Modeling expressive individual and collective behavior

Antonio Camurri, Gualtiero Volpe, Donald Glowinski,
Maurizio Mancini, Barbara Mazzarino, Giovanna Varni

Casa Paganini – InfoMus Lab, DIST – University of Genova
Viale F. Causa 13, 16145 Genova, Italy
www.infomus.org www.casapaganini.org
antonio.camurri@unige.it, gualtiero.volpe@unige donald@infomus.org,
maurizio@infomus.org barbara.mazzarino@unige.it, giovanna@infomus.org,

Abstract. This paper presents a brief survey of our recent research on analysis and modeling of human expressive behavior. In particular, the focus is on non-verbal expressive, emotional communication and research addresses expressive behavior of both individuals and groups of users. Starting from the analysis of expressive gesture in single users, we show how our work is evolving towards the analysis of expressive social interaction in a group. Social interaction and its expressive implications (e.g., emotional contagion, empathy) is a very relevant component for analysis of human behavior, since it provides significant information on the social context behavior is displayed. In this direction, this paper presents the techniques we developed for measuring (i) saliency and emerging behavior (i.e., how the behavior of a single user emerges from the group) and (ii) synchronization within a group of users (i.e., how much the behavior of the group is coherent). Research is performed in the framework of the EU-ICT Project SAME (www.sameproject.eu).

Keywords: expressive human behavior in both individuals and groups of users, expressive gesture, multimodal interactive systems, non-verbal communication, social interaction, social signals, context-awareness.

1 Introduction

This paper presents a survey of our recent research on analysis of expressive human behavior, featuring non-verbal communication and analysis of expressive gesture. We show how the focus is moving from the analysis of individual behavior, towards collective behavior, i.e., social interaction in (small) groups of users.

Indeed, whereas analysis of expressive behavior of individual users has received lot of attention in recent years, analysis of expressive collective behavior in social interaction is still often neglected. Nevertheless, social interaction is an extremely relevant component for analysis of human behavior, since it provides significant information on the social context where behavior is displayed. *Social intelligence* or social competencies, understood as the ability to deal effectively in interpersonal contexts, is a paradigmatic human ability, widely studied in psychology and more recently in neurophysiology, which is receiving a growing interest from the ICT communities, e.g., in the framework of social networks, networked media, Future Internet, user-centric media (e.g., see Laso-Ballesteros and Daras, 2008).

Current research on social interaction (e.g., see Vinciarelli et al. 2009), however, does not focus on the high-level emotional aspects, but rather on group cohesion and decision-making. In this framework, pioneering studies by Pentland (2007) investigated techniques to measure social signals in scenarios like salary negotiation and friendship. Particular attention was also directed to the recognition of functional roles (e.g., most dominant people) played during small-group meetings (e.g., Dong et al., 2007). These works are often based on laboratory experiments and do not address the more subtle aspects of social interaction such as emotional contagion and empathy. Empathy, in fact, has been studied mainly in the framework of synthesis of (verbal) dialogues by virtual characters and embodied conversational agents (see for example de Rosis et al., 2005; McQuiggan and Lester, 2007). The EU-ICT project SAME (www.sameproject.eu) has recently developed techniques for social active listening to music by mobile devices, i.e., for allowing a (small) group of users to mould collaboratively a pre-recorded music piece they are listening to (e.g., see Varni et al., 2009b, for the description of an application presented at Agora Festival, Ircam, Paris, in June 2009).

