ANTON NIJHOLT
University of Twente
Center of Telematics and Information Technology
Enschede, The Netherlands
anijholt@cs.utwente.nl

'Girlfriends and Strawberry Jam': Tagging Memories, Experiences, and Events for Future Retrieval

'Summer Guests' (Dutch television program in which well known people are interviewed, using television fragments, chosen by themselves, about what they find important in their life): Actress about a fragment of a documentary she wants to show: "It's so emotional, it's so ..". Interviewer: "But first tell us what it's about."

Introduction

In everyday life we interact with our friends and family members, with our colleagues and relatives, with people we arrange to meet and with people we meet by coincidence. We meet at home, in public transport, on the road, at parties and pubs, during sports and other recreational activities, in shops, offices and public spaces. We interact with the world around us. We experience a landscape, a canoe trip or a soccer game. In the latter example the experience can be from being in the audience, being the referee or being one of the players. In our home we can pick up a remote control, feel its weight and texture, and push buttons. Outside we can pick up a stone, admire its structure and experience dropping it in the water. When we walk we experience our body, its movements, our feet, our legs and our arms.

More and more, whatever activities we undertake, our environment becomes aware what we're doing and what we're doing and experiencing is recorded and archived for future use. Research in sensor technology, including cameras and microphones, aims at doing this in an unobtrusive way. We quote from [1]:

"According to the ISTAG vision statement (….), humans will, in an Ambient Intelligent (AmI) Environment, be surrounded by intelligent interfaces supported by computing and networking technology that is embedded in everyday objects such as furniture, clothes, vehicles, roads and smart materials - even particles of decorative substances like paint. AmI implies a seamless environment of computing, advanced networking technology and specific interfaces. This environment should be aware of the specific characteristics of human presence and personalities; adapt to the needs of users; be capable of responding intelligently to spoken or gestured indications of desire; and even result in systems that are capable of engaging in intelligent dialogue. Ambient Intelligence should also be unobtrusive - interaction should be relaxing and enjoyable for the citizen, and not involve a steep learning curve."

The technology promises to provide (real-time)support to our activities of daily living, our family life, our health and our safety concerns. These thoughts about AmI are not new, but in previous years we have seen an enormous amount of research devoted to the aims expressed in this quote. An extensive survey of how such ideas about AmI environments have given rise to all kinds of applications in the home environments can be found in [2]] and a survey on what is needed from a technological point of view (restricting ourselves to human-computer interaction issues) can be found in [3].

A second way of using recorded events and activities is recalling and re-experiencing certain events. Obviously, nowadays we are already using our own-made videos and photographs, notes that were taken, and diaries to recall certain events. But now, made possible by the Internet, Microsoft, Google, Yahoo and others are continuously extending their services to their users: chat environments, blogs, domain search engines, natural language access, picture search, sharing files, sharing music, sharing photographs, 2D maps and satellite images, annotated 3D Earth maps, etc. These companies have become interested in offering facilities for storage, user-friendly access and sharing of 'private' and social information. In current research projects we go a few steps further. In projects such as 'Electronic Chronicles', 'Memories for Life', and 'Lifelogs', or more general, the research efforts in 'digital prosthetic memories' [4], the idea is that for every person we have the possibility of continuously capturing, analyzing, interpreting and storing of temporal streams of data that concern that particular person. Developments in ambient intelligence, the commercial advantages companies see in offering support, and the research efforts that are under way will allow someone to build his or her own digital memory of (potentially) interesting events. And, as mentioned in [5], experience fusing can be a next challenge.

In this short paper we have some preliminary thoughts about tagging everyday life events in order to allow future retrieval of events or experiences related to events. Elaboration of these thoughts will be done in the context of the recently started Network of Excellence PetaMedia (Peer-to-Peer Tagged Media)[1] and the Network of Excellence SSPNet (Social Signal Processing), to start in 2009, both funded by the European Commission's Seventh Framework Programme. Descriptions of these networks will be given later in this paper.

