weak vs. strong).

3.2 Pain Intensity

Pain expression intensity is measured after the detection of Pain expression. The intensity of the Pain expression is measured based on the facial features deformation intensities. As one can see in Figure 1 the deformations of facial features is much more important in the third column (strong Pain expression) than in the second one (weak Pain expression). More importantly, the intensity of the deformation is subject dependant. Figure 1 shows that the facial feature deformation in the case of strong Pain expression of the first subject is lower than the facial feature deformation of subject 2 in the case of weak Pain expression. Then, the intensity of the facial features related to Pain intensities should be calibrated for each subject. Based on these considerations a 2-alternative model of Pain intensity is defined for each subject. To do that once Pain expression is detected, the facial feature deformation (and then the corresponding characteristic distance states (see section 3.1.1)) are measured by two new symbolic states: strong (St) if the absolute value of the corresponding deformation is higher than a predefined threshold and weak (Wk) if it is lower than a predefined threshold (see Figure 4). Based on this coding Table 3 summarizes the facial feature intensities states for strong and weak Pain expression intensities. In the current paper only the permanent facial features intensities are used.
Once the state intensity of each characteristic distance is defined a new 2-forced choice model based on the TBM is used. The belief on the intensities states $m_{Di}^{\Omega Inten_{Di}}$ of each characteristic distance is computed as described in section 3.1.2. The new frame of discernment is defined as $\Omega_{IntenDi} = \{St, Wk\}$ and the new power set is defined as $2^{\Omega_{IntenDi}} = \{St, Wk, St \cup Wk\}$ (the set of possible intensity states). St means that the distance intensity state is strong Wk means it is weak and $St \cup Wk$ corresponds to the doubt state between St and Wk.

Table 3. Rules table Pain intensities recognition

	D_1	D_2	D_3	D_4	D_5
Weak Pain	Wk	Wk	Wk	Wk	Wk
Strong Pain	St	St	St	St	St

The belief $m_{Di}^{\Omega Inten_{Di}}$ on each symbolic state from $2^{\Omega_{IntenDi}}$ is measured according to the model defined in Figure 4. The threshold values for strong St and weak Wk states are defined for each subject for each sensor according to the corresponding apex states in the case of weak and strong sequences leading to one model for each subject. In the current paper the threshold th_2 allows to associate the state St to the corresponding distance intensity. The value of th_2 is defined by expertise and corresponds to the maximum value the corresponding distance in the weak sequence. Then if the deformation of this distance is higher to its maximum value in the weak sequence it is considered as strong and the state St is associated to it. The threshold value th_1 allows associating the state Wk to the corresponding distance intensity. The value of this threshold corresponds to the median value of the corresponding distance in the case strong sequence. Then if the absolute value of the distance is higher to this threshold the distance intensity between th_1 and th_2 the current distance intensity is strong or weak ($St \cup Wk$). In the current paper the threshold values are defined by expertise however in a real application this thresholds can be defined based on human self-report of Pain intensities (see patient dataset in Figure 2).

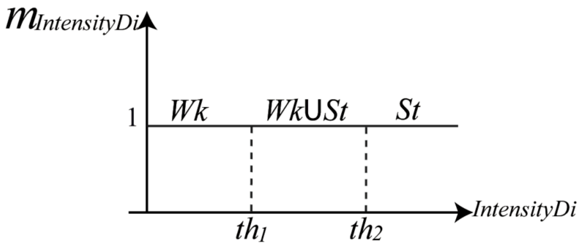

Figure 4. Model of belief assignment based on the characteristic distance motion intensity

Once the belief on the intensity states of the characteristic distances defined the corresponding Pain intensity is deduced based on Table 3. The fusion of the corresponding information is then made using the conjunctive rule of combination. If we consider two characteristic distances D_i and D_j with two the beliefs $m_{Di}^{\Omega'}$ and $m_{Dj}^{\Omega'}$ derived on the same frame of discernment, the joint belief $m_{Dij}^{\Omega'}$ is given using the conjunctive combination as:

$$m_{D_{i,j}}^{\Omega'}(A) = (m_{D_i}^{\Omega'} \oplus m_{D_j}^{\Omega'})(A) = \sum_{B \cap C = A} m_{D_i}^{\Omega'}(B) m_{D_j}^{\Omega'}(C) \quad (7)$$

where $\Omega' = \{StPain, WkPain, StPain \cup WkPain\}$, A, B and C denote propositions and subset of proposition from Ω'.

For example, if the belief resulted from the intensity of the distance D_1 is the doubt between strong and weak Pain as:

$$m_{D_1}^{\Omega\ IntensityD1} = m_{D_1}^{\Omega'}(StPain \cup WkPain)$$

and the belief resulted from the intensity of the distance D_2 is strong Pain as:

$$m_{D_2}^{\Omega\ IntensityD2} = m_{D2}^{\Omega'}(StPain)$$

the conjunctive combination of these two distance beliefs leads to the selection of strong Pain as:

$$m_{D_1}^{\Omega'}(StPain \cup WkPain) \oplus m_{D_2}^{\Omega'}(StPain) = StPain$$

4. PERFORMANCES

In this section we investigate how the proposed model performed on "spontaneous" Pain sequences. The simulation results were obtained on the 40 spontaneous Pain sequences (20 weak and 20 strong) obtained on the experimental condition described in section 2. Our simulations aimed at showing, the performances of the proposed model for Pain expression intensities recognition; and the comparison of the proposed model to human performances to prove its usefulness.

4.1 Human Performances

In order to compare our system performances to human observer (ground truth), an experiment was carried out for the classification of spontaneous Pain expression intensities by human observers. Due to the lack of spontaneous data for the six basic facial expressions, spontaneous Pain expression could not be identified among the six other expressions by human observers but only on a 2-forced choice (Neutral vs. Pain) from the same data.

First, human observers were asked to discriminate between Pain expression and Neutral expression. Secondly, they were asked to estimate the intensity of the recognized Pain expression by indicating if it was strong or weak. For the behavioral experiment, the stimuli corresponded to 80 videos of 3 facial expressions (2x20 Neutral, 20 weak and 20 strong Pain expressions) randomly interleaved in 4 separate blocks to the observers. 12 naïve observers (6 males and 6 females) were shown the videos and were asked to respond as accurately as possible in a 2-forced choice discrimination paradigm, by pressing the appropriate keyboard keys. The obtained results for Pain intensities classification by human observers revealed first their difficulty at recognizing Pain expression and secondly their difficulty to dissociate different Pain intensities (see Table 4).

Table 4. Human observer performances

	Neutral	Pain	
		Weak Pain	Strong Pain
Weak Pain	22	61	15
Strong Pain	28	57	15

Weak Pain recognition rates were 61% (with 22% as Neutral and 17% as strong Pain) and strong Pain recognition rates were 15% (with 28% as Neutral and 57% as weak Pain). These results reveal first a human difficulty to recognize spontaneous Pain expression and, secondly, to identify its intensities. The explanation is that this difficulty may be due to the great variability between the painful expressions displayed by different individuals. Another explanation is the fact that spontaneous facial expressions are often characterized by subtle changes of facial features that make their perception difficult [2].

4.2 Model Performances

The first step of the proposed model is the recognition of Pain expression from the six basic facial expressions and Neutral. The proposed model has already proves its robustness at identifying Pain from the six basic facial expressions and Neutral (77% of performances [3]). In order to compare the proposed model and the human performances only a 2-alternative choice between Neutral and spontaneous Pain facial expressions is made in the current paper. Then at each time after the recognition of Pain its intensity is then measured in a 2-alternative choice between weak and strong Pain intensities. Table 5 shows the obtained results. The average of the automatic classification of spontaneous Pain expression compare favorably (70%) with the human observers (74%) (see Tables 4, 5) and to the already reported classification rates on spontaneous Pain data in the state of the art (50.4% [7], to 72% [8]). The doubt state corresponds to the cases where the system hesitate between strong and weak Pain intensities (5.9% and 7.7%) (see Table 5). In these cases the system kept the doubt rather than making wrong decision. Considering an application in hospital context (waiting room, older people under camera monitoring), such information can be considered more than sufficient to alert somebody in charge of the patient by indicating a possibility of Pain (see Figure 2). A human observer (medical doctor or nurse) can then confirm or not this information.

Table 5. Model performances

	Ignorance (Pain, Neutral)	Pain		
		Weak Pain	Strong Pain	$StrongPain \cup WeakPain$
Weak Pain	32	60	2.1	5.9
Strong Pain	30	7.3	55	7.7

The obtained results show that compared to human observers the automatic recognition performances are comparable for weak Pain and higher in the case of strong intensity. This result can be explained by the fact that the proposed model is subject based and then takes into account the difference between human facial feature intensities. Indeed, as showed in Figure 1 Pain intensity is subject based and weak Pain expression leads sometimes much more facial features deformation than strong Pain of other subject.

5. CONCLUSION

We presented a model able to recognize Pain expression intensity (weak vs. strong) based on the facial feature deformations of patients faces in a medical application context. The obtained results compare favorably with human performances for weak and strong Pain intensities. Given the higher confusion of human observers due to individual differences. The obtained results are encouraging, opening promising perspectives to enhance the model performances.

6. REFERENCES

[1] Smets, P. 2000 Data fusion in the transferable belief model. Proc. of International Conference on Information Fusion, pages 21–33, July Paris, France.

[2] Hammal, Z., Couvreur, L., Caplier, A., and Rombaut, M. 2007 Facial expressions classification: A new approach based on Transferable Belief Model. International Journal of Approximate Reasoning, 46, 542–567.

[3] Hammal, Z., Kunz, M., Arguin, M., and Gosselin, F. 2008 Spontaneous pain expression recognition in video sequences. Proceedings of BCS International Academic Conference – Visions of Computer Science, 191-210.

[4] Kunz, M., Scharmann, S., Hemmeter, U., Schepelmann K. and Lautenbacher S. (2007) The facial expression of pain in patients with dementia. *Pain*. December, 133(1-3):221-8.

[5] Ekman, P., and Friesen, W. V. 1978 The facial action coding system (facs): A technique for the measurement of facial action. Palo Alto, CA: Consulting Psychologists Press.

[6] Hammal, Z., Eveno, N., Caplier, A. and Coulon, P.Y. Parametric models for facial features segmentation. Signal processing, 86:399–413, 2006.

[7] M. F. Valstar, M. Pantic, Z. Ambadar & J. F. Cohn. Spontaneous vs. posed facial behavior: automatic analysis of brow actions. Proc. ACM Intl. Conference on Multimodal Interfaces, pages 162–170, 2006.

[8] Littlewort, G.C., Bartlett, M.S. and Kang, L. 2007 Faces of Pain: Automated Measurement of Spontaneous Facial Expressions of Genuine and Posed Pain. Proc. ICMI, Nagoya, Aichi, Japan, November 12–15.

[9] Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., and Perrett, D. I. 1997 Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. Cognition, 63, 271–313.

Examining the Utility of Affective Response in Search of Personal Lifelogs

Liadh Kelly
Centre for Digital Video Processing
Dublin City University
Ireland
lkelly@computing.dcu.ie

Gareth J.F. Jones
Centre for Digital Video Processing
Dublin City University
Ireland
gjones@computing.dcu.ie

ABSTRACT

Personal lifelog archives contain digital records captured from an individual's daily life, for example emails, documents edited, webpages downloaded and photographs taken. While capturing this information is becoming increasingly easy, subsequently locating interesting items from within these archives is a significant challenge. One potential source of information to identify items of importance to an individual is their affective state during the capture of the information. The strength of an individual's affective response to their current situation can often be gauged from their physiological response. For this study we explored the utility of the following biometric features to indicate significant items: galvanic skin response (GSR), heart rate (HR) and skin temperature (ST). Significant or important events tend to raise an individual's arousal level, causing a measurable biometric response. We examined the utility of using biometric response to identify significant items and for re-ranking traditional information retrieval (IR) result sets. Results obtained indicate that skin temperature is most useful for extracting interesting items from personal archives containing passively captured images, computer activity and SMS messages.

1. INTRODUCTION

Advances in digital technologies mean that a wealth of personal information is now becoming available in digital format. This information can be gathered together and stored in a personal lifelog [2] [7] [12]. Personal lifelog archives can contain everything from items read, written, or downloaded; to footage from life experiences, e.g. photographs taken, videos seen, music heard, details of places visited, details of people met, etc, along with details of location and social context. Finding important relevant items from within these archives in response to user queries, or presenting interesting data to a subject browsing through their archive, poses significant challenges. Any additional information which can assist in identifying important items is thus potentially very important. Such information could be used in the re-ranking of information retrieval (IR) result sets, and for the promotion of interesting items when browsing a lifelog collection. One potential source of useful information is the user's biometric response associated with an item. In this study we explore three biometric responses associated with items, namely galvanic skin response (GSR), heart rate (HR) and skin temperature (ST).

Previous work has shown an individual's biometric response to be related to their overall arousal levels [13]. Significant or important events tend to raise an individual's arousal level, causing a measurable biometric response [16]. Events that can be recalled clearly in the future are often those which were important or emotional in our lives [6]. It has been demonstrated that the strength of the declarative or explicit memory for such emotionally charged events has a biological basis within the brain. Specifically involving interaction between the amygdala and the hippocampal memory system [5]. Variations in arousal level elicit physiological responses such as changes in heart rate or increased sweat production. Thus one way of observing an arousal response is by measuring the skin conductance response (SCR) (also referred to as the galvanic skin response (GSR)). The GSR reflects a change in the electrical conductivity of the skin as a result of variation in the activity of the sweat glands. It can be measured even if this change is only subtle and transient, and the individual concerned is not obviously sweating [6]. Curiously these biometric responses can be anticipatory of the consequences of a possible potentially risky action based on previous experiences, and may be observed before an individual is even consciously aware that the action may have significant consequences [4]. Arousal response can also be observed through skin temperature. With increased arousal levels, sympathetic nervous activity increases, resulting in a decrease of blood flow in peripheral vessels. This blood flow decrease causes a decrease in skin temperature [21]. Current technologies enable the capture of a number of biometric measures on a continuous basis. For example using a device such as the BodyMedia SenseWear Pro II armband [3] which can continuously record the wearer's GSR and skin temperature, or using the Polar Heart rate Monitor [19] which can continuously record the wearer's heart rate.

We propose that items or events which are important to an individual at the time they occurred may be useful to the individual again in the future, and further that such incidents are associated with emotional responses that can be detected by measuring an individual's biometric response when experiencing these events. Thus recording GSR, HR

5th Workshop on Emotion in HCI, British HCI Conference 2009, September 1, 2009, Cambridge, UK.

and ST as part of a lifelog may enable us to identify important items or events in a lifelog which would be most interesting for an individual to browse through in the future or which would be most important in a given information searching task.

In this paper we explore our hypothesis and report our findings to date which may guide future research in this area. We describe two studies: the first a study designed to explore the use of biometric data in detecting useful lifelog items for future browsing; the second is a study to investigate the utility of biometric response in re-ranking traditional information retrieval result lists. The next section describes the test-set gathered for these experiments. Section 3 presents our first study setup and results obtained. Our second study and results obtained are provided in Section 4. We then conclude the paper with a discussion of the findings and directions for future work.

2. TEST-SET

In order to explore our hypothesis, a suitable test-set must be available. As part of our ongoing work on personal lifelogs we are gathering long term lifelog collections from a small group of subjects. For the current investigation we augmented these for 3 subjects, for a 1 month period, with capture of their GSR, HR and ST data. For our current experiment we chose to examine: 1) whether these biometric data types can be useful in identifying important and memorable items which the subject may wish to view again in the future; and 2) the utility of these biometric data types in re-ranking the output of a user query driven IR result list. The lifelogs used for these experiments contained computer items accessed (web pages viewed, files created or accessed, emails sent and received, etc), SMSs sent and received, and images capturing an individual's activity.

The heart rate data was collected using a Polar Heart Rate Monitor [19]. The heart rate monitor is worn around the chest, and heart rate readings are transmitted to a watch worn on the subject's wrist. Software provided with the device generates reports, graphs and text files of the heart rate readings for data analysis. All other biometric data was collected using a BodyMedia SenseWear Pro2 armband [3]. The BodyMedia armband is worn on the upper arm and measures a range of psychological data. Data captured includes galvanic skin response along with transverse acceleration, longitudinal acceleration, heat flux, skin temperature, and energy expenditure. Software provided with the device generates graphs, reports and comma separated files of all the sensor output for data analysis. Our three subjects wore the heart rate monitor and BodyMedia devices for a one month period to capture biometric data for this investigation.