The major research challenge in our work consists of analyzing even the subtlest and most significant emotional expressions of human behavior in a social framework, such as empathy and emotional contagion. After recalling our research on analysis of expressive gesture in individual users, this paper presents two steps in the direction of novel computational models and techniques for analysis of collective expressive human behavior: (i) the techniques we developed for investigating the emergence of a single user within a group, i.e., for identifying salient behavior, and (ii) the techniques we developed for explicitly measuring synchronization (i.e., behavior coherence) within a group of users. The latter are based on analysis of complex systems, where each user is modeled as a component of a complex system. Finally, starting from this experience, we conclude with the research challenges we will face in the near future. These are in the direction of a deeper understanding of the mechanisms underlying complex phenomena such as empathy and emotional contagion, that are studied by considering ensemble music performance (e.g., a string quartet). Since the major role of expressive non verbal communication channels, we selected music performance and performing arts in general (e.g., dance) as an ideal test-bed for experiments and proof-of-concepts. However, results from research can be applied to many other application fields, including, for example, cultural heritage and museum and cultural applications, networked user centric media, and therapy and rehabilitation.

2 Analysis of expressive gesture of individual users

Our work on analysis of expressive individual behavior focuses on the non-verbal communication channels, with a particular focus on human full-body movement and gesture. Research is grounded on the concept of *expressive gesture*. Basing on the Kurtenbach and Hulteen's definition of gesture as "a movement of the body that contains information", a gesture can be said expressive since the information it carries is an expressive content, i.e., a content related to the emotional, affective sphere (Camurri et al., 2004). That is, expressive gesture is responsible of the communication of a kind of information (addressed as expressive content) that is different and independent, even if often superimposed, to a possible denotative meaning, and that concerns aspects related to affects. Expressive gesture as a key aspect of human behavior and in particular of expressive human behavior, became particularly relevant in recent years (e.g., see the post-proceedings of Gesture Workshops 2003-2009). Psychological studies have been a rich source for research on automatic analysis of expressive gesture since they identified which features are most significant (e.g., De Meijer, 1989; Wallbott, 1998; Boone and Cunningham, 1998). A further relevant source has been research in the humanistic tradition, in particular choreography. As a major example, in his Theory of Effort, choreographer Rudolf Laban (1947) describes the most significant qualities of movement. Starting from these sources, several systems for analysis of expressive gesture were developed (e.g., Camurri et al., 2003, 2005; Kapur et al., 2005; Bernhardt et al., 2007).

Our approach to analysis of expressive gesture grounds its bases on a multilayered conceptual model for multimodal analysis of affective, emotional content in human full-body movement and gesture (Camurri et al. 2004). The model builds on four different layers/steps following a bottom-up approach. Layer 1 includes techniques for pre-processing of data from different kinds of sensors such as video-cameras, on-body (e.g., accelerometers), physiological, and environmental sensors. Layer 2 extracts from the sensors data a collection of expressive motion features describing the movement being performed. Features are derived from the above-mentioned research in psychology and human sciences. The EyesWeb XMI platform for synchronized analysis of multimodal data streams (www.eyesweb.org) allows for the extraction of a wide collection of motion features from video and sensors data streams. Layer 3 deals with two major issues: segmentation of movement in its composing gestures and representation of such gestures in suitable spaces. Layer 4 is conceived as a conceptual network mapping the extracted features and gestures into conceptual structures. For example, an experiment was carried out for distinguishing among the four basic emotions (anger, fear, grief, and joy) in dance performances (Camurri et al., 2003; 2004). Machine learning techniques are usually employed at this layer ranging from statistical techniques (e.g., multiple regression and generalized linear techniques), to fuzzy logics or probabilistic reasoning systems (e.g., Bayesian networks), to various kinds of neural networks (e.g., classical back-propagation networks, Kohonen networks), support vector machines, and decision trees.

