Context-Aware Multimodal Sharing of Emotions

Maurizio Caon[1,2], Leonardo Angelini[1], Yong Yue[2],
Omar Abou Khaled[1], and Elena Mugellini[1]

[1] University of Applied Sciences of Western Switzerland, Fribourg
[2] University of Bedfordshire, Luton
{Maurizio.Caon,Leonardo.Angelini,Omar.AbouKhaled,
Elena.Mugellini}@hefr.ch, Yong.Yue@beds.ac.uk

Abstract. Computer mediated interaction often lacks of expressivity, in particular for emotion communication. Therefore, we present a concept for context-aware multimodal sharing of emotions for human-to-computer-to-human interaction in social networks. The multimodal inputs and outputs of this system are distributed in a smart environment in order to grant a more immersive and natural interaction experience. The context information is used to improve the opportuneness and the quality of feedback. We implemented an evaluation scenario and we conducted an observation study during some events with the participants. We reported our considerations at the end of this paper.

Keywords: affective computing, multimodal interaction, computer mediated communication, social sharing of emotions.

1 Introduction

The communication and information technology changed the way people communicate. In particular, Internet played a key role in this communication revolution that led to the concept of human-to-computer-to-human interaction (HCHI) as explained by Clubb in [1]. All the paralanguage or normal face-to-face nonverbal communication is missing in HCHI. In particular, the need of expressing emotions through messages became important. Indeed, there has been an attempt in messaging to integrate this kind of non-verbal communication that knew a large success: the emoticons [9]. After they have firstly been introduced by Scott E. Fahlman in 1982, they evolved and spread all over the world. In particular, some studies highlighted that people need to express emotions in popular communication systems as Facebook and Twitter, which are recently referred as "Social Awareness Streams" (SAS) [2]. In fact, shared emotional states have been observed creating engagement and participation among the users [2]. This phenomenon is due to the natural need of the human being of social sharing of emotion. In fact, Rime et al. showed that most emotional experiences are shared with others shortly after they occurred [10]. They pointed out that social sharing represents an integral part of emotional experiences. Since currently the technology is ubiquitous and the access to Internet is indispensable [12], people share their emotions in the SAS in order to openly communicate about the emotional

M. Kurosu (Ed.): Human-Computer Interaction, Part V, HCII 2013, LNCS 8008, pp. 19–28, 2013.

circumstances and their feelings and reactions. In this paper, we present a multimodal system that allows people to socially share text, images and their emotions. This system exploits multimodality since the use of natural means of communication facilitates the human interaction with the machine [3]. This statement is particularly valid when context information is taken into account to choose best modality in an opportune manner. The main contribution of this paper consists of introducing the concept of multimodal social interaction in affective computing for the HCHI. In particular, we wanted to explore the communication in SAS.

The rest of this paper is structured as follows: Section 2 analyzes the literature of the Affective Computing domain with focus on computer mediated communication; Section 3 explains the concept of the proposed system whose architecture is described in Section 4; Section 5 shows the implemented evaluation scenario and the discussion; finally, Section 6 reports the conclusion and the future work.

2 Related Work

Many psychologists argue that it is impossible for a person to have a thought or perform an action without engaging, at least unconsciously, his or her emotional systems [13]. This is true also for the interaction with computers. In fact, recent research in psychology and technology assesses that emotions play a key role also in every computer-related activity [14]. From these findings, a new domain emerged and is called affective computing. This name comes from Rosalind W. Picard, one of the first pioneers of this new and exciting domain (the fascinating story of the birth of affective computing has been narrated in the introduction of the first issue of the IEEE Transactions on Affective Computing [34]). She defined the affective computing as the "computing that relates to, arises from or deliberately influences emotions" [15]. This definition gives the idea of how wide this domain is; in fact, affective computing "can mean recognizing user affect, adapting to the user's affective state, generating 'affective' behavior by the machine, modeling user's affective states, or generating affective states within an agent's cognitive architecture" [11]. The first issue tackled by the researchers in affective computing was the emotion detection and recognition. Indeed, the main human communication modalities have been thoroughly studied in order to extract the emotional information (e.g., body gestures and motion [16], facial expressions [17] and speech recognition [18]). The aforementioned communication means are also the most important in the human judgment of behavioral cues [19]. In contrast, computers can capture other signals that the human sensory system completely ignores. For instance, specific sensors on the user can capture physiological signals that have been demonstrated being reliable for the emotion recognition (e.g., electromyography, electrocardiography, electrodermal activity, blood volume pulse, peripheral temperature and respiration as in [20], and electroencephalography as in [21]).

