Creative applications of human behavior understanding

Albert Ali Salah[1], Hayley Hung[2], Oya Aran[3], and Hatice Gunes[4]

[1] Boğaziçi University, Department of Computer Engineering,
Istanbul, Turkey
salah@boun.edu.tr

[2] Delft University of Technology, Intelligent Systems
2628 CD Delft, The Netherlands
H.Hung@tudelft.nl

[3] Idiap Research Institute, Social Computing Group
Martigny, Switzerland
oaran@idiap.ch

[4] Queen Mary University of London, School of Elect. Eng. & Computer Science
London, U.K.
hatice@eecs.qmul.ac.uk

Abstract. The role of computer science in the creative industries is becoming recognised as an important way to bring forward progress in both domains. There is a need for smarter applications that sense and adapt to their users in arts, creativity, entertainment and edutainment domains. Understanding human behavior in this area is challenging because it forces practioners to engage in both creative, and perhaps counterintuitively, analytic processes of understanding how people engage with creative phenomena. The systems constructed for this purpose promise to enhance and redefine the scope of creative industries significantly. This paper discusses scientific and technological factors that make this a challenging topic to address, provides a brief survey of related work in this area, and identifies active topics of research. Since arts, creativity, entertainment and edutainment all contribute to significant social and societal benefits, it is vital to tackle the problem of measuring and evaluating the success of automatic behavior analysis solutions as a social and human phenomenon.

1 Introduction

The important relationship between artistic creativity and science has existed for centuries, perhaps since the Ancient Greek mathematicians' scientific approach to aesthetics in what became the golden ratio, producing along the way people of genius like Leonardo Da Vinci. While there were periods in which arts and science drifted apart, the 20^{th} Century saw attempts to bring them together in different forms, for instance with the establishment of the art science journal Leonardo[1],

[1] http://www.leonardo.info/

international conferences such as Ars Electronica[1] and SIGGRAPH[2], institutes like Numediart[3] and special sessions in major conferences like the Interactive Arts Programme[4] in ACM Multimedia, which celebrates its 10^{th} year in 2013.

Obviously, science is instrumental in delivering artists new tools of artistic expression. More that than, it can provide systematic evaluation of the success of creative works, which remains an often taboo subject in the creative industries. For the consumer of art, the implications are very direct, as they expect to be stimulated intellectually, socially, or simply, to be entertained.

With arts and entertainment extending their scope to include societal themes, as well as with the proliferation of digital and computerized art forms, the intermingling of these domains increased. It is now commonplace to encounter artists that produce exclusively on digital medium, or interactive artworks that use state-of-the-art sensing technology to involve their spectators. With such technology becoming easily accessible, it is even possible for ordinary people with no specialized formation to grab a few sensors, and transform experiences, including daily living routines, into digital experiences, where data can be collected, analyzed, visualized, and acted upon [59]. In this paper, we investigate how human behavior understanding contributes to creative enterprises technologically.

Under the term *human behavior understanding* (HBU), we understand pattern recognition and modeling techniques to automatically interpret complex behavioral patterns generated when humans interact with machines or with other humans [46]. These patterns can involve actions, activities, attitudes, affective states, social signals, semantic descriptions, and contextual properties. Taken together, they define multimodal ways of enhancing human-computer, human-robot, and even human-human interactions [53].

We investigate the theme of HBU in relation to the arts, creativity, entertainment and edutainment under the following categories:

- HBU in the presence of passive entertainment systems: These analyze peoples' behavioral responses to stimuli from systems that do not require continuous input from a human.
- HBU in interactive systems: These systems inherently analyse people's behavior as input and respond accordingly.
- Multi-party HBU for systems that function as social spectacles: This refers to the analysis of the social element of audience participation in such systems where the awareness of reactions from others in the audience is part of the experience of the spectacle.

We proceed by looking at each of these categories in turn.

[1] http://www.aec.at/
[2] http://www.siggraph.org
[3] http://www.numediart.org/
[4] http://acmmm13.org/submissions/call-for-artworks/

2 HBU in the presence of passive entertainment systems

In its most traditional form, arts and entertainment are about providing visual, auditory, or haptic stimuli for a human observer. HBU can be used to observe the participant's implicit (e.g., physiology) and explicit behavior (e.g., emotional expressions), either to improve the quality of presentation, or to gain a better understanding of the participant's response.

