Towards Analysis of Expressive Gesture in Groups of Users: Computational Models of Expressive Social Interaction

Antonio Camurri, Giovanna Varni, and Gualtiero Volpe

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

Abstract. In this paper we present a survey of our research on analysis of expressive gesture and how it is evolving towards the analysis of expressive social interaction in groups of users. Social interaction and its expressive implications (e.g., emotional contagion, empathy) is an extremely relevant component for analysis of expressive gesture, since it provides significant information on the context expressive gestures are performed in. However, most of the current systems analyze expressive gestures according to basic emotion categories or simple dimensional approaches. Moreover, almost all of them are intended for a single user, whereas social interaction is often neglected. After briefly recalling our pioneering studies on collaborative robot-human interaction, this paper presents two steps in the direction of novel computational models and techniques for measuring social interaction: (i) the interactive installation *Mappe per Affetti Erranti* for active listening to sound and music content, and (ii) the techniques we developed for explicitly measuring synchronization within a group of users. We conclude with the research challenges we will face in the near future.

Keywords: expressive gesture analysis and processing, analysis of social interaction in small groups, multimodal interactive systems.

1 Introduction

This paper presents a survey of our research on analysis of expressive gesture and how it is evolving towards the analysis of social interaction in (small) groups of users.

Research on *expressive gesture* became particularly relevant in recent years (e.g., see the post-proceedings of Gesture Workshops 2003, 2005, and 2007). Psychological studies have been a fundamental source for 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, the choreographer Rudolf Laban (1947)

S. Kopp and I. Wachsmuth (Eds.): GW 2009, LNAI 5934, pp. 122–133, 2010.

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).

However, most of such systems classify gestures according to basic emotion categories or simple dimensional approaches. Moreover, almost all of the existing systems are intended for a single user, whereas social interaction is neglected.

Nevertheless, social interaction is an extremely relevant component for analysis of expressive gesture, since it provides significant information on the context expressive gestures are performed in. *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. Research in experimental psychology and neurosciences has shown that nonverbal communication, and in particular expressive gesture, is a key aspect of social interaction.

Current research on social interaction, 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., 2009, 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 conveyed by expressive gesture in a social framework, such as empathy and emotional contagion. After recalling our pioneering studies on collaborative robot-human interaction, carried out in the late Nineties, this paper presents two steps in the direction of novel computational models and techniques for measuring social interaction: (i) the interactive installation *Mappe per Affetti Erranti*, where we first investigated social interaction in a small group of users and explored how to use this information in a multi-user multimodal interactive system for active listening of sound and music content, and (ii) the techniques we developed for explicitly measuring synchronization within a group of users. Such techniques are based on the analysis of complex systems, where each user is modelled as a component of a complex system, and which have been applied both in experiments and applications in the framework of the SAME Project.

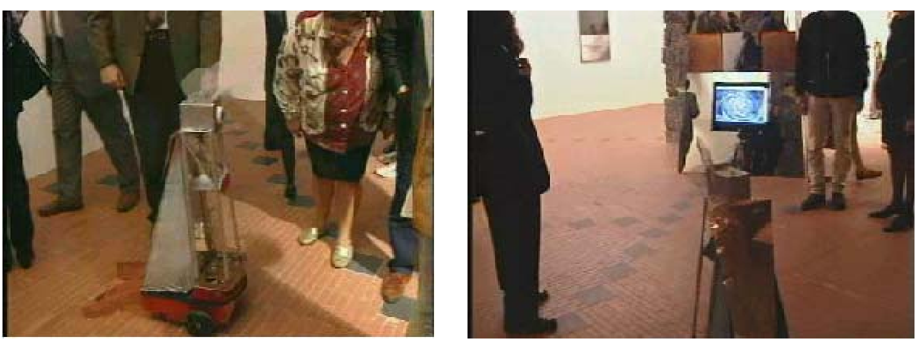
Finally, starting from this experience, we conclude with the research challenges we will face in the near future. These are in the direction of a deeper understanding of the mechanisms underlying phenomena such as empathy and emotional contagion, that are studied by considering ensemble music performance (e.g., a string quartet) as an ideal test-bed for experiments and proof-of-concepts.