In addition to the biometric data, our experimental lifelogs contained data of computer activity, SMSs sent and received, and a visual log of activities. Computer activity was recorded using the Slife package [24]. Slife monitors computer activity and records the event of a window being brought to the foreground. For each event it records: type of application (e.g. web, chat), document source (e.g. Microsoft Word), window title and begin and end time of the event. Window title, application and document source were used to determine extension type (e.g. pdf, doc). The textual content inside the window (e.g. the text of an email, web page or document being written) and path to each file were obtained using MyLifeBits [17] for two subjects who are users of Windows XP and using locally written scripts for the other one who uses Mac OS X. The subjects all used Nokia N95 mobile phones [18] to capture SMSs. Logs of SMSs sent and received were generated using scripts installed on N95s. A visual log of subjects' activities was created using a Microsoft Research SenseCam [8] [23]. The SenseCam is a digital camera, with fish-eye lens, worn around a subject's neck. This passively captures images approximately every 20 seconds. Image capture is triggered based on changes in sensor data captured by the device. For example, high acceleration values, passive infrared (body heat detector) as someone walks in front of the wearer or changes in light level. If no sensor has triggered an image to be captured, the camera takes one anyway after a period of approximately 30 seconds. When worn continuously, roughly 3,000 images are captured in an average day.

Lifelog items (i.e. computer items accessed, SMS sent and received, and SenseCam images) were also annotated with the following types of context data:

- Using time and date information functions were written to determine, the month, day of week, part of week e.g. weekend or weekday, hour, minute, second, and period of the day e.g. morning, afternoon, evening, night, in which the event took place.
- Events were annotated with geo-location. GPS data, wireless network presence and GSM location data was captured by constantly running the Campaignr software, provided to us by UCLA (USA) [9], on subjects N95 mobile phones. From which geo-location was derived using in-house scripts.
- Events were annotated with light status and weather conditions derived using date, time and geo-location information.
- People present were annotated to events. The Campaignr software also recorded co-present Bluetooth devices, from which people present can be uncovered [11], [14].
- Using mobile phone call logs, generated using freeware [15] installed on N95s, events were annotated with phone conversations which occurred when the event took place.

The context data types used in this study were: file name; extension type; date; time; month; day of week; weekday or weekend; morning, afternoon, evening or night; light status.

Lucene [1], an open source search engine, was used to index items and their associated context data into different fields (e.g. day of week field, etc). The StandardAnalyzer built into Lucene was used to index the content of items. This tokenizes the content based on a sophisticated grammar that recognises email address, acronyms, alphanumeric and more; converts to lowercase, and removes stopwords.

3. EXPERIMENT 1: LOCATING INTERESTING EVENTS

In this section we describe the setup of a study to examine whether biometric data can be useful in identifying important and memorable events (an event is a group of SenseCam images or computer and SMS items) which the subject may wish to view again in the future. The results of this study suggest that biometric data can be useful in detecting important lifelog events and highlights the types of events it is most beneficial for. We begin by describing the test-set used for this study and the experimental approach taken. The results obtained are then discussed.

3.1 Extracting Important Events

We postulate that important events from a lifelog archive are coincident with maximum observed GSR and HR readings and with minimum observed ST readings at the original time of event occurrence, and that these events would be most interesting for subjects during future archive browsing. In this study we investigate three types of biometric reading: GSR, HR and ST. The BodyMedia device samples the values from its inbuilt sensors at settable predefined intervals. Based on results from initial calibration experiments we set the device to capture GSR data once per second and ST data once every ten seconds. The maximum rate of HR data captured afforded by the Polar heart rate monitor device was once every five seconds.

Variations in biometric response occur all the time and can be caused by many things, for example changes in arousal level or changes in physical activity such as walking down a corridor or running. A problem in analysis of biometric data for the purposes of this experiment is to identify variation in biometric data which are likely to be the result of variations in arousal levels, as opposed to physical reasons. An additional source of data that can be inferred from captured biometric data using the BodyMedia armband is the energy expenditure (sample rate set to once per minute) of the individual. Energy expenditure correlates well with periods of physical activity. Thus measured energy expenditure can be used to differentiate between high GSR and HR biometric data levels and low ST biometric data levels, resulting from physical activity and those arising from events experienced from the environment. GSR, HR and ST data captured during periods of energy expenditure above the average energy level * α (α = empirically determined scalar constant) were removed from the data set. To determine correlation between item importance and GSR, HR or ST, we attempted to extract 10 max, 10 average and 10 min [1] GSR lifelog items/events; 10 max, 10 average and 10 min HR items/events; and 10 max, 10 average and 10 min ST lifelog items/events, this corresponds to 5 SenseCam and 5 computer/SMS items/events for each GSR, HR and ST level from each subject's lifelog. The procedure for extraction of these SenseCam and computer/SMS items/events was as follows:

1. Determining begin and end timestamps of max GSR and HR: Begin and end timestamps for periods in a subject's GSR/HR dataset where the GSR/HR level was greater than a preset threshold for an empirically determined number of seconds were recorded. (threshold = average of GSR/HR data * β, β = empirically determined scalar constant)

   Determine begin and end timestamps of max ST: Timestamps were obtained by taking periods where ST levels were less than a preset threshold for an empirically determined number of seconds. (threshold = average of ST data / β, β = empirically determined scalar constant)

   Determining begin and end timestamps of min GSR and HR: Timestamps were obtained by taking periods where GSR/HR levels were less than a preset threshold for an empirically determined number of seconds. (threshold = average of GSR/HR data / χ, χ = empirically determined scalar constant)

   Determining begin and end timestamps of min ST: Begin and end timestamps for periods in a subject's ST dataset where the ST level was greater than a preset threshold for an empirically determined number of seconds were recorded. (threshold = average of ST data * χ, χ = empirically determined scalar constant)

   Determining begin and end timestamps of average GSR, HR or ST: Timestamps were obtained by taking periods where GSR/HR/ST levels were greater than threshold1 and less than threshold2 for an empirically determined number of seconds. (threshold1 = average of GSR/HR/ST data - δ, δ= empirically determined scalar constant; threshold2 = average of energy expenditure data + σ, where $\sigma = \delta$)

2. Extracting items/events from the subject's lifelog: The begin and end timestamps from step 1 were used to extract SenseCam, and computer/SMS events as follows: *if* computer or mobile activity occurred between the begin and end timestamps, these items were extracted, *else if* SenseCam images occurred between the begin and end timestamps, these images were extracted.

3. Removing duplicates: An item/event may cause a max biometric response on one access to it and on a different access cause an average or min biometric response for example. Items were removed from all but their highest occuring threshold (e.g. if a computer file currently in the min collection to be presented to subject caused a max biometric response on a different occasion it was removed from the min collection).

On completion of this process, having expanded thresholds as far as possible, we had sets of <=45 SenseCam events and <=45 computer/SMS events from each subject's lifelog. These sets of events were used to test our hypothesis, as described in the next section.

3.2 Experiment Procedure

The goal of this research is to establish if max periods of GSR, HR and ST are good indicators of lifelog items/events which are most useful for presentation to subjects when browsing their personal information archives. Personal lifelog items of varying GSR, HR and ST were presented to subjects and questionnaires completed to determine if GSR, HR and ST corresponded with memorable-ness, significance of events and desire to view again. Post questionnaire interviews were then conducted. This section describes the details of these procedures.

We wished to establish the relationship between biometric response at time of item/event creation/access on subjects' desire to re-view lifelog items/events over the long-term. We

[1] Max GSR/HR = periods of high GSR/HR; max ST = periods of low ST; average GSR/HR/ST = periods of average GSR/HR/ST; min GSR/HR = periods of low GSR/HR; and min ST = periods of high ST.

thus waited for nine months after the test-set collection period to present subjects with a set of events taken from their lifelogs. A total of <=90 events generated using the technique described in Section 3.1 were presented to subjects in this set. The set included: for each of GSR, HR and ST <=5 computer or mobile phone activity events with the max GSR/HR/ST and <=5 SenseCam image events corresponding to the max GSR/HR/ST; and for comparison purposes similar sets of events with average GSR, HR, ST and min GSR, HR, ST. For each of average GSR, HR and ST the <=5 computer or mobile phone activity events and <=5 SenseCam events closest to the subjects average GSR, HR and ST were chosen (as described in previous section). For each of min GSR, HR and ST the <=5 computer or mobile phone activity events and <=5 SenseCam events closest to the subject's lowest min GSR/HR/ST were chosen (also described in previous section).

Each subject was presented with their set of <=45 computer/SMS events and <=45 SenseCam events. Subjects were aware that the sets presented to them contained events with varying associated biometric levels and of the specific hypothesis we wished to test. However, they were not aware of the biometric response associated with each event. The questionnaire was explained to subjects and sample answers provided.

The subjects completed Questionnaire 1 for these <=90 events (and returned the completed questionnaire to the investigator). Details of this questionnaire were as follows:

1. Is this event memorable? (4-point scale: 4 = very memorable, 3 = memorable, 2 = not very memorable, 1 = not memorable).

2. Was the event important to you at the time? (4-point scale: 4 = very important, 3 = important, 2 = not very important, 1 = not important).

3. Is the event important to you now? (4-point scale: 4 = very important, 3 = important, 2 = not very important, 1 = not important).

4. Is it interesting to see this event again? (4-point scale: 4 = very interesting, 3 = interesting, 2 = not very interesting, 1 = not interesting).

5. Would you want to view this event again? (3-point scale: 1 = yes, 2 = maybe, 3 = no).

The following sections discuss the findings of this study.

3.3 Experiment Results

For analysis purposes, for each of questions 1-4 we took a binary split of the 4-point scale; that is a score of 4 or 3 was taken as positive and a score of 2 or 1 was taken as negative. We then calculated the average number of positive scores for each question for SenseCam and for computer/mobile phone activity events. Figures 1 and 2 show the average number of positive scores for questions 1-4 across the 3 subjects.

These results suggest a certain level of correlation between memorableness, importance at time and GSR and ST levels associated with SenseCam images (Figure 1). However as can be seen in the graph, while memorableness and importance at time of events were well captured by max GSR, these factors did not correlate with perceived current importance or desire to view the images now. During informal interview many reasons for this were uncovered, for example an event, such as an important work meeting while well remembered and important at the time to a subject no longer held any relevance and the subject had no interest in viewing the images relating to this event. Good correlation was observed between ST levels associated with images for subjects perception of current importance of the images and for their interest in viewing them now. Suggesting that ST at time of passive image capture might be a better indicator than GSR of images to present a subject with in the future. The results observed for HR levels were very poor (Figure 1). These results might be partly explained by the unreliability of the HR recording device, which sporadically gave unreliable readings. While efforts were made during experimentation to remove HR readings that appeared erroneous, it is possible that some incorrect readings remained. Other devices can capture HR more reliably, we hope to explore their use in future experiments.

Question 5 on the questionnaire specifically examined if subjects anticipated wanting to view these images again in the future. Results are presented in Figure 3. Overall subjects would or might want to retrieve (yes and not sure in Figure 3) 40% of max GSR response images, this compares with 29% of average GSR response images and 20% of min GSR images. ST levels was a more reliable indicator of subjects possible desire to view images in the future. They would or might want to retrieve (yes and not sure in Figure 3) 45% of max ST images, this compares with 15% of average ST images and 13% of min ST images. Again poor results were observed for HR (see Figure 3).

During informal interview reasons for not positively rating all max biometric response images as memorable or important at time included lack of novelty, e.g. regular lunch date with same group of collegues, subjects stated that audio for such events might have helped them recollect the event and hence change their ratings. Additionally events considered mundane by the subject, such as cooking dinner, while receiving max biometric response, received low ratings in the questionnaire. While we have no way of knowing what would cause such events to receive a max biometric response, we postulate that it may be due to the thought process of the subject at the time, which is now not recalled or some brief shock factor such as a sudden noise for example. Generally speaking we found that SenseCam images with max biometric response, which did not focus on interaction with other people, were not interesting to view. Exceptions here were images depicting a novel vacation location for example. Image processing techniques could potentially be exploited to further enhance the set of SenseCam images presented to a subject for browsing by down-weighting the potential selection of SenseCam images where people are not present and by up-weighting novel locations. The lifelog archives we have created contain people present information and geo-location information, as discussed in Section 2, which would enable such enhancements.

Strong correlation was not observed between GSR or HR levels and memorabelness, perceived importance of computer and mobile activity events and desire to view (Figure 2). However, correlation between ST levels and how interesting the subject now considered events was observed (see Figure 2). Additionally, subjects would or might want to retrieve (yes and not sure in Figure 4) 64% of max ST images, this compares with 40% of average ST images and 40% of

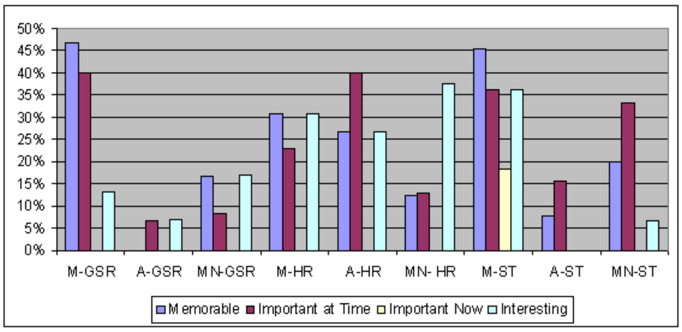

Figure 1: Percentage of SenseCam events for max (M), average (A) and min (MN) GSR, HR and ST the subjects considered memorable, important at the time, important now, and are interested in viewing now.

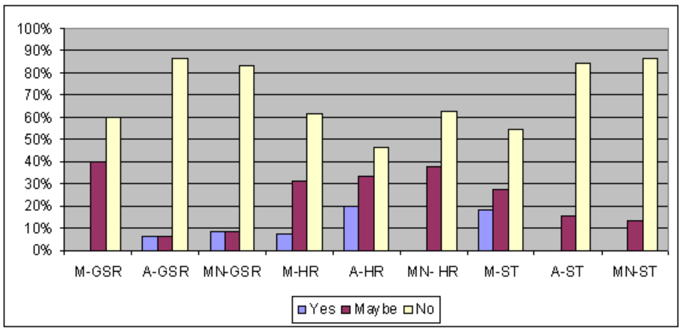

Figure 3: Percentage of SenseCam events for max (M), average (A) and min (MN) GSR, HR and ST the subjects would (Yes), might (Maybe) and would not (No) want to view again.

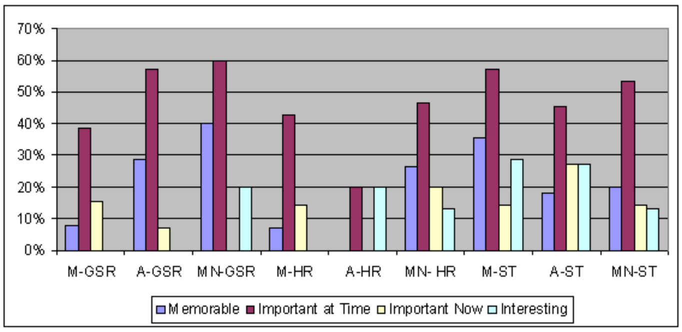

Figure 2: Percentage of Computer and SMS events for max (M), average (A) and min (MN) GSR, HR and ST the subjects considered memorable, important at the time, important now, and are interested in viewing now.

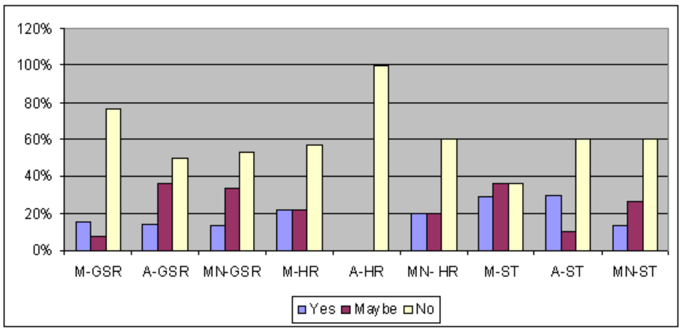

Figure 4: Percentage of Computer and SMS events for max (M), average (A) and min (MN) GSR, HR and ST the subjects would (Yes), might (Maybe) and would not (No) want to view again.

min ST images. No such correlation was observed for GSR or HR levels (Figure 4).

