

# DISEMBODIED VOICES: A KINECT VIRTUAL CHOIR CONDUCTOR

**Marcella Mandanici**

Music Conservatory of Brescia, Italy

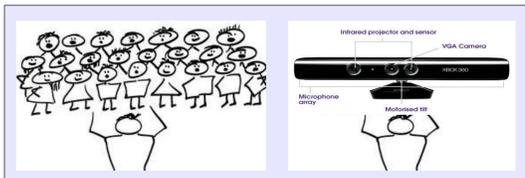
**Sylviane Sapir**

Music Conservatory of Como, Italy

## Overview

“Disembodied voices” is an interactive environment designed for an expressive, gesture-based musical performance.

- The motion sensor Kinect, placed in front of the performer, provides the computer with the 3D space coordinates of the two hands.
- The application is designed according to the metaphor of the choir director: the performer, through gestures, is able to run a score and to produce a real-time expressive interpretation.
- The software interprets the gestural data and controls articulated events to be sung and expressively performed by a virtual choir. Hence the name of the application: you follow the conductor’s gestures, hear the voices but don’t see any singer.



The “Disembodied Voices” metaphor

## Aim

This work is an attempt to explore natural interaction during different phases of a compositional process.

The idea is to get more awareness about the meaning of natural interaction respect to music activities

The hope is to build an intuitive user interface that supports musically skilled and less skilled users in order to deal with complex and structured musical materials.

## Our approach

### Hardware and software

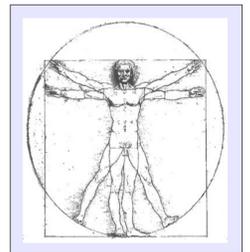
- Well-established hardware interface to track hand movements: **Kinect**
- Data communication protocol: **OSC**
- Real-time audio programming environment: **Max/MSP**

### Musical context

- Algorithmic composition based on Ligeti’s micro-polyphony paradigm
- Singing voice synthesis through sampling
- Arbitrary timbre variations through DSP (effects) used as “expressive” cues

### Interaction Design

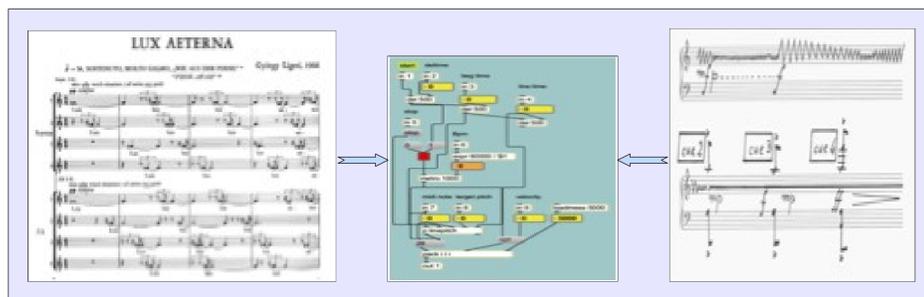
- Vitruvian man representation in a 3D space to cope with our body centred sensory-motor experiences of space.
- Choir conducting metaphor: attested conducting rules for right and left hands + arbitrary composer rules for expressive gestures acting on effects parameters.
- Two mapping levels:
  - High level for the control of the compositional process
    - ♦ Events triggering (discrete control)
    - ♦ Dynamics and density control (continuous control)
  - Low level for the control of effects sound processing parameters
    - ♦ one-to-one mapping and continuous control on effects parameters
- Adding visual feedback in order to fill up the lack of haptic feedback by displaying data on motion and musical score events



Starting from the analysis of Ligeti's micropolyphony

we developed the compositional algorithm model

and wrote an original score for the model



## Discussion and conclusion

The overall system is working but we’re just at the beginning

▪ Usability tests need to be done and collected gestural data still need to be analysed to check the “feel” of the whole system by different typology of users.

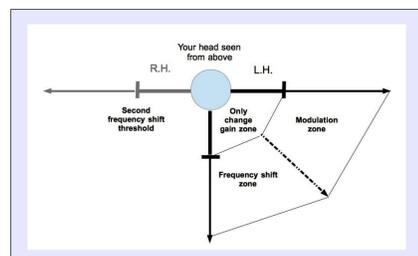
▪ Improvements will be done by:

- Extending the compositional algorithm controllability
- Working on voice synthesis and effects
- Working on higher hierarchy levels to better interact with structural aspects of the composition
- Further research need to be done for:
  - Handling different compositional algorithms
  - Exploring different interaction metaphors

## How to use “Disembodied voices”

1. Enter the play zone marked on the floor and take the “Ψ” pose
2. Use hands movements to conduct the performance as illustrated below
3. Exit the play zone by side

Overview of hands movements for conducting the performance



## A few references

- [1] J. Jarvlepp. “Pitch and texture analysis of ligeti’s lux aeterna”. [Online]. Available: <http://www.ex-tempore.org/jarvlepp/jarvlepp.htm>
- [2] J. Rován and V. Hayward, “Typology of Tactile Sounds and their Synthesis in Gesture-Driven Computer Music Performance,” Trends in Gestural Control of Music edited by Marcelo Wanderley Marc Battier, pp. 297–320, IRCAM, Paris, 2000.
- [3] M. O’Modhrain, “Playing by feel: Incorporating haptic feedback into computer-based musical instruments,” Ph.D. dissertation, Stanford University, November 2000.