2 Pioneering Studies: Collaborative Human-Robot Interaction

Our pioneering studies on analysis of expressive gesture in social interaction date back to the beginning of our Lab in 1984. In particular, in the late Nineties, such studies were carried out in the framework of human-robot interaction (e.g., Camurri et al., 2000). A major goal was (and partially still is) to explore paradigms of expressive social interaction between humans and robots in the framework of multimodal environments in music, theatre, museum exhibitions, and art installations.

In one of such works (Suzuki et al., 1998), we experimented with a small mobile robot on wheels (a Pioneer 1 from Stanford Research Institute) as a semi-autonomous agent capable of communicating with the visitors of a museum exhibit on contemporary art (*Arti Visive 2*, Palazzo Ducale, Genova, October 1998) by means of several channels, including sound and music, visual media, and style of movement (e.g., fast/slow, smooth/nervous, tail-wagging). The robot wandered in the museum exhibit as one of the visitors, a sort of medium between humans and machines inhabiting the exhibit area. Sensors allowed it to avoid collisions with the surrounding people and to observe artworks and visitors. The goal was to generate a musical and visual feedback for visitors and to interact with them (see Figure 1a). Feedback depended on a model of artificial emotions, simulating the internal emotional state of the robot. A basic idea for the visual output was to take the images the robot captured and to transform them in real time by "virtual mirrors" reflecting the internal state of the robot. For example, if the robot could not follow its path – one of its goal was to visit the exhibit – a mirror deformation in a sort of pulsing spiral emerged (see Figure 1b), accompanied by faster rhythmic music textures. More details on the models and the architecture we used can be found in (Suzuki et al., 1998; Camurri and Coglio, 1998, Camurri and Ferrentino, 1999).

In another experimental set-up, we developed a "theatrical machine" for the performance of the music piece *Spiral*, by K. Stockhausen, for one singer or player (in this performance, one trombone) endowed with a short wave radio. The radio, audio amplifier, and loudspeakers were installed on board of a robot navigating on the stage, thus creating effects of physical spatialization of sound due to the movement of the



(a) (b)

Fig. 1. Social interaction between a robot and a group of visitors at the art exhibition *Arti Visive 2* (Genova, Italy, October 1998): (a) the robot interacting with visitors, (b) an example of visual output.

(a) (b)

Fig. 2. Collaborative human-robot interaction in the framework of the multimedia performance *L'Ala dei Sensi* (Ferrara, Italy, November 1998): (a) robot/dancer interaction, (b) an example of visual feedback

robot during the performance (trombone: Michele Lo Muto, live electronics: Giovanni Cospito, Civica Scuola di Musica, Sezione Musica Contemporanea, Milano, June 1996). The movements of the robot were influenced by sound parameters and by the gesture of the trombone performer: for example, a high "energy" content by the trombonist's gesture and a high sound spectral energy were stimuli for the robot to move away from the performer. Smooth and calm phrasing and movements were stimuli attracting the robot near and around the performer. Further, the robot sound and music output were part of the interaction process, i.e., the expressive gesture nuances and the sound produced by the performer influenced the robot, and vice-versa.

For *L'Ala dei Sensi* (see Figures 2a and 2b), a multimedia performance about human perception (Director: Ezio Cuoghi; Choreographer and dancer: Virgilio Sieni, Ferrara, Italy, November 1998), we contributed two episodes involving interactive dance/music performance, making use of a small mobile robotic platform (a Pioneer 2 from Stanford Research Institute). The robot was equipped with sensors, an on-board video projector and a video-camera. Sensors allowed the robot to avoid collisions with the scenery and the dancers. In the main episode, initially the on-board video projector and the video-camera were directly controlled in real-time by the director (off-stage). He also used the images coming from the robot (the robot's point of view) to mix them in real time on a large screen. The director controlled the movements of the robot too. That is, the robot was a sort of passive companion of the dancer. At a certain point, the dancer plugged off the electric power cable of the robot. This was a specially important gesture: the robot came to life and a deeper dialogue with the dancer started. The dancer was equipped with two sensors on the palms of the hands. By acting on the first one, he was allowed to influence the robot towards one between two different styles of movement: a kind of "ordered" movement (aiming at a direct, constant speed movement) and a "disordered" type of movement. Through the second sensor the movement could be stopped and restarted. The dancer was also observed by a video-camera and his expressive gestures were a further stimulus for the robot, which was able to react by changing (morphing) its style of moving.