As showed, the scientific community intensively studied the techniques for the emotion recognition but this was only the beginning. Afterwards, the researchers in human-computer interaction broadened the spectrum of this domain exploring how the emotional information can be exploited to enhance the user experience. For example, the user's emotional state can be used to recommend the opportune video [22] or the best song [23] or for marketing purposes [29]. Another interesting application of affective computing can be found in [24] where the authors developed an emotion-aware game-based system for adaptive learning; they observed an enhancement of motivation and satisfaction of the students. Affective computing can also help the rehabilitation process, as proposed in [25], or just help to make the user happier [28].

The aforementioned applications refer to the role of emotion in the interaction between a user and the computer in order to make it more natural, but another important aspect of affective computing is related to HCHI. In fact, the *affective interactions* are not limited to the human-human interaction and the relation between the user and the system, but they refer also to computer mediated communication [26]. Nowadays, it is easy to find people who spend more time interacting with a computer than with other humans in face-to-face encounters [27] where affect not only creates richer interaction, but also helps to disambiguate meaning, allowing for more effective communication [14]. Hence, a current and important trend of affective computing focused on the transmission of emotions in order to fill the gap due to the lack of expressivity in HCHI, where all the paralanguage and non-verbal communication are excluded. In fact, some works focused on introducing a remote system for the communication of touch in order to create empathy between two users who were distant; an example is the "Huggy Pajama", which remotely reproduced the feeling of a hug [30]. The "keep in touch" project followed a similar concept, but it associated the haptic feedback to the voice for an enhanced experience of connectedness [31]. In particular, multimodal interfaces are considered suitable for the communication of emotions. A perfect example of this kind of application is the system for tele-home health care presented in [32]. This system sensed the user's emotional and affective states in order to remotely monitor him/her and opportunely respond to this information by a multimodal anthropomorphic interface agent. As the above-mentioned system, our work takes advantage of the rich expressivity of multimodal interfaces but applied to HCHI. In particular, our system has been designed to provide a feedback representing the emotional states of other members of a social group, exactly as the Emotishare project [33] but with context-aware multimodal representation of emotion in the ambient.

3 Concept

In the previous sections we showed how the scientific community highlighted the importance of emotional information communication in HCHI. With this concept, we address the representation of the users' emotional states in a social group. This system aims at allowing the users to share a message, a picture and their emotional state in an SAS, for instance Twitter. The SASs do not allow a variety of modalities and usually have also a limited number of characters for messages; therefore, the expressivity of emotions in SASs is limited. In order to enhance the user experience, the system provides a multimodal feedback of the shared content to all the followers. The feedback is displayed in the whole environment using different modalities in order to grant a fully immersive experience. In particular, the system should be able to handle the sound (i.e., the vocal synthesis of the message), to display images and to represent emotions. The emotions can be displayed in personal devices (i.e., as pictures) or in the ambient; the emotion state is displayed in the ambient via Aphrodite [7]. Aphrodite is a robotic painting that depicts Venus' head from Botticelli's "the Birth of Venus". This painting comes alive as the magic paintings in Harry Potters' saga; in fact, Aphrodite interacts with the user since she is able to mime the human facial expressions and to reproduce sounds using her artificial voice in order to communicate a certain number of emotional states. We adopted this approach since some researches demonstrated that anthropomorphic artificial agents are very effective for emotion communication [35].