Let us give the example of films, which are rich in audiovisual and emotional stimuli. People watching movies go through a variety of experiences, engineered to some extent by the director of the movie, and their expressive responses during the course of the movie can be analyzed using HBU techniques. For instance Smeaton and colleagues previously looked at heart rates, Galvanic skin response, and the amount of movement in spectators watching films in a special movie theatre equipped with smart chairs [50]. They found that group psychology affects people watching movies together, and they have highly correlated emotional responses. Similarly, Baumgartner et al. investigated emotions evoked by still pictures and classical music, and demonstrated that emotional responses are strongest when audio and vision are combined [3].

In this volume, Tarvainen et al. investigate stylistic features of movies in visual, aural and temporal modalities, and correlate these features with perceived and felt affect [52]. While it is difficult to sort through idiosyncratic variations of affective experiences of different subjects, this kind of analysis, beyond engineering an emotional experience, can serve pragmatic purposes, like summarization of a movie by its emotional peaks, or for customized recommender systems.

3 HBU in interactive systems

This topic refers to the role of human behavior as a trigger for interactive systems. Representative applications are the ones where a person directly interacts with a system that displays an artistic message, encourages creativity, play, or learning.

For some interactive art pieces, the humans' behavior towards the work becomes part of the narrative of the piece [51, 23]. In interactive arts, a key issue is the mode of interaction. HBU essentially provides artists with new tools, and with new interaction possibilities. In the absence of analysis, a clever setup may permit the use of simple sensors in a reactive application scenario. The Piano Stairs installation at Odenplan Station in Stockholm[1] is a well-known example of how an innovative behavior sensing system can be made to change people's daily routines: by converting an ordinary stair into a musical instrument via pressure sensors and loudspeakers, it was possible to get more people to take the stairs over the escalator.

Peter Beesley's *H*ylozoic Soil is an example of a responsive architectural geotactile space, made up of a network of micro-controllers, proximity sensors

[1] http://www.thefuntheory.com/

and shape-memory alloy actuators [5]. Its layers move in response to the presence of human occupants. Samadani et al. ilustrated that movement in such frond-like structures can easily create the perception of affective states in their users [47]. The Sissy (Sound-driven, Interactive, Self-conscious SYstem) installation at the STRP festival was made of 700 flip dots with white and black coloured sides, and it slowly responded to visitor movements that it sensed via its camera [8]. It was perceived as "shy," which is an unusual attribute for seemingly random dot patterns. In [35], an interactive playground was designed for children, indicating that through context-awareness, adaptability, and personalization, rich game experiences can be created.

In this volume, several works illustrate that responsive and interactive spaces can create stimulating and engaging experiences. Morgan and Gunes [36] introduce an installation designed to let people represent their "mood" by controlling the lights around the rim of the London Eye. Over 800 people participated, taking control of the lights using their heart rates and hand gestures. Affective and behavioral computing inspired techniques applied for analysing the physiological and motion capture data showed that during the interaction with the Eye people's heart and kinematic behavior differed according to the content style they were interacting with, and gestures were predominantly performed with extended arms. In [30], a pressure sensing floor prototype for detecting human actions is described. Once human presence is sensed at a certain location, it can be connected to an event in an interaction scenario. The authors describe several such applications, including a memory game, a virtual piano, foot painting, and a position-controlled media player.

A key insight in interaction design is that the nature of interaction between a person and a computer system is dictated both by the affordances of the system, and by the cognitive and physical limitations of the interacting person [44]. Accurate models of the latter can be used to steer the former for improved interaction. In this volume, Fourati et al. present an analysis of walking and turning tasks based on shoulder, hip and head activity [18]. Their results show that the emotional state expressed by the subject strongly determines the style of head gestures and walking and turning behavior. Marchegiani and Fafoutis conduct a behavioral study to investigate the listening capabilities of subjects under challenging conditions, with rock music playing in the background [32]. Their results indicate that the subjects' familiarity with the music acts as an emotional trigger that attracts the subject's attention. It has been previously shown that emotions play a significant role in driving user's attention in a system [39]. Subsequently, automatic analysis of emotions has an additional potential in shaping interactions, i.e. by helping to estimate the user's attention.

4 Multi-party and social HBU

HBU in the context of multiple receivers becomes an interesting social phenomenon in its own right. An entertaining public spectacle becomes a social event, where the response of other audience members is a vital part of the experience of an event, heightening its value to the people receiving it. In some cases the spectacle becomes talking point for strangers, which was referred to as triangulation by Whyte [58] and similarly as collaborative sense-making by Hook et al. [24]. The behavior of performance artists and audience as a collaboration in creating and eliciting spectacles in public spaces has also been studied by Gardair et al. [19]. Similar research has been carried out on understanding the social behavior of people in front of a painting, and as mentioned in Section 2, while watching movies.