3 The Interactive Collaborative Installation *Mappe Per Affetti Erranti*

More recently, we studied expressive gesture in social interaction in the framework of collaborative and active listening to sound and music content. With *active listening* we mean that listeners are enabled to interactively operate on (pre-recorded) music content through their movement and gesture, by modifying and moulding it in real-time while listening. Active listening is investigated in the framework of the EU-ICT Project SAME (www.sameproject.eu). *Mappe per Affetti Erranti* is a collaborative multi-user multimodal interactive system, where four users collectively collaborate in the active listening to a music piece, by their expressive movement and gesture.

Mappe per Affetti Erranti – literally "maps for wandering affects" – (Camurri et al., 2008) reworks and extends the concept of navigation and exploration of a virtual orchestra by introducing multiple levels: from the navigation in a physical space populated by virtual objects or subjects up to the virtual navigation in emotional spaces populated by different expressive performances of the same music piece. Users can navigate such emotional spaces by their expressive movement and gesture. Only social interaction and collaboration leads to a correct reconstruction of the music piece. In other words, while users explore the physical space, the (expressive) way in which they move and the degree of collaboration between them allow them to explore at the same time an emotional space, by means of transformations in real-time of the interpretation of the music content.

At the physical level, space is divided in several areas. A voice of a polyphonic music piece is associated to each area. The music content is coded as multiple independent audio channels. The presence of a user (even a single user, who can vary from a non-expert user to a dancer) triggers the reproduction of the music piece. By exploring the space, the user walks through several areas and listens to the single voices separately. If the user stays in a single area, she listens to the voice associated to that area only. If she does not move for a given time interval, music fades out and turns off. The user can mould the voice she is listening to in several ways. At a low level, she can intervene on parameters such as loudness, density, amount of reverberation, and spatial position. At a higher level, she can intervene on the expressive features of the music performance, i.e., change the interpretation. This is done through the navigation of an emotional space. The system analyzes the expressive intention the user conveys with her expressive gesture and translates it in a position (or a trajectory) in an emotional space. Like the physical space, such emotional space is also divided in several areas, each corresponding to a different performance of the same voice with a different expressive intention. Several examples of such emotional spaces are available in the literature, for example the spaces used in dimensional theories of emotion (e.g., Russell, 1980; Tellegen et al., 1999) or those especially developed for analysis and synthesis of expressive music performance (e.g., Juslin, 2000; Vines et al., 2005). Users can thus explore the music piece in a twofold perspective: navigating the physical space they explore the polyphonic musical structure; navigating the emotional space they explore music interpretation. A single user, however, can only listen to and intervene on a single voice at time: she cannot listen to the whole polyphonic piece.

Only a group of users can fully experience *Mappe per Affetti Erranti*. In particular, the music piece can be listened to in its whole polyphony only if a number of users at least equal to the number of voices is interacting with the installation. Each user controls the interpretation of the performance of the voice associated to the area she occupies: e.g., a shy, hesitant behavior causes the corresponding voice interpretation to be whispering, soft; if a user is moving aggressively and with sudden impulsive gesture, the corresponding voice becomes interpreted correspondingly (louder, with sudden peaks, etc.). The whole piece is performed coherently, that is with the same expressive intention, only if all the users are moving with the same expressive intention. Thus, the more users move with different, conflicting expressive intentions, the more the musical output is incoherent and chaotic. But the more users move with similar expressive intentions and in a collaborative way, the more the musical output is coherent and the music piece is listened to in one of its different expressive interpretations.