As an example of use and implementation of the conceptual model, in a recent application – the interactive collaborative installation *Mappe per Affetti Erranti*, presented at the Festival della Scienza of Genova (Camurri et al., 2008) – we analyzed the expressive gestures performed by the users in order to control in real-time the reproduction of four expressive performances of the same music piece (a choral by J.S. Bach). The four expressive performances corresponded to the following expressive intentions: Happy/Joyful, Solemn, Intimate/Shy, and Angry/Aggressive. These were associated to the same four expressive intentions classified from users' expressive gestures. Analysis of expressive gesture was performed by means of twelve expressive features: Motion Index (i.e., motor activity), computed on the overall body movement and on translational movement only; Impulsiveness, vertical and horizontal components of velocity of peripheral upper parts of the body; speed of the barycentre; variation of the Contraction Index; Space Occupation Area; Directness Index (inspired by the Space dimension of Laban's Effort Theory), Space Allure (inspired by the Pierre Schaeffer's Morphology), Amount of Periodic Movement, and Symmetry Index. Such features were computed in real-time for each single user (a maximum of four users could experience the installation at the same time). Further features were also computed on the whole group of users, such as for example, the contraction/expansion of the group and its cohesion. This perspective corresponds to R. Laban's *General Space* (Laban and Lawrence, 1947). Classification was performed following a fuzzy-logic like approach. Such an approach had the advantage that it did not need a training set of recorded movement and it was also flexible enough to be applied to the movement of different kinds of users (e.g., adults, children, elder people). Figure 1 shows *Mappe per Affetti Erranti* experienced by a group of users and a couple of dancers during a dance performance.

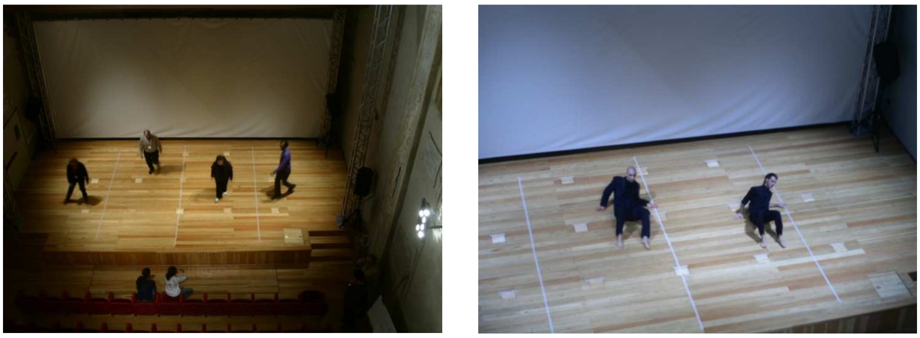

Fig. 1. *Mappe per Affetti Erranti* experienced by a group of four users (left) and by two dancers during a dance performance (right).

The conceptual model has been recently extended to take into account context-awareness and social interaction. Context-awareness enables changing and adapting processing and mapping at all levels to what a subject is doing (e.g., at home, walking, running, driving) and to the input/output devices that are available. Analysis of social interaction will be discussed in the following. Research in experimental psychology and neurosciences has shown that non-verbal communication, and in particular expressive gesture, is a key aspect of collective human behavior.

3 Emerging behavior of individual users: group cohesion and expressive behavior saliency

As a first direction towards analysis of collective expressive behavior we consider the emergence of a particularly salient behavior in the components of a group. Analysis of saliency can also provide information about group cohesion, since the detection of too many salient behaviors, i.e., of too many components of the group displaying a behavior which is different from the average behavior of the group, may indicate a scarce cohesion between the components of the group. In other words, analysis of salient behavior enables the identification of salient participants in a group or of participants breaking the cohesion of a group.

Behavior saliency with respect to expressive gesture is analyzed both in the spatial and in the temporal dimension starting from expressive features such as those mentioned in the previous section. Behavior saliency with respect to time is related to expressive features measured on an individual user assuming values and having a dynamics which are significantly different from the average values and dynamics measured on all the components of the group in a selected time window. Similarly, behavior saliency with respect to space is related to trajectories and kinematical features measured on an individual user assuming values and having a dynamics which are significantly different from the average values and dynamics measured on all the components of the group in the physical space the group occupies.