HBU with multiple receivers is not only about human-arts interaction, it also extends to situations where the focus is directed towards collaborative creativity and performance. Understanding the social behavior of people and enhancing the interaction between them when they are engaged in a creative activity, promises great potential for linking HBU with new media art and computational creativity research [37]. Recent works on the analysis of groups show that the collective intelligence of a group, while performing several cognitive tasks, ranging from brainstorming to video game playing, not only depends on the average member intelligence, but also on the way the members interact with each other [60]. Computer supported systems have been proposed to ensure the equal participation from the group members for various occasions such as collaborative learning [2] and social events [40]. In this volume, [42] proposes a real-time system that automatically analyzes the audio recordings of two-person dialog to assess several social constructs such as interest, agreement and dominance.

Several other works look at musical performance and focus on the analysis of interaction between the members in musical groups, investigating the emergence of several social constructs such as leadership [54] and dominance [20]. Varni et al. [54] also investigated the synchronization of the affective behavior within the group and showed that the synchronization and leadership measures are effective and reliable in supporting social active music listening applications, such as Sync'n Move [55].

Education and edutainment also involve interactions that are essentially of a social nature. The keys to successful education are ensuring engagement (or activation) of the student, and providing educational material that has the appropriate level for the student. Both tasks require real-time monitoring of the student and estimation of engagement and comprehension levels. Computer games for students have been recently used in education for ensuring better engagement and comprehension, and it has been shown that children are more engaged if multimodal interaction systems are enabled [25, 26]. Interactive tabletops has also been used to support collaborative work in education [13].

The interaction between the teacher and student is also of great importance for successful outcomes in education, and several works have investigated this interaction from the HBU perspective either in classroom settings to detect the influence of one party on another via automatic analysis of facial movements, face and hand gestures, and speech loudness [57] or in dyadic teacher-student interaction settings to detect helping and over-helping [11].

Recently, robots are being increasingly employed to engage children in educational scenarios [16]. There is significant evidence that robots can enhance learning [6], or can induce behavioral change, for instance in helping autistic individuals improve their social skills [10]. One such robotic application is described in this volume. The ALIZ-E Project[1] targets behavior change for children suffering from diabetes or obesity. In their work, Ros and Demiris propose to use robots as creative dance tutors [45]. In their experiments with 8 and 9-years-old children, the authors observe that a robot that displays non-verbal signals of interaction successfully engages children in a joint dance session, where children imitate the robot, and expect to be imitated by it. Similar engagement scenarios have been deployed for other socially assistive robotics applications, including fitness coaching for the elderly [21], and post-stroke rehabilitation [17], where the robot needs to observe and guide human subjects through imitation-based interactions. Towards the generation of multimodal behavior for a humanoid robot engaged in a collaborative task with a human partner, Mihoub et al. present a probabilistic approach based on hidden Markov models in this volume, using multimodal speech and gaze data from computer-mediated dyadic conversations [34].

Coaching systems come in many flavors, and robotic variants are relatively marginally used. Pfister and Robinson described a public speaking skills coach that is based on the analysis of affective states from nonverbal features of speech [41]. In this volume, Baur et al. present the NovA (NOnVerbal behavior Analyzer) system that automatically analyzes and interprets social signals from posture, gestures, facial expressions and others in the context of a job interview, which constitutes a starting point for social coaching in human-human interaction settings [4].

5 Open Challenges of HBU for Creative Applications

HBU for creative applications has an enormous potential that brings along a number of challenges. In what follows we attempt to focus specifically on three challenges that we group under the headings of behavioral experimentation, ecological validity, and data collection for analysis and validation.

5.1 Behavioral Experimentation

The rigors of experimental science, where conditions can be controlled for while varying one parameter, and the proof of a hypotheses is required based on measurable outcomes, may not be practically achievable "in the wild", outside of a clean lab environment. Sometimes, the real world imitates the lab. For instance, there can be striking similarities between galleries and exhibition spaces and the controlled lab environments. The gallery environment is purposefully designed to be clear of the clutter of the outside world so all attention can be focused

[1] www.aliz-e.org

on the artwork itself. Rooms are often curated into themes, which enables people to be exposed to a common aspect relating to a set of works, stimulating certain ideas while de-emphasizing others. Such controlled setups may also pose the Observer's Paradox problem identified by Labov as a situation where data gathering is influenced by the presence of the experimenter [28]. Even in "the wild" experimentation settings are not fully exempt from this dilemma - participants are still aware of placement of the sensors and their locations which tend to affect their (expressive) behavior [36].

