A Sensing Architecture for Empathetic Data Systems

Johannes Wagner
Florian Lingenfelser
Elisabeth André
Human Centered Multimedia
University of Augsburg
Augsburg, Germany
wagner@hcm-lab.de

Daniele Mazzei
Alessandro Tognetti
Antonio Lanatà
Danilo De Rossi
Department of Electrical systems and automation
University of Pisa
Pisa, Italy
mazzei@di.unipi.it

Alberto Betella[1]
Riccardo Zucca[1]
Pedro Omedas[1]
Paul F.M.J. Verschure[1,2]
[1] Universitat Pompeu Fabra
[2] Institució Catalana de Recerca i Estudis Avançats
Barcelona, Spain
paul.verschure@upf.edu

ABSTRACT

Today's increasingly large and complex databases require novel and machine aided ways of exploring data. To optimize the selection and presentation of data, we suggest an unconventional approach. Instead of exclusively relying on explicit user input to specify relevant information or to navigate through a data space, we exploit the power and potential of the users' unconscious processes in addition. To this end, the user is immersed in a mixed reality environment while his bodily reactions are captured using unobtrusive wearable devices. The users' reactions are analyzed in real-time and mapped onto higher-level psychological states, such as surprise or boredom, in order to trigger appropriate system responses that direct the users' attention to areas of potential interest in the visualizations. The realization of such a close experience-based human-machine loop raises a number of technical challenges, such as the real-time interpretation of psychological user states. The paper at hand describes a sensing architecture for empathetic data systems that has been developed as part of such a loop and how it tackles the diverse challenges.

1. INTRODUCTION

Over the past years, constant improvements in storage technology have led to increasingly large data sets in a wide range of specialist areas, such as astronomy, neuroscience, archaeology, history and economics [2]. Size and complexity make it increasingly difficult to make sense of the explored data, not only for novices, but also for experts. The Collective Experience of Empathic Data Systems (CEEDs) is a project funded by the European Commission FP7 that follows an unconventional approach: instead of exclusively relying on explicit user input, we exploit the power of the user's implicit signals in addition. While explicit user input refers to information the user intentionally communicates, implicit user input refers to signals which the user is not aware of, but which may be used by a system as information to adapt to the users' need.

To achieve this goal, users are immersed in a mixed-reality space, allowing them to explore complex data while freely moving around and interact with the physical and virtual world [3]. In addition, users are equipped with unobtrusive multi-modal wearable technologies, which allow them to navigate through and interact with the virtual world. However, beside explicit input (e.g. gestures, motion, verbal commands) the sensors are used to capture implicit responses (e.g. heart rate, electrodermal activity and gaze behaviour). Since only a small subset of sensory input of the human brain actually reaches conscious awareness, these otherwise hidden cues can help guide users' discovery of patterns and meaning within the datasets [10].

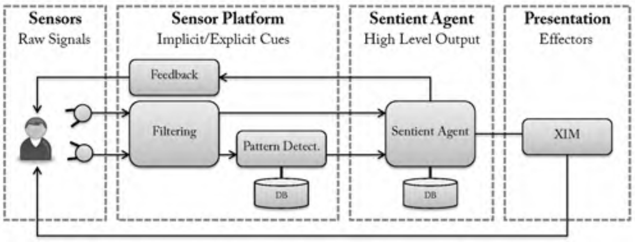
The potential of implicit human computer interaction has been pointed out in Schmidt's seminal paper [12]. Nevertheless, most work in the area of data exploration has focused on input the user provides explicitly to a system, see a recent survey by [9]. One of best explored implicit user signals is eye gaze as an indicator of user interest [6] or preferred choices [1]. However, eye gaze has usually not been considered in combination with other implicit signals. In the Callas project, a number of interactive art installations have been developed that respond to the multimodal emotional input of spectators [7]. However, the research did not focus on data exploration, but the experience of emotions that were intentionally expressed to control the installation.