Behavior saliency can be modeled starting from recent results obtained in the field of *computational attention*. The aim of computational attention is to automatically predict human attention based on different kinds of data such as sounds, images, video sequences, smell, or taste. In this framework, salient behavior is understood as a behavior capturing the attention of the observer. Whereas may models were provided for attention on still images, time-evolving two-dimensional signals such as videos have been much less investigated. Nevertheless, some of the authors providing static attention approaches generalized their models to the time dimension: e.g., Dhavale and Itti (2003), Yee and Pattanaik (2001), Parkhurst and Niebur (2004), Itti and Baldi (2006), Le Meur et al., (2006), Liu and Gleicher (2006). Motion has a predominant role and the temporal contrast of its features is mainly used to highlight important movements. Boiman and Irani (2005) provided an outstanding model which is able to compare observed movements with others from a video history or video database. Attention is related to motion similarity. The major problem of this approach is in its high computational cost.

In order to get an efficient model of motion attention and behavior salience, in collaboration with University of Mons we developed a three-level rarity-based model (Mancas et al., 2009). At the first level, motion features are compared in the spatial context of the current video frame; at the intermediate level, salient behavior is analyzed on a short temporal context; at the third level, computation of saliency is extended to longer time windows. An attention/saliency index is computed at each of the three levels based on an information theory approach.

We tested both the behavior saliency model and the attention/saliency indexes it provides during a dance master-class directed by choreographer Giovanni Di Cicco in the framework of Digifestival (Casa Paganini, Genova, Italy, November 2008). The feature taken into account here was Motion Index, measuring motor activity. In this dance application, the value of the saliency index in the temporal context (salient values of the Motion Index in a time windows of 3s), computed for each dancer, controlled the transparency of the silhouette of the dancer, which was extracted from the live video from an infra-red video-camera using a multi-blob tracking technique. The higher was the dancer’s saliency index, the more opaque was its silhouette. Figure 2 shows some results. On the left image, the dancer located in the middle stays still whereas the two others are running: his behavior is salient relatively to the others.

On the right image, the dancer, located in the right, is moving at a higher speed than the two others, thus having the most salient behavior.

Fig. 2. Two snapshots corresponding to two situations observed during the dance master-class at Digifestival. In both situations the silhouette which appears on the video in the background is the one of the dancer displaying the most salient behavior with respect to the two others.

Analysis of salient behavior can also help in making analysis of expressive gesture context-aware and adaptive. The detection of a salient behavior can result in setting the focus of analysis of expressive gesture on the individual displaying such a salient behavior. Moreover, analysis parameters can be dynamically changed on the basis of the emerging behavior, thus making analysis of expressive gesture customizable and personalized to each single user. Future research will go in this direction, extending models and techniques and applying and validating them with new data obtained from experimental scenarios (e.g., ensemble music playing). In particular, analysis will extend to measures of complexity, namely *Multi-scale Entropy* (MSE).

4 Towards analysis of collective behavior

Analysis of expressive gesture of individual users and detection of salient and emerging behavior are still not sufficient for fully characterizing expressive collective behavior. Key aspects for this are rather a measure of the degree of synchronization and, in particular, of emotional synchronization or empathy within the group.

Synchronization can be broadly referred to as a phenomenon occurring when “two or many systems adjust a given property of their motion to a common behavior, due to coupling or forcing” (Boccaletti et al., 2002). This definition covers different kinds of synchronization: from Phase Synchronization (PS), in which only the phases of the trajectories described by the systems in the phase space are locked, to Complete Synchronization (CS) in which the trajectories are almost identical.

To date, notwithstanding the large number of works on synchronization in many research fields (e.g., electronics, physics, medicine, psychology), there is a lack of studies focusing on this phenomenon in non stereotyped and non laboratory conditions and taking into account non-verbal expressive communication.