An interesting case can be made for sports games, which constitute rich settings particularly suited for measuring affective states like excitement, triumph, disappointment, joy, and pride in naturalistic conditions. Matsumoto's well-publicized research investigated spontaneous facial expressions of medal winners in olympic games [33], focusing on the moment when the gold medal winner triumphs, and the silver medal winner (despite a tremendous achievement) faces the agony of defeat. Analysis of such moments presupposes that the recorded individual is either not aware of a detailed analysis that will take place, or is in an emotional state that is so intense, that the cameras pointed at the subject do not matter much.

The trade-off then is in capturing detailed information about a phenomenon, as opposed to capturing the information naturally. A recent example that illustrates this is the Multimodal and Multiperson Corpus of Laughter in Interaction (MMLI), described in this volume [38]. This corpus captures different laughter types, in different contexts, and with full body movement information in addition to audio and facial video recordings. The subjects are equipped with 17-19 inertial sensors placed on Velcro straps, in addition to recordings with two personal microphones close to the mouth, one room microphone, four webcams, two high-rate color cameras, two RGB-D cameras and one respiration sensor. The MMLI scenario illustrates the data intensive recording extreme of the spectrum, providing a much clearer picture of what happens at the expense of a natural setting.

5.2 Ecological Validity

In recent years, ecological validity has become a prerequisite for research studies focusing on human behavior understanding. Given the level of subjectivity in creativity, entertainment, and edutainment, it is certainly not possible to talk about the ability of the studies conducted in this field, and their results, to *generalize* (i.e., external validity). Would these studies that are meant to encourage originality, ambiguity, and even incongruity at times, be able to claim ecological validity? How closely do the settings, the conditions and the methods used in these studies replicate the actual look, feel and procedure of the so-called *real-world*? Is it possible to conduct research studies in these fields without disrupting levels of creativity and playfulness? These are key questions that cannot be approached and discussed by researchers alone, requiring input from artists and practitioners.

Interactive and new media art technologies have the potential to take human behavior research into the wider world, serving as a valid and a genuine context in which to conduct user-centered tests [12]. When using technologies as an ecological testbed for research, there is a need to achieve a balance between the accuracy of hypotheses testing in experiments and the naturalness of the experience and situation. Ideally, the contextual and aesthetic aspects of the artwork should guide the participant with respect to the purpose of the experiment, without limiting the overall experience [12]. In order to strike a balance between the functional needs of a given artistic context and assess the artistic purpose of the technology, Deweppe et al. [12] advocate the use of new media festivals (e.g. Ars Electronica, Transmediale) as ecological testbed for the abovementioned type of research. In this volume, the London Eye Mood Conductor project, with its design, outdoor setup and participants appears to form a good example - it meets the functional and the experimental needs of the artwork as well as participant experience, providing an ecologically valid case study of human affective behavior in new media art [36].

5.3 Data Collection for Analysis and Validation

A major challenge for creating rigours experimental conditions is obtaining some level of "understanding" about the data through an evaluation process that can then be subsequently used for analysis and validation. Participant evaluation can be obtained either explicitly or implicitly. Explicit evaluations typically provide self-report measures obtained by interviewing participants explicitly using surveys and questionnaires, as well as some standard evaluation and labelling techniques (e.g., the Self Assessment Mannequin (SAM) [29]). Implicit evaluation instead is obtained by recording the participants' visual, auditive and physiological behavior while they are experiencing the interaction. When recording multiple cues for implicit evaluation, it is important to keep in mind that not all recorded cues might be correlated with the displayed stimuli equally. Due to this, different results may be obtained with different types of data.

When it comes to behavior analysis of people interacting with art installations, special circumstances apply. Needless to say, art installations are subjective phenomena, therefore, should ideally be evaluated by self-report. However, self-report measures have their own limitations: they may be inconsistent, variable, and depend on past experience. Obtaining participant self-report might be easier in indoor and gallery environments where variables such as entry, number of participants, and access can be monitored and controlled. This in turn allows for the systematic exclusion of confounding factors and generally make data analysis more straightforward and the results more compelling. If the digital artwork moves outdoors, attemtping to obtain implicit measures might be more appropriate. This certainly appears to be one of the reasons for the London Eye Mood Conductor project, introduced in this volume, to opt for implicit evaluation [36]. However, data with implicit measures also presents new challenges when it comes to validation, analysis and extraction of meaningful results. Tackling these challenges is a necessary step if human behavior understanding is to be fully incorporated into new media art applications.