Informal interviews with subjects revealed some explanations for subjects' ratings of max biometric computer events. Completing a tedious user experiment stood out as memorable to subjects due to being particularly mundane, however subjects did not want to see these items again. Other events, such as emails containing information which was very important to subjects at the time of capture were very memorable, but as this information was no longer relevent to subjects they had no desire to view it again. Investigation of low ratings of instant messaging (IM) activity items with a max biometric response, revealed that subjects could not recall the content of IM which if available may have changed their ratings of these items. This result suggests that capturing additional data of items with max biometric values would be useful in further experimentation.

3.4 Concluding Remarks

The results of this experiment provide some perliminary support for the use of biometric response (especially ST) at time of item capture in extracting interesting items from large multimedia lifelog collections. We are interested in establishing if the results presented in this section can be improved using alternate approaches. We conducted similar experiments to the one described in this section to select max, average and min events, namely: 1) removing all items/events from the possible items to present to a subject in the max, average and min categories which occurred in more than one of these categories; 2) presenting items to subjects which occurred near in time to the noted max, average or min biometric responses. Result analysis is still on-going. Future work will explore combining GSR, HR and ST levels in the extraction of interesting events.

4. EXPERIMENT 2: RE-RANKING RESULT LISTS USING BIOMETRIC RESPONSE

In this section we describe the setup of a study to examine the utility of GSR and HR biometric data at original time of the computer item or SMS access, in re-ranking the output of a user query driven IR result list. For this investigation we used the 1 month collections of computer activity and SMS data annotated with biometric data from subjects' lifelogs. The results of this initial investigation do not support the use of GSR and HR biometric response at time of item creation/access for re-ranking lifelog query result lists; however interesting observations were made, as discussed later. We begin this section with a description of the experiment approach taken, and follow with an analysis of the results obtained.

4.1 Experiment Procedure

4.1.1 Test Case Generation

If lifelogs are to be recorded and accessed over an extended period it is important that users are able to reliably retrieve content recorded in the distant past. It is clear that a user is likely to remember a significant amount of content and context data soon after an event occurred, however with time memory fades and it is anticipated that less will be remembered a substantial delay after the event occurred [10]. In this experiment we wished to mimic the 'real' re-finding requirements of individuals, and details they are likely to recall about required items as closely as possible. As such, in generating the test cases for this experiment the following query generation technique was used:

- After 8 months lifelog collection build up (5 months after the one month biometric data capture period) subjects were required to list lifelog retrieval tasks they might want to perform in the future. This list was extended by the subjects consulting their lifelogs to determine additional items they might want to retrieve in the future. Typical test cases generated in this manner were: 'show me all documents I created associated with conference X; show me the SMS message my friend Sarah sent me regarding the location in Paris we were to meet in'.
- Subjects then entered their list of task descriptions along with keywords and remembered context, e.g. word in file name; extension type; month; day of week; weekday or weekend; morning, afternoon, evening or night; light status, into a provided form.

50 test cases were generated by each subject using this technique. Of these test cases subject 1 had 22 tasks containing items which occurred during the biometric data capture period, subject 2 had 8 and subject 3 had 36. These subsets of the generated test cases were used in this experiment.

4.1.2 Result Set Generation

To test the performace of our re-ranking approach a set of relevant computer/SMS items is required. Pooling is a commonly used approch in IR for generating lists of relevant items for a query. With pooling the query is passed into different IR algorithms and the results from each IR algorithm presented to subjects for relevance rating. We created our pooled result lists by entering various combinations of remembered content (keywords) and context into the vector space model (VSM) [22] and BM25 [20]. VSM and BM25 are good standard IR algorithms. They rank a set of documents in reponse to a user query using different term weighting approaches. The query types entered into VSM and BM25 were:

1. Content only.
2. Context only, this incorporated the following fields: word in file name; extension type; month; day of week; weekday or weekend; morning, afternoon, evening or night; light status.
3. Content + extension type.
4. Content + context.

The Lucene implementation of the vector space model and an in-house developed implementation of BM25 for Lucene were used to process these queries. Query type three and four are straightforward concatenations of the content data from query one with various types of the context data from query two.

The results from each of the 8 IR techniques were pooled and presented to subjects for relevance judgment. That is, for each test case the subjects rated each of the retrieved items as relevant or irrelevant. These judged sets were used for determining the utility of our techniques.

4.1.3 Investigation

We wished to investigate if biometric response (specifically GSR and HR) at time of item access/creation could be used to re-rank the output of an IR in response to a user query. To do this we investigated adding a query independent biometric score (static score) to items queried content and items queried content + context based retrieval score. The queried context fields used were: word in file name; extension type; month; day of week; weekday or weekend; morning, afternoon, evening or night; light status. While VSM was found to enrich the pooled result lists generated in Section 4.1.2, comparison showed BM25 to perform better. Hence BM25 was used to obtain queried content and queried content + context retrieval scores in this experiment. Static biometric scores were obtained by dividing the GSR associated with an item by the energy expenditure at that time, this value was multiplied by an empirically determined constant. The maximum static biometric score obtained in this manner across all accesses to the item was chosen as the static score value to add to the BM25 score. We also examined static biometric scores created by dividing HR by energy expenditure, again this value was multiplied by an empirically determined constant and the maximum score obtained for the item chosen to add as a static score to the BM25 item score.

For this investigation the following IR algorithms were investigated:

1. Content only
2. Content + GSR static score
3. Content + HR static score
4. Content + Context
5. Content + Context + GSR static score
6. Content + Context + HR static score

Query type one represents a good current standard approach for retrieval using search engines. Results generated from this content only query were used as a baseline.

The goal of our experiment is to retrieve the correct file(s) at top rank for a given query task. To investigate the usefulness of our static biometric scoring approaches in the retrieval process, the biometric data capture month of each subject's lifelog was queried with their set of queries using the 6 IR algorithms. The queries were processed using an in-house developed version of BM25 for Lucene. To obtain a relevance score for an item the relevance scores obtained

Table 1: Average precision, precision @ 5 and @ 10 documents and R-precision results for 3 subjects.

IR Technique	AveP	P@5	P@10	Rprec
Content Only	0.268366667	0.235566667	0.187933333	0.255233333
Content Only+GSR	0.248433333	0.2203	0.177166667	0.223
Content Only+HR	0.237566667	0.213966667	0.1708	0.2141
Content+Context	0.3025	0.257666667	0.208466667	0.261866667
Content+Context+GSR	0.271566667	0.2349	0.195633333	0.225466667
Content+Context+HR	0.256766667	0.2402	0.184833333	0.206633333

for the items content and each of the items context types were summed, and the static biometric score added to this. In each case the rank of relevant items in the result set was noted. The next section discusses the results of this investigation.

4.2 Experiment Results

TREC_EVAL [25], a freely available IR evaluation tool was used to calculate the precision at 5 and 10 documents, average precision, and R-precision of retrieved events. R-precision is the recall result, it measures precision after the total number of relevent items for a query have been retrieved (e.g. if there were 9 relevant items for a query, precision would be calculated after 9 documents had been retrieved). Results obtained are shown in Table 1. As can be seen from the results adding a biometric static score to content or content + context IR scores does not prove useful. Substantial drop in results is noted across average precision, R-precision and precision at 5 and 10 documents with the addition of either static GSR or HR scores. These preliminary results suggest that biometric response at time of item creation/access is not useful for re-ranking lifelog text-based collections. We discovered that, for a given task, biometric response of relevant retrieved items serves as a good indicator of a user's engagement with a computer item at time of creation/access, however this does not necessarily mean that the item will be useful in the future. For example, an activity such as registering for a conference may cause a marked biometric response at the time, but this does not mean that an individual will find this registration form useful in the future, nor are they likely to consider it important if they perform a future query for information relating to the conference. Similarly, twittering about the frustrating Java coding task you are currently engaged in may cause a marked biometric response, but this twitter will not be considered relevant in a future query on how to code in 'Java'.

4.3 Concluding Remarks

Based on the results of this perliminary experiment, GSR and HR levels at time of item access do not seem useful in re-ranking user query driven IR result lists. The experiment in this section was conducted prior to Experiment 1. In light of the correlation between ST and importance and desire to view items observed in Experiment 1, ST may prove to be a more useful static biometric score. Experimentation to test this hypothesis is currently in progress.

5. CONCLUSIONS

In this paper we set out to explore the role of biometric response in lifelog item/event retrieval. We investigated whether items coincident with maximum observed biometric galvanic skin response (GSR) and heart rate (HR) and with minimum observed skin temperature (ST) readings were more important to subjects and whether this would mean they were most useful for presentation to subjects when browsing their personal information archives. From this preliminary study correlation between GSR and ST levels and SenseCam event importance was observed. The SenseCam event selection results are important since ability to extract interesting events from vast SenseCam collections is challenging but important if these archives are to have long-term use. From this preliminary study GSR and ST appear potential sources of evidence for selection of SenseCam images. Combining GSR and ST responses may prove to further improve this performance. Correlation between ST response and computer item and SMS importance was also observed. While these results are promising, it is acknowledged that this study was conducted on a limited number of subjects over a short period of time. Investigation using more participants over a longer timeframe is required to further test our conclusions.

Our second investigation looked at the utility of using GSR or HR level at time of item creation/access to boost results in a query driven IR result list. Our preliminary results suggest that biometric response at original time of item access, while indicating the current importance of, or engagement with an item does not reflect the future likelihood of this item being relevant to a given query. We are currently investigating if ST may prove more useful.

Developments in technology are enabling individuals to store increasing amounts of digital data pertaining to their lives. As these personal archives grow ever larger, reliable ways to help individuals locate interesting items from these multimedia lifelogs becomes increasingly important. The results of these experiments provide perliminary indication that biometric response, in particular ST, may serve as a useful tool for extracting interesting items from long-term multimedia lifelogs. Additionally, the observed correlation between biometric response and current engagement found in these studies suggests that biometric response could be useful in 'live' applications such as 'on-line recommender systems (e.g. shopping websites), web search systems, or archive search systems (e.g. searching through picture archives or document archives) where the subject's search could be guided by their biometric response to items suggested / viewed at each iteration of the recommendation / search process.

6. ACKNOWLEDGMENTS

We would like to acknowledge the test subjects for their contribution to this study.

This work is funded by a grant under the Science Foundation Ireland Research Frontiers Programme 2006, and supported by Microsoft Research under grant 2007-056.

7. REFERENCES

[1] Apache Lucene. Overview, Copyright ©2006 The Apache Software Foundation. http://lucene.apache.org/java/docs/.

[2] G. Bell. Challenges in Using Lifetime Personal Information Stores based on MyLifeBits.

[3] BodyMedia - http://www.bodymedia.com.

[4] A. Damsaio. *Descartes' Error Emotion, Reason and the Human Brain*. Bard, 1994.

[5] B. Ferry, B. Roozendaal, and J. McGaugh. Basolaterail amygdala noradrenergic influences on memory storage are mediated by an interaction between beta- and alpha1-adrenoceptors. *Journal of Neuroscience*, 19:5119–5123, 1999.

[6] M. Gazzaniga, R. Ivry, and G. Mangun. *Cognitive Neuroscience (Second Edition)*. Norton, 2002.

[7] J. Gemmell, G. Bell, and R. Lueder. MyLifeBits: A Personal Database for Everything. *Communications of the ACM. Personal Information Management*, 49:88–95, 2006.

[8] C. Gurrin, A. Smeaton, D. Byrne, N. O'Hare, G. Jones, and N. O'Connor. An Examination of a Large Visual Lifelog. In *4th Asia Information Retrieval Symposium (AIRS)*, pages 537–542, 2008.

[9] A. Joki, J. Burke, and D. Estrin. Campaignr - a framework for participatory data collection on mobile phones. Technical report, 2007.

[10] L. Kelly, Y. Chen, M. Fuller, and G. Jones. A study of remembered context for information access from personal digital archives. In *2nd International Symposium on Information Interaction in Context (IIiX)*, pages 44–50, 2008.

[11] V. Kostakos, E. O'Neill, and S. Jones. Social networking 2.0. In *Extended abstracts of the conference on Human factors in computing systems*, pages 3555–3560, 2008.

[12] L. Kelly and G.J.F. Jones. Venturing into the Labyrinth: the Information Retrieval Challenge of Human Digital Memories. In *Proceedings of the Workshop on Supporting Human Memory with Interactive Systems, at HCI 2007: The 21st British HCI Group Annual Conference*.

[13] P. Lang. The emotion probe: Studies of motivation and attention. *American Psychologist*, 50(5):372–385, 1995.

[14] B. Lavelle, D. Byrne, C. Gurrin, A. Smeaton, and G. Jones. Bluetooth Familiarity: Methods of Calculation,Applications and Limitations. In *MIRW 2007 - Mobile Interaction with the Real World, Workshop at the MobileHCI07: 9th International Conference on Human Computer Interaction with Mobile Devices and Services*, pages 55–58, 2007.

[15] LogExport - http://tinyhack.com/freewarelist/s603rd/2007/03/02/logexport/.

[16] J. McGaugh. *Strong memories are made of this Memory and Emotion: The Making of Lasting Memories*. Columbia University Press, 2003.

[17] MSR MyLifeBits Project - http://research.microsoft.com/barc/mediapresence/MyLifeBits.aspx.

[18] Nokia N95 mobile phone - http://www.forum.nokia.com/devices/N95.

[19] Polar Heart Rate Monitor - http://www.polarusa.com/.

[20] S. Robertson and K. S. Jones. Simple, proven approaches to text retrieval. Technical Report 356, University of Cambridge, 1994.

[21] R. Sakamoto, A. Nozawa, H. Tanaka, T. Mizuno, and H. Ide. Evaluation of the driver's temporary arousal level by facial skin thermogram-effect of surrounding temperature and wind on the thermogram. *IEEJ Trans. EIS*, 126(804-9), 2006.

[22] G. Salton and C. Buckley. Term-weighting approaches in automatic text retrieval. In *Information Processing and Management*, pages 513–523, 1988.

[23] MSR SenseCam - http://research.microsoft.com/sendev/projects/sensecam.

[24] Slife Labs - http://www.slifelabs.com.

[25] trec_eval - http://trec.nist.gov/trec_eval/.

Personality Based User Similarity Measure for a Collaborative Recommender System

Marko Tkalčič
University of Ljubljana Faculty of electrical engineering
Tržaška 25
Ljubljana, Slovenia
marko.tkalcic@fe.uni-lj.si

Matevž Kunaver
University of Ljubljana Faculty of electrical engineering
Tržaška 25
Ljubljana, Slovenia
matevz.kunaver@fe.uni-lj.si

Jurij Tasič
University of Ljubljana Faculty of electrical engineering
Tržaška 25
Ljubljana, Slovenia
jurij.tasic@fe.uni-lj.si

Andrej Košir
University of Ljubljana Faculty of electrical engineering
Tržaška 25
Ljubljana, Slovenia
andrej.kosir@fe.uni-lj.si

ABSTRACT

We propose a novel approach for calculating the user similarity for collaborative filtering recommender systems that is based on the big five personality model. Experimental results showed that the performance of the proposed measures is either equal or better (depending on the measure under evaluation) than the ratings based measures used in state-of-the-art collaborative recommender systems. This makes the proposed approach, with its benefits in terms of computational complexity, for calculating user similarities a very promising one for future collaborative recommender systems that will be more affect-oriented.

Categories and Subject Descriptors

H.1.2 [**Models and Principles**]: User Machine Systems; H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

Keywords

collaborative filtering, big five personality model, personality based user similarity measure, affective computing

1. INTRODUCTION

Recommender systems are gaining on importance as the amount of available multimedia content is expanding rapidly. The main drivers are digital television, video on demand and web2.0 services like flickr and youtube. It is very hard for end users to find suitable content in large databases where most of the content is irrelevant. This is why recommender systems that filter only relevant items for each user are important. Based on how recommendations are made Adomavicius and Tuzhilin [2005] differ between content based recommenders (CBR), collaborative filtering recommenders (CF) and hybrid recommenders (HR). CBR systems are based on user's inclinations towards specific attributes of the content item. The knowledge about the user is stored in data structures called user models. CF systems recommend items that similar users have liked in the past. HR use both CBR and CF approaches. The CF recommenders are further divided into memory-based and model-based recommenders. While memory-based recommenders work with the whole database of users' past ratings and preferences the model-based recommenders compile the whole database into a smaller data structure, the model [Pennock et al., 2000, Adomavicius and Tuzhilin, 2005].