Our approach considers the component of a group of interacting users as a complex system having as basic units the single users. It is well known that interacting units of a complex system are able to auto-organize and exhibit global properties, which are not obviously derived from their individual dynamics: synchronization is one of these properties. Each user is described by means of the time evolution of a N-dimensional state vector of behavioral expressive features. The state vector components may include, for example, position and velocity of joints or other body parts (e.g., center of mass of head or limbs), energy and amount of motion, or audio and physiological features. We refer to such multimodal features as *expressive Movement, Audio, Physiological* (eMAP) features. eMAP features are extracted using real-time, synchronized, multi-modal feature extraction techniques and are the inputs to the computational models explaining the processes underlying interpersonal creative communication.

We chose PS as one of the baseline low-level signals to indirectly measure more complex phenomena like empathy and dominance in small groups of subjects. Our hypothesis is that empathy occurs when synchronization of specific expressive features emerge. Nonetheless, our hypothesis considers this as a necessary, but not sufficient, condition: synchronization may emerge also in cases where empathy does not occur. Our work addresses PS exploiting the concepts of Recurrence, introduced by Poincaré in the late 19th century, Recurrence Plots (RP) (Eckmann et al., 1987) and Cross-Recurrence Plot (CRP), and their quantification by means of Recurrence Quantification Analysis (RQA) (Marwan et al., 2007; Zbilut et al., 1992). RP/CRP and RQA provide qualitative and quantitative information on systems' dynamics and their interrelations in terms of trajectories in the phase space, whereas RQA allows to quantify small-scale patterns in RP/CRP and provides quantitative information on the systems dynamics. We think that changes in the number of occurrences and strength of PS among users can be considered useful features toward evaluation of empathy.

In our research, we focused on joint music performance as an ideal test-bed for the development of models and techniques for measuring creative social interaction in an ecologically valid framework. Music is widely regarded as the medium of emotional expression par excellence. Moreover, ensemble performance is one of the most closely synchronized activity that human beings engage in: it is believed that this ability by individuals and groups to entrain to music is unique only to humans and that, unlike speech, music performance is one of the few expressive activities allowing simultaneous participation.

During the last three years, we focused on the analysis of famous string quartets and of duos of violin players. The ensembles Cuarteto Casal, Quartetto di Cremona, Quartetto Prometeo have been involved initially in feasibility studies (e.g., to study and understand which multimodal features can explain their expressive social behavior) and in experiments at our Centre and in occasion of their concerts at the Opera House of Genova. In addition, in collaboration with Ben Knapp (SARC, Queen's University, Belfast) and Carol Krumhansl we carried out measurements of duos of violinists participating in the International Violin Competition Premio Paganini in 2006, in the framework of the EU Summer School of the HUMAINE Network of Excellence. More recently, again in collaboration with the SARC team, we performed multimodal synchronized recordings of the Quartetto di Cremona. Figure 3 shows snapshots from the experiments.

Fig. 3. Experiments on joint music performance: (a) music duo performance in which the two musicians can communicate, also exchanging visual information, (b) the famous string quartet Quartetto di Cremona during the experiment. Each musician wears a white hat including a green passive marker and a 3-axis accelerometer, plus a 3-axis accelerometer on the back, and physiological sensors for heart rate, breath, ocular movements, and face muscles.

Using the PS approach described above, several results emerged: for example, in the case of a music duo performance, it was possible to evaluate how the visual and acoustic channels affect the exchange of expressive information during the performance and how positive emotion can affect the emergence of synchronization (Varni et al., 2008). Moreover, foundations for a criterion to distinguish between parallel and reactive empathic outcomes have been defined. Measures of the direction of PS confirmed the hypothesis on egalitarian distribution of dominance in a duo performance. Furthermore, preliminary results from the analysis of string quartets highlighted how the induction of a positive emotion in one of the musicians of the group resulted in an increased synchronization among musicians (in terms of heads movement), with respect to no emotion induction condition. In the same experiment, the SARC colleagues found high physiological synchronization with the structural changes in the music. Moreover, measures relating to performer mistakes, and the perceived difficulty of the music were found, which also strongly affect both intra- and inter-personal synchronization. This effect of emotion on synchronization (emotional synchronization) is an important issue that will be further explored in our research.

