Mappe per Affetti Erranti: a Multimodal System for Social Active Listening and Expressive Performance

Antonio Camurri, Corrado Canepa, Paolo Coletta, Barbara Mazzarino, Gualtiero Volpe

InfoMus Lab – Casa Paganini
DIST- University of Genova
Piazza Santa Maria in Passione 34
16123 Genova, Italy
+39 010 2758252

{toni, corrado, colettap, bunny, volpe}@infomus.org

ABSTRACT

This paper presents our new system *Mappe per Affetti Erranti* (literally Maps for Wandering Affects), enabling a novel paradigm for social active experience and dynamic molding of expressive content of a music piece. Mappe per Affetti Erranti allows multiple users to interact with the music piece at several levels. On the one hand, multiple users can physically navigate a polyphonic music piece, actively exploring it; on the other hand they can intervene on the music performance modifying and molding its expressive content in real-time through full-body movement and gesture. An implementation of Mappe per Affetti Erranti was presented in the framework of the science exhibition "Metamorfosi del Senso", held at Casa Paganini, Genova, in October – November 2007. In that occasion Mappe per Affetti Erranti was also used for a contemporary dance performance. The research topics addressed in Mappe per Affetti Erranti are currently investigated in the new EU-ICT Project SAME (Sound and Music for Everyone, Everyday, Everywhere, Every Way, www.sameproject.eu).

Keywords

Active listening of music, expressive interfaces, full-body motion analysis and expressive gesture processing, multimodal interactive systems for music and performing arts applications, collaborative environments, social interaction.

1. INTRODUCTION

Music making and listening are a clear example of a human activity that is above all interactive and social. However, nowadays mediated music making and listening is usually still a passive, non–interactive, and non-context sensitive experience. The current electronic technologies, with all their potential for interactivity and communication, have not yet been able to support and promote this essential aspect of music making and listening. This can be considered a degradation of traditional listening experience, in which the public can interact in many ways with performers to modify the expressive features of a piece.

The need of recovering such active attitude with respect to music is strongly emerging and novel paradigms of active

experience are going to be developed. With *active experience* and *active listening* we mean that listeners are enabled to interactively operate on music content, by modifying and molding it in real-time while listening. Active listening is the basic concept for a novel generation of interactive music systems [1], which are particularly addressed to a public of beginners, naïve and inexperienced users, rather than to professional musicians and composers.

Active listening is also a major focus for the new EU-ICT Project SAME (Sound and Music for Everyone, Everyday, Everywhere, Every Way, www.sameproject.eu). SAME aims at: (i) defining and developing an innovative networked end-to-end research platform for novel mobile music applications, allowing new forms of participative, experience-centric, context-aware, social, shared, active listening of music; (ii) investigating and implementing novel paradigms for natural, expressive/emotional multimodal interfaces, empowering the user to influence, interact, mould, and shape the music content, by intervening actively and physically into the experience; and (iii) developing new mobile context-aware music applications, starting from the active listening paradigm, which will bring back the social and interactive aspects of music to our information technology age.

In the direction of defining novel active listening paradigms, we recently developed a system, the Orchestra Explorer [2], allowing users to physically navigate inside a virtual orchestra, to actively explore the music piece the orchestra is playing, to modify and mold in real-time the music performance through expressive full-body movement and gesture. By walking and moving on the surface, the user discovers each single instrument and can operate through her expressive gestures on the music piece the instrument is playing. The interaction paradigm developed in the Orchestra Explorer is strongly based on the concept of navigation in a physical space where the orchestra instruments are placed. The Orchestra Explorer is intended for fruition by a single user.

Our novel multimodal system for social active listening *Mappe per Affetti Erranti* starts from the Orchestra Explorer and the lessons learned in over one year of permanent installation of the Orchestra Explorer at our site at Casa Paganini, and several installations of the Orchestra Explorer at science exhibitions and public events.

