PROVIDING FEEDBACK ON EMOTIONAL EXPERIENCES AND DECISION MAKING

Jana Machajdik, Julian Stöttinger, Ester Danelova, Martin Pongratz, Lukas Kavicky, Roberto Valenti, Allan Hanbury

September 2011

Technical Report # DISI-11-476

In "Proceedings of the IEEE AFRICON", Livingstone, Zambia, September 13-15, 2011.

Providing feedback on emotional experiences and decision making.

Jana Machajdik*, Julian Stöttinger¶, Ester Danelova‖, Martin Pongratz‡, Lukas Kavicky*, Roberto Valenti†, Allan Hanbury§

*Computer Vision Lab, Vienna University of Technology, jana@caa.tuwien.ac.at, lkavicky@gmail.com
¶Dept. of Information Engineering and Computer Science, University of Trento, julian@disi.unitn.it
‖Department of Psychology, Masaryk University, danelova@fss.muni.cz
†Intelligent System Lab Amsterdam, University of Amsterdam, rvalenti@science.uva.nl
‡Institute of Computer Technology, Vienna University of Technology, pongratz@ict.tuwien.ac.at
§Information Retrieval Facility, Vienna, Austria, a.hanbury@ir-facility.org

***Abstract*—We present a novel lifelog system concept created to provide a human user with feedback on their conscious and unconscious emotional reactions and encourage the process of self-reflection by looking into an *affective mirror*. The emotion of the user is deduced from biometric data and enhanced by affective sound analysis and facial expression recognition of faces of the weares conversational partners. These high-level analyses also offer information about the social context of the user. All of the recorded data is processed using a novel modification of the spreading-activation theory which is used to model human-like associative thinking.**

***Index Terms*—affective computing, lifelog, visual diary, human digital memory, tools for self-reflection, emotional decision making, affective mirror**

I. Introduction

In this paper we address another necessity of human life that keeps us sane, apart from interpersonal relationships: The immanent need of self-reflection and experience revision, as a key feature of our intrapersonal relationship. This is an area not tackled by the new media so far. We trust in our digital systems when it comes to organizing our daily lives, storing our most precious memories or to defining ourselves with prestigious gadgets. Self-reflection is not part of this boom. And emotions keep flying by, either unnoticed (unrecorded) or verbalized transcribed to a very limited code. While modern psychology comes up with ideas how to become whole again and deepen our sensitivity and mindfulness concerning our emotions and bodily reactions, computer science (paradoxically taking lead in providing opportunities for personalized self-expression) goes the exact opposite way. We therefore need a system that encourages self-reflection, especially one that gives us feedback on our emotional reactions and decision-making processes. Such a system could work both ways: provide users with feedback on their individual processes and help machines deepen an understanding of human emotional decision making. The proposed system will provide visual, audio, and biometric data and log it in a synchronized way (converging information from various sources just like human brain does on a much more complex level). With this data, we do not only have an idea what, where and when we see and hear in our life, but we are also able to identify emotionally arousing moments of our day. This system results in a lifelog storing data we perceive with our ears and eyes, enhanced with information that does not usually pass the threshold of our consciousness: We are able to retrieve our feelings (in their most fundamental form of physical reactions). With the combination of this multi-modal information, when certain experiences can be matched not only by their visual perception, but by the sensual experience we actually encountered, we will be able to create new associations and come to realize hidden relations: This is crucial as people tend to be quite dishonest to themselves when it comes to their very private feelings. The system cannot be lied to and is completely honest. Sentences like "I have never felt so happy before" or "this is the worst job I've ever had" can now be measured and verified. Even long-term questions like "Do I really love him/her or am I just afraid to be alone?" will be tackled and implicit beliefs challenged. It is made possible by a combination of three recent developments: 1.) Lifelogs and their specialized devices to store personal data over a long period of time; 2.) information theory to match multi-modal features of very different sources and distributions; and 3.) free associative machine learning to identify patterns without prior knowledge of connections that are being searched for.

The paper is structured as follows. The following section II gives an overview over current diaries and lifelog systems. Section III explains the inherent need of this concept by a use case scenario. Data capturing is covered in Section IV, while Section V gives the concepts how to process gathered multi-modal data. Section VI concludes.

