GSPAT: LIVE SOUND SPATIALISATION USING GESTURAL CONTROL

Balandino Di Donato

Integra Lab
Birmingham City University
Birmingham, United Kingdom
balandino.didonato@mail.bcu.ac.uk

Jamie Bullock

Integra Lab
Birmingham City University
Birmingham, United Kingdom
jamie.bullock@bcu.ac.uk

ABSTRACT

Sound spatialisation is an important component in interactive performances as well as in game audio and virtual or mixed reality systems. HCI practices are increasingly focused on creating a natural user experience and embodied interaction through gestural control. Body movements that coincide with sounds consist of both performed 'sound producing' gestures, and ancillary and communicative movements. Thus, different gestural typologies may relate to the same audio source. Furthermore, gestures may depend on the context in which they have been expressed; in other words, they can carry different semantic or semiotic meanings in relationship to the situation and environment or reality in which they have been enacted. In order to explore these research themes, we are developing gSPAT: a software and hardware system able to drive live sound spatialisation for interactive audio performance using gestural control based on human-meaningful gesture-sound relationships. The ultimate aim is to provide a highly natural and musically expressive sound spatialisation experience for the performer. Here we describe three experiments conducted to explore the possible directions for the future of gSPAT's development. The tests employ a range of practice-based and ethnographic research methods to establish applicability, naturalness and usability across a range of approaches to the interaction design of the system.

1. BACKGROUND

Gestural control for live electronic music is widely used by performers, composers and researchers in their artistic and academic works. Mapping strategies have already been utilised by Tanaka [1] and Wanderley [2]. The constant development of tools for interactive performance, for instance MuBU [3] and Harmonic Motion [4], gestural recognition and machine learning systems for audio interactive programming environment like ml-lib [6] and devices such as Kinect[1], Leap Motion[2] and Myo[3], confirm that the research community and the industry is likely to exploit technology, which allows users to engage with seamless experiences between the physical and the virtual worlds. [7] However, particularly in the domain of traditional western concert music, not all performers have the technical knowledge to set-up and use such powerful tools.

The objective of our work, is to find ways in which gestural data can be mapped and interpreted as spatialisation parameters using a standard commercial gestural controller; such that the results are musically meaningful to both performer and audience. To achieve these results, it is extremely important to find out the relationship between the meaning carried by the gestures and the sound spatialisation.

[1] https://www.microsoft.com/en-us/kinectforwindows/
[2] https://www.leapmotion.com
[3] https://www.thalmic.com/en/myo/

2. EXPERIMENTS

gSPAT is currently at an experimental stage, and we are engaged in an exploratory research process incorporating interactive audio programming environments, gestural recognition systems and possible applications in interactive performances. Three experiments have been conducted employing a practice-based methodology, the Myo armband as gestural controller, a MIDI[4] and OSC[5] bridge developed in Processing, Integra Live[6] and Pure Data as audio software.

2.1. Sound spatialisation through gestural mapping using Myo gesture recognition system

The first experiment[7] (Fig. 1a) entailed the evaluation of sound signal elaboration and its spatialisation using mapping strategies based on the user's gestures as captured by the Myo Gesture Recognition system and processed in Integra Live. The performer was able to process an audio file by making a 'hand pose' of a fist and moving his arm. A pitch transposition was controlled by the pitch value from the Myo's IMU sensor and a stereo panning was controlled by the yaw value. By changing hand pose to a 'fingers spread' position, the pitch value and yaw value were switched to drive the relative speed and the stereo pan of a second audio file. Thus the pose of the hand (spread vs clenched) was used to change the operating mode of the system.

[4] https://github.com/balandinodidonato/myo-processing/tree/dev/examples/e6_IMUemg_data_MIDI
[5] https://github.com/balandinodidonato/myo-processing/tree/dev/examples/e5_IMUemg_data_OSC
[6] http://integra.io/
[7] https://vimeo.com/119328623

2.2. Sound spatialisation through gestural recognition

Whilst the Myo controller provides a rich set of IMU and EMG data, it is only configured to recognise a limited set of gestures. The second experiment[8] (Fig. 1b) therefore looks at possibilities to expand the Myo's gestural palette through the implementation of machine learning system within the Pure Data interactive environment. Raw gestural data was captured from the Myo and used to train a Support Vector Machine (SVM) classifier system implemented in Pd using ml-lib[9]. The system parameters were derived through an experimental process of testing and refinement working in close collaboration with a professional singer. Following the training of the model, the developed system was utilised for VoicErutseG v0.1. An interactive performance where the singer spatialised her own voice through an octophonic audio system. [10] (Fig. 1c).

2.3. Mixed reality trough EMG and IMU data mapping

The third experiment[11] (Fig. 1d) was based upon the usability of Myo's EMG data in applying mapping strategies to control the amplitude of eight sinusoidal oscillators. Following this initial test, the EMG data were used to drive the amplitude modulation on a pre-recorded sound of crumpled paper[12] in order to construct an audio-based mixed reality system where the user can crumple paper through a natural hand movement and then 'throw' it into an imaginary 'bin'.