In this paper we focus on memory-based CF systems. CF recommenders are based on the presumption that when the similarity between two users is high both users will like similar items. The similarity measure is thus a crucial part of any CF system.

1.1 Related Work and Problem Statement

State of the art multimedia recommender solutions (both CF and CBR) work on a very non-affective, technical basis as can be deducted from the overviews of Adomavicius and Tuzhilin [2005] and Lew et al. [2006]. CBR systems exploit metadata fields (e.g. genre) to build user and item profiles upon which the recommendations are made. On the other hand, CF recommenders rely solely on past user ratings to build the recommendations.

The strict technical approach in recommender systems that ignores the users' affective experiences during the consumption of multimedia content is odd because the entertainment industry is based on *giving people emotions*. In fact, Masthoff and Gatt [2006] treat the user's satisfaction when consuming multimedia content as an affective state. While there have been some research efforts in CBR systems to exploit emotive responses of users during content consumption for producing better recommendations [González et al., 2004, Lekakos and Giaglis, 2006, Shan et al., 2009, Tkalčič et al., 2009] CF recommenders have not received any such attention. The problem that we are addressing is the clear lack of affective elements in state of the art CF recommender systems.

The main reason probably lies in the fact that, by their nature, CF recommenders ignore content and user metadata where affective information could be stored. They are based on the presumption that *close* users (user *closeness* is calculated using a user similarity measure) have similar preferences for multimedia content. So, if users u_1 and u_2 are close and user u_1 has liked the content item h it is highly probable that user u_2 is going to like the item h as well. Thus the only place where one could make good use of emotions in CF systems is the user similarity measure.

One possible solution would be to construct a user similarity measure based on users' past emotive responses on same items. This would require to have an automatic emotion detection system. Such algorithms do exist [Donato et al., 1999, Picard and Daily, 2005, Zeng et al., 2009] but there is another problem: the similarities between users would need to be calculated on a regularly basis which is computationally very expensive. Each time a user consumes an item new information are added to the usage history which is the basis for the calculation of the similarity. This implies that after each update of the usage history the similarities between users should be recalculated.

1.2 Proposed Solution

We propose to use a similarity measure that yields, for each user u, a list of close neighbours that have in common a similar emotive response pattern to content items. We suggest to exploit end users' personalities to build such a similarity measure. According to McCrae and John [1992] personality tries to explain the individual differences in emotive reactions to common stimuli. So personality does affect the emotive response of users during multimedia content consumption. The big five personality model [McCrae and John, 1992, John and Srivastava, 1999] appears to be a promising instrument since the user's personality is described by a tuple of five numerical values which are convenient for computer calculations. Especially the extraversion and neuroticism factors seem to be tightly connected with individuals' emotive responses [Yik et al., 2002]. Despite the appearance that personality and emotions are two distinct theoretical and empirical streams substantial evidence has been found that they are tightly connected [Carver et al., 2000, Luminet et al., 2000, Davidson, 2001]. The main advantage of the proposed solution is that personality doesn't change with time. Thus once we have the big five tuple for a user we can calculate the neighbours in advance and have the list ready at any time which is a big improvement in terms of computational complexity.

In addition to the introduction of elements of affective computing in CF recommenders, the proposed solution has three advantages over standard memory based CF recommenders: (i) it solves the new user problem by introducing an initial questionnaire, (ii) it has lower computational requirements for the calculation of similarities between users and (iii) it lowers the impact of the sparsity problem as the calculation of similarities does not depend on ratings. The main drawback is that the initial questionnaire could be annoying for users.

Our hypothesis is that the CF recommender system's performance is not significantly lower when using the personality based measure than using the standard rating based measures.

1.3 Paper Outline and Notations

We provide the argumentation how personality affects emotions in section 2. In section 3 we give a description of the CF recommender. The experimental procedure along with the evaluated similarity measures is described in section 4. The results of the experiment are given in section 5. These results are further discussed in section 6.

The notations used throughout this paper are given in table 1.

2. PERSONALITY IMPLICATIONS ON EMOTIONS

Westen [1999] states that personality refers to the enduring patterns of thought, feeling, motivation and behaviour that are expressed in different circumstances. According to McCrae and John [1992] the big five factor model of personality is a hierarchical organization of personality traits in terms of five basic dimensions: extraversion (E), agreeableness (A), conscientiousness (C), neuroticism (N) and openness (O). These are the most important ways in which the individuals differ in their enduring emotional, interpersonal, experiential, attitudinal and motivational styles. The order of the factors is important as the first two (E and A) account for the largest percentage of variance in personality [John and Srivastava, 1999]. Each of the factors encompasses more specific traits.

Yik et al. [2002] reported that all five factors influence feelings and emotional behaviour. This is especially evident for the E and N superfactors while there is a nonzero correlation between the other three factors (O, C and A) and affect behaviour.

Johnson [2009] provided a description of the five factors and their subfactors. In the following of this section we give an overview based on Johnson [2009].

The E factor tells the degree of engagement with the external world (in case of high values) or the lack of it (low values). The subfactors of E are friendliness, gregariousness, assertiveness, activity level, excitement-seeking and cheerfulness. Extrovert people (high score on the E factor) tend to react with enthusiasm and often have positive emotions while introverted people (low score on the E factor) tend to be quiet, low-key and disengaged in social interactions.

The N factor refers to the tendency of experiencing negative feelings. People with high N values are emotionally reactive. They tend to respond emotionally to relatively neutral

Sign	Description
U	set of users
H	set of items
$u \in U$	single user
$h \in H$	single item
$e_L(u,h) \in \Omega_L$	Likert rating given to item h by user u
Ω_L	set of discrete Likert values (1 to 5)
$e(u,h) \in \Omega$	binary rating given to item h by user u
$\Omega = \{C_0, C_1\}$	set of binary ratings/classes (C_1 =relevant / C_0 =non-relevant)
$d(u_i, u_j) \in [0, \infty)$	distance function between users u_i and u_j
$sim(u_i, u_j) = \frac{1}{1+d} \in (0, 1]$	similarity measure between users u_i and u_j
$\vec{b}_i = (b_{i1}, \ldots, b_{i5})$	vector of big five values for the user u_i
$F_j(\vec{b}) = \lambda_j \vec{b}^\top . \vec{v}_j$	principal components of the big five values
$d_R(u_i, u_j)$	rating based distance function between users u_i and u_j
$d_E(\vec{b}_i, \vec{b}_j)$	big five euclidian distance function between users' u_i and u_j respective big five vectors $\vec{b}_i$ and $\vec{b}_j$
$d_{PCA}(\vec{b}_i, \vec{b}_j)$	weighted big five euclidian distance function between users' u_i and u_j respective big five vectors $\vec{b}_i$ and $\vec{b}_j$. The weights are the PCA components
$\bar{e}_L^{NN}(u,h)$	average rating of user's u nearest neighbours for the item h in the Ω domain
$\bar{e}_L^P$(h)	average rating of all users for the item h in the Ω_L domain
$\hat{e}_L(u,h) \in [1, 5]$	estimated numerical rating of item h for user u
$\hat{e}(u,h) \in \Omega$	estimated binary rating of item h for user u
k	number of neighbours
m	threshold
n	number of neighbours who have rated item h
$H_R(u)$	set of items relevant for the user u

Table 1: Notations used throughout the article

stimuli. They are often in a bad mood which strongly affects their thinking and decision making. Low N scorers are calm, emotionally stable and free from persistent bad mood. The subfactors are anxiety, anger, depression, self-consciousness, immoderation and vulnerability.

The distinction between imaginative, creative people and down-to-earth, conventional people is described by the O factor. High O scorers are typically individualistic, non conforming and are very aware of their feelings. They can easily think in abstraction. People with low O values tend to have common interests. They prefer simple and straightforward thinking over complex, ambiguous and subtle. The subfactors are imagination, artistic interest, emotionality, adventurousness, intellect and liberalism.

The C factor concerns the way in which we control, regulate and direct our impulses. People with high C values tend to be prudent while those with low values tend to be impulsive. The subfactors are self-efficacy, orderliness, dutifulness, achievement-striving, self-discipline and cautiousness.

The subdomains of the A factor are trust, morality, altruism, cooperation, modesty and sympathy. The A factor reflects individual differences in concern with cooperation and social harmony.

3. CONTENT FILTERING SCENARIO

Let us have a set of I users $U = \{u_1, u_2, \ldots u_I\}$ that are using our CF recommender system. We also have a set of J multimedia items $H = \{h_1, h_2, \ldots h_J\}$. Each time a user u consumes an item h she/he is required to give an explicit rating to the item which we denote as $e_L(u,h)$. The rating values e_L are taken from a set of discrete Likert values Ω_L which ranges from 1 to 5. The Likert values are mapped to binary ratings $e(u,h)$ from the set of possible values (classes)

$$\Omega = \{C_0, C_1\} \tag{1}$$

where $C_0 =' non-relevant'$ and $C_1 =' relevant'$ by applying the mapping $\Omega_L \rightarrow \Omega$:

$$e(u,h) = C_0 : e_L(u,h) <= 3 \tag{2}$$

$$e(u,h) = C_1 : e_L(u,h) > 3. \tag{3}$$

The usage history of a memory based CF system can be represented as a table of all item ratings given by users until the observed moment (see Tab. 2 for a hypothetical example).

	h_1	h_2	...	h_J
u_1	4	2		1
u_2		2		
u_3	2			1
u_4		3		
u_5	3			
u_6	1			2
u_7		5		3
...				
u_I		2		4

Table 2: Usage and rating history table. As all users have not rated all items there are several empty entries.

The user u then chooses an item h to consume and then gives it a rating which is stored in the usage history table.

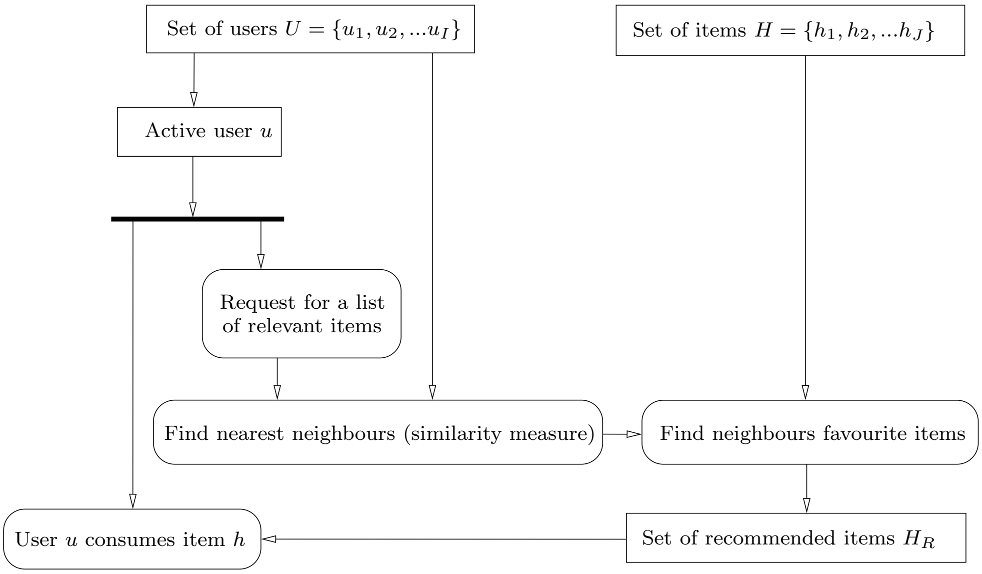

Figure 1: CF recommender system scenario.

3.1 Rating Prediction in Content Filtering Recommenders

The recommending procedure is centred around a single user u (see Fig. 2). When user u requests the CF a list of recommended items the system calculates binary rating predictions $\hat{e}(u, h)$ from the set Ω for all the items that have not been viewed by the user. The list of recommended items $H_R(u) \subset H$ is composed by all items that fall in the C_1 class (relevant items),

$$H_R(u) = \{h : \hat{e}(u, h) = C_1\}. \tag{4}$$

The procedure starts by calculating the list of k nearest neighbours for the user u. The similarity between user u and all other users is calculated using a user similarity measure $sim(u, u_i)$ where $u_i \in U \backslash \{u\}$.

The k users with the highest values $sim(u, u_i)$ are chosen as the k nearest neighbours.

For each item h two ratings that take values from Ω_L are calculated: the user's u nearest neighbours' average rating $\bar{e}_L^{NN}(h, u)$ and the overall average rating $\bar{e}_L^P(h)$. Both average ratings are aggregated into a single numerical rating prediction value $\hat{e}_L$ using the true Bayesian estimate [imd, 2009]

$$\hat{e}_L(u, h) = \frac{n}{n+m}\bar{e}_L^{NN}(u, h) + \frac{m}{n+m}\bar{e}_L^P(h) \tag{5}$$

where n represents the number of neighbours who have rated the item h while m represents the threshold value. If we set m to a lower value then $\hat{e}_L(u, h)$ is more dependent of $\bar{e}_L^{NN}(u, h)$ and vice versa, if we set m to a value close to n then $\hat{e}_L(u, h)$ is more dependent on $\bar{e}_L^P(u, h)$.

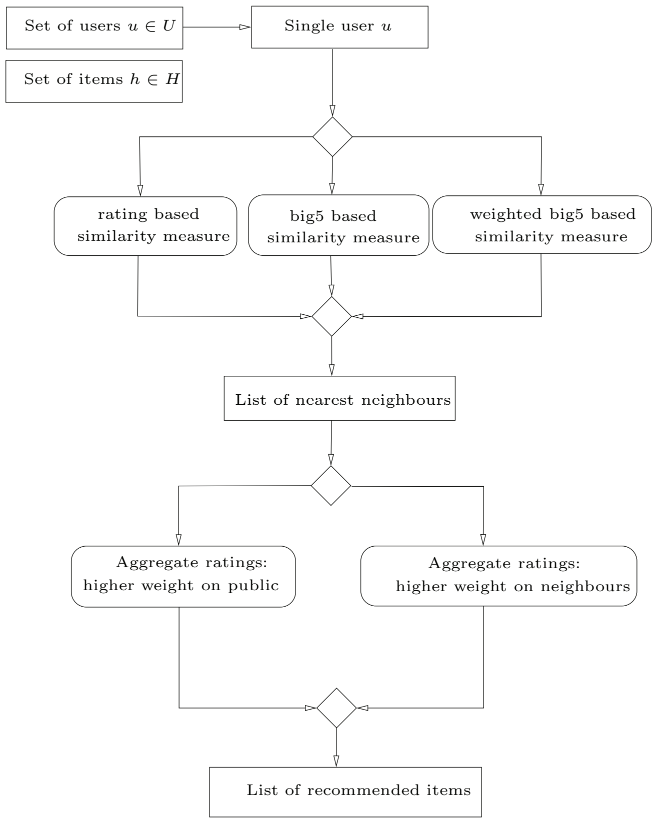

Figure 2: Rating prediction procedure for a selected user u

The estimated rating $\hat{e}_L$ is a numerical value with $\hat{e}_L \in [0, 5]$. The final estimation of the item's binary rating is calculated by thresholding $\hat{e}_L$

$$\hat{e}(u,h) = C_0 : \hat{e}_L(u,h) <= 3 \tag{6}$$

$$\hat{e}(u,h) = C_1 : \hat{e}_L(u,h) > 3 \tag{7}$$

After this procedure has gone through all non viewed items and has thus calculated all the rating predictions the CF recommender compiles a personalized list of recommended items according to equation (4) which is a reduced and thus manageable subset of relevant items $H_R(u)$ among all available items H.

4. MATERIALS AND METHODS

We performed a two stage experiment. First we acquired the dataset needed and then we used it in the CF recommender system. We used the precision P, recall R and F measure to assess the performance of the recommender system with the evaluated similarity measures and thus to validate (or reject) our hypothesis.