II. State of the Art

It all began with diaries - records with a sequence of entries arranged chronologically - created to report on what has happened over the course of a day (or another period of time). Next to descriptions of events, personal diaries usually focus on the writer's thoughts and feelings. Originally in handwritten format (in the form of books or notebooks), the diary transformed from paper to electronic formats concurrently with the daily use of personal computers. Along with new media, new forms of diaries developed. Adding photos, drawings and visual souvenirs gave diaries the form of *photo albums* and *visual diaries*. While the "paper version" was traditionally intended to remain private (or with a limited circulation among friends or relatives), this aspect widely changed as Internet became commonly accessible. Many people adopted the new medium to chronicle their lives with the added dimension of an audience - in form of *blogs*. Still, the main point of a traditional diary (public or private) is that the content is actively generated by the user, consciously for adding to the diary. However, documenting a life takes discipline and effort and hence the amount of information in the traditional diary is constrained either by the capacity of the medium or the writer. An individual is forced to consciously filter the content he/she writes down. And not knowingly, they also filter the content unconsciously (we cannot record what we are unaware of). The issue of necessity (and inevitability) to limit information that the user wants to have preserved is being addressed by new technologies. First envisioned by Vannevar Bush in 1945 [6], *lifeLogs* present a possibility to capture and store a whole lifetime of a person's personal information digitally, so that it can be retrieved whenever needed. In recent years, this idea was revived in an experiment of Gordon Bell [2], who started the initiative by scanning in all of his paperwork, photographs, medical records and other personal data. Today's lifelog systems go even further and passively collect and store a variety of different signals from the user, throughout the day – from tracking the use of ones personal computer or mobile phone, including GPS, to taking a photo every few seconds. The reason why diaries are (or can be) useful is not solely to remember events. They can also serve as a means of self-reflection. Self-reflection is an important process of human functioning leading to further self awareness and development (it is a prerequisite of change). However, this function of a diary is little exploited. The reason for this is that most diaries (and now lifelogs) are *WORN - write once, read never* [3]. In most cases, this is due to our limited capacity to create and read diaries. One problem is that in traditional diaries, too little information being recorded and the parts that they looking for in retrospect might be missing. Lifelogs, on the other hand, present the exact opposite problem. The captured information is too much for a human to sift through and ends with the same result, namely not finding what is needed. In both cases the recorded information will be useless.

Recently there are several projects creating lifeLog hardware or applications. Some of them are meant to play the role of a personal assistant, who knows about our plans and preferences and is able to remind us of important issues. However, we are most interested in systems which actively take the fact that we are emotional beings into account.

The Affective Diary [25] offers an aesthetic user interface that encourages users to reflect on their day and to evaluate their emotional reactions measured by wearable bio-sensors and visualized as colorful blobs. It also collects mobile activity data and enables the user to add their own content (photos or comments). Apart from the blob-like visualization of biometric data, the diary does not do any interpretation — that is left to the user. Affective Diary shows that even relatively simple, but well conceived setups are useful to animate humans to reflect on their lives.

The SCOPE [18] system shows a unique security application of a lifelog-like system. Hereby, data from mobile context and the heart rate and motion of the user are processed to recognize emergency situations of children. Passive camera is only triggered to show the exact circumstances of such events. The image data is not processed, neither is user input possible — the system is completely automated and passive.

LifeLog using SenseCam [17] is a large-scale project aiming to incorporate all plausible sensors, mobile data and activity monitoring. In theory, SenceCam is suitable for a wide range of applications. However, to be made useful, the enormous amounts of data first have to be structured. So far, the main effort is to interpret visual data from passive camera. However, incorporation of the other data is disconnected, and their potential therefore is largely untapped. An experiment on including biometric data was made [16]. The iClips project [8] focuses on creating automated annotation and linking of multi-modal data within lifelog archives to enable effective search. The Idetic project [24] searches for new visualization concepts for lifelog data, trying to come away from the classic time-line concept.