Since communicating using different interaction modalities can be considered as an innate characteristic of human beings, we provided multiple input modalities. The user can communicate his/her emotional state via text, adding an emoticon, using a picture and performing a facial expression. The system manages several distributed devices (as smartphones, PCs, interactive surfaces et cetera) and can recognize the user's facial expression.

The system coordinates inputs and outputs according to the context. The contextual information is used to choose the opportune means of communication in order to improve the user experience.

4 System Architecture

The system presented in this paper has been developed adopting two frameworks: Inter-Face and NAIF (as depicted in Fig. 1).

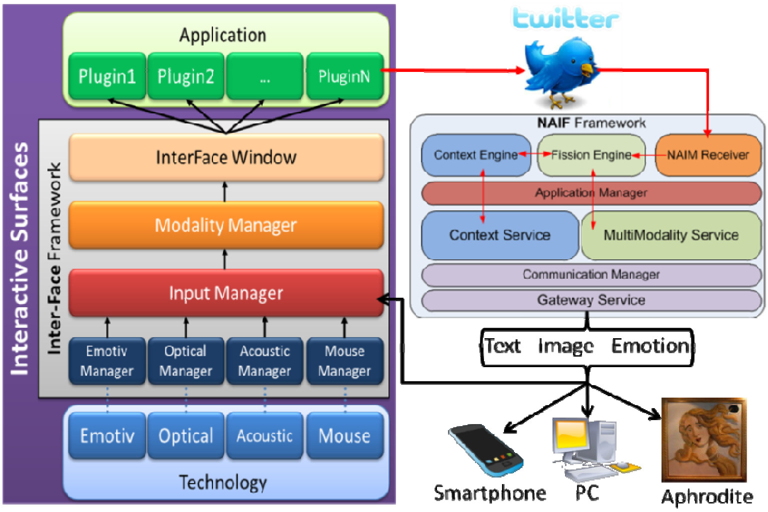

Fig. 1. System architecture integrating the structures of the Inter-Face and NAIF frameworks

The Inter-Face framework has been conceived in order to guarantee the compatibility among multiple heterogeneous technologies by providing a runtime environment for the dynamic adaptation of an application to a surface. In fact, Inter-Face allows transforming any object into an interactive surface [4]. Moreover, it integrates the drivers for the management of the Emotiv EPOC neuroheadset [5]. The NAIF framework aims at interconnecting heterogeneous devices in smart environments. These devices can offer resources to display information to the user and can contribute to the gathering of contextual information. Context information is managed by NAIF [6] and can be used to deliver opportune feedback to the user using the best available modalities. Moreover, each device can send communication intents to the framework, which automatically performs the context-aware multimodal fission of the message content.

Our system allows the user to interact with several surfaces distributed in the environment (e.g., tables, walls, TVs, smartphones et cetera); then, he/she can share a text, an image and his/her personal emotional state. In fact, the user can select the image in the specific application and type the text through touch interaction on the surfaces; in addition, he/she can decide to press the button for the emotion detection via the EPOC neuroheadset. The Emotiv EPOC senses the facial expression through the surface electromyography (sEMG) and then the recognized user's emotional state is added to the message. Afterwards, the user presses the "tweet" button to share his/her message on Twitter. Inter-Face supplies another plugin dedicated to the reception of messages "tweeted" by the people "followed" by the user. These plugins are showed in Fig. 2.