Different sets of sensors can be used in tandem when observing people. In [49], the authors use the audio modality and investigate the usage of functional meanings of discourse particles for emotion recognition. They present methods to extract the pitch-contour of the discourse particles and link them with emotions. While audio is seen as the more reliable modality in recognition of affect, the visual modality is predominantly used for identification of people, particularly from their faces [14]. Face analysis is also used to classify facial actions, which can give clues about affective states of a person [61, 15], and for gaze estimation, which can indicate the focus of attention of a person during interactions. In [31], an eye-gaze framework is proposed for capturing the dependence of eye-gaze on visual stimulus, intent and person.

When detailed information of individuals is not available, it may still be possible to sense collective emotions via the analysis of more general audiovisual cues. In one of the earliest papers on automatic detection of excitement in videos, Hanjalic used overall motion activity (measured at video frame transitions), the rhythm (via the changes in shot lengths along the video), and the energy (computed from the audio track of a video) to higlight segments with high expected excitement levels [22]. In this volume, Conigliaro et al. use image segmentation to measure excitement levels of groups of people during the 2013 IIHF Ice Hockey U18 World Championship [9]. This kind of analysis can be used for automatically segmenting groups of people among the spectators or players of a game, to monitor student engagement in edutainment applications, or to change game experience as the excitement levels are translated into visualizations and sonifications that encourage the spectators to be more active.

For interactive systems that sense human body positions, the recently developed RGB-D cameras (like Microsoft Kinect) quickly became standard. Thanks to specialized hardware support, RGB-D sensors make tracking a human body skeleton accurately in real-time possible, after which a number of gestures can be automatically recognized. Such devices made a whole new set of entertainment applications possible [48]. One example in this volume is the work on iconic gestures used by first person shooter games (e.g. crouch, shoot, throw, kick, etc.) and metaphoric gestures (e.g. raising hands to increase volume) that are recognized via rapid random forest feature selection and gentle AdaBoost classifiers on such depth camera images [7]. In [56], Wang et al. present an approach for real-time continuous emotion recognition from body movements. They extract both low level and high level 3D postural features to represent dynamics of body movements and train a random forest classifier to obtain their model, which has been tested on continuous Kinect data.

State-of-the-art gesture classification approaches (including random forest based methods) require large sets of training samples for accurate representation of each gesture class. For this reason, innovative methods are sought to help in training generalizable and flexible models. In [27], a 3D hand model is is used to synthesize many hand scans taken from different angles to train a hand gesture recognition system. In this volume, a transfer learning procedure is proposed to take advantage of the fact that some human body poses are shared among actions, and key poses can be trained from external sources [43]. Most action and activity recognition systems target a specific set of behavior classes (e.g. pre-determined gaming gestures for a gesture-sensing game, as in [7]), which reduces the number of classes and simplifies the training procedure. Avcı and Passerini use correlations in multiple sensor streams to detect change points, and to automatically segment behavior classes [1]. While this is a more challenging scenario, the idiosyncratic variations in behavior classes make it also a potentially more powerful approach.

6 Conclusions

This paper highlights research on analysis of human behavior under its multiple facets (expression of emotions, display of complex social and relational behaviors, performance of individual or joint actions, etc.), with particular attention to interactions in arts, creativity, entertainment and edutainment.

With advances in pattern recognition and multimedia computing, it has become possible to analyze human behavior via multimodal sensors, at different time-scales and at different levels of interaction and interpretation. This ability opens up enormous possibilities for multimedia and multimodal interaction, with a potential of endowing the computers with a capacity to attribute meaning to users attitudes, preferences, personality, social relationships, etc., as well as to understand what people are doing, the activities they have been engaged in, their routines and lifestyles. Re-defining the relationship between the computer and the interacting human, moving the computer from a passive observer role to a socially active participant role and enabling it to drive different kinds of interaction has implications across multiple domains, including arts, creativity, entertainment and edutainment. In these domains in particular, modes of interaction are highly variable and so observing, analyzing and interpreting behaviors systematically is a challenge.