In the current instance of the system of *Mappe per Affetti Erranti* (Figure 3), the physical map is composed by four areas. Tenore and soprano voices are associated to the central areas and contralto and basso to the lateral ones. Four expressive performances of the same music piece are available: Happy/Joyful, Solemn, Intimate/Shy, and Angry/Aggressive. These are associated to the same four expressive intentions classified from users expressive gestures. Analysis of expressive gesture is performed by means of twelve expressive features: Quantity of Motion, 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 descriptors are computed in real-time for each user. Further descriptors are computed also on the whole group of users: e.g., the contraction/expansion of the group and its cohesion. This perspective corresponds to Rudolf Laban's *General Space* (Laban, 1947). Classification is performed following a fuzzy-logic like approach. Such approach has the advantage that it does not need a training set of recorded movement and it is also flexible enough to be applied to the movement of different kinds of users (e.g., adults, children, elder people).

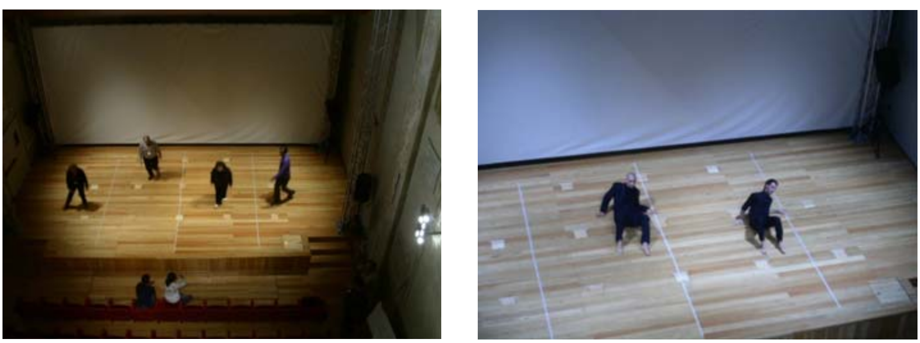


Fig. 3. *Mappe per Affetti Erranti* experienced by a group of four users (left) and by two dancers during a dance performance (right)

In *Mappe per Affetti Erranti*, synchronization has a particular importance. Only synchronization in the expressive intentions conveyed by the users allows the correct reconstruction of the music piece.

Techniques for the real-time management of the synchronization of the four musical voices have been developed. At the audio level, each audio file is manually or semi-automatically segmented in phrases and sub-phrases. Changes in the expressive intention detected from movement trigger a switch to the corresponding audio file at a position, which is coherent to the position reached by that expressive interpretation as a result of the movement of the users. In such a way, we obtain a continuous resynchronization of the single voices depending on the expressive intentions conveyed by the users.

The system analyzes the expressive gesture of each single user. Synchronization emerges as the result of the collective behavior of the users. As for the audio processing and real-time synchronization of the several voices, we developed a series of techniques to align the audio channels following the phrasing, the changing of interpretation of the recorded voices, and the mutual behavior of the users. For example, when a user imitates the expressive behavior of another, her associated voice changes to the same interpretation as the other one (the leader, at this moment): the two voices then synchronize, by aligning in real-time the second voice to the leader's voice, keeping into account the phrasing for having naturalness and continuity in the perceived audio. The synchronization is therefore modeled at a "sub-phrasing" level. A sudden, musically unrelated synchronization of two voices just at the time instant in which the system detects a change of behavior would be perceived unnatural. Let us consider the two-user example mentioned above. The first is moving joyful/happy and the second shy/intimate/hesitant, with the corresponding rendering of the music interpretations joyful and soft/whispering, respectively. The music results not synchronized: an intimate performance is usually slower than a joyful one. If the second user changes her behavior from shy to joyful, her corresponding voice, once concluded the current musical sub-phrase, will shift from the shy to the joyful audio file, and the time instant where it continues to play the audio will be the same of the other (leading) voice. This synchronization driven by the expressive gesture and by the semantics of the musical content causes a short delay (a few hundreds of ms, corresponding to sub-phrase average duration), but proved to be the best strategy: it is perceived as natural by users, and it also follows the natural motor task strategy that would be adopted by a singer to manage sudden changes in interpretation (the most natural way is to change interpretation at the starting of the next sub-phrase, allowing a preparation and re-programming of the next motor task of singing). We considered other synchronization strategies, e.g., time-warping of voice music signals, but these demonstrated to be unnatural both from the singer and listener perspective. Further, they destroy the naturalness of expressive interpretation, and are computationally much more intensive.