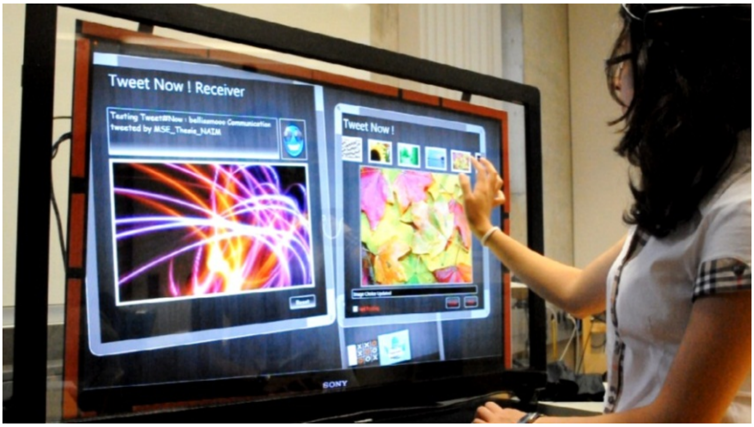

Fig. 2. A user wearing the EPOC neuroheadset and interacting with a touch-enabled TV with Inter-Face running the Twitter application

5 Evaluation Scenario and Discussion

The system presented in the previous section has been implemented in a smart environment, i.e., a smart living room (as depicted in Fig. 3). The NAIF gateway runs on a dedicated server and interconnects all the devices present in the environment. In this scenario, the user can send tweets using his/her smartphone, or laptop, or interacting with the smart surfaces distributed in the environment as the TV and the table.

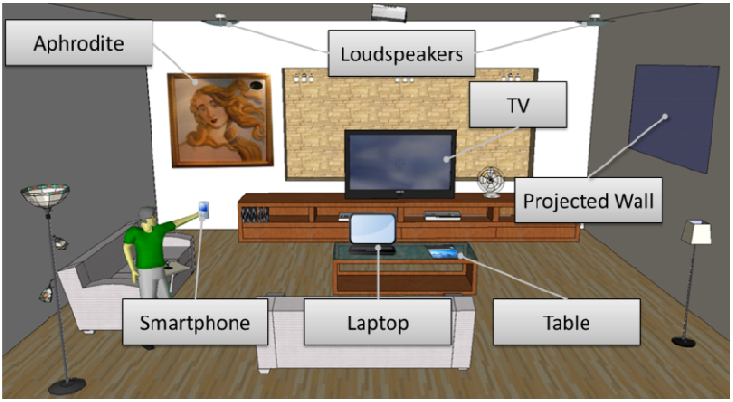

Fig. 3. Study of the evaluation scenario: all the input and output devices are highlighted

The user's emotional state is captured whenever he/she wears the Emotiv EPOC neuroheadset; in fact, the emotion is directly connected to the user's facial expression that is recognized through the sEMG signals. Messages are automatically composed using a specific protocol as described in [6]. The set of available emotions are 5 and they are directly connected to 5 different facial expressions: smile for happiness ":-)", frown for sadness ":-(", wink for trust ";-)", laugh for ecstasy ":-D" and clench for anger ":-@". The facial expressions have been mapped on Plutchik's wheel of emotions [36]. Fig. 4 shows the emotional states represented by Aphrodite.

Fig. 4. Aphrodite can mime several facial expressions that are mapped on the Plutchik's wheel of emotions: a) is the neutral expression with no associated emotion; b) represents the smile for happiness; c) is frown for sadness; d) is frown and glum for anger; e) is wink for trust; f) is laugh for ecstasy.

NAIF takes advantage of several output generators: text messages can be communicated either via Text-To-Speech (TTS) in loudspeakers or via TV speakers;

emotions are shared via emoticons in the displays or via Aphrodite; images are shown in the available displays. In this scenario, the Twitter receiver is implemented in different devices: besides the aforementioned Inter-Face plugin (in the TV and the interactive table), other applications are running in Microsoft Windows (in the laptop and the projected wall).

Output modalities are chosen according to context information. For instance, if the luminosity of the room is high, the system chooses displaying information on the TV instead of the projected wall in order to grant the best quality of feedback (this scenario is shown in Fig. 5).

