Tele-media-art – Web-based Inclusive Teaching of Body Expression

Elizabeth Simão Carvalho
UAb/CIAC
University Aberta
Lisbon, Portugal
elizabeth.carvalho@uab.pt

Adérito Fernandes Marcos
UAb/CIAC
University Aberta
Lisbon, Portugal
aderito.marcos@uab.pt

***Abstract*— The Tele-Media-Art project aims to promote the improvement of the online distance learning and artistic teaching process applied in the teaching of two test scenarios, doctorate in digital art-media and the lifelong learning course "the experience of diversity" by exploiting multimodal telepresence facilities encompassing the diversified visual, auditory and sensory channels, as well as rich forms of gestural / body interaction. To this end, a telepresence system was developed to be installed at Palácio Ceia, in Lisbon, Portugal, headquarters of the Portuguese Open University, from which methodologies of artistic teaching in mixed regime - face-to-face and online distance - that are inclusive to blind and partially sighted students. This system has already been tested against a group of subjects, including blind people. Although positive results were achieved, more development and further tests will be carried in the future.**

Keywords—component, formatting, style, styling, insert *(key words)*

I. Introduction

The Telemedia-art project aimed to develop a distance, web and asynchronous learning platform that, using low cost motion capture technologies, feasible for home use, enables blind students to enjoy classes that address the reflection on poetics of the body articulated with the practice of body movement.

The main goal of this project is to develop a teaching platform with video recordings of the body movement that can be accessed and used by both students and teachers. These recordings are made by one (or more) motion capture Kinect camera. Based on the input received, sounds related to the movement of six reference points in the body and its spatial projections are produced in parallel with the kinematic 3D representation of the skeleton of the user.

II. Background

It [3] is arguable that there are really no models of e-learning per se – only e-enhancements of existing models of learning. Technology can play an important role in the achievement of learning outcomes but it is not necessary to explain this enhancement with a special account of learning. Rather, the challenge is to describe how the technology allows underlying processes common to all learning to function effectively. A true model of e-learning would need to demonstrate on what new learning principles the added value of the 'e' was operating. Where, for example, the 'e' allows remote learners to interact with each other and with the representations of the subject matter in a form that could simply not be achieved for those learners without the technology, then we may have a genuine example of added value.

A way of adding value to e-learning should be also thought in terms of going beyond the scope of traditional and already in use technology. Exploring how each channel of information can be used, may result in an interesting solution for both blind and non-blind students of body expression.

For instance, Dimitrius et al. [1] proposed a novel system that automatically evaluates dance performances against a gold-standard performance and provides visual feedback to the performer in a 3D virtual environment. The system acquires the motion of a performer via Kinect-based human skeleton tracking, making the approach viable for a large range of users, including home enthusiasts. Unlike traditional gaming scenarios, when the motion of a user must be kept in synch with a pre-recorded avatar that is displayed on screen, the technique described in this paper targets online interactive scenarios where dance choreographies can be set, altered, practiced and refined by users. In this work, we have addressed some areas of this application scenario. In particular, a set of appropriate signal processing and soft computing methodologies is proposed for temporally aligning dance movements from two different users and quantitatively evaluating one performance against another. Such approach would be interesting to be considered in any e-learning system that intends to teach body expression.

Other relevant channel, especially for the blind ones, is the sound. Converting spatial [4] information to non-speech audio for blind people has been explored for many years. Combinations between the pitch and stereo to represent line graphs with two data series: each series was represented using pitch, while stereo position was used to separate the series or pitch and stereo panning to represent two-dimensional map and table data, are some of the existing approaches. Depth studies have already been performed on the sonification mappings for blind and sighted readers of auditory graphs [5], finding that both blind and sighted individuals typically applied similar mappings between audio pitch and other variables, such as size and velocity.