4.1 Dataset Acquisition

The subjects involved in the dataset acquisition phase were $N_U = 52$ (15 males) with average age of 18.3 years (SD = 0.56). Each subject was shown a sequence of 70 images taken from the IAPS database [Lang et al., 2005]. The images were carefully chosen to equally cover the widest possible area in the value-arousal space of induced emotions. After viewing each image the subjects were required to rate the image on a Likert scale from 1 (meaning *I don't like it at all*) to 5 (meaning *I like it very much*) in order to advance to the next image. This procedure yielded a full usage history matrix (same structure as table 2 but without missing values).

In order to assess the big five personality values each subject was required to fill in a big five questionnaire. We used the IPIP questionnaire with 50 items [ipi, 2009]. Each of the five factors were covered by 10 items in the questionnaire. The subjects had to describe how accurately each statement describes her/him on a scale from 1 (meaning *very inaccurate*) through 3 (*neither inaccurate nor accurate*) to 5 (*very accurate*). Half of the statements had a positive relation to the describing factor and the other half had a negative relation. Each answer was corrected according to the relation to the factor and distinctly summed and normalized in order to yield separate sums for each factor. Table 3 shows an excerpt of our dataset.

4.2 Content Filtering Recommender Implementation

The acquired dataset was used in the CF recommender system developed by Kunaver et al. [2007].

We evaluated the CF using three different user similarity measures: (i) a standard, rating based measure (see equation (8)), (ii) an Euclidian big five based measure (see equation (9)) and a (iii) weighted Euclidian big five based measure (see equation (10)). After calculating the k nearest neighbours using these measures the item rating prediction (for all items in the dataset) was calculated by combining the average ratings of the neighbours and the average rating of all users (which we denote as *public*). We aggregated both average ratings into the final item rating estimation using the equation (5) with two weight configurations: (i) by putting more weight on the neighbours' average rating $\bar{e}_L^{NN}$ and (ii) by putting more weight on all users' average rating $\bar{e}_L^{P}$ (public). This yielded six CF recommender experiment runs:

- Rating based measure with more weight on neighbours
- Rating based measure with more weight on public
- Euclidian big five based measure with more weight on neighbours
- Euclidian big five based measure with more weight on public
- Weighted Euclidian big five based measure with more weight on neighbours
- Weighted Euclidian big five based measure with more weight on public

We chose $k = 5$ for the number of neighbours which is a rough 10% of all users in our dataset.

The three similarity measures $SIM = \{sim_R, sim_E, sim_{PCA}\}$ were calculated using three different distance measures $D = \{d_R, d_E, d_{PCA}\}$ respectively. For the respective similarity measures $sim \in SIM$ and distance measures $d \in D$ we set $sim = \frac{1}{1+d}$.

We calculated the rating based distance d_R between two users u_i and u_j based on past ratings of both users to all items except the observed one $h_k \in H \backslash \{h\}$

$$d_R(u_i,u_j)^2 = \sum_k (e_L(u_i,h_k) - e_L(u_j,h_k))^2 \tag{8}$$

and the big five distance measures using the users' respective big five vectors $\vec{b}_i$ and $\vec{b}_j$

$$d_E(\vec{b}_i,\vec{b}_j)^2 = \sum_l |b_{il} - b_{jl}|^2 \tag{9}$$

$$d_{PCA}(\vec{b}_i,\vec{b}_j)^2 = \sum_l |F_l(\vec{b}_i) - F_l(\vec{b}_j)|^2 \tag{10}$$

where the weights $F_l(\vec{b}) = \lambda_l \vec{b}^\top . \vec{v}_l$ are the result of the principal component analysis which yields

$$d_{PCA}(\vec{b}_i,\vec{b}_j)^2 = \sum_l |\lambda_l (\vec{b}_i - \vec{b}_j)^\top . \vec{v}_l|^2 \tag{11}$$

A study from Kunaver et al. [2007] showed that the best performance of a CF recommender in terms of F-measure is yielded when one of the two average ratings, $\bar{e}_L^{NN}(h)$ and $\bar{e}_L^{P}(h)$, has a higher weight than the other. In terms of equation (5) we had the parameter n set at $n = 5$ (because of the full dataset all neighbours have rated all items) and we evaluated two m parameter values $m \in \{1, 4\}$.

	BIG5 values					Content ratings					
	E	A	C	N	O	h_1	h_2	h_3	...	h_{J-1}	h_J
u_1	3,2	2,7	2,9	3,5	2,9	4	3	1	...	1	3
u_2	2,1	3,5	3,1	3,4	3,6	2	4	4	...	4	1
u_3	3,2	3,0	2,8	3,2	3,1	2	3	2	...	1	3
...	...	...	...	...	...	...	...	...	...	...	...
u_I	3,3	3,0	3,4	3,9	3,2	4	3	4	...	2	3

Table 3: The acquired dataset

4.3 Evaluation Methodology

Recommender systems are usually assessed based on how well they distinguish items that are relevant for a specific user from those that are non-relevant (see Herlocker et al. [2004]). The performance of this binary classification problem is described with the confusion matrix of correctly and incorrectly classified instances. Focusing on the relevant items a classifier yields four groups: (i) true positives (TP) are items that are relevant to the user and have been correctly classified as relevant, (ii) true negatives (TN) are items that have been correctly classified as non relevant, (iii) false positives (FP) are items that are non-relevant but have been misclassified as relevant and (iv) false negatives (FN) are items that are relevant but have been misclassified as non relevant. Following Herlocker et al. [2004] we calculated three numeric values that describe each a certain aspect of the CF recommender's performance: precision, recall and F-measure.

Precision P is the rate of truly relevant items among all the items classified as relevant by the CF system

$$P = \frac{TP}{TP + FP} \tag{12}$$

Recall R is the rate of returned relevant items to all relevant items

$$R = \frac{TP}{TP + FN} \tag{13}$$

The F measure combines precision and recall in a single numerical value

$$F = \frac{2PR}{P + R} \tag{14}$$

We calculated the P, R and F values for each single user $u \in U$ over all items in the dataset H and for each of the six CF combinations.

In order to prove (or reject) the hypothesis set in Sec. 1.2 we performed the one way analysis of variance to determine whether the differences of mean values of the F measure for the six CF combinations are significantly different or not.

5. RESULTS

In terms of mean values of P, R and F the big five based approaches performed better than the ratings based approaches. The mean values of P, R and F are reported in Tab. 4. The ANOVA analysis (with the significance level $p < 0.05$, we do not report the respective p values here) of the F-measure further showed that all the big five based approaches perform significantly better than the distance based measure with higher weight on the neighbours and are statistically equivalent to the distance based measure with higher weight on the public. The box plot of the values of the F-measure is shown in Fig. 3.

similarity measure	aggregation mode	P	R	F
rating based	neighbours	0.6666	0.5895	0.6268
rating based	public	0.7042	0.7401	0.7232
big5	neighbours	0.6309	0.8533	0.7062
big5	public	0.7093	0.8068	0.7442
weighted big5	neighbours	0.6455	0.8398	0.7165
weighted big5	public	0.7104	0.8064	0.7450

Table 4: Mean values of P, R and F for the different combinations of measures and weighting.

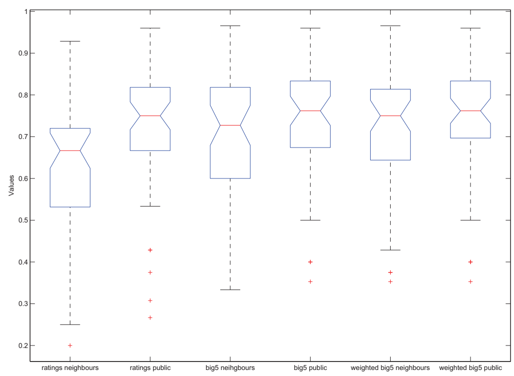

Figure 3: Box plots of the F measure results for six combinations of measures and weighting.

6. DISCUSSION AND FUTURE WORK

We performed an offline experiment of a memory based CF recommender system that relies on end users' personality parameters to determine the nearest neighbours, which is a crucial part of the recommending procedure. We compared four personality based similarity measures and two rating based similarity measures. The CF recommender system's performance results showed that the personality based measures were statistically equivalent or superior (the mean value of F was significantly higher) to the rating based measure which is used in most state-of-the-art CF recommender systems. The proposed approach has several advantages compared to the rating based similarity measures: (i) it is less computationally intensive because the similarities need to be calculated only when a user joins the system while in rating based similarity measures they need to be updated each time a user adds a new rating, (ii) according to the analysis of variance it is statistically equivalent to state-of-the-art rating based methods and (iii) it includes personality and consequently affect in the domain of CF recommenders which are very user oriented. This makes the proposed approach an excellent candidate for more efficient and more affective oriented CF recommender systems.

Unfortunately the proposed system still requires explicit user feedback in the form of ratings which can be annoying for end users. The fact that we introduced a user similarity measure that does not rely on ratings is not enough to give up the explicit feedback. The proposed measure allows the system to find users that have common multimedia interest which is reflected by their common emotive responses based on their personality values. But we still need to know which items are relevant for specific groups of users and this is the reason why we still need to have explicit ratings. Automatic methods for the detection of users' satisfaction (through the analysis of users' facial actions, gestures etc.) would be a valid alternative.

We must also be aware of the fact that personality is not the only parameter that influences our affective responses to stimuli. According to Westen [1999] the human behaviour (and thus emotional dynamics) is better described by *if-then* patterns where situational context plays an equal part as does personality. So in addition to the personality parameters we would also need to know the context (the *if* of the *if-then* pattern) in which the user is during the consumption of the item to improve the performance of the recommender system.

We acknowledge that the biggest drawback of the proposed approach is the need for an initial questionnaire to determine the big five values for each user. Such questionnaires are usually annoying and turn away potential users from using the system. For future work we propose to change the way the questionnaire is implemented into a more friendly and funny thing that can work as an attractor. A similar approach, called photo profiling, has been taken by Berger et al. [2007] that implemented a game where each user (of a personalized tourism destination recommender) had a set of images and she/he had to drag each image into a bin (like it / don't like it). In this way the system calculates two parameters that describe the potential tourist: the pack and the kick factors.

There are several other issues to address in the near future: (i) search for new measures that are based on personality/affect, (ii) include context awareness in the recommender system and (iii) perform more exhausting testings.

References

Imdb top 250. http://www.imdb.com/chart/top, June 2009.

International personality item pool: A scientific collaboratory for the development of advanced measures of personality traits and other individual differences. http://ipip.ori.org/, June 2009.

G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):734–749, 2005. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1423975.

Helmut Berger, Michaela Denk, Michael Dittenbach, Andreas Pesenhofer, and Dieter Merkl. Photo-based user profiling for tourism recommender systems. In Giuseppe Psaila and RolandWagner, editors, *E-Commerce and Web Technologies*, volume 4655/2007, pages 46–55. Springer Berlin / Heidelberg, 2007.

C.S. Carver, S.K. Sutton, and M.F. Scheier. Action, emotion, and personality: Emerging conceptual integration. *PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN*, 26(6):741–751, JUN 2000. ISSN 0146-1672.

R.J. Davidson. Toward a biology of personality and emotion. In Damasio, AR and Harrington, A and Kagan, J and McEwen, BS and Moss, H and Shaikh, R, editor, *UNITY OF KNOWLEDGE: THE CONVERGENCE OF NATURAL AND HUMAN SCIENCE*, volume 935 of *ANNALS OF THE NEW YORK ACADEMY OF SCIENCES*, pages 191–207. NEW YORK ACAD SCIENCES, 2 EAST 63RD ST, NEW YORK, NY 10021 USA, 2001. ISBN 1-57331-310-6.

Gianluca Donato, MArian Stewart Bartlett, Joseph C. Hager, and Paul Ekmanand Terrence J. Sejnowski. Classifying facial actions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21(10):974–989, October 1999.

G. González, B. López, and de la Rosa J. LL. Managing emotions in smart user models for recommender systems. In *Proceedings of 6th International Conference on Enterprise Information Systems ICEIS 2004*, volume 5, pages 187–194, April 2004.

Jonathan L. Herlocker, Joseph A. Konnstan, Loren G. Terveen, and John T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactionson Information Systems*, 22(1), January 2004.

Oliver P. John and Sanjay Srivastava. The big five trait taxonomy: History, measurement, and theoretical perspectives. In Lawrence A. Pervin and Oliver P. John, editors, *Handbook of Personality: Theory and Research*, pages 102–138. Guilford Press, New York, second edition, 1999. URL http://www.uoregon.edu/\~{}sanjay/pubs/bigfive.pdf.

John A. Johnson. Descriptions used in ipip-neo narrative report. http://www.personal.psu.edu/faculty/j/5/j5j/IPIPNEOdescriptions.html, June 2009.

Matevž Kunaver, Tomaž Požrl, Matevž Pogačnik, and Jurij Tasič. Optimisation of combined collaborative recommended systems. *International Journal of Electronic Communications*, 61:433–443, 2007.

P.J. Lang, M.M. Bradley., and B.N. Cuthbert. International affective picture system (iaps): Affective ratings of pictures and instruction manual. technical report a-6. Technical report, University of Florida, Gainesville, FL, 2005.

George Lekakos and George M. Giaglis. Improving the prediction accuracy of recommendation algorithms: Approaches anchored on human factors. *Interacting with computers*, 18:410–431, 2006.

Michael S. Lew, Nicu Sebe, Chabane Djeraba, and Ramesh Jain. Content-based multimedia information rtrieval: State of the art and challenges. *ACM Transactions on multimedia computing*, 2(1):1–19, February 2006.

O Luminet, E Zech, B Rime, and H Wagner. Predicting cognitive and social consequences of emotional episodes: The contribution of emotional intensity, the Five Factor Model, and alexithymia. *JOURNAL OF RESEARCH IN PERSONALITY*, 34(4):471–497, DEC 2000. ISSN 0092-6566. doi: {10.1006/jrpe.2000.2286}.

Judith Masthoff and Albert Gatt. In pursuit of satisfaction and the prevention of embarrassment: affective state in group recommender systems. *User Modeling and User-Adapted Interaction*, 16(3-4):281–319, September 2006.

Robert R. McCrae and Oliver P. John. An introduction to the five-factor model and its applications. *Journal of Personality*, 60(2):p175 – 215, 06 1992. ISSN 00223506.

David Pennock, Eric Horvitz, Steve Lawrence, and Lee C. Giles. Collaborative filtering by personality diagnosis: A hybrid memory- and model-based approach. In *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence, UAI 2000*, pages 473–480, Stanford, CA, 2000. URL `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.38.6498`.

Rosalind Picard and Shaundra Briant Daily. Evaluating affective interactions: Alternatives to asking what users feel. In *CHI Workshop on Evaluating Affective Interfaces: Innovative Approaches*, Portland, OR, April 2005.

Man-Kwan Shan, Fang-Fei Kuo, Meng-Fen Chiang, and Suh-Yin Lee. Emotion-based music recommendation by affinity discovery from film music. *Expert Syst. Appl.*, 36(4):7666–7674, 2009. ISSN 0957-4174. doi: http://dx.doi.org/10.1016/j.eswa.2008.09.042.

Marko Tkalčič, Jurij Tasič, and Andrej Košir. Usage of affective parameters in a content-based multimedia recommender system. *Submitted to the Elsevier International Journal of Human-Computer Studies*, 2009.

Drew Westen. *Psychology: Mind, Brain and Culture*. John Wiley and Sons, Inc, 2nd edition, 1999.

Michelle Yik, James A. Russell, Chang-kyu Ahn, Jose Miguel Fernandez Dols, and Naoto Suzuki. Relating the five-factor model of personality to a circumplex model of affect : a five-language study. In R. R. McCrae and J. Allik, editors, *The Five-Factor Model of Personality Across Cultures*, pages 79 – 104. Kluwer Academic Publishers, New York, 2002.

Zhihong Zeng, Maja Pantic, Glenn I. Roisman, and Thomas S. Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(1):39–58, Jan. 2009. ISSN 0162-8828. doi: 10.1109/TPAMI.2008.52.

Affective Awareness in Computer-Supported Collaborative Learning

Ulises Xolocotzin-Eligio
University of Nottingham
B3 Exchange Building, University of Nottingham, Jubilee Campus,
Nottingham, NG8 1BB
+44(0)115 8467672
lpxux@nottingham.ac.uk

ABSTRACT

This paper describes the benefits of integrating the affective awareness in the design and evaluation of CSCL technology. The role of the affective awareness during collaboration is outlined, highlighting the aspects of empathy, emotional convergence and biases in social perception. This is integrated with HCI research to illustrate how collaborative technologies could improve with the implementation of affective awareness and some solutions to do so. It is concluded that a framework to integrate affective awareness in CSCL should consider the social context and the ethical aspects involved.