The paper at hand explains the architecture of the CEEDs sensing system as part of an experience-based human-machine loop for data exploration. First we go into several problems that emerge when trying to extend a mixed-reality space with explicit and implicit user reactions. Solutions to these problems and how they have driven the design of the CEEDs sensing system will be discussed afterwards. Finally, an application that is currently implemented using the described architecture is shortly introduced.

2. CHALLENGES

A system meant to support and guide the user in a mixed-reality space must not interrupt the created experience. This starts from the sensors, which should be beyond the awareness of the user as far as possible. Though there is growing market for wireless sensor devices, the development of small, wearable and unobtrusive sensors that can be worn without discomfort is still in an early state. A way to re-

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in:
AH'13, March 07 – 08 2013, Stuttgart, Germany.



Sensor	Signal	Cues	Type	Meaning
Eye tracker	Gaze behaviour	Fixations	Implicit	Interest detection
e-Shirt	Acceleration	Energy		Activity detection
	Respiration	Breathing rate		Arousal detection
	Electrocardiogram	Heart rate		
Glove	Electrodermal activity	Peaks/Slopes		
	Forearm orientation		Explicit	Manual manipulation
	Finger position	Grab gestures		
Kinect	Body tracking			

Figure 1: Sketch of the sensing architecture (top) and a listing of sensors along with derived signals/cues and their meanings (bottom).

duce the physical load of the user is by combining several sensors within a single device, e.g. think of a data glove that measures both, finger position and electrodermal activity. The physical closeness of the sensors, however, may introduce certain signal artifacts, which have to be taken into account in order to not trigger false detections. Likewise, the user should be able to move in a natural way, i.e. the system should not restrict the mobility of the user, which also disturbs the measured signals.

Another issue when processing multiple signals in parallel refers to the synchronization between the modalities. Only by proper alignment of the data streams it becomes possible to relate one input source to another and create a coherent picture of the interaction.

Finally, the delay between user input (explicit and implicit) and presentation should be prompt enough to guarantee a fluent interaction between user and system. Hence, processing and interpretation of the recorded signal must happen in (near) real-time, while also a communication layer is required to allow the different subsystems to exchange messages in an efficient way.

3. ARCHITECTURE

In Figure 1 a sketch of the system architecture is presented along with a summary of sensors and signals together with their explicit or implicit meanings. It puts the user into an interactive mixed-reality space, while constantly reacting to its explicit and implicit responses and adapting in real-time the mixed-reality space according to the derived user state. In this way the loop between the user and the system is closed.

3.1 Sensor Devices

3.1.1 Glove

Figure 2: Newly developed wearable sensor devices. From left to right: e-Shirt, eye tracker and glove.

A multi-parameter sensing glove for the simultaneous acquisition of hand gestures and electrodermal activity (EDA) has been built [14]. Measuring both signals with a single sensor has the advantage that only one hand of the user is occupied, while the ergonomic design of the glove encourages natural movements and gestures. Both EDA and deformation signals are acquired and elaborated on-line by using the dedicated wearable and wireless electronic unit. The forearm orientation (in terms of Euler angles pitch, roll and yaw) is fixed in the dorsal part of the forearm close to the wrist. Finger position tracking is performed by means of textile deformation sensors. EDA signal acquisition is performed by using three electrodes made of conductive textile material on the index, medium and middle fingertip.

3.1.2 e-Shirt

The Smartex e-shirt system is a piece of garment that is meant for acquisition of electrocardiography (ECG), respiration and tri-axial accelerometry data [11]. Note, that again three types of signal types are obtained with a single sensor. The according sensors are integrated into the textile of the shirt and therefore not visible to the user. In a front pocket, a small electronic device is responsible for power supply and streaming of acquired data.