We developed a real-time implementation of these techniques, resulting in the EyesWeb XMI Social Signal Processing Library (Varni et al., 2009a), which is employed in the framework of the SAME Project to develop applications for social active music listening experiences. In particular, the *Sync'n'Move* application prototype, based on EyesWeb XMI and its extensions to Nokia S60 mobile phones, enables users to experience novel form of social interaction based on music and gesture (Varni et al., 2009b). Users move rhythmically (e.g., dance) while wearing their mobiles. Their PS is extracted from their gesture (e.g., using the accelerometers data from the mobiles) and used to modify in real-time the performance of a pre-recorded music. More specifically, every time users are successful in synchronizing among themselves, music orchestration and rendering is enhanced; whereas in cases of low synchronization, i.e., poor collaborative interaction, music gradually corrupts, looses sections and rendering features, until it becomes a very poor audio signal.

5 Conclusion

This paper presented a survey of our research on analysis of expressive individual and collective behavior, grounding on our previous work on expressive gesture processing and evolving towards analysis of social interaction in (small) groups of users.
Research on measuring quantitative and qualitative interpersonal communication and social behavior in groups and on the supporting computational models, techniques, and tools is still a broadly unexplored field. In this scenario, our future research will include: investigation on the key factors driving interpersonal synchronization in a group and determining the feeling of group cohesion; how emotional, physical, and social contexts can affect interpersonal and intrapersonal synchronization in one or more modalities; the identification of specific functional roles inside a group (e.g., the leader). Further, another recent research project concerns the social active experience of audiovisual content, in the framework of museum and cultural applications: here we analyze expressive individual and collective behavior in order to enhance the experience of museum visitors.

Acknowledgements. This research is partially supported by the EU 7FP ICT SAME Project (www.sameproject.eu). We thank our colleagues at InfoMus – Casa Paganini Corrado Canepa, Paolo Coletta, and Alberto Massari for their important contributions to research.

References

Bernhardt, D., Robinson, P., 2007. Detecting Affect from Non-stylised Body Motions. LNCS 4738:59, Springer Verlag.

Boccaletti, S., Kurths, J., Osipov, G., Valladeres, D.L., Zhou, C.S., 2002. The Synchronization of Cahotic Systems. Phys. Rep. 366, (1-2): 1-101.

Boiman, O., Irani, M., 2005. Detecting Irregularities in Images and in Video, International Conference on Computer Vision (ICCV).

Boone R. T., Cunningham J. G., 1998. Children's decoding of emotion in expressive body movement: The development of cue attunement. Developmental Psychology, 34:1007-1016.

Camurri, A., Lagerlöf, I., Volpe, G., 2003. Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques. Intl. J. Human-Computer Studies, Elsevier, 59:213–225.

Camurri, A., Mazzarino, B., Ricchetti, M., Timmers, R., Volpe, G., 2004. Multimodal analysis of expressive gesture in music and dance performances. In Camurri, A. and Volpe G., (Eds.), Gesture-based Communication in Human-Computer Interaction, 20-39, Springer Verlag, LNAI 2915

Camurri, A., De Poli, G., Leman, M., Volpe, G., 2005. Toward Communicating Expressiveness and Affect in Multimodal Interactive Systems for Performing Art and Cultural Applications. IEEE Multimedia, 12(1): 43-53, IEEE CS Press.

Camurri, A., Canepa, C., Coletta, P., Ferrari, N., Mazzarino, B., Volpe, G., 2008. The Interactive Piece The Bow is bent and drawn. In Proc. 3rd ACM Intl. Conf. on Digital Interactive Media in Entertainment and Arts (DIMEA 2008), 376-383, Athens, Greece.