Categories and Subject Descriptors

K.3 [**Computers and Education**]: Collaborative Learning

General Terms

Human Factors, Theory

Keywords

Affective awareness, Computer-supported Collaborative Learning, Social-affective computing

1. INTRODUCTION

The extensive arguments in favor of a social interactional approach for affective computing [1] and the works compiled in [2] indicate an increased interest in the interplay between affect and social factors in HCI. Nevertheless, the frameworks for the social dimension of affect have yet to be completed. Accordingly, this paper uses the term 'affective awareness' in reference to the capacity to understand and react to the emotions of others in a social situation. The aim is to highlight the importance of affective awareness between users in the design of collaborative technology, focusing on the field of Computer-Supported Collaborative Learning (CSCL).

Current views of evaluation and design of technology that aim to support collaborative learning have paid little attention to affective awareness between users. This is probably due to the inherent complexity of the interplay between affect and social interaction in collaboration. The first section presents a short overview of how the affective awareness works during collaboration. This is integrated with examples of research with collaborative technologies to highlight the opportunities that supporting affective awareness can bring to the user's experience. This elaboration is taken into the field of CSCL, where some research, including my own, is reviewed to explain how affective awareness works during interactions around collaborative learning technology. In the third section some ideas of how to implement affective awareness taken from various domains of HCI are invoked and its implications for CSCL are invoked. The concluding section explains some future steps in the integration of affective awareness in CSCL.

5th Workshop on Emotion in HCI, British HCI Conference 2009, September 1, 2009, Cambridge, UK, Copyright 2009 ACM

2. AFFECTIVE AWARENESS AND COLLABORATION

People engaged in collaboration constantly take into account the actions, intentions and emotions of others to sustain their joint actions [3]. Moreover, what accounts for 'good' and successful collaboration is a twofold motivation of the participants: to understand the ideas of, and be understood by, the other [4]. It has been suggested that the collaborators' mutual perception of these motivations underlies the 'natural' attraction for shared experiences that characterizes the development of humans as individual and species [5].

People engaged in collaboration understand the emotions of one another in two ways, as actors and observers, which from a first person perspective is explained as follows:

Actor: My emotions influence the actions and emotions of my partner. Being aware of this is a powerful drive for my affective experience. That is, I will align my emotions according to the expected reaction of my partner [6].

Observer: The emotions of my partner influence my actions and emotions. My partners' emotions are a potential clue to react appropriately to her intentions. It is known for example, that people cooperate more or less as a function of the known emotional state of a cooperator [7, 8].

This interlocking of actions and emotions involves mechanisms of social affective understanding grouped in this paper under the label of *affective awareness*. The affective awareness serves for collaborators to adequately perceive and react to the emotions of one another. An adequate reaction is that which aligns collaborators' intentions and behaviors in order to achieve the shared goal. It is expected for example, that collaborators provide mutual support to each other [9]. Three features of the affective awareness are highlighted below: affective convergence, empathy and biases in social perception.

2.1 Affective Convergence in Collaboration

A shared intentionality seems to provide an adequate ground for convergent emotional experiences to emerge from collaboration. Comparisons between cooperation and competition -a social situation in which the participants' intentions are opposed rather than shared- suggest differential effects in the tendency to express and experience the emotions of others known as *emotional contagion* [10, 11]. The expectations and awareness about the intentions and actions of cooperators lead to emotional convergence. Conversely, these expectations lead to emotional divergence in a competitive situation [12].

2.2 Empathy in Collaboration

To be empathic means to adequately understand and react to the emotions of other person. In collaboration, an adequate emotional reaction to the others' emotions does not necessarily mean that such an affective response has to be similar –leading to affective convergence. It can also be complementary. In either case, being empathetic has instrumental utility in cooperation [13].

Emphasis on emotions is important since people understand emotions differently to other kinds of mental states, e.g. thoughts. The understanding of emotions demands a wider view of the other. A person pays much more attention to verbal cues (e.g. content of words and intonation) than to non-verbal cues (e.g. facial expressions). when inferring others' thoughts, but uses verbal cues and nonverbal cues evenly when inferring other people's feelings [14].

2.3 Interpersonal Perception Biases

Natural biases in the interpersonal perception of emotions limit the affective awareness. One is the *false consensus effect*. This is a tendency of people to see their own behaviors, judgments and emotions as appropriate to existing circumstances. That is, people think that others in the same situation feel the same. This is strengthened when social affiliation is involved, e.g. in group work [15]. Another bias is the *illusion of transparency*. This is a tendency for people to overestimate the extent to which others can read their internal states.[16]

3. AFFECTIVE AWARENES IN COLLABORATIVE TECHNOLOGY

There are potential benefits in the integration of the affective awareness in the design and evaluation of collaborative technologies. Doing so can improve user experience by enhancing interpersonal perception and supporting coordination. The following examples show the opportunities to support affective awareness in collaborative scenarios.

The first example is the project described in [17]. The authors were concerned with situational awareness and the feeling of presence in mixed collaborative environments. They found that co-located collaboration and remote collaboration can be equally effective in operational terms when providing adequate resources –video communication and shared interfaces. Nevertheless, user experiences were different in regards of mutual awareness. Users focused more on the task and bothered less about the other in remote collaboration. Probably, participants could have felt a stronger presence and more interpersonal awareness have they been aware of the emotions of one another. In turn, a more shared experience of the collaborative experience could have been expected.

In a second example [18], authors compared the performance of a collaborative task in a multi-touch table and a multiple mice interface. They found that participants working around the multi-touch table showed more coordination but also more interferences, although these were quickly resolved in negotiations. Both coordination and negotiation indicate a constant interpersonal awareness. The authors concluded that rather than aiming to prevent interference, technology should aim to support the negotiation that resolves it. In this scenario, it is possible that more agile negotiation can be achieved if the interface allows users to be aware of the other's emotions. In fact, some research [19] suggests that emotion expression and understanding of the other influences the outcomes of negotiation

The examples above illustrate the potential of supporting affective awareness to frame shared emotional experiences and enhance empathy during collaboration. The following section deals with the role of affective awareness in CSCL.

4. AFFECTIVE AWARENESS IN CSCL

Affect has received an increasing attention in the CSCL field. However, the importance of affective awareness has received less attention than other, more individual aspects of affect. The potential benefits of doing so have been already pointed out. It has been theorized that the positive affect experience that arises from collaboration may precipitate useful cognitive activity [20]. This is partially confirmed by some empirical research. In [21], the author found that fluent and emotionally animated collaborations are associated with productive forms of peer collaboration –shared associative brainstorming- that improve the quality of collaborative story writing. It has also been reported that during scientific reasoning, collaborations between friends generate more evaluation and justification of proposals and more effective problem solving than collaborations between acquaintances [22]. This in turn, may be explained in part by the affective link between friends.

I have collected quantitative and qualitative data, e.g. self reports and screen recordings of dyads using collaborative technology to analyze how they understand the emotions of one another. The study reported in [23] compared the effects in the social affective understanding of collaborators when using a concept mapping tool and a collaborative learning computer game. Questionnaire data indicated that in both conditions collaborators strongly projected their own emotions onto their partners, which indicated they assumed a shared emotional experience. Nevertheless, such sharing was not always true, which explained the low accuracy of collaborators at inferring the emotions of the other. This is probably an expression of the false consensus effect, which in turn affects the empathic accuracy between collaborators. Further qualitative analysis indicated that those collaborators who in fact reported more similar emotions also showed more coordinated interaction, especially at playing the computer game. So it is possible to think of a technological solution to diminish the bias in the interpersonal perception to generate that affective convergence that benefits collaborations around the learning technology.

I invoked the examples above to illustrate the relevance of the affective awareness in CSCL, and show the potential benefits of designing for it in order to improve the quality of the interaction between users of CSCL technologies. Some ideas of how to do so can be taken from other HCI domains. Examples are provided in the following section.

5. IMPLEMENTING AFFECTIVE AWARENESS

There are examples of technological solutions to support affective awareness between users that can be applied to collaborative learning scenarios. One is the game *Emotional Flowers* [24]. This game uses the facial emotional expressions of users as input. In a wall screen, users' facial expressions are represented as flowers. The flowers grow and blossom with positive emotion expressions and decrease with negative expressions. The aim is to grow the flowers. The social nature of the game implies that users are aware of their own and others' emotions. This generates empathic emotional reactions between players, which in turn is associated with increasing interaction and interest for one another. Other examples can be found in the attempts to assist the social-affective difficulties of individuals with autism [25]. These projects aim to enhance empathy with wearable technologies for affect sensing and recognition. Although this technology has been developed in the context of autism research, it may well be used to support empathy between collaborators in learning or work settings. Another example of technology for affective awareness in face-to-face interaction is the *Subtle Stone* [26], a handheld device that enhances communication and interpretation of emotions between teacher and students. This project showed how users are motivated to communicate their emotions when the resources to do so with privacy are provided.

The examples above illustrate some possibilities of how to design for affective awareness. However, the differences between CSCL and other domains of HCI need to be taken into account when thinking about the support of affective awareness. A framework to determine the requirements for technology to support affective awareness in CSCL and its difference with other HCI domains needs to be constructed. To do so it will be important to recognize the nature of the social context, and the role that affective awareness plays on it. In scenarios of collaborative learning and collaborative work, we would like technology to frame complementary and convergent emotional experiences aligned with the sharing of intentions and coordinated actions. In other kinds of social contexts, the affective awareness plays a different role. Say, for example, a competitive game scenario such as the one described in the emotional flowers game, in which knowing about the emotions of others may well be undesirable. Moreover, affective awareness tools may well be used to deceive an opponent or to monitor the own affect.

But the treatment of affective awareness also requires attention to ethical concerns. As a private phenomenon, emotion and its communication are a delicate matter for users. In thinking about the implementation of affective awareness it is necessary to recognize the limits of technology to 'expose' the emotions of the users to other users.

6. CONCLUSIONS

The affective awareness is important during collaboration because of its role in the social affective experience of participants and its utility for coordinated action. The invoked examples of HCI research illustrate opportunities for the affective awareness to be supported, in order for the sharing of emotions and emotional convergence to be instrumental to the quality of the interaction. In CSCL, the reviewed research suggests that the quality and outcome of collaborations might be improved with technology aimed to generate an affective link between collaborators or to correct the natural biases in the interpersonal perception of emotions. The current solutions for implementing affective awareness developed in HCI domains such as collective gaming and clinical and educational scenarios have potential to be implemented in CSCL. Nevertheless, it is necessary to consider that the affective awareness plays different roles depending on the social context. In the future, these aspects need to be considered in the production of heuristics and evaluation criteria for affective awareness in CSCL.

7. REFERENCES

[1] K. Boehner, R. DePaula, p. Dourish, and P. Sengers, How emotion is made and measured, International Journal of Human-Computer Studies, vol. 65, pp. 275-291, 2007.

[2] N. S. Shami, T. H. Jeffrey, P. Christian, M. Michael, and M. Regan, Measuring affect in HCI: going beyond the individual, in CHI '08 extended abstracts on Human factors in computing systems Florence, Italy: ACM, 2008.

[3] H. Moll and M. Tomasello, Cooperation and human intelligence: The Vygotskian intelligence hypothesis, Philosophical transactions - Royal society of London series B biological sciences, pp. 693-648, 2007.

[4] D. Schwartz, The productive agency that drives collaborative learning, in Collaborative learning: cognitive and computational approaches, P. Dillenbourg, Ed. New York: Elsevier Science/Pergamon, 1998, pp. 197-218.

[5] M. Tomasello, M. Carpenter, J. Call, T. Behne, and H. Moll, Understanding and sharing intentions: The origins of social cognition, Behavioral and brain sciences, pp. 675-735, 2005.

[6] B. Parkinson, Emotions in direct and remote social interaction: getting thorugh the spaces between us, Computers in Human Behavior, vol. in press, 2007.

[7] T. Ketelaar and W. T. Au, The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction, Cognition and Emotion, vol. 17, pp. 429-453, 2003.

[8] C. D. Batson and N. Ahmad, Empathy-induced altruism in a prisoner's dilemma II: what if the target of empathy has defected?, European Journal of Social Psychology, vol. 31, pp. 25-36, 2001.

[9] M. Bratman, Shared Cooperative Activity, The philosophical review, vol. 101, pp. 327-341, 1992.

[10] B. Wild, M. Erb, M. Eyb, M. Bartels, and W. Grodd, Why are smiles contagious? An fMRI study of the interaction between perception of facial affect and facial movements, Psychiatry Research: Neuroimaging, vol. 123, pp. 17-36, 2003.

[11] E. Hatfield, J. Cacioppo, and R. Rapson, Emotional Contagion, Current Directions in Psychological Science, vol. 2, pp. 96-100, 1993.

[12] J. Lanzetta and B. G. Engils, Expectations of cooperation and competition and their effects on observer's vicarious emotional responses, Journal of Personality and Social Psychology, vol. 56, pp. 543-554, 1989.

[13] A. D. Galinsky, W. Maddux, D. Gilin, and J. B. White, Why It Pays to Get Inside the Head of Your Opponent: The Differential Effects of Perspective Taking and Empathy in Negotiations, Psychological Science, vol. 19, pp. 378-384, 2008.

[14] J. Hall and M. M. Schmid, Sources of Accuracy in the Empathic Accuracy Paradigm, Emotion, vol. 7, pp. 438-446, 2007.

[15] G. Marks and N. Miller, Ten years of research on the false-consensus effect: An empirical and theoretical review, Psychological Bulletin, vol. 102, pp. 72-90, 1987.

[16] T. Gilovich, K. Savitsky, and V. H. Medvec, The illusion of transparency: Biased assessments of others' ability to read one's emotional states, Journal of Personality and Social Psychology, vol. 75, pp. 332-346, 1998.

[17] A. Pauchet, F. Coldefy, L. Lefebvre, S. L. D. Picard, A. Bouguet, L. Perron, J. Guerin, D. Corvaisier, and M. Collobert, Mutual awareness in collocated and distant collaborative tasks using shared interfaces, in 11th IFIP International Conference on Human-Computer Interaction, Rio de Janeiro, BRAZIL, 2007, pp. 59-73.

[18] E. Hornecker, P. Marshall, N. S. Dalton, and Y. Rogers, Collaboration and Interference: Awareness with Mice or Touch Input, in ACM Conference on Computer Supported Cooperative Work, San Diego, CA, 2008, pp. 167-176.

[19] W. Steinel, G. A. Van Kleef, and F. Harinck, Are you talking to me?! Separating the people from the problem when expressing emotions in negotiation, Journal of Experimental Social Psychology, vol. 44, pp. 362-369, 2008.

[20] C. Crook, Motivation and the ecology of collaborative learning, in Rethinking Collaborative Learning, R. Joiner, K. Littleton, D. Faulkner, and D. Miell, Eds. London: Free Association Press, 2000.

[21] E. Vass, Exploring processes of collaborative creativity--The role of emotions in children's joint creative writing, Thinking Skills and Creativity, vol. 2, pp. 107-117, 2007.

[22] M. Azmitia and R. Montgomery, Friendship, transactive dialogues, and the development of scientific reasoning, Social Development, vol. 2, pp. 202-221, 1993.

[23] U. Xolocotzin-Eligio, S. Ainsworth, and C. Crook, Mutual influence and emotions in Computer Supported Collaborative Learning in 13th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI), Aug 25-29 2009, Amsterdam, 2009.

[24] R. Bernhaupt, A. Boldt, T. Mirlacher, D. Wilfinger, and M. Tschlegi, Using emotion in games: emotional flowers, in International conference on Advances in computer entertainment technology, Salzburg, Austria, 2007, pp. 41-48.

[25] R. Kaliouby, R. Picard, and S. Baron-Cohen, Affective computing and autism, Annals of the New York Academy of Sciences, vol. 1093, pp. 228-248, 2006.

[26] M. Alsmeyer, R. Luckin, and J. Good, Developing a novel interface for capturing self reports of affect, in CHI '08 extended abstracts on Human factors in computing systems Florence, Italy: ACM, 2008.