3.1.3 Eye Tracker

A wearable eye gaze acquisition system – permitting free user movement – sends the gaze point as two–dimensional coordinates, which are referred to the top left corner of the viewed scene image [8]. It is equipped with a lightweight wireless camera ($20g$, $2 \times 2 \times 2cm$), and an inertial motion unit that provides three rotation Euler-angles of the head. An open issue with the system is the "absolute head position" problem, which refers to the problem of mapping the recorded scene to actual points in the application. A solution could be to include markers in the virtual scene that can be detected in the view of the camera. Currently, the obtained gaze data is used to extract general gaze behaviour such as fixations or rapid eye movement.

3.2 Sensing Platform

The sensing platform is responsible to capture and process in real-time raw sensor data delivered by the devices. It is implemented using the Social Signal Interpretation (SSI) framework[1] [16].

Each sensor stream is provided on an own channel to enable specific preprocessing steps, e.g. in case of the glove

---

[1] http://openssi.net

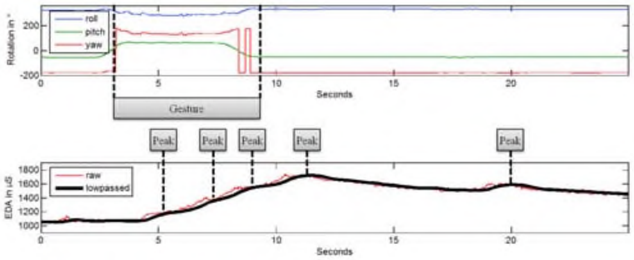


Figure 3: A synchronized recording of hand position and EDA recorded with the glove. Since hand movements are likely to cause artifacts in the EDA signal we ignore peaks if a gesture is detected at the same time. Here, the first three peak events.

forearm orientation, finger position and EDA come on separate channels. The first task taken over by SSI is the synchronization of incoming streams. This is important since devices are not synchronized per se. Synchronization is achieved by first establishing a stable connection to all connected sensors. As soon as a stable connectivity is warranted the framework starts to simultaneously buffer data streams. In regular intervals the number of samples received in each buffer is compared to a common clock and possibly adjusted in case a shortfall or surplus is observed.

After synchronization, each type of signal is given a tailored treatment to separate information relevant for the interaction from non-meaningful parts. In case of the glove this means that random movements have to be sort from interactive gestures. At the same time those movements may interrupt the measurement of the EDA signal (both sensors are part of the glove), which has to be taken in account during analysis of physiological data. For this purpose SSI offers filter methods like moving average or Butterworth filtering that can be applied in real-time. Such techniques help smoothing the signal and suppressing frequency bands that include noise. Some of the pre-processed signals – like the rotation of the forearm – are made available as a continuous stream for the purpose of continuous manipulation, e.g. to rotate a selected object in the virtual space.

Other information is only of interest during certain activities, e.g. when the fingers of the user are in certain positions to each other or a R-wave signals a heart beat in the ECG. Such patterns are detected by template matching or by applying certain thresholds that act as processing gates. Often these thresholds have to be calculated in an adaptive manner to react to fluctuating baselines and compensate differences between users. Since these patterns no longer occur at regular intervals and are of variable length, they are handled in event lists. A filter mechanism allows listening components to access events of interest and to remove events due to the presence of others, e.g. peaks detected from EDA are removed during a grab event (Figure 3). Likewise, the probability of particular events is lowered under certain circumstances, e.g. detection of heart rate is less reliable when the user is walking around derived from the acceleration signal.

3.3 Sentient Agent

The sentient agent, based on the Distributed Adaptive Control cognitive architecture [15], orchestrates the interaction between the user and the CEEDs application. It comprises an internal model of the user, and it possesses its own interests, intentions, and personality. This agent acts as an independent component in the architecture, receiving the inputs from the sensing platform and sending messages to the application, in order to optimize the content delivered to the user during data exploration. User inputs may be either explicit manipulations or implicit reactions. The explicit orders are meant for controlling the application, implicit reactions are used to alter the application's parameters in order to tailor the experience to the user. As possible user reactions can at this point emerge from a changed application environment, the experience-loop between user and application scenario is closed by the sentient agent. Both implicit and explicit user reactions are also stored in a connected database and accessed by the agent. This way former user experiences can be used as additional input to update the user model and for the agent's decisions.