Experiences

In human-computer interaction, we see the emergence of research on designing and generating experiences. While originally this research was mostly associated with the design of products, we now see experience design associated with environments, services, and entertainment. In particular the idea of ambient intelligence with its

[1] http://www.petamedia.eu/

sensor-equipped environments that on the one hand are meant to capture and understand events, activities and behavior of its human inhabitants, and on the other hand are meant to use this understanding to provide multi-media support and experiences to their human inhabitants, has helped researchers to think about generating and evaluating user experiences. Clearly, it is extremely hard to design and evaluate 'experiences'. For example, how should one design a game such that a gamer will become immersed in the game? And how should one perform a quantative evaluation of such immersion?

In daily life we have experiences. There is the experience of meeting a good colleague, a friend or a loved one. There is the experience of being together with an audience during a performance, being together with the performer, or being together with both. There is the experience of being the performer: feeling the pleasure of writing a nice paragraph, experiencing the exertion of running a marathon, or feeling emptiness after successfully finishing a particular task. But there is also the experience of feeling disappointed because of events or behavior beyond our control. Moods and emotions are associated with experiences. They are caused by the experiences and they interact with the experiences described in this and the previous paragraphs.

Obviously, there is a different kind of experience that should be mentioned that also receives attention in computer science research. It concerns being able to apply heuristics and knowing about 'best practices' in order to tackle a certain problem or task. This view of experience became part of computer science with the advent of expert systems in the 1980s and, many years later, with research interest in (domain-dependent) knowledge representation, knowledge management and access to fuzzy knowledge that has been extracted, for example by using machine learning algorithms, from recorded histories of interactions, activities, and events, in their particular contexts. Although related to the other topics mentioned here, for example, how do we know, express, retrieve and evaluate certain knowledge that may be needed in order to get our system working in the desired way, we will not explicitly comment on this view and application.

Memories and Experiences

We all have memories (a fuzzy memory of first meeting a particular person, being successful on a particular occasion, being at a particular place, driving a car in a city unknown to us), we all know and have learned things (eight times four is thirty-two, social conventions), and we all remember things from the past that help us managing new and otherwise problematic situations. We have memories of events and experiences. Fully recalling an experience requires similar mind, brain and body experience (activity) to that experienced during the original event. Probably we do not always mean this when we talk or think about re-experiencing events or re-calling events in which we are interested because of the experience associated with them. Retrieving the experience is a possible issue, re-experiencing the event by being exposed to the retrieved or a similar event is another issue.

Looking at the first issue, we need access to a formal representation of the detected and stored experience. It is not quite clear what can be done to obtain such a representation. Should we look at physiological (including brain activity) information obtained from physiological sensors and should we attempt to directly generate brain and other physiologically displayed activity from it? Or, taking the point of view of our second issue, should we try to expose our human 'experience sensors' to a retrieval or re-generation of the events (including everything that makes up a context) that designed our experience at that time? Obviously, annotating experiences has to be guided by such viewpoints.

From recent research we may conclude that until now most of the research has tackled the capturing of events in the physical world, the formal representation of such events and a straightforward multimedia presentation of recorded events. Research attempts aim at storing, fusing, manipulating, integrating, and summarizing recorded events. Many of these attempts can be found in conferences, workshops and papers devoted to projects such as Memories for Life, Hypermem, and MyLifeBits [6,7,8,9]. Nevertheless, also in these projects the presentation and regeneration of retrieved results in an immersive mixed and virtual reality environment is becoming an important research issue. Retrieval in the sense of 're-experiencing' an event taking into account the conditions of being immersed in the event again is seldom discussed. Re-experience in an immersive way (re-immersion) one of your weddings or another important event in your life is one of the many examples where the use of mixed and virtual reality has been proposed [6,7,9,10]. Recalling events in your 'ubiquitous home' (although it is, as usual, called experience retrieval) is discussed in [11]. In the latter paper we also see so an interesting and high-level 'why' retrieval question appear: "How did the strawberry jam that I bought last week finish in 4 days?" What needs to be retrieved, understood, and presented, in order to answer such and similar, maybe more emotionally oriented questions ("Why did my girlfriend leave me?")?

Community Aspects and Narrative

In addition to the (many) issues we introduced above two more issues that need to be mentioned. Both of them are also mentioned in O'Hara [8,12].