7 Acknowledgments

The work of Albert Ali Salah is supported by Boğaziçi University project BAP 6531. The work of Hatice Gunes is partially supported by the EPSRC MAPTRAITS Project (Grant Ref: EP/K017500/1). The work of Oya Aran is supported by the Swiss National Science Foundation (SNSF) Ambizione fellowship under the project "Multimodal Computational Modeling of Nonverbal Social Behavior in Face to Face Interaction"(SOBE).

References

1. Avci, U., Passerini, A.: A fully unsupervised approach to activity discovery. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 77–88. Springer, Heidelberg (2013)

2. Bachour, K., Kaplan, F., Dillenbourg, P.: An interactive table for supporting participation balance in face-to-face collaborative learning. Learning Technologies, IEEE Transactions on 3(3), 203–213 (2010)
3. Baumgartner, T., Esslen, M., Jäncke, L.: From emotion perception to emotion experience: Emotions evoked by pictures and classical music. International Journal of Psychophysiology 60(1), 34–43 (2006)
4. Baur, T., Damian, I., Lingenfelser, F., Wagner, J., André, E.: NovA: Automated analysis of nonverbal signals in social interactions. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 158–169. Springer, Heidelberg (2013)
5. Beesley, P.: Hylozoic soil. Leonardo 42(4), 360–361 (2009)
6. Benitti, F.B.V.: Exploring the educational potential of robotics in schools: A systematic review. Computers & Education 58(3), 978–988 (2012)
7. Bloom, V., Makris, D., Argyriou, V.: Dynamic feature selection for online action recognition. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 64–76. Springer, Heidelberg (2013)
8. Boerman, H., Bergman, T., Pieters, J., Reitsma, L., van den Hoven, E.: Sissy: an interactive installation with a personality. In: OZCHI Workshop Proceedings - The Body In Design. pp. 1–4 (2011)
9. Conigliaro, D., Setti, F., Bassetti, C., Ferrario, R., Cristani, M.: ATTENTO: ATTENTion Observed for automated spectator crowd monitoring. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 100–109. Springer, Heidelberg (2013)
10. Dautenhahn, K., Werry, I.: Towards interactive robots in autism therapy: Background, motivation and challenges. Pragmatics & Cognition 12(1), 1–35 (2004)
11. D'Errico, F., Leone, G., Poggi, I.: Types of help in the teacher's multimodal behavior. In: Salah, A., Gevers, T., Sebe, N., Vinciarelli, A. (eds.) Human Behavior Understanding, Lecture Notes in Computer Science, vol. 6219, pp. 125–139. Springer Berlin Heidelberg (2010)
12. Deweppe, A., Diniz, N., Coussement, P., Leman, M.: A methodological framework for the development and evaluation of user-centered art installations. Journal of Interdisciplinary Music Studies 5(1), 19–39 (2011)
13. Dillenbourg, P., Evans, M.: Interactive tabletops in education. International Journal of Computer-Supported Collaborative Learning 6(4), 491–514 (2011), http://dx.doi.org/10.1007/s11412-011-9127-7
14. Dornaika, F., Bosgahzadeh, A., Raducanu, B.: Efficient graph construction for label propagation based multi-observation face recognition. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 122–133. Springer, Heidelberg (2013)
15. Drira, H., Ben Amor, B., Daoudi, M., Berretti, S.: A dense deformation field for facial expression analysis in dynamic sequences of 3D scans. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 146–157. Springer, Heidelberg (2013)
16. Druin, A., Hendler, J.A.: Robots for kids: Exploring new technologies for learning experiences. Morgan Kaufmann (2000)
17. Eriksson, J., Mataric, M.J., Winstein, C.J.: Hands-off assistive robotics for post-stroke arm rehabilitation. In: Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on. pp. 21–24. IEEE (2005)
18. Fourati, N., Pelachaud, C.: Head, shoulders and hips behaviors during turning. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 221–232. Springer, Heidelberg (2013)