In higher education, web-based learning environments are commonly used to support students' reflection on competencies. Web-based learning [3] environments offer practical benefits such as accessibility of practices and flexibility in updating information. However, these technologies are not optimal for dance education because most of the time they only include written sources. Dance students need to reflect on their dance practice. Video has become increasingly popular in professional learning because of its unique ability to capture the richness and complexity of practices for later analyses. That means that any distance platform that intends to support the learning and teaching of dance, or any variation of it, like body expression, should include videos.

III. Requirements

As an artistic area to be initially focused, it was considered the teaching of corporal expression. Based on the scenario of a class in this area, the following pedagogical and cognitive requirements were identified (in order to allow a good follow-up of the class), as the most significant ones to be supported by the Tele-media-art system:

- The existence of a subject-teacher;
- The existence of a student or more, remotely located, accompanying the class in a synchronous way or not, and with varied age;
- It is important that the student does the lesson, and does not follow or simply watch it;
- Sensory development is important, taking advantage of different channels of information: audio, visual and tactile;
- It should allow inclusion and non-differentiation among students - especially by visual limitation;
- Allow the student to choose the best way to accompany the lesson - by which channel, synchronous or not;
- Work of basic concepts (depending on the meaning, language or medium): movement, time, silence, colour, gesture, sound and space;
- Division by the senses / languages / means:
- Vision: colour, luminosity, displacement, spatiality;
- Hearing: clean and concrete sounds, silences, differentiation between acute, medium and severe (to be able to associate to levels of movement), volume;
- Tact: textures, thermal sensations and movements;

On the other hand, it is necessary that Tele-Media-Art be supported by an e-learning environment. We aim to give participants the possibility of monitoring an artistic performance at a distance, with some degree of intervention through an e-learning platform (namely the UAb's Moodle).

The scenario considered was in an e-learning environment and in mobile applications (Moodle Mobile). A group of students / trainees can contribute to a hybrid narrative that converges multimedia elements (words, images, sounds, and video) on the performance recorded with the help of the Tele-Media-Arte system. The created object will be a video recording that represents an event to which other words, images, sounds, with new meaning are linked. From this it is possible to develop a didactic resource in e-learning.

IV. The Telemedia-Art System

Before proceeding to considerations regarding the outcome of the tests of this system with blind volunteers it is important to clarify its context of use in the poetic classes of the movement and the performative body in interaction with technologies Multimedia. In a generic way, the learning of dance as an artistic field can be considered under two strands: a strand of execution, which aims to develop technical and sensor-motion capacities and a strand of exploration/creation of movement, which aims to develop the perception of movement and performative and creative capacities.

In the context of the multimedia arts course, it is not intended to address a strand of execution, but rather a strand of artistic awareness in relation to the body and its movement, where it focuses on the development of the exploration of capacities, perception and creation of movement in its interaction with multimedia technologies. However, given that this is a pilot and innovative project in its methods and purposes, it is important to verify primary technical skills but with regard to the capacity recognition of the form of the movement without the use of vision and of the touch.

Thus, the tests carried out needed to be done under an analytical aspect of the movement (which allows to examine the movement at the anatomical level, at spatial level as directions, levels, displacements and spatial design, and minimally at the level of time and dynamics), considering the realization of basic movements and not its poetic or artistic aspect. In other words, the tests carried out mainly considered a strand of movement execution with the intention that the student can accurately reproduce the teacher's movements.

The purpose of the tests is to make it possible for blind students to percept the teacher's movement through sound, compared the sound feedback of their movements with the sound feedback of the teacher's movements.

A. *Prototype*

The Telemedia-art system consists of 4 modules. Figure 1 illustrates its general architecture, where we can see the modules that allow the teacher to record the lesson in video and sound, the movements being decoded into sound and also represented in a kinematic skeleton. The system is accessed through Moodle and the class can be made available to students. On the student side, in addition to accessing the information provided by the modules, it is also possible to perform a video recording, and send to the platform for further analysis by the teacher. It is a completely web-based application.