Mappe per Affetti Erranti extends and enhances the Orchestra Explorer in two major directions. On the one hand it reworks and extends the concept of navigation by introducing multiple levels: from the navigation in a physical space populated by virtual objects or subjects (as it is in the Orchestra Explorer) up to the navigation in virtual emotional spaces populated by

different expressive performances of the same music piece. Users can navigate such affective spaces by their expressive movement and gesture. On the other hand, Mappe per Affetti Erranti explicitly addresses fruition by multiple users and encourages collaborative behavior: only social collaboration allows 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 affective, emotional space.

Section 2 presents the concept of Mappe per Affetti Erranti; Section 3 focuses on the specific aspect of expressive movement analysis and describes the model we designed for navigating the affective space; Section 4 and 5 illustrate the implementation of an installation of Mappe per Affetti Erranti developed for the science exhibit "Metamorfosi del Senso" (Casa Paganini, Genova, Italy, October 25 – November 6, 2007). Conclusions summarize some issues and future work that emerged from such installation.

2. CONCEPT

The basic concept of Mappe per Affetti Erranti is the collaborative active listening of a music piece through the navigation of maps at multiple levels, from the physical level to the emotional level.

At the physical level space is divided in several areas. A voice of a polyphonic music piece is associated to each area. The presence of a user (even a single user) 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 users stays in a single area, she listens to the voice associated to that area only. If the user 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. For example, by opening her arms, the user can increase the density of the voice (she listens to the two or more voices in unison). If she moves toward the back of the stage the amount of reverberation increases, whereas toward the front of the stage the voice becomes drier.

At a higher level the user can intervene on the expressive features of the music performance. This is done through the navigation of an emotional, affective space. The system analyzes the expressive intention the user conveys with her expressive movement and gesture and translates it in a position (or a trajectory) in an affective, emotional space. Like the physical space, such affective, emotional space is also divided in several areas, each one corresponding to a different performance of the same voice with a different expressive intention. Several examples of such affective, emotional spaces are available in the literature, for example the spaces used in dimensional theories of emotion (e.g., see [3][4]) or those especially developed for analysis and synthesis of expressive music performance (e.g., see [5][6][7]).

Users can thus explore the music piece in a twofold perspective: navigating the physical space they explore the polyphonic musical structure; navigating the affective, emotional space they explore music performance. A single user, however, can only listen to and intervene on a single voice at time: she cannot listen to the whole polyphonic piece with all the voices.

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. Moreover, since each user controls the performance of the voice associated to the area she occupies, the whole piece is performed 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 pieces is listened to in one of its different expressive performances.

Mappe per Affetti Erranti can therefore be experienced at several levels: by a single user who has a limited but still powerful set of possibilities of interaction, by a group of users who can fully experience the installation, by multiple groups of users. In fact, each physical area can be occupied by a group of users. In this case each single group is analyzed and each participant in a group contributes to intervene on the voice associated to the area the group is occupying. Therefore, at this level a collaborative behavior is encouraged among the participants in each single group and among the groups participating in the installation.

The possibility of observing a group or multiple groups of users during their interaction with Mappe per Affetti Erranti makes this installation an ideal test-bed for investigating and experimenting group dynamics and social network scenarios.

3. EXPRESSIVE MOVEMENT ANALYIS

This section focuses on a specific and relevant aspect of Mappe per Affetti Erranti, i.e., how the system analyses the expressive intentions conveyed by a user through her expressive movement and gesture. Such information is used for navigating the affective, emotional space and for controlling the expressive performance of a voice in the polyphonic music piece.

Expressive movement analysis is discussed with reference to an implementation of Mappe per Affetti Erranti we recently developed (see Section 4). In such implementation we selected four different expressive intentions: the first one refers to a happy, joyful behavior, the second one to solemnity, the third one to a intimate, introverted, shy behavior, the fourth to anger. In order to make description easier we will label such expressive intentions as Happy, Solemn, Intimate, Angry. Please note, however, that we consider the reduction to such labels as a too simplistic way of describing very subtle nuances of both movement and music performance. In fact, we never described Mappe per Affetti Erranti to users in terms of such labels. Rather, we provided (when needed) more articulated descriptions of the kind of expressive behavior we (and the system) expected and we let users to discover themselves the installation step by step.