19. Gardair, C., Healey, P.G., Welton, M.: Performing places. In: Proceedings of the 8th ACM Conference on Creativity and Cognition. pp. 51–60. ACM, New York, NY, USA (2011), http://doi.acm.org/10.1145/2069618.2069629
20. Glowinski, D., Coletta, P., Volpe, G., Camurri, A., Chiorri, C., Schenone, A.: Multi-scale entropy analysis of dominance in social creative activities. In: Proceedings of the international conference on Multimedia. pp. 1035–1038. MM '10, ACM, New York, NY, USA (2010), http://doi.acm.org/10.1145/1873951.1874143
21. Görer, B., Salah, A., Akın, H.: A robotic fitness coach for the elderly. In: Int. Joint Conf. on Ambient Intelligence (2013)
22. Hanjalic, A.: Adaptive extraction of highlights from a sport video based on excitement modeling. Multimedia, IEEE Transactions on 7(6), 1114–1122 (2005)
23. Hara, K., Takemura, N., Iwai, Y., Sato, K.: Table-Top Interface Using Fingernail Images and Real Object Recognition. In: Keyson, D., Maher, M., Streitz, N., Cheok, A., Augusto, J., Wichert, R., Englebienne, G., Aghajan, H., Kröse, B. (eds.) Ambient Intelligence, Lecture Notes in Computer Science, vol. 7040, pp. 21–30. Springer Berlin Heidelberg (2011)
24. Höök, K., Sengers, P., Andersson, G.: Sense and sensibility: evaluation and interactive art. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 241–248. CHI '03, ACM, New York, NY, USA (2003), http://doi.acm.org/10.1145/642611.642654
25. Jovanovic, M., Starcevic, D., Minovic, M., Stavljanin, V.: Motivation and multimodal interaction in model-driven educational game design. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on 41(4), 817–824 (2011)
26. Kannetis, T., Potamianos, A.: Towards adapting fantasy, curiosity and challenge in multimodal dialogue systems for preschoolers. In: Proceedings of the 2009 international conference on Multimodal interfaces. pp. 39–46. ICMI-MLMI'09, ACM, New York, NY, USA (2009), http://doi.acm.org/10.1145/1647314.1647324
27. Keskin, C., Kıraç, F., Kara, Y.E., Akarun, L.: Hand pose estimation and hand shape classification using multi-layered randomized decision forests. In: Proc. ECCV, pp. 852–863. Springer (2012)
28. Labov, W.: Language in Use, chap. Field Methods of the Project in Linguistic Change and Variation, pp. 28–53. Prentice-Hall (1984)
29. Lang, P.: The cognitive psychophysiology of Emotion: Anxiety and the anxiety disorders. Lawrence Erlbaum, Hillside, NJ (1985)
30. Lombardi, M., Pieracci, A., Santinelli, P., Vezzani, R., Cucchiara, R.: Human behavior understanding with wide area sensing floors. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 110–121. Springer, Heidelberg (2013)
31. Ma, K.T., Sim, T., Kankanhalli, M.: VIP: An unifying formal framework for computational eye-gaze research. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 207–220. Springer, Heidelberg (2013)
32. Marchegiani, L., Fafoutis, X.: A behavioral study on the effects of rock music on auditory attention. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 15–26. Springer, Heidelberg (2013)
33. Matsumoto, D., Willingham, B.: The thrill of victory and the agony of defeat: spontaneous expressions of medal winners of the 2004 athens olympic games. Journal of personality and social psychology 91(3), 568–581 (2006)
34. Mihoub, A., Bailly, G., Wolf, C.: Social behavior modeling based on incremental discrete hidden Markov models. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 170–181. Springer, Heidelberg (2013)