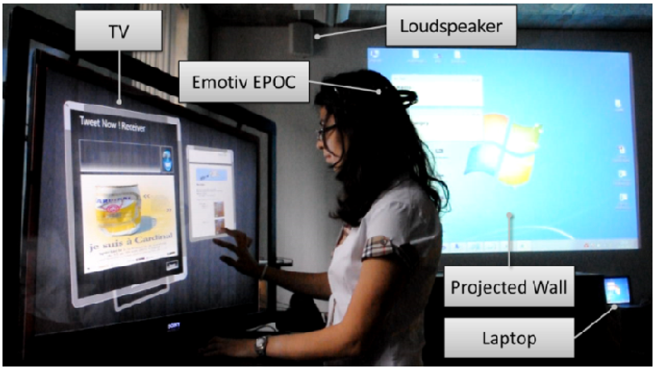

Fig. 5. A user wearing the EPOC neuroheadset and interacting with the touch-enabled TV using the Twitter application

The luminosity information is employed also for the reasoning about emotion sharing: in low light conditions, the robotic painting may not be visible, thus the emotion is displayed in the projected wall or the TV or the table (as depicted in Fig. 6). Similarly, if the room is noisy, audio feedback is avoided. In the current prototype, luminosity and noise information is retrieved through the popular Phidgets sensors [8]. Another context factor is based on the availability of the devices. Only the connected and functioning devices are taken into account during the multimodal fission and if they are not busy performing other tasks. Moreover, messages can be addressed to a specified user: "tweet" messages containing the @User tag can be delivered privately to the specified user according to his/her position. In fact, the user can identify him/herself near to a personal device using an RFID tag. Thus, all these messages are delivered only to the specific user's device in a discrete manner.

The system has been presented during a public event called "Socialize" [39] and other private demonstrations. During these events, we conducted an observation study on the evaluation scenario with the participants. The attenders of the "Socialize" event were young people aged between 20 and 35 years and enthusiast for social network technologies. Although the users showed an initial skepticism, they demonstrated a large interest about our system. The participants were very engaged while experiencing the interaction scenario; they appreciated the possibility of choice for the modality of input (even if the Emotiv EPOC resulted being very cumbersome) and of output. Finally, the users were particularly fascinated by Aphrodite and her capability of reproducing human emotions, who further engaged them.

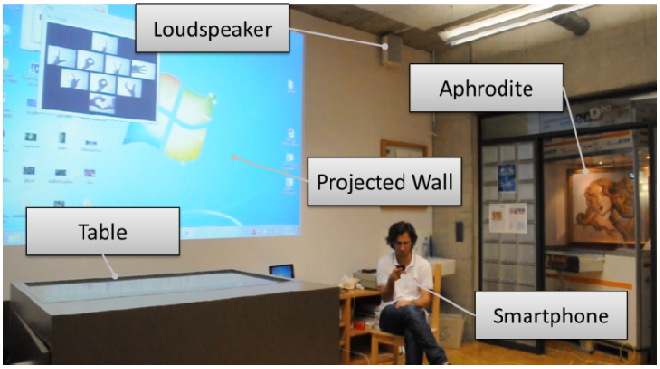

Fig. 6. A user interacting with the system via smartphone application

6 Conclusion and Future Work

In this paper, we presented the architecture of a system for context-aware multimodal HCHI. This system focuses on the computer mediated communication of emotional states via an SAS, i.e., Twitter, with pervasive multimodal feedback in the ambient in order to provide a more immersive user experience. An integrated multimodal fission engine evaluates the context information (e.g., luminosity level, noise level, user identification et cetera) to opportunely enhance the quality of feedback. The user can interact on several devices distributed in the smart environment through gestures and facial expression or traditional graphical user interfaces.