Such four expressive intentions were select since they are different and characterized enough to be easily conveyed and recognized by users. Furthermore, they are examples of low/high positive/negative affective states that can be easily mapped on existing dimensional theories of emotion (e.g., valence-arousal or Tellegen's space).

3.1 Feature extraction

In our current implementation, analysis of expressive gesture is performed by means of twelve expressive descriptors: 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, Space Allure,

Amount of Periodic Movement, Symmetry Index. Such descriptors are computed in real-time for each user. Most of descriptors are computed on a time window of 3 s. In the context of Mappe per Affetti Erranti, we considered such time interval as a good trade-off between the need on the one hand of having an enough responsive system, and the need on the other hand to give the users a time long enough for displaying an expressive intention.

Quantity of Motion (QoM) provides an estimation of the amount of overall movement (variation of pixels) the videocamera detects [8]. Quantity of Motion computed on translational movement only (TQoM) provides an estimation of how much the user is moving around the physical space. Using Rudolf Laban's terminology [9][10], whereas Quantity of Motion measures the amount of detected movement in both the Kinesphere and the General Space, its computation on translational movements refers to the overall detected movement in the General Space only. TQoM, together with speed of barycentre (BS) and variation of the Contraction Index (dCI) are introduced to distinguish between the movement of the body in the General Space and the movement of the limbs in the Kinesphere. Intuitively, if the user moves her limbs but does not change her position in the space, TQoM and BS will have low values, while QoM and dCI will have higher values.

Impulsiveness (IM) is measured as the variance of Quantity of Motion in a time window of 3 s, i.e., a user is considered to move in an impulsive way if the amount of movement the videocamera can detect on her changes considerably in the time window.

Vertical and horizontal components of velocity of peripheral upper parts of the body (VV, HV) are computed starting from the positions of the upper vertexes of the body bounding rectangle. The vertical component, in particular, is used for detecting upward movements that psychologists (e.g., Boone and Cunningham [11]) identified as significant indicator of positive emotional expression.

Space Occupation Area (SOA) is computed starting from the movement trajectory integrated over time. In such a way a bitmap is obtained, summarizing the trajectory followed along the considered time window (3 s). An elliptical approximation of the shape of the trajectory is then computed [12]. The area of such ellipse is taken as the Space Occupation Area of the movement trajectory. Intuitively, a trajectory spread over the whole gets high SOA values, whereas a trajectory confined in a small region gets low SOA values.

Directness Index (DI) is computed as the ratio between the length of the straight line connecting the first and last point of a given trajectory (in this case the movement trajectory in the selected 3 s time window) and the sum of the lengths of each segment composing the trajectory.

Space Allure (SA) measures local deviations from the straight line trajectory. Whereas DI provides information about whether the trajectory followed along the 3 s time window is direct or flexible, SA refers to waving movements around the straight trajectory in shorter time windows. In the current implementation SA is approximated with the variance of the DI in a time window of 1 s.

The Amount of Periodic Movement (PM) gives a preliminary information about the presence of rhythmic movements. Computation of PM starts from QoM. Movement is segmented in motion and pause phases using a threshold on QoM [13]; inter-onset intervals are then computed as the time elapsing from the beginning of a motion phase and the beginning of the

following motion phase. The variance of such inter-onset intervals is taken as an approximate measure of PM.

Symmetry Index (SI) is computed from the position of the barycenter and the left and right edges of the body bounding rectangle. That is, it is the ratio between the difference of the distances of the barycenter from the left and right edges and the width of the bounding rectangle:

$$SI = \frac{||x_B - x_L| - |x_B - x_R||}{|x_R - x_L|}$$

where x_B is the x coordinate of the barycentre, x_L is the x coordinate of the left edge of the body bounding rectangle and x_R is x coordinate of the left edge of the body bounding rectangle.

3.2 Classification

In order to classify movement with respect to the four expressive intentions, the values of the descriptors are quantized in five levels: Very Low, Low, Medium, High, Very High.

Starting from previous works by the authors (e.g., [8][13][14]) and results of psychological studies (e.g., [11][15][16][17]), we characterized each expressive intention with a combination of levels for each descriptor. The Happy intention is characterized by high energy, upward and fluent movement; almost constant values in the kinematical descriptor with periodical peaks are associated with Solemn behavior; low energy, contracted, and localized movements are typical of a shy behavior; impulsiveness, high energy and rigid movement are associated to Anger. Table 1 summarizes such characterization.

