Proceedings of the 3rd Workshop on Intelligent Music Production, Salford, UK, 15 September 2017

MYOSPAT: A SYSTEM FOR MANIPULATING SOUND AND LIGHT THROUGH HAND GESTURES

Balandino Di Donato and James Dooley

Integra Lab
Birmingham Conservatoire
{balandino,james}@integra.io

ABSTRACT

MyoSpat is an interactive audio-visual system that aims to augment musical performances by empowering musicians and allowing them to directly manipulate sound and light through hand gestures. We present the second iteration of the system that draws from the research findings to emerge from an evaluation of the first system [1].
MyoSpat 2 is designed and developed using the Myo gesture control armband as input device and Pure Data as gesture recognition and audio-visual engine. The system is informed by human-computer interaction (HCI) principles: tangible computing and embodied, sonic and music interaction design (MiXD). This paper reports a description of the system and its audio-visual feedback design. We present an evaluation of the system, its potential use in different multimedia contexts and in exploring embodied, sonic and music interaction principles.

1. INTRODUCTION

Performing with technology is often synonymous with learning new skills that are at odds with musical skills, potentially having negative and 'disruptive' effects [2]. MyoSpat is a gesture-controlled electronic interaction system that aims overcome the 'disruptive', 'highly complex' nature of live electronic processing experienced by many performers, providing them with an opportunity for new expressive ideas. Bullock et al. [3] identify divergences between the performer and technology caused by the lack of familiarity with complex systems. This can create in a dislocation between the performer's gestures and the musical result. Lippe [4] emphasises the importance of allowing musicians to interact confidently with technology in order to present a musical and expressive performance. With MyoSpat, we underline the importance of embodying music [5], empowering performers to express their musical ideas through gestural control over any electronic part in performance. Visual feedback can enhance the gesture-sound relationship, playing a significant role in guiding the user's actions during performance [6] and strengthening the perception of auditory feedback [7]. MyoSpat gives musicians control over sound manipulation through a direct connection between hand gestures and audio-visual feedback, whilst making the newly learnt gestures as intuitive and complementary to instrumental technique as possible. Using motion tracking, we are able to efficiently map hand gestures to audio-visual responses. By also tracking biodata – such as EMG, EEG, blood flow and heartbeat – it is possible to establish stronger action-sound relationships [8] that produce deeper understandings of the dynamics and mechanisms embedded in these actions [9]. Systems using myograph data have emerged over the past two decades [10, 11], with a number of recent works utilising the Myo armband, demonstrating its reliability and appropriateness as an expressive gestural controller for musical applications [12,13].

2. THE SYSTEM

Developed through an iterative design cycle, MyoSpat's design utilises context-based, activity-centred and emphatic design approaches: interactions between users and mediating tools are positioned within the motives, community, rules, history and culture of those users [14].
MyoSpat 2 (outlined in Fig. 1) uses: (i) Myo armband as an input device to track hand-gestures; (ii) Myo Mapper[1] to extract and convert data from the Myo into Open Sound Control (OSC) messages; (iii) Pd with the ml-lib[2] externals for the gesture recognition and audiovisual signal elaboration and spatialisation; and (iv) Arduino for converting serial data into DMX signals that control lighting effects.

2.1. Interaction Design

MyoSpat 2's interaction design (IxD) draws on mimetic theories, embodied simulations [15] and metaphorical actions [16]. This approach facilitates a direct connection between the gestural interaction and the audio-visual feedback, as sound is treated like a tangible object that can be grasped and modified through continuous interactions. MyoSpat's IxD aims to (i) create a gesture vocabulary that enables interaction with the system through meaningful gestures for both performer and audience; (ii) produce a clear and strong connection between gestures and audiovisual feedback; and (iii) allow musicians to use the system through natural interactions. The term natural here refers to the contextu-

[1] http://www.balandinodidonato.com/myomapper/
[2] https://github.com/cmuartfab/ml-lib

alised interaction with physical and virtual objects conditioned from previous knowledge [17].