The first issue concerns the social community aspect of sharing experiences and having experiences that can be considered as group experiences rather than individual experiences. Experiences are personal and experiences are shared. People participating in the same event can have quite different experiences (consider a basketball game having opposing supporter groups in the audience, or consider thousands of people that experience an earthquake), but their experiences are also shared. Statistics inform us that every minute in this world hundreds of babies are born; statistics does not inform us about personal experiences related to newborn children. Similarly,

people playing computer games have shared and individual game experiences. As mentioned in [8] '… applications can be envisaged that integrate individual memories into group memories, such as family memories, company or corporate memories …).'

The second issue concerns the recall of memories and a possible role of narrative. Clearly, events have duration, meaning we are talking about representations that take this time feature into account. That is, a representation of events should take care of starting conditions that put a user in the desired 'right mind'. Retrieval is context-dependent, for example, on the one hand depending on context available and used to store, represent and interpret certain information, on the other hand depending on the context in which a user asks the system to recall a particular experience. Moreover, there is a story, a sequence of events, we are trying to remember or recall, or put another way, and only by connecting disparate events to a storyline we are able to remember them. Useful observations on memory organization, reconfiguration and recall can be found in [12].

Sensor Technology and Implicit Tagging

Sensor technologies allow us to register information from events experienced by individuals. Cameras, microphones, and other sensors can tell us about the behaviour of a particular inhabitant in a smart environment and the events that take place in the same environment. There is also the viewpoint of a particular user, obtained from his or her wearable information and communication technology devices that capture information that is unique for this particular user, allowing him or her to create a personal life log. In three successive EU projects (M4, AMI, AMIDA[2]) in which we participated the aim was to capture and understand multi-party interaction behaviour, in particular interaction behaviour in the context of small group meetings. Tools for browsing captured meeting events and meeting behaviour of participants have been developed. Clearly, efficient and intelligent browsing requires automatic analysis and tagging of events and behaviour. There is a quite smooth transition from browsing to retrieval and from retrieval to on-line support of meeting activities. In these projects, tools have been developed for automatic tagging of multi-party interaction behaviour. Moreover, using machine learning algorithms it was possible to extract 'high-level' information ("Who was the most dominant person during the meeting.") from 'low-level' features that could be automatically detected.

One project that continues this line of research is the earlier mentioned Network of Excellence SSPNet (Social Signal Processing). Here we look at nonverbal cues that show interest, agreement, sympathy, confusion, focus of interest, et cetera, in face-to-face and multi-party interaction. Among the research topics we are involved with are the detection of nonverbal cues that voluntarily or involuntarily show the user's interest in what he or she is perceiving. Detecting of such cues is done by audio analysis, by video analysis or by audiovisual analysis where information from the two modalities is fused in order to disambiguate signals. The nonverbal signals are obtained from face, gaze, head movements, body posture and acoustic signals related sighs, laughter, crying, pain, utterances of surprise, et cetera. Visual focus of attention can be derived from gaze or combinations of these features. Sometimes high-level information can be obtained from combinations of these features, often by using machine learning techniques, such that attitudes and affective states of the user can be obtained.

In addition to these ways of detecting the affective states related to interest, frustration, boredom, et cetera, we also measure physiological information from a user. Partly this is done with traditional means (skin conductivity, respiration, blood pressure, et cetera). However, the main efforts here involve brain-computer interfacing (BCI). Measuring (controlled) 'event-related potentials' is one of the obvious ways to detect user interests when confronted with displayed multimedia content. Measuring physiological signals with the aim to detect the affective state of someone engaged in a game is done in the Dutch national Smart Mix project BrainGain. Measuring gamer experience with the aim to adapt the game to what becomes known from the gamer during the interaction is one of the aims. In various related applications we have looked at situations where the user may be a student learning in a virtual reality environment, a gamer engaged in a game environment, or a meeting participant from who we want to read his interest in the topics that are discussed.

The message here is, we can learn about the user's interests and preferences by looking at his or her nonverbal behaviour. We are developing various ways to detect and interpret this nonverbal behavior. Our aim is to introduce new ways of real-time tagging of events by measuring a person's nonverbal behavior while looking at these events and while these events are being recorded. This can be called 'implicit tagging'. Moreover, we can look at applications where only those events are recorded that match the interests of the user as shown in his or her nonverbal behavior. Clearly, this should be done in such a way that also the user profile (knowing about personality, interests, preferences, background, and activity-related context) are taking into account when making such a decision. And clearly, in such a situation other users may take different decisions and therefore may help to obtain different viewpoints on the way a particular event should be tagged.