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 also is flexible enough to be applied to the movement of different kinds of users (e.g., adults, children, elder people).

The expressive intention EI referred to a time window covering the last 3 seconds of movement is computed as:

$$EI = \arg\max_h \sum_{k=1}^{N} w_k f_{hk}(x_k) \quad \text{with } h = 1..M$$

where M = 4 is the number of expressive intentions, N = 12 is the number of motion descriptors, w_k is the weight of the k-th motion descriptor, x_k is the value of the k-th motion descriptor and f_{hk} is a function applied to the k-th motion descriptor. The function being applied depends on the expected level for such motion descriptor if the h-th expressive intention were detected. In this first implementation weights have been selected empirically while developing and testing the system.

A Gaussian function is applied for motion descriptors whose values are expected to be High, Medium, or Low:

$$f_{hk}(x_k) = A \, e^{-\frac{(x_k - \mu_{hk})^2}{\sigma_{hk}^2}}$$

where A is the amplitude (set to 1), μ_{hk} is the expected value for descriptor k, when the user shows expressive intention h and the variance σ_{hk}^2 is used for tuning the range of values for which the descriptor can be considered to be at the appropriate level (High, Medium, Low).

A sigmoid is applied for Very High or Very Low descriptors:

$$f_{hk}(x_k) = \frac{1}{1 + t_{hk} e^{\alpha_{hk}(x_k - \beta_{hk})}}$$

where β_{hk} is used for tuning the range of values for which the descriptor can be considered to be at the appropriate level (Very High or Very Low), α_{hk} controls the steepness of the sigmoid and t_{hk} controls the type of sigmoid, i.e., $t_{hk} = 1$ if the descriptor is expected to be Very Low and $t_{hk} = -1$ if the descriptor is expected to be Very High (inverse sigmoid tending to 1 for high values).

Intuitively, the output of the Gaussian and sigmoid functions applied to motion descriptors is a measure of how much the actual value of a motion descriptor is near to the value expected for a given expressive intention. For example, if a motion descriptor is expected to be Low for a given expressive intention and its expected value is 0.4, a Gaussian is placed with its peak (normalized to 1) centered in 0.4. That motion descriptor will therefore provide the highest contribution to the overall sum if the real value is in fact the expected value. As a consequence the highest value for the sum is obtained by that expressive intention whose expected values for descriptors according to Table 1 best match the actual computed values.

Table 1. Expected levels of each motion descriptor for the four expressive intentions

Motion descriptor	Happy	Solemn	Intimate	Angry
QoM	High	Low	Low	High
TQoM	High	Low	Low	High
IM	Medium	Low	Very low	Very high
VV	High	Low	Low	Medium
HV	High	Medium	Low	High
BS	Not relevant	Not relevant	Low	Medium
dCI	Medium	Low	Low	Very high
SOA	Not relevant	Not relevant	Low	High
DI	Medium	High	Low	Low
SA	Low	Low	Medium	Low
PM	High	Very high	Low	Very low
SI	Medium	Medium	Not relevant	Low

4. THE INSTALLATION AT THE SCIENCE EXHIBITION "METAMORFOSI DEL SENSO"

Mappe per Affetti Erranti was presented the first time at the science exhibition "Metamorfosi del Senso", held at Casa Paganini, Genova, Italy, on October 25th – November 6th, 2007. The exhibition was part of "Festival della Scienza", a huge international science festival held in Genova every year.

Mappe per Affetti Erranti was installed on the stage of the 250-seats auditorium at Casa Paganini, an international center of excellence for research on sound, music, and new media, where InfoMus Lab has its main site. The installation covered a surface of about 9 m × 3.5 m. A single vidocamera observed the whole surface from the top, about 7 m high, and at a distance of about 10 m from the stage (we did not use sensors or additional videocameras in this first experience). Four loudspeakers were placed at the four corners of the stage for audio output. A white screen covered the back of the stage for the whole 9 m width: this was used as scenery since the current implementation of the installation does not include video feedback. Lights were set in order to enhance the feeling of immersion for the users and to have a homogenous lighting of the stage.