Figure 4: User interaction in the XIM.

3.4 The eXperience Induction Machine (XIM)

Finally, an appropriate stage for presenting the application to the user is needed. For this purpose the described system is connected to the eXperience Induction Machine (XIM), which is a multiuser mixed-reality space covering a surface area of $5.5 \times 5.5m$ equipped with a number of sensors and effectors, see [3] and Figure 4.

4. APPLICATION

As an example application that is currently developed with the CEEDs architecture we will shortly introduce the Neuroscience Showcase. Neuroscience is one of the scientific fields that in the last decades has generated the most extensive number of data. For this reason we developed a 3D real-time visualization system to graphically represent the massive connectivity of brain networks in the XIM.

On a first stage a Neuroscience application was designed interfacing the XIM visualization with large–scale computational neuronal networks developed with the Iqr simulation software [4]. With this version we were able to show in real-time the neuronal activity of the running simulation and allowed the active exploration of the intricate organization and connectivity of the networks through real-time interaction in the XIM [5]. We are extending the Neuroscience application to represent the human brain structural *connectome* [13], a large dataset that stems from the analysis of the connectivity of the human neocortex. So far, this still follows a classical approach of visualising a complex dataset in a mixed reality environment.

On a second stage we now focus on enriching this classic visualisation approach with explicit and implicit interaction and have an intelligent sentient agent to close the loop between user experience and data presentation. Therefore, we use the components and their high-level output provided by the CEEDs sensing architecture (see Figure 1). Navigation in the 3D model of the neuronal connectivity is now possible by explicit orientation signals and grab gestures from the serial glove, accompanied by user tracking with a Kinect (Microsoft). At the same time, some of the properties of the represented model (e.g. the level of complexity of the connectivity network) are modulated depending on the current user's level of arousal measured through the glove and the e-shirt sensors. From analysis of the gaze behaviour – especially the detection of fixation events – we try to infer the level of interest the participant is having in the presented excerpt of the data. Based on these signals the application can thus zoom closer into that area of interest or guide the user's attention to other areas of the model.

5. CONCLUSION AND FUTURE WORK

In this paper, we introduced the CEEDs sensing architecture, a sensing architecture for empathetic data systems. We motivated our system with the need of giving users a novel way of experiencing large and complex datasets. Data is presented to the user inside a mixed reality space in a well-arranged and handy fashion (e.g. connectome data of brain connectivity as an animated, three-dimensional model).

In contrast to conventional cave system we go beyond explicit user interaction by making use of implicit reactions to guide users' discovery of patterns and meaning within the datasets. In order to not interfere the naturalness of the interaction, user behaviour is measured using unobtrusive sensor devices. To this end, we have described newly developed wearable sensor technology and how it reduces the physical load of the user by combining as many sensors as possible within few devices. This, however, requires certain treatment of the captured signals in order to remove artifacts caused by user movements and due to the physical closeness of the sensors. Therefore, we described a sensor platform able to synchronize various signal streams and apply fast signal processing and cue extraction for acceptable reaction timings. Explicit and implicit cues detected by the sensor platform are collected by the sentient agent, which includes the inner logic to react to the explicit user behaviour, as well as, his implicit responses and variates the presentation to the users needs – creating a loop between the system, presented data and the user.

Future work is mainly concerned with the evaluation of the system. The possibilities of experiencing huge datasets in this novel fashion are certainly huge but are still under exploration. To further intensify the user experience we plan to include direct physical feedback using a haptic device made of dielectric elastomer smart materials. It could be integrated into the glove, e.g. to create the sensation of actually touching a virtual object when grabbing it.

6. ACKNOWLEDGMENTS

The work described in this paper is funded by EU under research grant CEEDs (FP7-ICT-2009-5).