De Rosis, F, Cavalluzzi, A., Mazzotta, I., Novielli, N., 2005. Can embodied conversational agents induce empathy in users? In Proceedings of AISB 05, Virtual Social Character Symposium.
De Meijer, M., 1989. The contribution of general features of body movement to the attribution of emotions. J. of Nonverbal Behavior, 13, 247-268.
Dhavale, N., Itti., L., 2003. Saliency-based multi-foveated mpeg compression. IEEE 7th International Symposium on Signal Processing and its Applications.
Dong, W., Lepri, B., Cappelletti, A., Pentland, A., Pianesi, F., Zancanaro, M., 2007. Using the influence model to recognize functional roles in meetings. In Proc. 9th Intl. ACM Conf. on Multimodal Interfaces, 271-278.
Eckmann, J.P., Kamphorst, S.O., Ruelle, D., 1987. Recurrence plots of dynamical system. Eurpph. Lett., 5:973-977.
Itti, L., Baldi, P., 2006. Bayesian Surprise Attracts Human Attention. Advances in Neural Information Processing Systems, Vol. 19 (NIPS 2005), 1-8, MIT Press, Cambridge.
Kapur, A., Kapur, A., Virji-Babul, N., Tzanetakis, G., Driessen, P.F., 2005. Gesture-Based Affective Computing on Motion Capture Data. LNCS 3784:1, Springer Verlag.
Laban R., Lawrence F.C., 1947. Effort. Macdonald&Evans Ltd., London.
Laso-Ballesteros, I., and Daras, P. (Eds.), 2008. User Centric Future Media Internet, EU Commission.
Le Meur, O., Le Callet, P., Barba, D., Thoreau, D., 2006. A Coherent Computational Approach to Model Bottom-Up Visual Attention, IEEE Trans. on PAMI, 28(5): 802-817, 2006
Liu, F., Gleicher, M., 2006. Video Retargeting: Automating Pan-and-Scan. ACM Multimedia.
Parkhurst, D.J., Niebur, E., 2004. Texture contrast attracts overt visual attention in natural scenes. European Journal of Neuroscience, 19:783–789.
Pentland A., 2007. Social signal processing. IEEE Signal Processing Magazine, 24(4):108-111.
Mancas, M., Glowinski, D., Volpe, G., Camurri, A., Bretéché, P., Demeyer, J., Ravet, T., Coletta, P., 2009. Real-Time Motion Attention and Expressive Gesture Interfaces. Journal On Multimodal User Interfaces (JMUI), Springer Verlag, Berlin/Heidelberg.
Marwan, N., Romano, M.C.,Thiel, M., Kurths,J., 2007. Recurrence plots for the analysis of complex systems. Phsysics Reports, 438: 237-329.
McQuiggan S.W., Lester, J.C., 2007. Modeling and evaluating empathy in embodied companion agents. Intl. J. Human-Computer Studies, 65(4):348-360.
Varni, G., Camurri, A., Coletta, P., Volpe, G., 2008. "Emotional entrainment in music performance", in Proc. of 8th Intl. Conf. on Automatic Face and Gesture Recognition FG08.
Varni, G., Camurri, A., Coletta, P., Volpe, G., 2009a. Toward Real-time Automated Measure of Empathy and Dominance. In Proc. 2009 IEEE Intl. Conf. on Social Computing SocialCom, Vancouver, Canada.
Varni, G., Mancini, M., Volpe, G., Camurri, A., 2009b. Sync'n'Move: social interaction based on music and gesture. In Proc. 1st Intl. ICST Conference on User Centric Media, Venice, Italy, December 2009.
Vinciarelli, A., Pantic, M., Bourlard, H. 2009. Social Signal Processing: Survey of an emerging domain. Image and Vision Computing Journal, 27(12).
Wallbott H.G., 1998. Bodily expression of emotion. Eur. J. Soc. Psychol, 28:879–896.
Yee, H., Pattanaik, S., 2001. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. IACM.
Zbilut J., Webber Jr, C.L., 1992. Embeddings and delays as derived from quantification of recurrence plots. Phys. Lett. A, 5:199-203.