The music piece we selected is "Come again" by John Dowland for four singing voices: contralto, tenore, soprano, and basso. With the help of singer Roberto Tiranti and composer Marco Canepa we chose a piece that could be soundly interpreted with different expressive intentions (i.e., without becoming ridiculous) and could result interesting and agreeable for non expert users. We asked professional singers to sing it with the four different expressive intentions Happy, Solemn, Intimate, and Angry. The piece was performed so that changes in the interpretation could be perceived even by non-expert users.

The physical map is composed by four rectangular, parallel areas on the stage. Tenore and soprano voices are associated with the central areas, contralto and basso to the lateral ones. This allows an alternation of female and male voices and attracts users toward the "stronger" voices, i.e., the central ones.

Navigation in the affective, emotional space is obtained with the techniques for expressive movement analysis and classification discussed in Section 3. As for music performance, each recorded file was manually segmented in phrases and sub-phrases. Changes in the expressive intention detected from movement triggers 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 other users/groups. In such a way we obtain a continuous resynchronization of the single voices depending on the expressive intentions conveyed by users.

In occasion of the opening of "Metamorfosi del Senso", choreographer Giovanni Di Cicco and his dance ensemble designed and performed a contemporary dance performance on Mappe per Affetti Erranti. In such a performance, dancers interacted with the installation for over 20 min, repeatedly moving from order to chaos. The public of the dance performance counted more than 400 persons in 3 days. Figure 1 shows a moment of the dance performance and a group of users experiencing Mappe per Affetti Erranti. The installation was experienced by more than 1500 persons during "Metamorforsi del Senso", with general positive and sometimes enthusiastic feedback.

5. IMPLEMENTATION: THE EYESWEB XMI OPEN PLATFORM AND THE EYESWEB EXPRESSIVE GESTURE PROCESSING LIBRARY

The instance of Mappe per Affetti Erranti we developed for the exhibit "Metamorfosi del Senso" was implemented using a new version of our EyesWeb open platform [13][18]: EyesWeb XMI (for eXtended Multimodal Interaction). The EyesWeb open platform and related libraries are available for free on the EyesWeb website www.eyesweb.org.

With respect to its predecessors, EyesWeb XMI strongly enhances support to analysis and processing of synchronized streams at different sampling rates (e.g., audio, video, data from sensors). We exploited such support for the synchronized processing and reproduction of the audio tracks in "Come Again". The whole installation was implemented as a couple of

EyesWeb application (patch): the first one managing video processing, extraction of expressive features from movement and gestures, navigation in the physical and affective space; the second one devoted to audio processing, real-time audio mixing, and control of audio effects. Every single component of the two applications was implemented as an EyesWeb sub-patch. The two applications ran on two workstation (Dell Precision 380, equipped with two CPUs Pentium 4 3.20 GHz, 1 GB RAM, Windows XP Professional) with fast network connection.

Extraction of expressive descriptors and models for navigating physical and expressive spaces were implement as EyesWeb modules (blocks) in a new version of the EyesWeb Expressive Gesture Processing Library.

Figure 1. Mappe per Affetti Erranti: on the top a snapshot from the dance performance; in the bottom a group of user interacting with the installation.

6. CONCLUSIONS

From our experience with Mappe per Affetti Erranti, especially at the science exhibit "Metamorfosi del Senso", several issues emerged that need to be taken into account in future work.

A first issue is related to the expressive movement descriptors and the modalities of fruition of Mappe per Affetti Erranti. The installation can be experienced by a single user, by a group, or by multiple groups. However, the expressive descriptors have been defined and developed for analyzing movement and expressive intention of single users. To what extent can they be applied to groups of users? Can we approximate the expressive

intention of a group as a kind of average of the expressive intentions conveyed by its components or more complex group dynamics have to be taken into account? Research on computational models of emotion, affective computing, and expressive gesture processing usually focus on the expressive content communicated by single users. Such group dynamics and their relationships with emotional expression are still largely uninvestigated.