The next step, still missing in *Mappe per Affetti Erranti*, is to explicitly measure the synchronization of the group.

4 Measuring Synchronization in a Group of Users

Starting from the results obtained in *Mappe per Affetti Erranti*, the focus of our research is shifted to the development of methodological frameworks, computational models, and algorithms for the analysis of creative communication within groups of users in terms of synchronization.

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 behaviour, 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 from 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 all 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, coordinates 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, multimodal 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 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 (Poincaré, 1890), 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 give 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, 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 activities that human beings engage in: it is believed that this

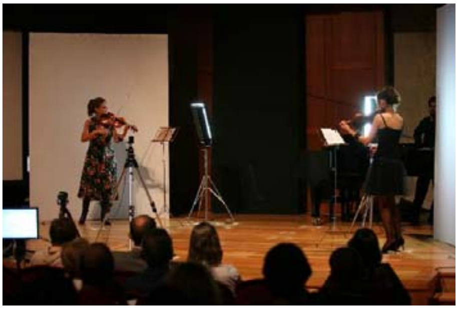


Fig. 4. Experiments on joint music performance: (a) music duo performance in which the two musicians can communicate, also exchanging visual information, (b) the famous string quartet Quartetto di Cremona during the experiment. Each musician wears a white hat including a green passive marker and a 3-axis accelerometer, plus a 3-axis accelerometer on the back, and physiological sensors for heart rate, breath, ocular movements, and face muscles.

ability from individuals and groups to entrain to music is unique only to humans and that, unlike speech, musical 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 on 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, University of Belfast) and Carol Krumhansl (Cornell University) 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 colleagues, we performed multimodal synchronized recordings of the Quartetto di Cremona. Figure 4 shows snapshots from the experiments.

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. Furthermore, measures of the direction of PS confirmed the hypothesis on egalitarian distribution of dominance in a duo performance . Further, 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; while in cases of low synchronization, i.e., poor collaborative interaction, the 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 gesture and how it is evolving towards the analysis of social interaction in (small) groups of users. However, to date, research on measuring quantitative and qualitative interpersonal communication in groups and on supporting models and tools is still a broadly unexplored field. In this scenario, our future research directions 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. Further, another recent research project concerns the social active experience of audiovisual content, in the framework of museum and cultural projects: it is an extension of the concept of active music listening, and we are exploring directions to enhance the visiting 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, Donald Glowinski, Maurizio Mancini, Alberto Massari, and Barbara Mazzarino for their important contributes to research. We are grateful to Ipke Wachsmuth and Stefan Kopp for the precious suggestions.

References

Bernhardt, D., Robinson, P.: Detecting Affect from Non-stylised Body Motions. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 59–70. Springer, Heidelberg (2007)

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

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

Camurri, A., Coglio, A.: An Architecture for Emotional Agents. In: IEEE Multimedia, October-December 1998, pp. 24–33. IEEE CS Press, Los Alamitos (1998)

Camurri, A., Ferrentino, P.: Interactive environments for music and multimedia. Multimedia Systems 7, 32–47 (1999)

Camurri, A., Coletta, P., Ricchetti, M., Volpe, G.: Expressiveness and Physicality in Interaction. Journal of New Music Research 29(3), 187–198 (2000)