III. Mirror, mirror on my screen what was the most emotional event of all?

A question of why are emotions important forces itself to our minds. Perhaps we should ask differently: how are emotions important to human functioning? The aspect of feeling is an inherent part of our inner architecture just like any other phenomena mentioned (perception, motivation or thinking). In fact viewing these phenomena (processes) as separate is a largely simplified way since subsystems of human functioning are interconnected and can hardly be explored by themselves. According to Frijda [14] the notion of emotion serves to "resolve discrepancies between what people do or feel and the events surrounding them, between the immediate cues for why they do what they do an what they actually do, between what they do and what they say"

After a nice holiday trip, pictures are loaded onto a personal computer and reviewed by the user. Visualizing adventures and experiences of the past days can be a very intimate and heart-warming experience. A private photography collection grows over time and is very hard to maintain. There are however many software products out there to make this task easier: iPhoto[1] and Picasa[2] are able to detect faces in the collection and allow you to order certainpictures according to the persons appearing (compare Figure 1 (a)). The user can thus follow the history they share with the person. This principle is recently adapted to "browse friendships" on facebook[3] showing mutual visual and text content. The software products additionally use

[1] http://www.apple.com/ilife/iphoto/
[2] http://picasa.google.com/
[3] http://www.facebook.com/

(a) Face detection and matching. (b) GPS or manual annotation. (c) Matching of tags and appearance.

Fig. 1. Automated associations in successful software products: iPhoto 11 detects and learns known faces (a) and maps GPS data to Google maps(b); Google image swirl (c) connects pictures according to text tags, categories and visual appearance.

GPS data of the pictures and are able to map them onto Google maps[4]. This is shown in Figure 1(b). The local information can be manually entered as well. Google image-swirl[5] is not focussed on the order of private photo collections, but connects public images and puts it in a graph based representation based on text tags, tag categories and visual matching (compare Figure Figure 1(c)).

Our approach uses these technologies, combines them and takes the idea further by *adding unconscious emotional data*. The user uploads collected data to a computer and initializes semi-automated processing. Images are being characterized by their emotional content, the people appearing and their emotions (based on facial expressions). Familiar faces are added to the library, for unknown faces, the system asks the user about the persons name. Visual content is then aligned with biomedical data: The most salient content - either visually or emotionally - is fed into the system. Grand majority of tedious material is deleted or archived right away. During the recording time, the user has an opportunity to press a button to manually mark a “special” moment.

Having a salient subset of the gathered data defined, the remaining data have to be associated to prior material. Persons are linked to pictures of them in the past. Salient emotions are connected and by that, new associations of pictures are drawn. This leads to an *affective mirror* [22]: It gives us an idea how we function emotionally. Typically, we are not fully aware, which situations are emotionally arousing. Strong emotions are only partially conscious. The system gives us a snapshot of our feelings, associated with locations, and persons attending within the whole data-set of the past.

Another important feature is the insight on *how we react to emotions of other people - and what we feel at these moments* compared to the emotions of the counterpart. The user gets an objective report on the way empathy occurs in his daily life. Up to a certain degree, comparing detected emotions from others and the users very own emotions allows conclusions to be drawn on how other people see the user or what they actually show when the user is not fully paying attention.

To take a closer look at their own emotions will open cast a new perspective at most life decisions the users make. Biological data show *our emotional state and the way it is expressed, even when we are not aware of it* - the application enables us to become conscious of them while reviewing the data. Again, with the gathered data recording facial expressions of our conversation partners in the images an indicator is given of *how emotional expressions of other people influence our own decisions*. Regarding a larger picture, statistics over the emotional data indicates in what general emotional states the user usually finds themselves and discloses possible patterns in their emotional processing and further behavior. Combining the visual and biometric data, *changes in the mood over the years* can be evaluated as well as *evolving patterns*.

The interpretation of this data is left only on the user himself. Reflecting on their processes can be a very intimate and sometimes tough experience and each user can come to different conclusions and create their own narrative. The user will be able to automatically group situations connected to a certain emotional state which can be reviewed together. Parallels that are not obvious at first glance can be made apparent and a whole new perspective of the situation can be drawn. This leads all to a very private leisure time activity, which does not affect any other persons, but allows the user to get in closer contact with themselves.

Table I describes the key aspects of human functioning according to general psychology (e.g. [23]). Its goal is to illustrate some of their partial processes expressed in the most general terms and draw elementary parallels as well as show differences between human functioning and capabilities of current lifelog systems. The distinction and contribution of our system is also made clear.