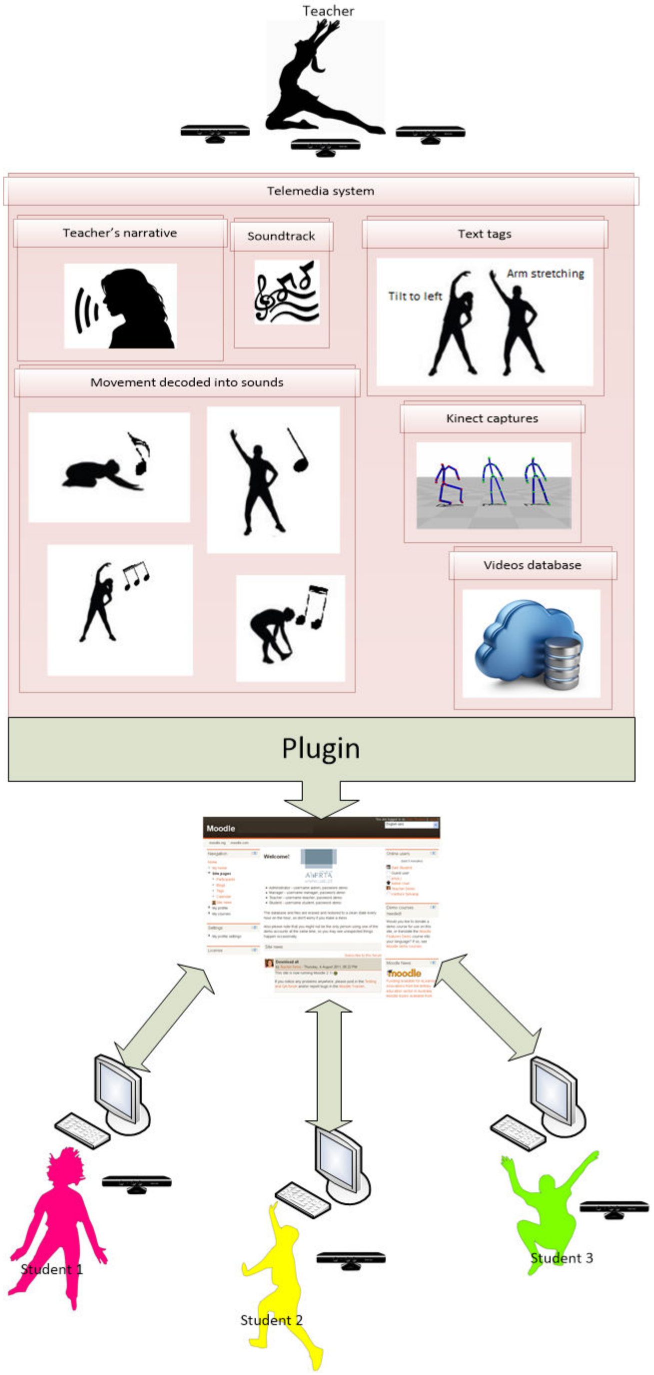

Figure 1. General architecture of the Telemedia-art system

Several tools were used for the implementation of the prototype, notably the following: Laravel 5.4.18, Popcorn.js, RecordRTC, Kinect – Server & Websockets, OSC, PKG – node package, FFMPEG and Bootstrap 3.0.

