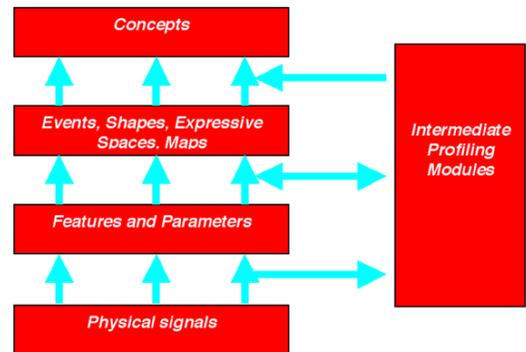


paper



megase

a multisensory expressive gesture applications system environment for artistic performances

Antonio Camurri, DIST-University of Genoa, music@dist.unige.it
 Giovanni De Poli, DEI-University of Padova, depoli@dei.unipd.it
 Marc Leman, IPEM-University of Gent, Marc.Leman@rug.ac.be

Abstract

This paper introduces the EU IST three-year project MEGA (Multisensory Expressive Gesture Applications). MEGA is centered on the modeling and communication of expressive and emotional content in non-verbal interaction by multi-sensory interfaces in shared interactive mixed reality environments. In particular, the project focuses on music performance and full-body movements as first class conveyors of expressive and emotional content. Main research issues include the analysis of expressive gestures (i.e. analysis of the expressive content conveyed through full body movement and musical gestures), the synthesis of expressive gesture (i.e. the communication of expressive content through computer-generated expressive gesture, such as music performances, movement of virtual characters, expressive utilization of visual media), the strategies for mapping the data coming from multimodal analysis onto the different output channels. Applications concern new generation of musical instruments (for professional as well as home consumer), applications based on enhanced Collaborative VEs embedding expressive communication, e.g. for theatre and interactive artistic performances.

Keywords: expressiveness, emotion, music, dance, CVEs, multimedia systems for artistic performance.

Project URL:

<http://www.megaproject.org>

1. Introduction

Research and industrial communities nowadays have a strong interest in non-verbal multi-modal communication in man-machine interaction. In music research, activities are concentrated on the development of audio-mining technology and interactive music systems for theater applications.

In this paper, we focus on the problem of non-verbal multi-modal communication in relation to theater applications. This work is based on the three-year MEGA IST-project no.20410, started on Nov 2000. MEGA stands for Multi-sensory Expressive Gesture Applications. MEGASE refers to the MEGA System Environment for artistic applications. The conceptual framework, which is at the core of the research activities in MEGA, is presented. A special focus is on (non-verbal) expressive gesture in audio and human movement. Research in MEGA is based on:

- Implementation of humanistic theories of non-verbal communication;
- Analysis/synthesis of artistic expression in dance and music;
- Applications envision artistic performances, using audio, video and sensing technologies.

This approach has some connections with emotion research and Kansei research. Emotion research is known as Affective Computing in the USA (cf. Bates/CMU, B.Hayes-Roth/Stanford, Picard /MIT, OCC model), while sensitive processing research is called Kansei Information Processing in

Intermediate paradigms can be defined, such as the Orchestra Conducting, where a director performs certain actions whose effect is sometimes almost reactive by orchestra performers, but with an additional level of cognitive interpretation and different time scales. In general Imitative behavior is in between these two paradigms.

The dialog paradigm may embed imitative or reactive forms of non-verbal multimodal expressive gesture communication.

3.2 The Layered Gesture Architecture

The recognition of features and sequences of features relevant for expressive gestures, in particular also for affective and sensitive computing, is an important aspect. The analysis of expressive gestures is based on a multi-layered gesture architecture. It aims at integrating in a multi-modal perspective analysis of audio, video, tactile sensor signals.

Layer 1 – Physical Signals

Layer 2 – Low-level features and statistical parameters:

For example, in the audio domain, this layer refers to features such as Tempo, Tempo variability, Sound level, Sound level variability, Spectral shape, Articulation, Articulation variability, Attack velocity, Mean Pitch, Duration contrast, Degree of accent on structural important notes, Periodicity, Dynamics, Roughness, Tonal tension. Similar features can be defined for movement analysis.