IV. Data capture

To implement the system described above, we capture the following signals, mimicking the human perception:

- *What we see:* Wearable “always-on” cameras automatically take images or videos, while the user wears them.
- *What we hear:* Sounds we hear, music we listen to, things that people tell us, what we say (and how): voice sound, pitch, intensity etc. is captured using a microphone.
- *How we feel:* Wearable sensors measuring bio-signals, such as heart rate, galvanic skin response, skin temperature or body motion are available. They can measure the level of emotional arousal of the wearer.
- *Where we are:* Using GPS the user can determine and save his/her location.

[4]http://maps.google.com/
[5]http://image-swirl.googlelabs.com/

<table>
<tr><td>Phenomena</td><td>Humans</td><td>State-of-the-Art lifelogs</td><td>Proposed affective lifelog</td></tr>
<tr><td>Sensation</td><td>Visual, Auditory, Somatosensory, Gustatory and Olfactory systems → transduction to neural activity → coding stimulus intensity and quality</td><td colspan="2">Wearable sensors: camera, audio, bio-sensors (GSR, temperature, motion), GPS, computer use, mobile phone use,... (various subsets of these)</td></tr>
<tr><td>Attention</td><td>Selectivity (ignoring unimportant stimuli), concentration (choosing one stimulus), distribution (conducting several automatic processes at once), capacity (up to approximately 7 stimuli coming from one sensual source), stability</td><td colspan="2">Total capture and storage without distinction, unfiltered, uninterrupted, stable, theoretically unlimited number of sensors at one time</td></tr>
<tr><td>Perception</td><td>Organization of sensory information → direct vs. constructive perception, perceptive field organization laws, bottom-up and top-down processes, perceptual illusions, individual differences (expectation, cognitive style, defenses, priming, preference etc.)</td><td>Data from different devices is disconnected, modes are partially incompatible, different sample frequencies, time synchronized, basic features are extracted</td><td rowspan="2">Based on automated multi-model recognition, associative machine learning is used to find connections between multi-media data.
These are then weighted by the emotional importance of the current situation. An interactive associated data space to browse the gathered data by multiple human-inspired associations and connections.</td></tr>
<tr><td>Learning</td><td>Habituation, sensitization, imprinting, associative learning (classical and instrumental conditioning), social learning, internalization etc.</td><td>Similarities and thresholds are used to filter, tag and retrieve data; face detection</td></tr>
<tr><td>Memory</td><td>Short-term, working, long-term, episodic, semantic, nonconscious, targeted vs. open recall, alternated memory (eg. By suggestion), distortions of memory: omission, commission</td><td colspan="2">Theoretically unlimited, not differentiated, data structure is strictly defined</td></tr>
<tr><td>Thinking</td><td>Heuristics, induction, deduction, magic thinking etc.</td><td>very limited</td><td>Allows for re-interpretation of experiences by the user. Provides feedback mechanisms to integrate affective appraisal provided by the user. Appraisal values are propagated by retrospective update mechanisms to provide flexibility and intelligent human-like associative retrieval and analysis.</td></tr>
<tr><td>Language</td><td>Icons, indexes, symbols etc.</td><td colspan="2">Bits, numbers, tags, links, commands etc.</td></tr>
<tr><td>Motivation</td><td>Ambitions (achievements, recognition, exhibition), relations to objects (acquisition, order, retention, construction), preserving social status (in avoidance, dependence, counteraction), power (dominance, deference, autonomy, contrariance, aggression, abasement), relating to people (affiliation, rejection, nurturance, succorance, play), information exchange (cognizance, exposition)</td><td>passive log, prone to the write-once-read-never effect</td><td>The adaptation of the system enables new possible interpretations of past experience. Encourages self reflection and gives new viewpoints of the user's life.</td></tr>
</table>

TABLE I

DIFFERENCES BETWEEN HUMAN PERCEPTION, CURRENT LIFELOGS AND THE PROPOSED SYSTEM.

Other sensors or data sources can be added to the system to provide more context information. Such data are e.g. copresent Bluetooth devices that may indicate presence of other people in up to 10m proximity, mobile activity (Call logs, SMSs) that can provide information about communication, and of course monitoring computer activity, viewed documents and activity on social networks could be considered.

V. PROCESSING

In this section, novel concepts of how to achieve the system described above are presented. Parts, that have often been discussed in other publications on lifelog systems, such as technical details of data capture [7], event segmentation [11], location recognition [4] etc. are omitted.