Pervasive Emotion Computing and a General Emotion Harmonization Model

Jiehan Zhou
MediaTeam, Department of Electrical and Information Engineering, University of Oulu
jiehan.zhou@ee.oulu.fi

Changrong Yu
English Philology, University of Oulu
changrong.yu@oulu.fi

Jukka Riekki
Intelligent System Group, Department of Electrical and Information Engineering, University of Oulu
jukka.riekki@ee.oulu.fi

ABSTRACT

Emotions - the experience of disappointed, anger, and joy results in an action tendency, e.g. requesting experts, talking with family. Appropriate and timely reactions to emotions harmonize people welfare by damping negative emotions or broadening positive emotions in normal. These reactions are being able to be facilitated by mobile services, e.g. instant text messaging, video call, in context of pervasive computing. This paper generalizes mobile services-enabled emotion harmonization as a novel concept of pervasive emotion computing (PEC). To approach this paradigm, this paper defines concepts related to pervasive emotion computing and overviews related work with emotion approaches in social science and emotion-oriented computing in computer science first. Pervasive emotion computing is conceptualized with proposing an emotion harmonization model and a service-enabled PEC framework. Combining conversation analysis techniques, PEC research approach is presented for realizing emotion-aware digital movie guidance service (EDMG) prototype. The result shows that EDMG service empowers people viewing experiences through interacting emotion-aware information.

Categories and Subject Descriptors

Categories and Subject Descriptors: I.2.7 [**Artificial Intelligence**]: Natural Language Processing – *Conversation analysis*; K.4 [Computer and Society]: Applications –*Emotion recognition*

General Terms: *Emotion harmonization model*

Keywords

Emotion regulation and mediation

1. INTRODUCTION

Anyone can become angry – that is easy. But to be angry with the right person, to the right degree, at the right time, for the right purpose, and in the right way – this is not easy (Aristotle, 384 BC-322 BC). When there are not instants of pleasure, anger, sorrow or joy, the mind may be said to be in a state of equilibrium. When those feelings have been stirred, and they act in their due degree, there ensues what may be called a state of harmony (Confucius, 551 BC- 479 BC). Aristotle's and Confucius's philosophical enquiries into virtue tell us that emotional intelligence, that is, skills to perceive, assess, and manage the emotions of one's self and of others, plays a vital role in human experience. A loss of emotions can have damaging effects; it can cause family separation, youth drug addiction, crime, or otherwise poor decision-making. On the other hand, a proper treatment of emotions can have benign effects; helping individuals, for example, to enhance parenting, to improve their health, to repair difficult relationships, or to succeed in their career.

Emotions and emotional intelligence have been studied by biologists, psychologists, anthropologists, sociologists, social constructionists, social psychologists, and other experts. The research results have been applied to training, developing and enhancing human emotional intelligence in order to fight against growing common social problems like crime, youth violence, drug abuse, terrorism, and the fragmentation of families.

Information and communication technology (ICT) permeates our modern society and has a strong effect on human life. In industry ICT enables companies to conduct project management, resource allocation, product design, product development and product maintenance efficiently. For individual users, ICT enables various services like email, electronic calendars, e-auctions, and downloading music and videos from the Internet. However, since the beginning of the 1990's, researchers have had the vision that ICT could do much more. ICT could disappear into the environment and small wearable devices. This invisible technology could offer useful services that support everyday activities whenever and wherever the user needs them – and it would offer the services in a user-friendly way, unobtrusively and in a natural fashion. This vision has several names and definitions that closely resemble each other; it can be called ubiquitous computing, pervasive computing, or ambient intelligence (AmI)[1-3]. However, human emotions are rarely included in these visions and the majority of the research in this field ignores human emotions.

We argue that what today's information and communication technology can do with facilitating people emotion harmonization and how? Motivated by answering the question, we propose a novel concept of pervasive emotion computing (i.e. regulation and mediation), which aims to ubiquitously damp negative emotion reaction and broaden positive emotion reaction by providing mobile emotion-aware services, e.g. video call, conference call, instant text message, multimedia message, virtual communities, various e-services - anytime, anyplace, any devices, and on any networks. For example, a real time community coordinated maintenance service can dispel a housewife's distress on a broken washing machine. Another example, Sophia's emotion is changed from negative to positive by chatting with her grandma through consuming a video call.

Obviously, pervasive emotion regulation and mediation will significantly alleviate the boundaries of location, time, cost and people in emotional communication and harmonization. The existing approaches and methodologies on emotion theories, emerging mobile services, pervasive service computing set a consolidate research foundation for advancing and realizing pervasive emotion computing. This paper has twofold contributions: (1) conceptualize pervasive emotion computing with proposing a research approach; (2) design and implement an emotion-aware digital movie guide service (EDMG). The remainder of the paper is organized as follows. Section 2 conceptualizes pervasive emotion computing. Section 3 proposes our research approach to pervasive emotion computing. Section 4 and 5 present experience on EDMG design and evaluation. Future work and conclusion are given in Section 6.

2. PERVASIVE EMOTION COMPUTING

Emotions can basically be classified by two types, positive and negative emotions. The rough distinction between these two emotions has been proposed by the cognitive approach [4]. We can regard them as polarized; there is a dividing line where one type of emotion changes into the other type. Negative emotions express an attempt or intention to exclude. Negative emotions are fueled by an underlying fear of the unknown, a fear of the actions of others, and a need to control them or stop them to avoid being harmed. Negative emotions lead to many problems in wellness, professional work, and social interaction, e.g. irritation, fear, despair.

For addressing pervasive human emotional experience, a few of emotional examples from children development are given as follows: (1) Sophia, a kid in the elementary school gets depressed because she cannot solve a math problem; (2) She is angry because she cannot blow the music instrument - recorder well. (3) She is sad because she loses her favorite book in a trip. (4) She is excited while receiving a special sports award in a city hall. (5) She feels bored while listening to a long wordy speech. (6) She is unhappy when she misses her grandma. Let's see how the emotions of Sophia are regulated and mediated in the real world. A likely investigation might be as follows: (1) She becomes happy after solving the math problem through instructions of her teacher. (2) She calms down when she can blow recorder with right guidance. (3) She cheers up after she gets another nice book. (4) She expects to receive an even higher sports award in the parliament next time. (5) She is released when the wordy talk finishes. (6) She is joyful after making a video call with her grandma. The above examples address that Sophia's negative emotions could be transformed into positive ones by offering proper help and services at proper time. One assumption is set that our emotional competency (e.g. regulation and mediation) can be enhanced by offering the proper emotional service and the proper service can initiate positive actions.

2.1 Pervasive Emotion Computing Model

Our observations suggest emotional experience is associated with knowledge acquisition. The emotion course in an emotional experience is divided into three phases, i.e. emotion reaction, emotion harmonization (i.e. regulation or mediation), and emotion reestablishment. During phase 1, emotion reaction results from the perception and recognition of differentiations between observed knowledge space and acquired knowledge space by the reactor. If the mapping is positive, then emotion reaction is regarded as positive, vice versa. The intensity of emotion reaction relies on the variation degree. If higher the degree is, emotion reaction (positive and negative) is more intensive. During phase 2, a series of actions will be organized by the reactor self or others for reducing or eliminating (for negative emotion reaction) or augmenting or broadening (for positive emotion reaction) the differentiation perceived during phase 1. These actions are called emotion-aware actions. The process of organizing actions by the reactor itself is called emotion regulation. The process of organizing actions by others is called emotion mediation. Both emotion regulation and mediation usually occur at the same time or in a temporal order. An important point is that emotion-aware actions can be facilitated by services. Here services are clearly defined as computer systems, especially mobile services in term of pervasive emotion computing, which can be consumed by the user and benefits the user in knowledge accumulation. During phase 3, emotion reactions is damped or broadened by emotion regulation and mediation. The original emotion reaction vanishes into reestablishing acquired knowledge, and a new knowledge space is established. And the reactor stays in a harmony state with a new acquired knowledge.

Thus it can be seen that pervasive emotion computing aims to harmonize people emotion reaction by providing mobile services at the right time, to right people, to the right degree, and in the right way. We have classified emotion harmonization into emotion regulation and mediation. Therefore, pervasive emotion computing can be categorized into two sub-tasks, i.e. pervasive emotion regulation and emotion mediation. Each of these two tasks can further be divided into two sub tasks - service discovery and development as follows. For service discovery, it means there exist desired services in the service registry; contrarily, for service development, it means there exists no desired services in the service registry.

(1) To discover emotion-aware services by the reactor self.
(2) To develop emotion-aware service by the reactor self.
(3) To discover emotion-aware services via communities.
(4) To develop emotion-aware services via communities.

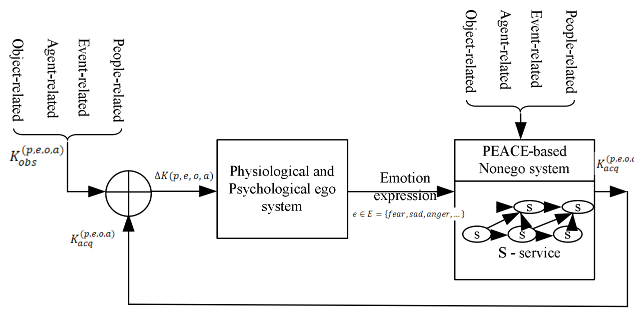

Figure 1. Emotion regulation and mediation model

Figure 1 depicts our proposed emotion harmonization model. The emotion harmonization model mainly consists of two parts. One is the inertia people physiological and psychological system, i.e. ego, which has a dominant element in dealing with emotion reaction, and stores acquired knowledge K_{acq}. Acquired knowledge K_{acq} distinguishes from observed knowledge K_{obs}. Both acquired knowledge K_{acq} and observed knowledge K_{obs} are captured and indexed based on PEACE model, which is specified in section 2.2. Another is the living environment, i.e. nonego system. In term of pervasive emotion computing, the nonego system is regarded as a set of services and a set of operators on services, i.e. $NES = \{S, \varphi\}$. The harmonization procedure can be broken down into three major steps: (1) Computing the differentiations between acquired knowledge K_{acq} and observed knowledge K_{obs} . The PEACE related variation yields a type of emotion $e \in E$. Emotion e has at least two attributions of type and intensity. Step 1 corresponds to emotion reaction phase. (2) Harmonizing emotion. To balance the differentiation, ego system will interact with nonego system through carrying out actions. Each service enchants and facilitates every one action. The important point is that the acquired knowledge K_{acq} is accumulated and value-added and ultimately reaches observed knowledge K_{obs} through consuming emotion-aware services. Step 2 corresponds to emotion harmonization phase. (3) Eventually, emotion reaction is damped and acquired knowledge K_{acq} is transited to a new one. Next emotion harmonization will be based on the transited K_{acq} . Figure 2 depicts the dynamics of emotion harmonization for a certain emotion e_i.

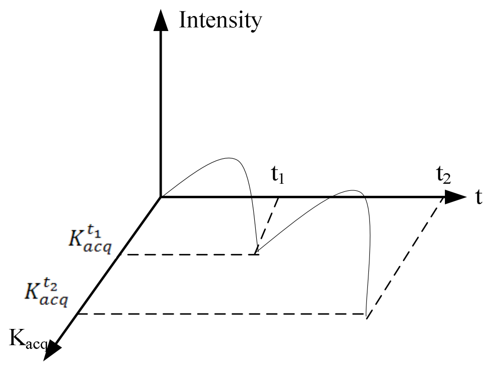

Figure 2. Emotion harmonization scheme for emotion

2.2 PEACE Metamodel for Recognizing Emotion

Emotion is disturbed or stirred by the differentiation between observed knowledge and acquired knowledge and reacted by a series of actions for accumulating acquired knowledge. When the differentiation becomes eliminated, emotion disturbance vanishes and acquired knowledge obtains a step transition. Thus it is seen that emotion reaction and harmonization always go with knowledge perception and knowledge acquisition. We call knowledge involving emotion harmonization as emotion-aware knowledge. It is very important to have a metamodel for emotion-aware knowledge perception and knowledge acquisition, further to harmonize emotion. We propose a PEACE meta model for unifying emotion-aware knowledge management. PEACE is a short name of the meta model for capturing emotion and emotion-aware knowledge, i.e. People-related, Event-related, Agent-related, and Object-related. Their specifications are briefly given as follows:

People-related emotion results from reactions to the differentiation between acquired knowledge and observed knowledge about people. Take a simple example, Sophia is unhappy when she misses her grandma.

Event-related emotion results from reactions to the differentiation between acquired knowledge and observed knowledge about events. Take a simple example, Sophia was disappointed when she missed yesterday's talent show party held by her school.

Agent-related emotion results from reactions to the differentiation between acquired knowledge and observed knowledge about agents. Agent refers to animal and man-made systems (e.g. computer, washing machine, car, etc.). Take simple examples, Sophia feels sad when her lovely pet dog died. Sophia's mum feels so bad when she couldn't fix her laptop.

Object-related emotion results from reactions to the differentiation between internal and external knowledge about objects. Take a simple example, Sophia is sad because she loses her favorite book in a trip.

Based on PEACE model, the above Sophia's emotional experience examples given in section 2.1 fall into the following types: (1) and (2) belong to Agent-related emotional experience, (3) belongs to object-related emotional experiences, (4) and (5) belong to event-related emotional experience, (6) belongs to people-related emotional experience.

3. RESEARCH APPROACH

PEC research approach is derived from the AmE vision [5, 6]. The first aim is to suggest an emotion model for expression and analysis of emotions in English conversation. The emotion model will be used for analyzing emotional information in artificial or real life data. Finally incorporating with pervasive service computing environment, computer can read the emotional information and provide corresponding mobile services in order to harmonize people's emotions. Figure 3 illustrates the AmE-driven emotion approach, which consists of three main building blocks:

Emotion modeling. Conversation is one of the main channels for communicating emotion. The expression and analysis of emotion in English conversation has not been systematically studied. However, amounts of existing research on emotion by biologists, psychologists, anthropologists, sociologists, social constructionists, social psychologists, and linguists set foundations. In the PEC study, a broad discourse perspective and the quadruple emotional state recognition model will be highlighted and used for detecting appropriate services (e.g. mobile voice and video call) underlying a certain emotional experiences.

Computer-aided emotion analysis and annotation. The emotion expression and analysis in English conversation will be studied by integrating computer technology and existing video annotation technology. The emotion model will be encoded from the emotion annotations occurring in naturally occurring conversation. There are two ways of creating emotion annotation, one by semi-automation, another one by computer automation.

Emotion-aware service computing. It is assumed that appropriate and real-time responses can regulate and mediate people's emotion experiences. Advance mobile services (e.g. voice call, video call, short message service (SMS), multimedia message service (MMS) make this assumption viable. In this block, people's emotional experiences will be regulated by communicating with matched appropriate mobile services in a pervasive service computing environment.

In the case of emotion-aware DMG service (EDMG), we applied cognitive emotion theory [7, 7] for tracking emotion information in dialogue, utilized Anvil [8] for emotion annotation. The annotation file will be read and integrated by our developed EDMG service. In this way, a preliminary prototype is implemented for enhancing the user's movie viewing experience in context of emotion. The following sections present the details on EDMG service design and implementation.