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

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

Camurri, A., Canepa, C., Coletta, P., Ferrari, N., Mazzarino, B., Volpe, G.: 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), Athens, Greece, pp. 376–383 (2008)

De Meijer, M.: The contribution of general features of body movement to the attribution of emotions. J. of Nonverbal Behavior 13, 247–268 (1989)

De Rosis, F., Cavalluzzi, A., Mazzotta, I., Novielli, N.: Can embodied conversational agents induce empathy in users? In: Proceedings of AISB 2005, Virtual Social Character Symposium (2005)

Dong, W., Lepri, B., Cappelletti, A., Pentland, A., Pianesi, F., Zancanaro, M.: Using the influence model to recognize functional roles in meetings. In: Proc. 9th Intl. ACM Conf. on Multimodal Interfaces, pp. 271–278 (2007)

Eckmann, J.P., Kamphorst, S.O., Ruelle, D.: Recurrence plots of dynamical system. Eurpph. Lett. 5, 973–977 (1987)

Juslin, P.N.: Cue utilization in communication of emotion in music performance: relating performance to perception. Journal of Experimental Psychology: Human Perception and Performance 26(6), 1797–1813 (2000)

Kapur, A., Kapur, A., Virji-Babul, N., Tzanetakis, G., Driessen, P.F.: Gesture-Based Affective Computing on Motion Capture Data. In: Tao, J., Tan, T., Picard, R.W. (eds.) ACII 2005. LNCS, vol. 3784, pp. 1–7. Springer, Heidelberg (2005)

Laban, R., Lawrence, F.C.: Effort. Macdonald&Evans Ltd., London (1947)

Marwan, N., Romano, M.C., Thiel, M., Kurths, J.: Recurrence plots for the analysis of complex systems. Phsysics Reports 438, 237–329 (2007)

McQuiggan, S.W., Lester, J.C.: Modeling and evaluating empathy in embodied companion agents. Intl. J. Human-Computer Studies 65(4), 348–360 (2007)

Pentland, A.: Social signal processing. IEEE Signal Processing Magazine 24(4), 108–111 (2007)

Poincaré, H.: Sur la problèmedes trois corps et les equations de la dynamique. Acta Math 13, 1–271 (1890)

Russell, J.A.: A circumplex model of affect. J. of Personality and Social Psychology 39, 1161–1178 (1980)

Suzuki, K., Camurri, A., Hashimoto, S., Ferrentino, P.: Intelligent Agent System for Human-Robot Interaction through Artificial Emotion. In: Proc. IEEE Intl. Conf. on Systems Man and Cybernetics SMC 1998, San Diego, USA (1998)

Tellegen, A., Watson, D., Clark, L.A.: On the dimensional and hierarchical structure of affect. Psychological Science 10(4), 297–303 (1999)

Varni, G., Camurri, A., Coletta, P., Volpe, G.: Emotional entrainment in music performance. In: Proc. of 8th Intl. Conf. on Automatic Face and Gesture Recognition FG 2008 (2008)

Varni, G., Camurri, A., Coletta, P., Volpe, G.: Toward Real-time Automated Measure of Empathy and Dominance. In: Proc. 2009 IEEE Intl. Conf. on Social Computing SocialCom, Vancouver, Canada (2009a)

Varni, G., Mancini, M., Volpe, G., Camurri, A.: Sync'n'Move: social interaction based on music and gesture. In: Proc. 1st Intl. ICST Conference on User Centric Media, Venice, Italy, December 2009 (2009b)

Vines, B.W., Krumhansl, C.L., Wanderley, M.M., Ioana, M.D., Levitin, D.J.: Dimensions of Emotion in Expressive Musical Performance. Ann. N.Y. Acad. Sci. 1060, 462–466 (2005)

Wallbott, H.G.: Bodily expression of emotion. Eur. J. Soc. Psychol. 28, 879–896 (1998)

Zbilut, J., Webber Jr., C.L.: Embeddings and delays as derived from quantification of recurrence plots. Phys. Lett. A 5, 199–203 (1992)