The function of each framework is summarized as follows:

- Laravel – Responsible for generating all the environment HTML, CSS, Javascript and PHP of the platform. From a technical standpoint, due to the articulation with the RecordRTC module, it was not taken advantage of the controllers of the Laravel;
- PopCorn. js – Used in video query videos on the part of student and teacher. It has the function of ensuring that by changing the "timeline" of the main video, the remaining channels follow this change;
- RecordRTC – Video recording framework by PHP. This framework is compatible with Chrome browser and Firefox;
- Kinect – Server & WebSockets – The official Microsoft Kinect SDK 2.0 has been resorted to accessing Kinect data;
- OSC – protocol used to connect Javascript to Puredata. The Puredata receives the x, y and z coordinates of the 6 chosen joints – head, pelvis, right wrist, left wrist, right ankle, left ankle;
- PKG – Node package – used to compile OSC into a Windows executable;
- FFMPEG – Used to remove the "timestamps" of the videos, add audio and video in a single video file, add time to the student's video, to ensure that the "timestamp" of the teacher's video corresponds to the student's "timestamp", and to process the high videos Quality of Kinect for an acceptable standard for web streaming;
- Bootstrap 3.0 – to speed up the development and design process, you chose to use the "Bootstrap" grid and functions to ensure that application can be used on monitors of different sizes.

At this point, the Telemedia-art can only run in Chrome due the existing incompatibility between Kinect and other browsers. CSS file has to be changed to ensure compatibility.

The hardware and software requirements are: Windows 8 or higher 64-bit operating system, NodeJS 64-bit 6.10.2, SDK Kinect 2.0, Puredata 0.47.1, Chrome Browser version 62.0.3202.94 (64 bit), Puredata sound algorithm, application to connect to Kinect, Package PKG of the OSC framework, USB 3.0 port, KinectV2 and integration with Moodle.

Moodle proved to be quite restrictive in the integration of Kinect, as the development of compatible modules required more time. It was chosen to develop an external and integrated application in Moodle by IFRAME.

Moodle allows URL variables to be sent to external applications. It took advantage of this functionality to obtain the identifiers of each user to ensure that the teacher and the student had the information filtered for each case.

The Telemedia-art can be accessed through the Moodle platform by any student, just as their computer is connected to a Kinect to be able to take advantage of its full features. The environment setup is shown in figure 2, where we can see the filed-of-view and the ideal location between the Kinect and the end-user.

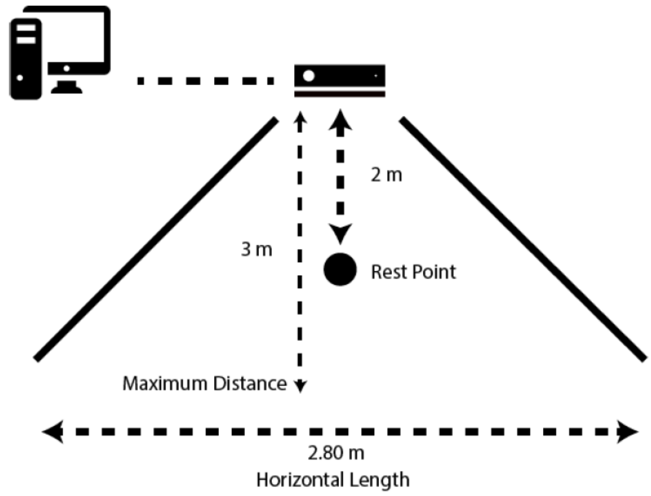

Figure 2. Environment setup

B. Tests and Evaluation

The Telemedia-art system is a multichannel system, of experimental and methodological character in the field of distance education of the Portuguese Open University. We can receive information and/or knowledge in various and different channels, such as: videos, audio files recorded with description for blind and non-blind people, music files, text, static image files, and also a channel that at was developed to enable the registration of movements of the body through the Kinect sensor and PureData programming (kinematic skeleton).

Sound constructions were generated from six points of recognition: the upper part of the head, the wrists, the pelvis and the ankles. For each movement is generated a different sound that is a coarse approximation of a Theremin. This sound is part of a point 0, relative to a "neutral" position (vertical body extended arms along the body, parallel legs stretched together, right head), and varies in volume, pitch and frequency, according to X, Y and Z axes of the spatial variations of the movement. Figure 3 and 4 illustrate how sound is mapped according the X or Y direction of the body movement.