New input modalities will be integrated in the future development of this system. For example, adding vision-based face expression recognition during the interaction with devices that have embedded cameras (e.g., smartphone, laptop, TV et cetera). Recently, more and more novel products provide new ways for the emotion detection as the Affectiva Q-Sensor wristband [37] and the Interaxon headset [38]. Therefore, exploiting the flexibility of the adopted frameworks, these new devices will be integrated in order to gather more information about the users' emotional state and to provide a less cumbersome interface. Moreover, NAIF framework could benefit of the information of the sensed user's emotional state to further improve the quality and the opportuneness of feedback. The future version of this system will include also a virtual avatar representing Aphrodite for ubiquitous representation of emotions.

References

1. Clubb, O.L.: Human-to-Computer-to-Human Interactions (HCHI) of the communications revolution. Interactions 14(2), 35 (2007)
2. Kivran-Swaine, F., Naaman, M.: Network properties and social sharing of emotions in social awareness streams. In: Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, p. 379 (July 2011)
3. Dumas, B., Lalanne, D., Oviatt, S.: Multimodal interfaces: A survey of principles, models and frameworks. In: Lalanne, D., Kohlas, J. (eds.) Human Machine Interaction. LNCS, vol. 5440, pp. 3–26. Springer, Heidelberg (2009)

4. Mugellini, E., Abou Khaled, O., Pierroz, S., Carrino, S., Chabbi Drissi, H.: Generic Framework for Transforming Everyday Objects into Interactive Surfaces. In: Jacko, J.A. (ed.) HCI International 2009, Part III. LNCS, vol. 5612, pp. 473–482. Springer, Heidelberg (2009)
5. Emotiv EPOC neuroheadset, `http://www.emotiv.com/apps/epoc/299/` (last visited: March 2013)
6. Perroud, D., Angelini, L., Khaled, O.A., Mugellini, E.: Context-Based Generation of Multimodal Feedbacks for Natural Interaction in Smart Environments. In: Proceedings of the Second International Conference on Ambient Computing, Applications, Services and Technologies, pp. 19–25 (2012)
7. Aphrodite project, `https://project.eia-fr.ch/aphrodite/Pages/Introduction.aspx` (last visited: March 2013)
8. Phidgets sensors, `http://www.phidgets.com/` (last visited: March 2013)
9. Hwang, H.C., Matsumoto, D.: Nonverbal Behaviors and Cross-Cultural Communication in the New Era. Language and Intercultural Communication in the New Era, 116 (2013)
10. Rimé, B., Finkenauer, C., Luminet, O., Zech, E., Philippot, P.: Social Sharing of Emotion: New Evidence and New Questions. European Review of Social Psychology 9, 145–189 (1998)
11. Hudlicka, E.: To feel or not to feel: The role of affect in human–computer interaction. International Journal of Human-Computer Studies 59, 1–32 (2003)
12. Hoffman, D.L., Novak, T.P., Venkatesh, A.: Has the Internet become indispensable? Communications of the ACM 47, 37–42 (2004)
13. Picard, R.W.: Does HAL cry digital tears? Emotions and computers. In: Hal's Legacy: 2001's Computer as Dream and Reality, pp. 279–303 (1997)
14. Brave, S., Nass, C.: Emotion in human–computer interaction. In: The Human–Computer Interaction Handbook, pp. 81–93 (2003)
15. Picard, R.W.: Affective Computing. The MIT Press (1997)
16. Kleinsmith, A., Bianchi-Berthouze, N.: Affective Body Expression Perception and Recognition: A Survey. IEEE Transactions on Affective Computing, 1 (2012)
17. Sandbach, G., Zafeiriou, S., Pantic, M., Yin, L.: Static and dynamic 3D facial expression recognition: A comprehensive survey. Image and Vision Computing 30, 683–697 (2012)
18. Anagnostopoulos, C.N., Iliou, T., Giannoukos, I.: Features and classifiers for emotion recognition from speech: a survey from 2000 to 2011. Artificial Intelligence Review, 1–23 (2012)
19. Pantic, M., Sebe, N., Cohn, J.F., Huang, T.: Affective multimodal human-computer interaction. In: Proceedings of the 13th Annual ACM International Conference on Multimedia, p. 669 (2005)
20. Canento, F., Fred, A., Silva, H., Gamboa, H., Lourenco, A.: Multimodal biosignal sensor data handling for emotion recognition. In: Proceedings of 2011 IEEE SENSORS, pp. 647–650 (2011)
21. Schaaff, K., Schultz, T.: Towards emotion recognition from electroencephalographic signals. In: Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, pp. 1–6 (2009)
22. Benini, S., Canini, L., Leonardi, R.: A Connotative Space for Supporting Movie Affective Recommendation. IEEE Transactions on Multimedia 13, 1356–1370 (2011)
23. Dornbush, S., Fisher, K., McKay, K., Prikhodko, A., Segall, Z.: XPOD - A Human Activity and Emotion Aware Mobile Music Player. In: Proceedings of the 2nd Asia Pacific Conference on Mobile Technology, Applications and Systems, pp. 1–6 (2005)