A. My own emotions: Emotional arousal from bio-signals

Strong emotions or feelings are not a purely spiritual phenomenon. In fact affect has various bodily manifestations which can be objectively measured. In lab settings, the emotional state of a human can be derived relatively precisely from measuring heart rate (+ other heart statistics), brain activity, galvanic skin response (GSR), respiration, blood volume, hormone levels, etc. It is not feasible to measure all of these signals in everyday life. A wearable device that is acceptable for everyday use has to be non-intrusive, comfortable, robust and easy to use without attracting attention. Currently, only heart rate monitors and sensors measuring galvanic skin response comply with these requirements.

Among these two signals, the galvanic skin response measured on palms, feet and forehead is the one which is has a stronger relationship to emotional arousal than heart rate. Therefore, GSR sensor armbands worn on the inner side of the wrist, such as e.g. the Q Sensor by Affectiva, are so far the simplest functional choice for measuring affect in daily life. Similar GSR devices (e.g. the Biomedia SenseWear) have been used before in similar set-ups [25] [16].

However, based on heart rate or galvanic skin response signals alone, we only can deduce the state of emotional arousal, not emotional valence. I.e. we can say that the user is feeling a strong emotion, but from the biological measurement alone, we are unable to detect whether it is joy or anger. Therefore we will also introduce other means of extracting emotional information from other sensor. These are described in the next sections.

B. Emotions of conversation partners: Facial expression recognition and eye gaze aversion

Besides personal emotional information, additional information about the people we encounter during our day life can be retrieved by using a passive always-on camera.

Just by looking at one persons face, it is in fact possible to deduce his feelings and his state of mind. There are several cues which could be analyzed, namely attention, facial expressions and gaze aversion. By combining this kind of information with the one obtained by the proposed system, it is be possible to extract valuable patterns, such as empathy, how people look at you, how you react to their presence and they react to yours, how their mood influences you, their attention towards you, and even deceit.

However, recent studies involving of small resolution webcams are yielding very promising results. For instance, the system in [28] is capable of analyzing the gaze of a person in real time, starting from monocular video images. The system combines head pose and eye locations information together to improve the final gaze estimation results.

Regarding the detection of facial expression, [27] also use a webcam based, initialization free system to track facial movements. The system tracks the movement of the face muscle and encodes them into the Facial Action Coding System (FACS) [13]. These deformations can be fed into a classifier in order to extract the most probable expression [29]. However, facial expression and gaze are not the only features which can be extracted by analyzing faces. It is said that eyes are the mirror of the soul, because much of the state of mind of one person can be seen through the eyes. Gaze aversions, for instance, often indicate deceit of the action of thinking. People stressed by your presence will perform more gaze aversion than people which are comfortable around you [12]. Furthermore, the work in [10] focuses on the dynamics of eye region to discriminate between fake and real smiles. The main idea comes from Ekmans studies on the field [13], which indicate that spontaneous smiles are formed by the contraction of both the zygomatic major (which raises the corners of the mouth) and the orbicularis oculi (which forms crows-feet around the eyes) muscles, where posed smiles involve only the zygomatic major muscle. Therefore, combining the facial expression classification algorithms with this information could be used as a tool to indicate deceit in the person we are communicating with.

C. Emotion classification enhancement using sound analysis

To enhance emotion classification and at the same time provide more context information, we use the audio signal. However, since many people object to their speech being recorded and replayed, we don't provide the possibility to replay the audio signal directly, but only to display the results of the audio-processing. In [30] an approach to exploit the audio signal for use in a "memory prosthesis" has been demonstrated. Using the speech recognition systen IBM ViaVoice[6] transcripts of conversations are recorded along with other information, like the gender of the speaker, speaker id, laughter, voice intensity etc. Since speech recognition is error-prone, a confidence value is given for each word. Words with high confidence are highlighted in the visualization of the text. Vemuri [30] additionally implemented a phonetic text search to further improve the retrieval within texts with errors. Further improvement can be achieved by weighting words within conversations based on their importance. One way to derive importance is e.g. by analyzing the intensity with which a word was spoken, if laughter occurred in short succession, if the word's frequency within the conversation is high, etc. However, it is also possible to classify speech according into basic emotions (under certain circumstances) [21]. The effects of emotion in speech tend to alter pitch, timing, voice quality and articulation of the speech signal [5]. An emotional speech recognizer can classify statistical measures of these acoustic features into classes that represent different basic affective states [31]. Adopting these methods will lead to improved memory cues and emotional feedback. Furthermore, mood analysis of the music the user is listening to (using tools like Sony's SensMe[7]) can add a further interesting dimension to the system.