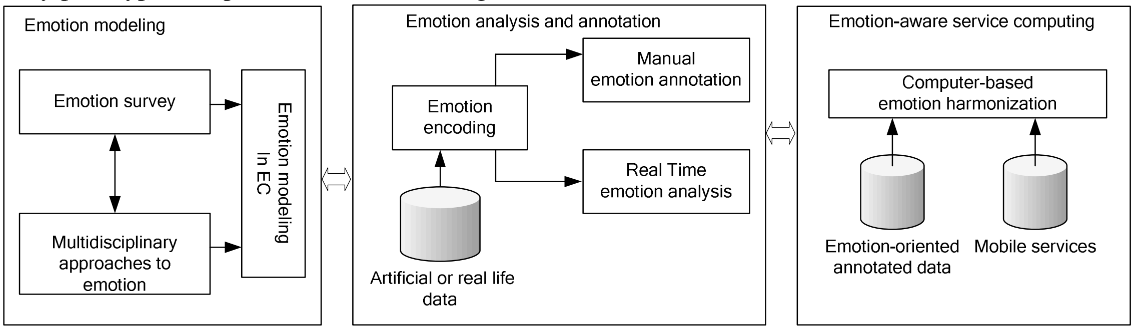

Figure 3. AmE-driven approach to pervasive emotion computing

4. DESIGN AND IMPLEMENTATION ON EDMG SERVICE

4.1 Emotion-aware Movie Viewing

This section highlights EDMG service-enhanced movie viewing experience as follows: (1) User Alice is viewing a movie with Windows Media Player 11. (2)She requests EDMG service as a user interface plug-in. (3) Alice presses the pause button while watching a scene clip of the movie, EDMG service is trigged and popups a data view window. (4) The data view window presents relevant emotional information about the scene in a chronological order, including the actors' emotion reactions of disappointed, distressed, etc. (5) Alice enjoys the extensive information by clicking links. These links are directed to external services, such as Web pages and Web applications. The functional requirements for EDMG service are divided into two parts:

1) Functional requirements for processing EDMG guide files. EDMG guide files provide additional emotion information on movie in order to extend user's viewing experience, especially emotion-awareness. Anvil can export manually annotated emotion information as an Anvil file. However Anvil file structure is different from the designed guide file structure used by EDMG service. Anvil file structure is specified in[8]. EDMG file structure is specified in section 4.4. Then a translator is needed for format transformation so that EDMG service can read Anvil file.
2) Functional requirements for developing EDMG. It should be easy to use EDMG service. Enabling and disabling EDMG service should be done through Media Player, without the need to access operating system functions. The user must be able to choose the EDMG guide file for relevant information of a movie. EDMG service must parse the EDMG guide file in order to find out any possible errors and the correctness of the file. Otherwise, the program would most likely function unexpectedly. As for the usability of EDMG service, there are four things to consider. First, a time buffer must be implemented. If the user would be expected to react and pause the movie in time to view information about an event, for example, she/he would often be forced to rewind the movie. This is not exactly a user friendly case. To avoid this, a buffer must be implemented to take into account information entries in close temporal proximity. Second, the entries must be shown in a Graphical User Interface (GUI), in the order they appear in EDMG guide file. Third, the quality of the GUI must be considered. This includes resizing the data view window and changing the theme (e.g. colours, etc).

4.2 EDMG Service Architecture

Figure 4 depicts the sequence diagram for EDMG service-enhanced user viewing experience, which consists of ten interacting components. The EDMG service starts when the user (component 1) views a movie and requests EDMG service (component 2). EDMG service opens a EDMG guide file (component 3). This guide file could be created either by writing it by hand, or using a tool chain intended to create these files. The tool chain consists of two programs: 3.1 Anvil Annotation tool, and 3.3 Anvil - EDMG Translator Tool. The annotation created with Anvil can be exported as an annotation file (3.2), which can be translated into the timing file for EDMG (3.4) (See section 4.4). As the file is opened, the EDMG service parses the EDMG guide file, and reads time-based emotion information. The time-based emotion information is only read once from the EDMG guide file during EDMG service. The broken lines refer to the one-time actions. After the EDMG service initialization, the control is given back to the user. After the user pauses the movie whenever she/he likes, a message is sent to Media Player (component 4) to pause the movie and retrieve the time-based emotion information. The EDMG service will then instruct the operating system (component 5) to draw a new window (i.e. emotion data view window (component 6)) to show emotion information. The emotion information includes labels for emotions, such as sadness, joy, etc. The user can interact with the Web browser (component 7). The user can continue to watch the movie after Web browsing.

4.3 Emotion Annotation Steps

This section presents how we annotate emotions in English conversation (Figure 5). The experiment consists of the following five main steps:

Step 1 - Choosing a tool. Our experiment aims to apply the initially summarized emotion model and theory in English conversation to create emotion annotation with AV data. Generic video annotation gives some extra information to the AV data. Extra information concerns events, people, and objects that are happening on the AV data. Information on emotions is focused on our experiments. In this way, other users or software systems can easily view and read emotion related information. These emotion and motivation related information could be used, for example, in the emotion intelligence training purpose to offer emotional information to the trainees, e.g. children emotional development. Nowadays there are many different AV annotation programs, e.g. YouTube Video Annotation[9], Anvil [10]. In our experiment, we use Anvil, which is a free video annotation program.

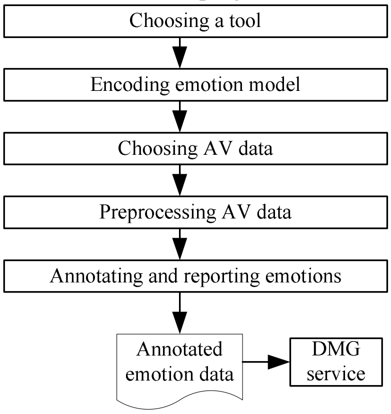

Figure 5. Steps of the experiment on emotion annotation with Anvil

Step 2 - Encoding emotion model. To make use of computer annotation systems, the extracted emotion model needs to be encoded with a certain computer language. In our experiment, we use the XML (eXtensible Mark-up Language) language to specify the emotion annotation. The specified emotion file consists of seven tracks based on emotion cues of emotional lexical words, emotional syntax, prosody, facial expression, gesture, laughter and sequential positioning. After settling down the seven tracks, we assign each track the forty eight emotion elements [11] as emotion labeling.

Step 3 - Choosing AV data. In the initial stage, the available data in our experiment comes from a clip of movie, 'FRIENDS - Last Scene In The Coffee House' (the video is available at http://www.youtube.com/watch?v=u3Qh4oB8u0o).

Step 4 - Preprocessing AV data. Sometimes AV data is too long and does not confirm to the required formats of the annotation software. In this situation, AV data needs to be preprocessed before making annotation. In our experiment, Simplified Universal Player Encoder & Renderer (SUPER) software[12] is used for format conversion. An open source program VirtualDub [13] is used for splitting AV data.

Step 5 - Creating and reporting annotation (Figure 6). After designing the annotation scheme and preprocessing AV data, it is ready to create a new annotation. In our experiment, first Anvil asks you to open the AV data you want to create annotations. Second Anvil asks you to open the emotion specification file. Third, Anvil provides four main windows [10] for creating annotation.

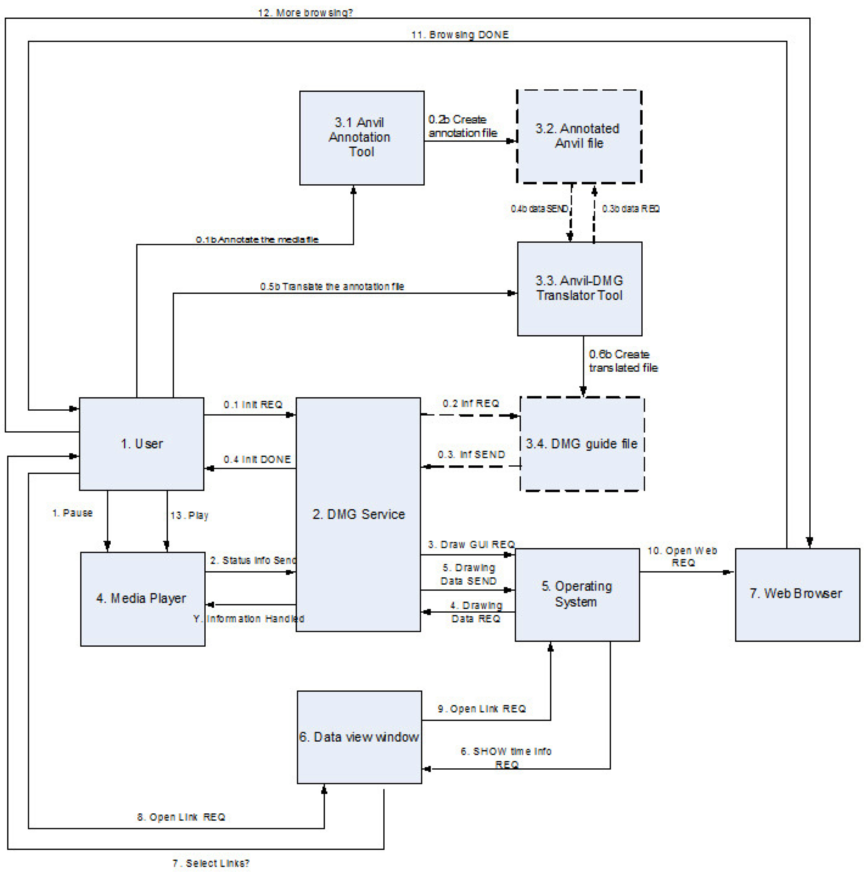

Figure 4. The sequence diagram of the EDMG service

Finally the annotated data on emotion will be input to EDMG and translated into EDMG guide file format given in section 4.4.

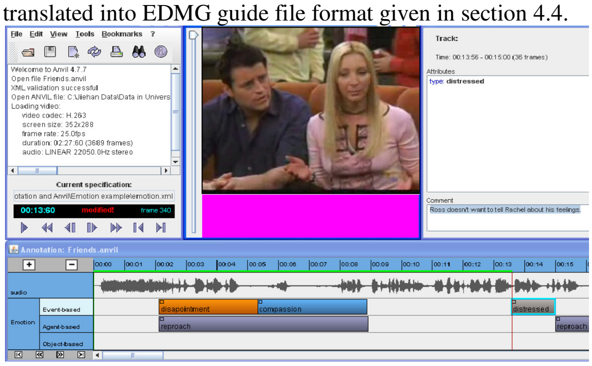

Figure 6. GUI for emotion annotation with Anvil

4.4 EDMG File Structure

Movie guide information is stored in the EDMG guide file, which consists of many information entries. Each entry has the following structure based on time-coding scheme:

Entry structure = starting time:ending time:type:value:link

Entry object = 00001:00007:Car:Aston Martin:http://www.astonmartin.com

Where,

1) 5 digits are for the stating time.
2) The colon is to separate the beginning time from the next 5 digits.
3) The next 5 digits are for the ending time.
4) The next is type information which can be for example car, actor, etc. The links that has the same type will be grouped under the title type.
5) The value of a type starts after colon.
6) The value is one of a type, which can be for example a model of a car like Aston Martin, a name of an actor, etc.
7) The link information starts after a colon.
8) An entry definition ends up the link information, which is passed to Web browser.

5. SERVICE IMPLEMENTATION AND EVALUATION

5.1 Implementation Techniques

The EDMG service is implemented as a plug-in for Windows Media Player 10/11. It can work in Windows XP or Vista. The EDMG service applied Windows plug-ins packets. Microsoft offers starting packets for different kind of self made plug-ins. The Visual Studio 2008 (referred to as Visual Studio in the sequel) is the recommended programming environment for creating EDMG service. Windows Software Development Kit (SDK) offers the necessary libraries and tools for coding work in Visual Studio. Visual Studio supports many coding languages. Visual C++ is selected as programming language. Active Template Library (ATL) template based C++ classes are used to support programming. Web links are shown by using Windows Media Player core interface. C++ programming is also used in parsing EDMG files. The EDMG service runs on Windows XP/Vista operating system.

The EDMG service demo version has the following constraints: The EDMG service is only compatible with Windows Media Player 11 and XP/Vista operating system at the writing moment.

5.2 Evaluation

Figure 7 illustrates the GUI of the EDMG service demo version in Windows Vista operating system. The window on the left is the Windows Media Player 11's main window and next to it, on the right side, is the EDMG emotion data view window. In the emotion data view window, the time coded entries are listed as link buttons in a chronological order. The first link button refers to the entry which has come first on screen; in this case it is 'disappointment'. As a default, the links will direct to http://www.emotionalcompetency.com/. The EDMG service demo is deployed and tested for home entertainment at the moment.

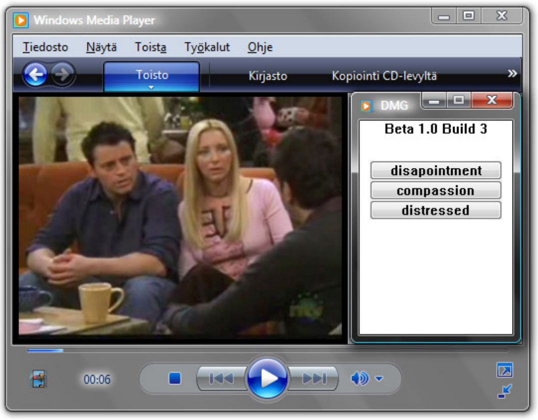

Figure 7. The graphical user interface of EDMG

6. CONCLUSION

Enquiring about 'what today's information and communication technology can do with facilitating people emotion harmonization and how?' we propose the novel concept of pervasive emotion computing. To approach PEC paradigm, firstly pervasive emotion computing is conceptualized with a generalized emotion harmonization model and PEACE model. Focusing on enhancing the user's viewing experience in context of emotion, Emotion-aware Digital Movie Guide (EDMG) service is designed and implemented by integrating Anvil emotion annotation. The demo shows that with EDMG user's viewing experience is enhanced and extended by interacting emotion information.

ACKNOWLEDGEMENT

This work was carried out in the project of SOPSCC (Pervasive Service Computing: A Solution Based on Web Services), funded in the Ubiquitous Computing and Diversity of Communication (MOTIVE) program by the Academy of Finland. Mr. Jaakko Moisanen is acknowledged for implementing EDMG service.

REFERENCES

[1] Ruyter, B. de and Aarts, E. 2004. Ambient intelligence: Visualizing the future," in AVI '04: Proceedings of the Working Conference on Advanced Visual Interfaces, 2004, pp. 203-208.

[2] Riekki, J., Sanchez, I. and Pyykkonen, M. 2008. Universal remote control for the smart world. in UIC '08: Proceedings of the
[3] Want, R. and Pering, T. 2005. System challenges for ubiquitous & pervasive computing. In ICSE '05: Proceedings of the 27th International Conference on Software Engineering, pp. 9-14.
[4] Fredrickson B.L. 1998. What good are positive emotions? Review of General Psychology: Special Issue: New Directions in Research on Emotion, vol. 2, pp. 300–319.
[5] Zhou, J., Yu, C., Riekki, J. and Kärkkäinen, E. 2007. AmE Framework: a Model for Emotion-aware Ambient Intelligence, 2007. The Second International Conference on Affective Computing and Intelligent Interaction (ACII2007): Lisbon, Portugal, pp.163-169.
[6] Yu Changrong and Zhou Jiehan. 2008. Research on service-mediated emotion computing and communication," in The First Finnish Symposium on Emotions and Human Technology Interaction, Tampere Finland, pp. 48-52.
[7] Ortony, Andrew, Clore Gerald L. and Collins, Allan. 1988. The Cognitive Structure of Emotions. Cambridge University.
5th International Conference on Ubiquitous Intelligence and Computing, pp. 563-577.
[8] Kipp Michael. 2003. ANVIL-the video annotation research tool. http://www.anvil-software.de/download.html. Retrieved by 1st July 2009.
[9] YouTube. 2009. YouTube Video Annotation. http://www.youtube.com/t/annotations_about. Retrieved by 1st July 2009.
[10] Kipp Michael. 2003. Anil 4.0 annotation of video and spoken language, user manual. University of the Saarland, German Research Center for Artificial Intelligence, Germany.
[11] Douglas-Cowie, E., Devillers, L., Martin, J-C., Cowie, R., Savvidou, S., Abrilian, S. and Cox, C. 2005. Multimodal databases of everyday emotion: Facing up to complexity. In proceedings of InterSpeech.
[12] SUPER. 2009. SUPER. http://www.erightsoft.com/SUPER.html. Retrieved by 1st July 2009.
[13] VirtualDub. 2009. VirtualDub, http://www.afterdawn.com/guides/archive/cut_avi_with_virtualdub.cfm. Retrieved by 1st July 2009.

Emotion in HCI – Real World Challenges

Proceedings of the 2009 International Workshop

Editors

Christian Peter, Elizabeth Crane, Lesley Axelrod, Harry Agius, Shazia Afzal, Madeline Balaam

Content

Affective computing is well recognised as an interdisciplinary field of research combining, bringing together aspects of human-computer interaction, design, psychology, sociology and the arts with the common goal of developing technology that serves the human in more sensible and sentient way. With the community growing ever faster and the number of projects, ideas, problems and challenges still rising faster than that of possible solutions, an annual meeting of those affected by such developments has proven to be a good anchor point in the course of the scientific year. This volume provides an account of the fifth workshop on emotion in human-computer interaction, held in the frame of the British HCI group's annual conference in 2009.

ISBN: 978-3-8396-0088-7