Expressive Control of Music and Visual Media by Full-Body Movement

Ginevra Castellano[1], Roberto Bresin[2], Antonio Camurri[1], Gualtiero Volpe[1]

[1]InfoMus Lab, DIST - University of Genova
Viale Causa 13, I-16145, Genova, Italy
+390103532201
{ginevra.castellano, antonio.camurri, gualtiero.volpe}@unige.it
[2] KTH Speech Music and Hearing
Lindstedtsvägen 24, 100 44 Stockholm, Sweden
+46 8 790 7876
roberto@kth.se

ABSTRACT
In this paper we describe a system which allows users to use their full-body for controlling in real-time the generation of an expressive audio-visual feedback. The system extracts expressive motion features from the user's full-body movements and gestures. The values of these motion features are mapped both onto acoustic parameters for the real-time expressive rendering of a piece of music, and onto real-time generated visual feedback projected on a screen in front of the user.

Keywords
Expressive interaction; multimodal environments; interactive music systems

1. INTRODUCTION
In the music computing community there is an increasing attention on identifying new paradigms of expressive interaction with machines. One of the most consolidated directions is that of interactive music systems, i.e. systems able to process expressive gestures for generating and controlling musical signals [2], such as hyper- and virtual musical instruments [6].

Attention of the music computing community is increasingly focusing on the development of interaction metaphors that take into account full-body movements and gestures at different levels of abstraction, and the interaction with an active environment with evolution and dialogue capabilities. This leads to Multimodal Environments (MEs) enabling multimodal user interaction by exhibiting real-time adaptive behaviour. We refer in particular to immersive environments enabling communication by means of full body movements, such as in dancing, singing, and playing. The present work is a contribution to the investigation of MEs with focus on expressive communication. We present a system allowing users to express themselves through their full-body movement by real-time controlling the generation of an audio-visual feedback.

2. SYSTEM ARCHITECTURE
The system presented in this work is based on the integration of two different systems: EyesWeb[1] [4], a platform for multimodal analysis and for development of interactive systems and MEs, and pDM[2] [5], a software for expressive music performances. Both systems work in real-time.

The system acquires input from a video camera, processes information related to the expressivity of human full-body movement, extracts expressive indicators, classifies movement according to basic emotions (sadness, serenity, happiness, fear and anger) and controls both the expressivity of the music performance and the visual feedback generation (see Figure 1).

Body movements analysis and virtual feedback generation were realized using the EyesWeb platform. The expressive music performance rendering was implemented with pDM.

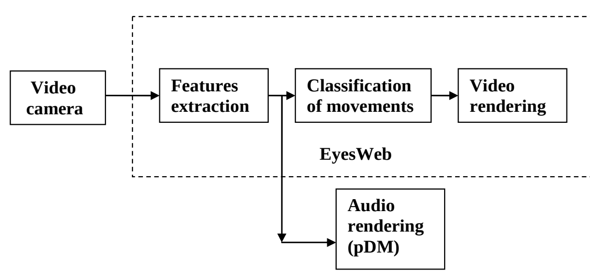

Figure 1. Overview of the system.

2.1 Set up
In order to facilitate tracking of the participants' body, the system runs in a square room with diffuse lights. A computer running EyesWeb and pDM is connected through a sound card (Edirol FA-101) to wireless headphones. A video camera (Panasonic WVCP450/G), located in front of the user, is used for capturing

[1] http://www.eyesweb.org

[2] http://www.speech.kth.se/music/performance/download/

her full-body movement and for providing EyesWeb with the visual input (25fps non interlaced). A video-projector projects the visual feedback on the wall in front of the user.

3. ANALYSIS OF USER'S EXPRESSIVITY

Movement and gesture are key elements in communicating expressivity. In our system we consider two expressive motion cues: the Quantity of Motion (QoM) and the Contraction Index (CI). Both cues are global indicators of human movement; QoM is correlated with the user's energy, and CI with the space occupied by the user. The calculated values of QoM and CI are used for identifying the basic emotions used by the user. Results are used for the control of musical and visual outputs.