D. Modeling and learning associations from multi-modal life-log data

As has been mentioned above, a main challenge of life-logging systems is to deal with a huge amount of data and extract the relevant material in order to supply the user with a reasonable and compact summary of the day. It is also necessary to mimic the human way to select and associate important scenes of the day. Psychoanalysis offers models for both required tasks. In general the importance of an event is expressed through the intensity of the affect caused by it [9]. For events with high affect associations covering inputs from the different modalities (acoustic, tactile, visual) are built and a multi-modal representation of the event is stored in memory. In addition to this binding of the different modalities inputs the multi-modal event is associated to similar events already stored in the memory. The base for evaluation of their similarity can be either multi-modal (e.g. smell), based on low-level features (shape, color, movement, sound intensity) or high-level (words, faces , voice intonation).

An important attribute of the human memory is that information previously considered as non-relevant or suppressed information can be assigned to be relevant based on a new experience. This means that associations can retrospectively alter the memory and the internal judgment of an event importance. In addition to non-conscious or affective associations, which are stronger, also rational associations are built by humans.

Rational associations usually are of factual character. They can be modeled by semantic webs by explicitly defining the complete ontologies (all important semantic classes and their

[6] http://www-01.ibm.com/software/voice/

[7] http://www.sonyericsson.com/cws/support/phones/detailed/whatissenseme/

mutual relationships)[20][19]. Modeling semantic ontologies is an ongoing research topic and is not the focus of this paper. However, we are interested in affective associations. As opposed to rational thinking, affective associations are often seemingly random but they have a strong effect on our preferences and therefore also implicitly our decisions and behavior [15].

To implement human-like affective associative thinking we employ the spreading-activation theory [1]. When working with factual memory, nodes represent semantic concepts (such as John, living room, etc.), connected by predefined associative links which are strengthened based on the frequency of use or *activation*.

Considering the technical details: The captured multimodal data will be stored in an ontology implemented using the OWL language and related technologies. The ontology will facilitate the modeling of concepts representing heterogeneous sensor information as well as detected entities such as persons, objects, places etc. and their relations. The scalability of the technology enables to later add additional data from call logs, SMSs and social networks to be integrated into the ontology. It will be possible to handle inconsistent, incomplete or missing information (caused be the malfunction of one of the sensors or because of discharge of its batteries). From all the sensory information (and a knowledge base of basic facts), the system will derive a model of similar, related and antagonistic concepts [26]. The user will be able (and encouraged) to alter the findings of this process using a feedback mechanism.

VI. Conclusions

We have presented a system concept created to provide a human user with feedback on their conscious and unconscious emotional reactions. We encourage the process of self-reflection by looking into an *affective mirror*. Self-reflection is an adult learning process and as such it is considered prosperous. The concept is strongly focused on the emotional autobiographical experiences.

The proposed system thus contradicts the current trend towards social networks which constantly draw our attention towards interaction with other people. This system let the user re-interpret his past life in a manner it has not been possible before. The system will provide an interpretation of daily experiences by means of free associations of prior experiences which represents a new way of active self-mirroring of the user. Based on knowledge of bodily reactions, connections are drawn that possibly weren't so obvious before. Usage of the system and subsequent browsing through new associations of our "parallel mind" could become a completely new leisure time activity. At night, people will not browse facebook pictures of loose acquaintances but rather explore their own feelings inside out. Contrary to all media experiences people recently enjoy in their leisure time, this system will play an introspective and meditative role and "bring us back to ourselves".

References

[1] J. R. Anderson. A spreading activation theory of memory. *Verbal Learning and Verbal Behavior*, 22:261–295, 1983.

[2] G. Bell. A Personal Digital Store. *Communications*, 44:86–91, January 2001.

[3] G. Bell and J. Gemmell. *Total Recall: How the E-Memory Revolution Will Change Everything*. Dutton Adult, 2009.

[4] M. Blighe and N. O'Connor. MyPlaces: Detecting Important Settings in a Visual Diary. In *CIVR*, 2008.

[5] C. Breazeal and L. Aryananda. Recognition of affective communicative intent in robot-directed speech. *Autonomous Robots*, 12:83–104, 2002.