35. Moreno, A., Reidsma, D., van Delden, R., Poppe, R.: Socially aware interactive playgrounds: Sensing and inducing social behavior. IEEE Pervasive Computing 12(3), 40–47 (2013)
36. Morgan, E., Gunes, H.: Human nonverbal behaviour understanding in the wild for new media art. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 27–39. Springer, Heidelberg (2013)
37. Morgan, E., Gunes, H., Bryan-Kinns, N.: Measuring affect for the study and enhancement of co-present creative collaboration. In: Proc. of Int'l Conference on Affective Computing and Intelligent Interaction (ACII) (2013)
38. Niewiadomski, R., Mancini, M., Baur, T., Varni, G., Griffin, H., Aung, M.S.: MMLI: Multimodal multiperson corpus of laughter in interaction. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 182–193. Springer, Heidelberg (2013)
39. Öhman, A., Flykt, A., Esteves, F.: Emotion drives attention: detecting the snake in the grass. Journal of Experimental Psychology: General 130(3), 466–478 (2001)
40. Otsuka, Y., Inoue, T.: Designing a conversation support system in dining together based on the investigation of actual party. In: Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on. pp. 1467–1472 (2012)
41. Pfister, T., Robinson, P.: Real-time recognition of affective states from nonverbal features of speech and its application for public speaking skill analysis. Affective Computing, IEEE Transactions on 2(2), 66–78 (2011)
42. Rasheed, U., Tahir, Y., Dauwels, S., Dauwels, J., Thalmann, D., Thalmann, N.: Real-time comprehensive sociometrics for two-person dialogs. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 194–206. Springer, Heidelberg (2013)
43. Rodríguez Martínez, M.F., Medrano, C., Herrero, E., Orrite, C.: Transfer learning of human poses for action recognition. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 89–99. Springer, Heidelberg (2013)
44. Rogers, Y., Preece, J., Sharp, H.: Interaction Design. Wiley & Sons 2011 (2007)
45. Ros, R., Demiris, Y.: Creative dance: an approach for social interaction between robots and children. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 40–51. Springer, Heidelberg (2013)
46. Salah, A.A., Gevers, T., Sebe, N., Vinciarelli, A.: Challenges of human behavior understanding. In: HBU. Lecture Notes in Computer Science, vol. 6219, pp. 1–12. Springer (2010)
47. Samadani, A.A., Gorbet, R., Kulić, D.: Gender differences in the perception of affective movements. In: Salah, A.A., Ruiz-del Solar, J., Mericli, C., Oudeyer, P.Y. (eds.) HBU 2012. LNCS, vol. 7559. pp. 65–76. Springer, Heidelberg (2012)
48. Schouten, B.A., Tieben, R., van de Ven, A., Schouten, D.W.: Human behavior analysis in ambient gaming and playful interaction. In: Salah, A.A., Gevers, T. (eds.) Computer Analysis of Human Behavior, pp. 387–403. Springer (2011)
49. Siegert, I., Hartmann, K., Philippou-Hübner, D., Wendemuth, A.: Human behaviour in HCI: Complex emotion detection through sparse speech features. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 244–255. Springer, Heidelberg (2013)
50. Smeaton, A.F., Rothwell, S.: Biometric responses to music-rich segments in films: The cdvplex. In: Content-Based Multimedia Indexing, 2009. CBMI'09. Seventh International Workshop on. pp. 162–168. IEEE (2009)
51. Snibbe, S.S., Raffle, H.S.: Social immersive media: pursuing best practices for multi-user interactive camera/projector exhibits. In: Proceedings of the SIGCHI

Conference on Human Factors in Computing Systems. pp. 1447–1456. CHI '09, ACM, New York, NY, USA (2009), http://doi.acm.org/10.1145/1518701.1518920
52. Tarvainen, J., Westman, S., Oittinen, P.: Stylistic features for affect-based movie recommendations. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 52–63. Springer, Heidelberg (2013)
53. Turk, M.: Multimodal interaction: A review. Pattern Recognition Letters (2013)
54. Varni, G., Volpe, G., Camurri, A.: A system for real-time multimodal analysis of nonverbal affective social interaction in user-centric media. IEEE Transactions on MultiMedia 12(6), 576–590 (2010)
55. Varni, G., Mancini, M., Volpe, G., Camurri, A.: Sync'n'Move: Social Interaction Based on Music and Gesture. In: Daras, P., Ibarra, O. (eds.) User Centric Media, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 40, pp. 31–38. Springer Berlin Heidelberg (2010)
56. Wang, W., Enescu, V., Sahli, H.: Towards real-time continuous emotion recognition from body movements. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 233–243. Springer, Heidelberg (2013)
57. Watanabe, E., Ozeki, T., Kohama, T.: Extraction of relations between behaviors by lecturer and students in lectures. In: Automatic Face Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on. pp. 945–950 (2011)
58. Whyte, W.H.: The Social Life of Small Urban Spaces. Project for Public Spaces Inc (2001)
59. Wolf, G.: The data-driven life. The New York Times 28 (2010)
60. Woolley, A.W., Chabris, C.F., Pentland, A., Hashmi, N., Malone, T.W.: Evidence for a Collective Intelligence Factor in the Performance of Human Groups. Science 330(6004), 686–688 (Oct 2010), http://dx.doi.org/10.1126/science.1193147
61. Yüce, A., Arar, N.M., Thiran, J.P.: Multiple local curvature Gabor binary patterns for facial action recognition. In: Salah, A.A., Hung, H., Aran, O., Gunes, H. (eds.) HBU 2013. LNCS, vol. 8212. pp. 134–145. Springer, Heidelberg (2013)