24. Tsai, T.W., Lo, H.Y., Chen, K.S.: An affective computing approach to develop the game-based adaptive learning material for the elementary students. In: Proceedings of the 2012 Joint International Conference on Human-Centered Computer Environments, p. 8 (2012)
25. Mihelj, M., Novak, D., Munih, M.: Emotion-aware system for upper extremity rehabilitation. In: Proceedings of Virtual Rehabilitation International Conference, pp. 160–165 (2009)
26. Paiva, A.: Affective interactions: towards a new generation of computer interfaces. In: Paiva, A.M. (ed.) Affective Interactions. LNCS (LNAI), vol. 1814, pp. 1–8. Springer, Heidelberg (2000)
27. Picard, R.W., Cosier, G.: Affective intelligence — the missing link? BT Technology Journal (1997)
28. Tsujita, H., Rekimoto, J.: Smiling makes us happier. In: Proceedings of the 13th International Conference on Ubiquitous Computing (2011)
29. McDuff, D., Kaliouby, R.E., Picard, R.W.: Crowdsourcing Facial Responses to Online Videos. IEEE Transactions on Affective Computing 3, 456–468 (2012)
30. Teh, J.K.S., Cheok, A.D., Peiris, R.L., Choi, Y., Thuong, V., Lai, S.: Huggy Pajama. In: Proceedings of the 7th International Conference on Interaction Design and Children, p. 250 (2008)
31. Wang, R., Quek, F., Tatar, D., Teh, K.S., Cheok, A.: Keep in touch. In: Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems, p. 139 (2012)
32. Lisetti, C., Nasoz, F., LeRouge, C., Ozyer, O., Alvarez, K.: Developing multimodal intelligent affective interfaces for tele-home health care. International Journal of Human-Computer Studies 59, 245–255 (2003)
33. Willis, M.J., Jones, C.M.: Emotishare: emotion sharing on mobile devices. In: Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers (2012)
34. Picard, R.W.: Affective Computing: From Laughter to IEEE. IEEE Transactions on Affective Computing 1, 11–17 (2010)
35. Moser, E., Derntl, B., Robinson, S., Fink, B., Gur, R.C., Grammer, K.: Amygdala activation at 3T in response to human and avatar facial expressions of emotions. Journal of Neuroscience Methods 161, 126–133 (2007)
36. Plutchik, R.: The Nature of Emotions. American Scientist 89(4), 344–350 (2001)
37. Affectiva Q-Sensor, `http://www.affectiva.com/q-sensor/` (last visited: March 2013)
38. Interaxon headset, `http://www.interaxon.ca/` (last visited: March 2013)
39. Socialize event at Fribourg, `http://www.socialize-network.com/29-novembre-2012/` (last visited: March 2013)