These tests were performed with very simple and basic movements and with primary spatial displacements that cannot go beyond the camera's field of view, e.g.: Raise the right or left arm to the shoulder height or above the head, tilt the head, circular movement of the basin, walk back and forth.

During the testing sessions the blind volunteers were able to identify and reproduce most of the simple movements, but as soon as the movements became slightly more complex, there were clear difficulties in identifying and playing the movement through the sound. Some of these difficulties could be improved by the use of sounds referring to different musical instruments rather than a single type of sound. Clearly this would improve not only the ability to recognize sounds but also the aesthetic experience.

The main problem encountered was not because of the way things were developed and implemented but because of a limitation imposed by the Kinect system itself – some movements are not properly recognized by it and in result, some movements and sounds are not well mapped.

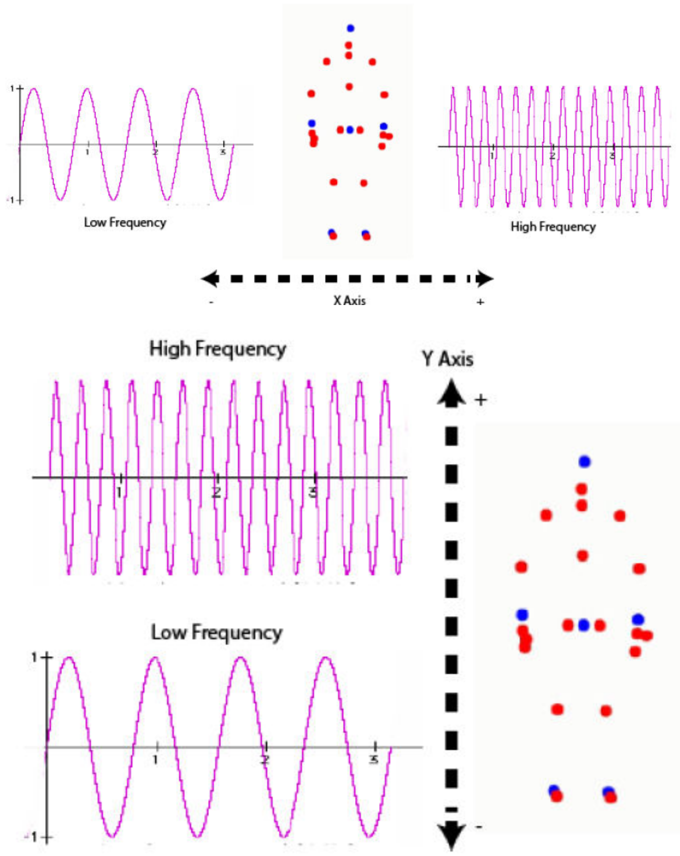

Figure 3. Sound mapping to the body movement direction

The technical capacity of the equipment demonstrated a strong limitation on the recognition of more complex movements such as: simultaneous movements of the trunk and upper and/or lower limbs; movements in which the arms and/or legs pass through the front or behind the trunk; movements in which the body rotates on its own axis; movements in which the spatial projection amplitude is reduced; spatial movements outside the camera's field of view and changes in spatial orientation of the body relative to point 0 (e.g. movements on the floor, inverted movements or movements directed to a different front of position 0).

Since motion sensors have their spatial reference located on the body spatial movement axes (which must be initiated from a static position), they are not able to detect properly the rotations of the articulations that do not have spatial displacements. It is also important to add that the bounced movements become inadequately recognizable due to the poor technical ability to process very fast movements and variation of speeds. Similarly, dynamic variations, strength and tension become hardly recognizable.

With regard to students without disabilities, Telemedia-art presented in general good results, allowing the student to perform the activity without great difficult. Figure 4 shows a snapshot of the tests performed by one blind subject.