[6] V. Bush. As We May Think. *The Atlantic Monthly*, 176(1):101–108, July 1945.

[7] D. Byrne, L. Kelly, and G. Jones. *Handbook of Research on Mobile Software Engineering: Design, Implementation and Emergent Applications*, chapter Multiple Multimodal Mobile Devices: Lessons Learned from Engineering Lifelog Solutions. IGI Publishing, 2010.

[8] Y. Chen and G. J. F. Jones. Augmenting human memory using personal lifelogs. In *Proceedings of the 1st Augmented Human International Conference*, 2010.

[9] A. Damasio. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. Harcourt, 1999.

[10] H. Dibeklioglu, R. Valenti, A. A. Salah, and T. Gevers. Eyes do not lie: Spontaneous versus posed smiles. In *MM*, 2010.

[11] A. Doherty and A. F. Smeaton. Automatically Segmenting Lifelog Data Into Events. In *WIAMIS*, 2008.

[12] G. Doherty-Sneddon and F. G. Phelps. Gaze aversion: A response to cognitive or social difficulty? *Memory & Cognition*, 2005.

[13] P. Ekman and W. V. Friesen. Felt, false, andmiserable smiles. *Nonverbal Behavior*, 1982.

[14] N. Frijda. *Handbook of emotions*, chapter The psychologists' point of view. Guilford Press, 2 edition, 2000.

[15] A. S. Ghuman and M. Bar. The Influence of Nonremembered Affective Associations on Preference. *Emotion*, 2006.

[16] L. Kelly and G. Jones. An Exploration of the Utility of GSR in Locating Events from Personal Lifelogs for Reflection. In *iHCI*, 2010.

[17] H. Lee, A. Smeaton, N. O'Connor, G. Jones, M. Blighe, D. Byrne, A. Doherty, and C. Gurrin. Constructing a SenseCam Visual Diary as a Media Process. *Multimedia Systems*, 2008.

[18] S. Lee, J. Sohn, A. Usami, and M. Hamanaka. Development of Wearable Device by Kid's friendly Design for Kid's Safety. In *Human Computer Interaction*, 2010.

[19] M. Y. Lim, R. Aylett, and W. C. Ho. Spreading Activation An Autobiographical Memory Retrieval Mechanism for Social Companions. *Intelligent Virtual Agents*, 2010.

[20] P. Mika. Ontologies Are Us: A Unified Model of Social Networks and Semantics. In *The Semantic Web*. Springer Berlin / Heidelberg, 2005.

[21] M. Pantic, N. Sebe, J. F. Cohn, and T. Huang. Affective multimodal human-computer interaction. In *Proceedings of the 13th annual ACM international conference on Multimedia*, MM, pages 669–676, New York, NY, USA, 2005. ACM.

[22] R. W. Picard. *Affective Computing*. MIT Press, Cambridge, MA, USA, 1997.

[23] A. Plhakova. *Ucebnice obecne psychologie*. Praha : Academia, 2003.

[24] S. Selke and I. Just. The idetic project. *http://www.idetic.de/*, 2010.

[25] A. Ståhl, K. Höök, M. Svensson, A. S. Taylor, and M. Combetto. Experiencing the Affective Diary. *Personal Ubiquitous Comput.*, 13:365–378, June 2009.

[26] A. Troussov, J. Judge, and D. Botvich. Mining socio-semantic networks using spreading activation technique. In *KASW*, 2008.

[27] R. Valenti, N. Sebe, and T. Gevers. Facial expression recognition: A fully integrated approach. In *Workshop on Visual and Multimedia Digital Libraries*, 2007.

[28] R. Valenti, N. Sebe, and T. Gevers. Visual gaze estimation by joint head and eye information. In *ICPR*, 2010.

[29] R. Valenti, N. Sebe, T. Gevers, and I. Cohen. Machine learning techniques for face analysis. *Springer*, 2008.

[30] S. Vemuri, C. Schmandt, and W. Bender. iremember: a personal, long-term memory prosthesis. In *CARPE*, New York, NY, USA, 2006.

[31] S. Yilmazyildiz, W. Mattheyses, Y. Patsis, and W. Verhelst. Expressive speech recognition and synthesis as enabling technologies for affective robot-child communication. In *Advances in Multimedia Information Processing*. 2006.