Another issue concerns the robustness of the selected expressive movement descriptors with respect to different analysis contexts. For example, the kind of motion a user performs when she stays inside an area in the space is often different under several aspects from the motion she performs when wandering around the whole space. Motion inside an area is characterized by movement of limbs, the amount of energy is mainly due to how much limbs move, the expressive intention is conveyed through movement in the Kinesphere. Walking is instead the main action characterizing motion around the space, the amount of energy of walking is much higher than the amount of energy associated with possible other movements of arms, the expressive intention is conveyed through the walking style. The system should be able to adapt to such different analysis contexts and different sets of motion descriptors should be developed either specifically for a given context or robust to different contexts.

Future work will also include refinements to the classifier and formal evaluation with users. As for the classifier, it encompasses many parameters (e.g., weights, parameters of the functions applied to the movement descriptors) that need to be fine tuned. In this first installation such parameters have been set empirically during tests with dancers and potential users. However, a deeper investigation based on rigorous experiments would be needed in order to individuate a minimum set of statistically significant descriptors and to find for them suitable values or range of values for parameters. Formal evaluation with professional and non expert users is needed for a correct estimation of the effectiveness of the installation and its usability.

Such future work will be addressed in the framework of the EU-ICT Project SAME (www.sameproject.eu), focusing on new forms of participative and social active listening of music.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] Rowe, R. *Interactive music systems: Machine listening and composition*. MIT Press, Cambridge MA, 1993.

[2] Camurri A., Canepa C., and Volpe G. Active listening to a virtual orchestra through an expressive gestural interface:

The Orchestra Explorer. In Proceedings of the *7th Intl. Conference on New Interfaces for Musical Expression (NIME-07)* (New York, USA, June 2007).

[3] Russell J.A. A circumplex model of affect, *Journal of Personality and Social Psychology*, 39 (1980), 1161-1178.

[4] Tellegen A., Watson D., and Clark L. A. On the dimensional and hierarchical structure of affect. *Psychological Science*, 10, 4 (Jul 1999), 297-303.

[5] 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 (2000), 1797-1813.

[6] Canazza S., De Poli G., Drioli C., Rodà A., and Vidolin A. Audio Morphing Different Expressive Intentions for Multimedia Systems, *IEEE Multimedia*, 7, 3 (200), 79 – 83.

[7] 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 (2005), 462-466.

[8] Camurri A., Lagerlöf I., and Volpe G. Recognizing Emotion from Dance Movement: Comparison of Spectator Recognition and Automated Techniques, *International Journal of Human-Computer Studies*, 59, 1-2 (2003), 213-225, Elsevier Science.

[9] Laban R., and Lawrence F.C. *Effort*. Macdonald & Evans Ltd., London, 1947.

[10] Laban R., *Modern Educational Dance*. Macdonald & Evans Ltd., London, 1963.

[11] Boone R. T., and Cunningham J. G. Children's decoding of emotion in expressive body movement: The development

of cue attunement, *Developmental Psychology*, 34 (1998), 1007-1016.

[12] Kilian J. *Simple Image Analysis By Moments*, Open Computer Vision (OpenCV) Library documentation, 2001

[13] Camurri A., De Poli G., Leman M., and Volpe G. Toward Communicating Expressiveness and Affect in Multimodal Interactive Systems for Performing Art and Cultural Applications, *IEEE Multimedia Magazine*, *12,1* (Jan. 2005), 43-53.

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

[15] Wallbott H. G. Bodily expression of emotion, *European Journal of Social Psychology*, 28 (1998), 879-896.

[16] Argyle M., *Bodily Communication*, Methuen & Co Ltd, London, 1980.

[17] De Meijer M., The contribution of general features of body movement to the attribution of emotions, *Journal of Nonverbal Behavior*, 13 (1989), 247-268.

[18] Camurri A., Coletta P., Demurtas M., Peri M., Ricci A., Sagoleo R., Simonetti M., Varni G., and Volpe G. A Platform for Real-Time Multimodal Processing, in Proceedings *International Conference Sound and Music Computing 2007 (SMC2007)* (Lefkada, Greece, July 2007).