4. VISUAL FEEDBACK GENERATION

Visual feedback was designed to respond to a specific expressive motor behaviour of the users. The colour assumed by the projected silhouette is associated to the emotions communicated by the users' movement and which are previously identified by the system. Users can see their own silhouette projected on a wall in front of them with different colours, depending on their movement expressivity.

Accordingly to theories on movement and emotions from psychology [7] and previous studies on expressivity [3], we defined a correspondence between emotions and QoM and CI movement characteristics as presented in Table 1

.

Table 1. Relations between emotions and movement characteristics.

Emotion	Movement characteristics
Sadness	Slow (low QoM), contract (high CI)
Serenity	Slow (low QoM), expanse (low CI)
Happiness	Fast (high QoM), expanse (low CI)
Anger	Very fast (very high QoM), expanse (low CI)
Fear	Fast (high QoM), contract (high CI)

Starting from results from previous studies on the association of colours to expressive music performances [1], we associated specific colours (see some examples in figures 2 and 3) to the emotions communicated by the user's movements; violet was associated to sadness, pink to serenity, yellow to happiness, red to anger, blue to fear.

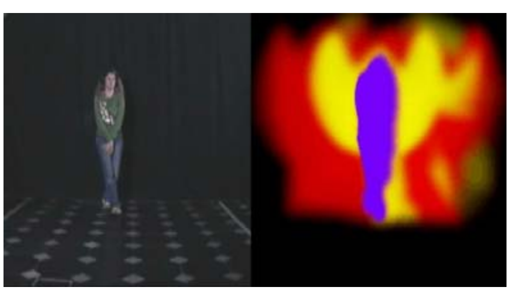

Figure 2. Example of visual feedback related to sadness (violet), that was preceded by anger (red) and happiness (yellow).

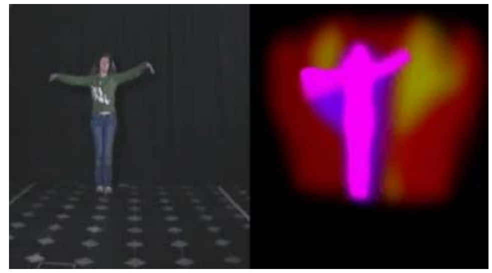

Figure 3. Example of visual feedback: serenity (pink), following sadness (violet).

5. MUSIC PERFORMANCE RENDERING

QoM and CI are used for controlling the real-time expressive performance of a piece of music. QoM and CI are mapped onto tempo and sound level respectively. In this way the user's performance gets faster for faster movements (i.e. higher QoM values) and louder for the user is expanding her body (i.e. lower CI). Articulation is controlled by mapping the emotion detected in the user's movements to the articulation parameters in pDM. More *legato* articulation for emotions with lower activity and more *staccato* articulation for those characterized by higher energy.

7. ACKNOWLEDGEMENTS

The research work has been realised in the framework of the EU-IST Project HUMAINE (Human-Machine Interaction Network on Emotion), a Network of Excellence (NoE) in the EU 6th Framework Programme (2004-2007) and the COST287 Action ConGAS (Gesture CONtrolled Audio Systems).

8. REFERENCES

[1] Bresin, R.: What is the color of that music performance?, In proceedings of the International Computer Music Conference- ICMC 2005, 367–370, 2005.

[2] Camurri, A.: Interactive Dance/Music Systems, Proc. Intl. Computer Music Conference ICMC-95, pp.245-252, The Banff Centre for the arts, Sept.3-7, 1995, Canada, ICMA-Intl.Comp.Mus.Association.

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

[4] Camurri A., Coletta P., Massari A., Mazzarino B., Peri M., Ricchetti M., Ricci A., Volpe G.: Toward real-time multimodal processing: EyesWeb 4.0, in Proc. AISB 2004 Convention: Motion, Emotion and Cognition, Leeds, UK, March 2004.

[5] Friberg, A.: pDM: an expressive sequencer with real-time control of the KTH music performance rules. Computer Music Journal, 30(1), 37-48, 2006.

[6] Machover, T. and Chung J.: Hyperinstruments: Musically intelligent and interactive performance and creativity systems. *Proc.ICMC'89*, 186-190, 1989.

[7] Wallbott, H.G.: Bodily expression of emotion. European Journal of Social Psychology, Eur. J. Soc. Psychol. 28, 879-896, 1998.