3. CONCLUSIONS

Findings gathered from these experiments reflect positively on the usability of the Myo armband as gestural controller for the development of a gesturally-controlled sound spatialisation system. In particular, the experiments revealed an immediate usability of the Myo data in Integra Live through direct mapping (Experiment 1); incorporating a machine learning layer resulted in a powerful tool to smooth out discrepancies between human gestures and computer interpretation (Experiment 2). EMG data gathered from sound producing gestures enacted on imaginary physical objects used to process the pre-recorded sound associated with the gesture (Experiment 3) can be a simple and engaging solution for the development of an interactive audio system usable for interactive performance and more widely in mixed reality environments.

Future work wil entail the collection of gestural, audio, video and ethnographic data, which will be fundamental to gaining knowledge regarding the relationship between human gesture and sound spatialisation. Next steps include implementing machine learning within Integra Live to improve ease-of-use and ease-of-setup, and the use of Dolby Digital audio system and and K-Array KW8 speaker for the sound spatialisation.

4. ACKNOWLEDGMENT

The performance VoiceErutseG v0.1 has been performed by Vittoriana De Amicis (Fig. 1c) during the Frontiers Festival 2015.

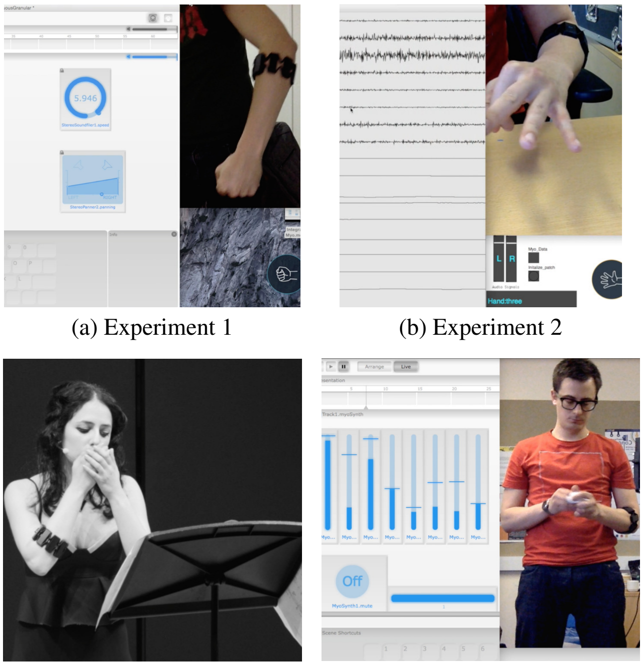
[8] https://vimeo.com/128388288

[9] https://github.com/cmuartfab/ml-lib

[10] http://www.balandinodidonato.com/research/voicerutseg-v0-1/

[11] https://vimeo.com/127913259

[12] https://vimeo.com/128395577



(a) Experiment 1 (b) Experiment 2

(c) VoicErutseG v0.1 (d) Experiment 3

Figure 1: Screenshots and pictures realised during the experimental stage.

5. REFERENCES

[1] A. Tanaka, "Mapping out instruments, affordances, and mobiles," in *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010), Sydney, Australia*, vol. I, 2010, pp. 88 – 93. [Online]. Available: http://www.nime.org/proceedings/2010/nime2010_088.pdf

[2] M. M. Wanderley, "Mapping Strategies in Digital Musical Instruments: Concepts and tools," in *Crativit Instrumentale, Crativit Ambiente: Colloque Art-Science Technologie*, A. Luciani, O. Tache, and N. Castagn, Eds., 2013, pp. 183–190.

[3] N. Schnell, A. Robel, D. Schwarz, G. Peeters, and R. Borghesi, "Mubu friends assembling tools for content based real-time interactive audio processing in max/msp," in *International Computer Music Conference (ICMC). Montreal, August 2009*, 2009.

[4] T. Murray-Browne and M. D. Plumbley, "Harmonic motion: A toolkit for processing gestural data for interactive sound," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2014, pp. 213–216. [Online]. Available: http://nime2014.org/proceedings/papers/273_paper.pdf

[5] N. Gillian and J. A. Paradiso, "The gesture recognition toolkit," in *Journal of Machine Learning Research 15*, 2014, pp. 3483–3487. [Online]. Available: http://jmlr.csail.mit.edu/papers/volume15/gillian14a/gillian14a.pdf

[6] A. Momeni and J. Bullock, *ml-lib*. https://github.com/cmuartfab/ml-lib, 2014.

[7] E. O. Boyer, L. Vandervoorde, F. Bevilacqua, and S. Hanneton, "Touching sounds: Audio virtual surfaces," 2015.