Figure 4. Testing the system with a blind subject

V. CONCLUSIONS AND FUTURE WORK

To appreciate in depth the scope and overall goal of this project it is essential to take into consideration that the body movement is characterized by properties related to: spatial characteristics such as directions, levels, spatial displacements, positioning in relation to the observer, temporal characteristics such as dynamics, speeds and rhythms, bodily characteristics such as, strengths, muscular tensions and bodily morphology, formal characteristics such as movement dimensions, shapes and relationships between the various segments and body joints, and symbolic characteristics such as gesture and relationship with elements around them, etc.

According to the considerations mentioned above regarding the test report, the motion recognition system by sound articulated with the Kinect cameras is quite limiting so that a blind person can recognize the formal characteristics of the dance movement and is virtually inefficient in recognizing any other features. However, information on non-formal motion characteristics can be transmitted orally, using even metaphorical descriptions.

Throughout the project it was possible to compare the motion recognition capabilities provided by the Kinect system and the motion recognition capabilities provided by the XSENS system. The XSENS system demonstrated a notable superiority to Kinect, as it makes a movement recognition from the topographic space of the body identifying a large number of joints, and has no spatial boundaries, thus allowing a formal and temporal recognition of almost a perfect motion. However, in the context of distance learning, the XSENS system would have to be something acquired by each student. Since at present the purchase of a XSENS system has a very high cost, this would not be financially bearable by most students, so at the moment, this is not a solution conniving with the mode of teaching practiced in the Portuguese Open University (UAb).

Finally, it is important to highlight another approach to use this technology. Being a technology that produces possibilities of creating/manipulating sounds in conjunction with the creation of movement, it can be used not merely with the motion recognition function for blind students, but as a creative tool (for all students) that allows a singular way to develop the relationship between sound and body movement.

This technology offers the possibility of exploring a creative manipulation of sound through movement, that is, it provides the body to become creator and producer of sound compositions. This characteristic provides an exceptional possibility of practicing the perception of kinesthesia between the ear and proprioceptive, and thus opens a new space for the exploration of other creative and artistic manifestations, providing still a new substrate of research and reflection.

In order for this mode of use to be practicable it is only necessary, in the first phase, to change the sound sources of the system, being that, as mentioned above, instead of a single type of sound it can work with variations of different types of sound.

Some possible improvements include:

- The attempts of vocal instructions should be erased (in the teacher's space). Currently every time a new statement is written, a player is added to the sound;
- The "Loading" bar of the video should indicate the percentage to which the process is concluded;
- Allow students to see other students ' videos;
- Simplify the selection of lessons by the student, without having to go through a list of classes;
- The point cloud created by the input of the Kinect should be replaced with an Avatar (3d model);
- The video captured with the kinematic skeleton should be recorded preferably in a text file with the values of the X, Y, and Z coordinates. In the class consultation, this file should be read and reproduced by the system;
- Propose a new design for the pages;
- Replacement of the Kinect by a module implemented with automatic motion detection algorithm, which uses the computer's own camera.

ACKNOWLEDGMENT

This project was financed by Calouste Gulbenkian Foundation under Grant number 142793.

REFERENCES

[1] Alexiadis, Dimitrios S., et al. "Evaluating a dancer's performance using kinect-based skeleton tracking." Proceedings of the 19th ACM international conference on Multimedia. ACM, 2011.

[2] Beetham, Helen, and Rhona Sharpe, eds. Rethinking pedagogy for a digital age: Designing for 21st century learning. routledge, 2013.

[3] Leijen, Äli, et al. "Streaming video to enhance students' reflection in dance education." Computers & Education 52.1 (2009): 169-176.

[4] Oh, Uran, Shaun K. Kane, and Leah Findlater. "Follow that sound: using sonification and corrective verbal feedback to teach touchscreen gestures." Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility. ACM, 2013.

[5] Walker, Bruce N., and Lisa M. Mauney. "Universal design of auditory graphs: A comparison of sonification mappings for visually impaired and sighted listeners." ACM Transactions on Accessible Computing (TACCESS) 2.3 (2010): 12.