Conclusions and Future Work

In [2] we discussed the possible use of the AMI approach and the AMI technologies in other settings than meetings. In this short paper we introduced the problem of retrieving and re-experiencing events. Various viewpoints were introduced, including the use of physiological information, the retrieval of multimedia information, the use of immersive mixed and virtual reality, and storytelling (narrative). In various EU or 'Grand Challenges' projects we now see attempts to tackle problems associated with retrieving and re-experiencing of events. The aim of this note was to give a rough sketch of this field of research and its problems and how existing research,

[2] http://www.amiproject.org/

among other things research on capturing and interpreting human multi-party interaction behaviour, can help to tackle the problems related to implicit tagging in this field.

Acknowledgements This work is supported by the European IST Programme Project FP6-033812 (Augmented Multi-party Interaction). The EU IST Programme PetaMedia (Peer-to-Peer) provides an other context for our research. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

References

[1] IST Advisory Group (ISTAG). Ambient Intelligence: from Vision to Reality. In Riva, G., Vatalaro, F., Davide, F., Alcañiz, M. (Eds.) Chapter 4 in Ambient Intelligence. IOS Press, 2005, 45-68.
[2] A. Nijholt. Google Home: Experience, Support and Re-Experience of Social Home Activities. Information Sciences. Special Issue on Ambient Intelligence. Vol 178/3, Elsevier, Amsterdam, 2008, 612-630.
[3] M. Pantic, S. Pentland, A. Nijholt, & T. Huang. Human computing and machine understanding of human behavior: A survey. In: Artificial Intelligence for Human Computing. Lecture Notes in Artificial Intelligence 4451, M. Pantic, S. Pentland, A. Nijholt, & T. Huang (Eds.), Springer-Verlag, Berlin, 2007, 47-71.
[4] V. Kalnikaité, V. and S. Whittaker. Software or wetware?: discovering when and why people use digital prosthetic memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA, April 28 - May 03, 2007). CHI '07. ACM, New York, NY, 71-80.
[5] G. Pingali. Experience fusion: the next challenge for CARPE. In *Proceedings of the 3rd ACM Workshop on Continuous Archival and Retrival of Personal Experences* (Santa Barbara, California, USA, October 28 - 28, 2006). CARPE '06. ACM, New York, NY, 1-1.
[6] B.H. Prananto, I.-J. Kim, H.-G. Kim. Multi-Level Experience Retrieval for the Personal Lifelog Media System. IEEE International Conference on Signal-Image Technology & Internet-based Systems, 2007, 162-169.
[7] N. Correia, L. Alves, J. Santiago, L. Romero. Storing and replaying experiences in mixed environments using hypermedia, in: Memory and Sharing of Experiences Workshop, Pervasive 2004 Conference, Vienna, Austria, April 2004.
[9] N. Correia, L. Alves, R. Sa', J. Santiago, L. Romero. HyperMem: a system to store and replay experiences in mixed reality worlds, in: 2005 International Conference on Cyberworlds, November 23–25, Singapore, 83–90.
[9] K. O'Hara and N.l Shadbolt (Eds.). Road Map for Memories for Life Research. 1st January, 2008.
[10] P. Turner, S. Turner, D. Tzovaras, Reliving VE day with schemata activation, in: M. Slater (Ed.), Proceedings 8th International Workshop on Presence: Presence 2005, London, 2005, 33–38.
[11] G.C. de Silva, B. Oh, T. Yamasaki, K. Aizawa. Experience Retrieval in a Ubiquitous Home. Proceedings of the 2nd ACM workshop on Continuous Archival and Retrieval of Personal Experiences (CARPE'05), November 11, 2005, Singapore, 35-44.
[12] K. O'Hara, R. Morris, N. Shadbolt, G.J. Hitch, W. Hall, and N. Beagrie. Memories for life: a review of the science and technology. J. R. Soc. Interface, 2006.