Layer 3 – Events, Shapes, Phrase, Expression Spaces, Maps:

- Expressive features modeling: e.g. HMMs, sequences of parameters of a physical model
- Expressive gesture spaces modeling: e.g., energy-velocity spaces)

Layer 4 – Concepts and structures:

- Taxonomies of concepts related to emotion expression (e.g., using probabilistic reasoning systems). Laban's Effort, Schaeffer's Morphology, and OCC emotion theory are examples of taxonomies at this level.

An important aspect concerns the personalization of this architecture with respect to gesture, context, and user, obtained by Intermediate Profiling Modules (figure 1). This mechanism makes the "archetypical" parameters of the architecture more flexible and personalised, by keeping track of

- their evolution over time given specific contexts, and
- different biasing due to "personality" and focus of interest, etc.

Preliminary examples of results from research work on the extraction of expressive cues and can be found as for audio in (Leman 2001) and (De Poli et al 2001) and for movement in (Camurri Mazzarino Trocca and Volpe 2001 – these proceedings).

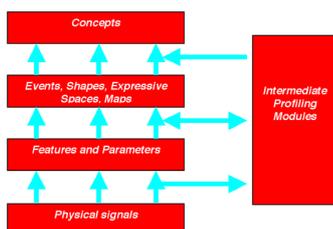


Figure 1- The MEGA Layered Architecture

4. The MEGA System Environment

The MEGA System Environment is the main scientific/technological result and concrete product of this project. It will be used along the project in public events, including museum exhibits and artistic performances. The MEGA System Environment consists of a real-time environment formed by the following main components:

- the EyesWeb real-time open software architecture for music and multimedia performance (developed by DIST): EyesWeb is the basic kernel for the MEGASE
- a number of various EyesWeb libraries (for expressive audio analysis, movement analysis, mapping etc), developed by MEGA partners.

MEGASE includes and is connected to a number of proprietary systems (e.g. Director Musices from KTH and Dovre from Telenor) and with commercial systems (e.g. Opcode Max/MSP).

Evaluation and validation of libraries modules for analysis of expressiveness are performed basically by comparing the output from such modules with spectators ratings (Uppsala University).

During the development of methods and algorithms, various tools are also used, such as the IPEM ToolBox.

5. Performances, Demos and Prototypes of New Multimedia Applications

Even if the MEGA project is only going to conclude his first year, a number of outputs are coming out. Public artistic performances already used some of its preliminary results, in particular the EyesWeb platform and the new libraries. They include:

- Concert of tap dance at the Museum of Modern Art in Gent SMAK, 1 September 2001; A tap-dancer is "observed" by floor-microphones and by a video-camera. He controls visual outputs and live electronics (IPEM).
- Ballet "Messaggero Muto" (music by F.Giomi, with the Dance Company of Virgilio Sieni), 15 July 2001. In this performance a dancer interacts with a his visual clone (implementation in EyesWeb, DIST);
- Concert of Roberto Doati: an actress dialogs in real-time with another audio-clone. The clone has a voice depending on the voice of the actress and her facial (lips) movements (DIST). (figure 2)

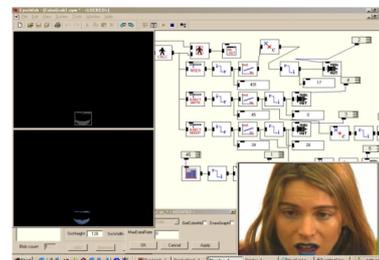


Figure 2- an example of a patch at run-time implementing a simple movement analysis of the lips of an actress, based on a videocamera. The patch has been developed for a piece by the composer Roberto Doati. The two black windows on the left show intermediate outputs from the lip analysis. The right window shows the original image. The patch generates MIDI signals related to lip movement, which are used to re-create in real-time an "alter ego" of the actress (Francesca Faiella) with which the actress dialogues.