7. REFERENCES

[1] Bee, N., Prendinger, H., Nakasone, A., André, E., and Ishizuka, M. *Autoselect: What you want is what you get: Real-time processing of visual attention and affect.* In Perception and Interactive Technologies, International Tutorial and Research Workshop, PIT 2006, Kloster Irsee, Germany, June 19-21, 40–52.

[2] Bell, G., Hey, T., & Szalay, A. (2009). *Beyond the Data Deluge.* Science, 323(5919), 1297–1298.

[3] Bernardet, U., Inderbitzin, M., Wierenga, S., Väljamäe, A., Mura, A., & Verschure, P.F.M.J. (2008). *Validating presence by relying on recollection: Human experience and performance in the mixed reality system XIM.* In L. G. Anna Spagnolli (Ed.), 11th Annual International Workshop on Presence Padova Italy (Vol. 54), 16–18.

[4] Bernardet, U., & Verschure, P.F.M.J. (2010). *iqr: a tool for the construction of multi-level simulations of brain and behaviour.* Neuroinformatics, 8(2), 113–134.

[5] Betella, A., Carvalho, R., Sanchez-Palencia, J., Bernardet, U., & Verschure, P.F.M.J. *Embodied Interaction with Complex Neuronal Data in Mixed-Reality.* Virtual Reality International Conference (VRIC 2012).

[6] Buscher, G., Dengel, A., Biedert, R., & Elst, L. V. *Attentive documents: Eye tracking as implicit feedback for information retrieval and beyond.* ACM Trans. Interact. Intell. Syst. 1, 2 (Jan. 2012), 9:1–9:30.

[7] Gilroy, S.W., Cavazza, M., Chaignon, R., Mäkelä, S.-M., Niranen, M., André, E., Vogt, T., Urbain, J., Billinghurst, M., Seichter, H., & Benayoun, M. *E-tree: emotionally driven augmented reality art.* In ACM Multimedia (2008), 945–948.

[8] Lanatá, A., Armato, A., Valenza, G., & Scilingo, E.P., 2011. *Eye tracking and pupil size variation as response to affective stimuli: A preliminary study.* 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth), 78–84.

[9] von Landesberger, T., Kuijper, A., Schreck, T., Kohlhammer, J., van Wijk, J.J., Fekete, J.D., & Fellner, D.W. *Visual analysis of large graphs: State-of-the-art and future research challenges.* Comput. Graph. Forum 30, 6 (2011), 1719–1749.

[10] Lessiter, J., Miotto, A., Freeman, J., Verschure, P.F.M.J., & Bernardet, U. (2011). *CEEDs: Unleashing the Power of the Subconscious.* Procedia Computer Science, 7, 214-215.

[11] Paradiso, R., Loriga, G., Taccini, N., (2005). *A wearable health care system based on knitted integrated sensors.* IEEE Transactions on Information Technology in Biomedicine, vol. 9, 337-344.

[12] Schmidt, A., *Implicit human computer interaction through context.* Personal and Ubiquitous Computing 4, 2/3 (2000), 191–199.

[13] Sporns, O., Tononi, G., & Kötter, R., *The Human Connectome: A Structural Description of the Human Brain.* PLoS Comput Biol 1 (2005), no. 4, e42.

[14] Tognetti, A., Bartalesi, R., Lorussi, F., De Rossi, D., 2007. *Body segment position reconstruction and posture classification by smart textiles.* Transactions of the Institute of Measurement and Control, SAGE Publications, vol. 29, n. 3-4, 215–253.

[15] Verschure, P.F.M.J. (2012). *Distributed Adaptive Control: A theory of the Mind, Brain, Body Nexus.* Biologically Inspired Cognitive Architectures.

[16] Wagner, J., Lingenfelser, F., & André, E., (2011). *The Social Signal Interpretation Framework (SSI) for Real Time Signal Processing and Recognition.* Proceedings of Interspeech 2011.