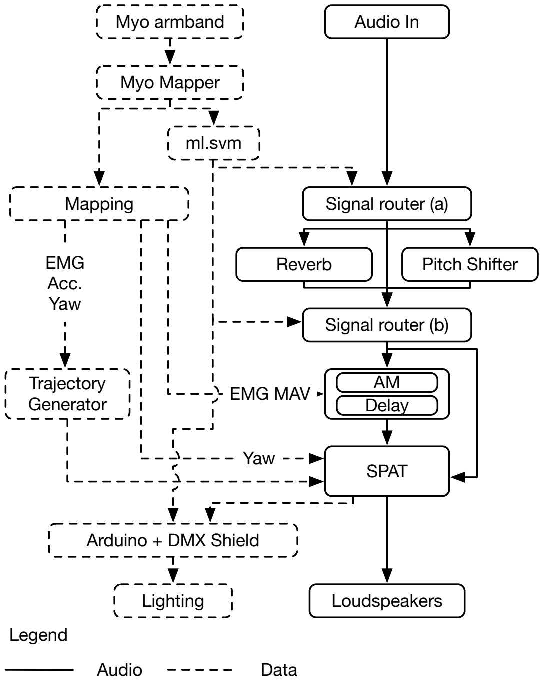

Figure 1: MyoSpat implementation.

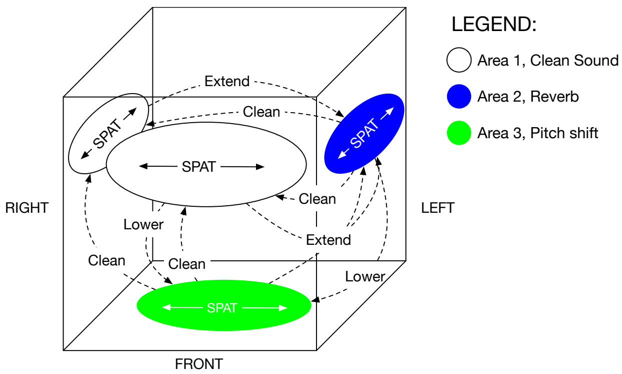

Figure 2: MyoSpat interactive areas, when the Myo armband is worn on the left arm.

The MyoSpat gesture library includes six gestures. The *clean* gesture is performed by orienting the arm towards the front of the body and/or inwards towards the chest (Fig. 2, area 1). It allows users to obtain a clean sound and to set the lighting system colour to white. The *extend* gesture is performed by orienting the arm outwards (Fig. 2, area 2), allowing users to apply a long reverb to the sound. This gesture sets the colour of the lights to blue. The *lower* gesture is performed by lowering the arm towards the ground (Fig. 2, area 3). It enables the user to pitch shift the sound one octave lower, setting the colour of the lights to green. The *crumpling* gestures are performed by hand movements that repeatedly contract the forearm's muscles, thus generating fluctuations in EMG data, as taken from previous experiments with sound design in mixed realities. This gesture applies amplitude modulation (AM) followed by delay effects to the audio signal. As the forearm's muscles become more contracted AM depth and delay feedback are increased, whilst delay time is shortened. Here the lights have a strobing effect, where strobe time and intensity are directly related to delay time and intensity. The *pointing* gesture allows the user to spatialise the sound around the audience. This gesture is performed by pointing at the location from where the sound should come from. This user interaction is an implementation of the same type of gesture described in [18]. The brightness of each light is adjusted relative to the spatial position of the sound. The *throwing* gesture involves a rapid movement of the arm, as if throwing an object, enabling the user to spatialise the sound through a circular trajectory. Once the gesture is detected by the system, the duration of the audio effect is determined by mapping arm acceleration and Myo's EMG mean absolute value, and trajectory direction is determined by the yaw value. The brightness of each light is dynamically adjusted in relation to the spatial position of the sound. This gesture is inspired from previous works on approaches controlling and visualising the spatial position of 'sound-objects' [19]. The relationship between sound and gesture refers to metaphor and mimetic theories [20] embedded in the movements performed when the hand moves towards each of the three areas, and not the pose assumed by the hand once it has reached one of these areas. When performing the *extend* gesture, users move their arm outwards, thus extending the area that the body covers within the space. We try to represent the expansion of the body within the 3D space extending the sound with a long reverb. We associate the *lower* gesture with a pitch shift one octave lower, as a result of lowering the arm.

2.2. Gesture recognition

Machine learning is used to recognise the *clean*, *extend* and *lower* gestures. Specifically, we use the Support Vector Machine (SVM) algorithm from the *ml-lib* library for Pd. Gesture identification probabilities are used for routing signals towards the pitch shifter, the reverb or to obtain a clean sound. *Crumpling*, *pointing* and *throwing* gestures are recognised and controlled through direct mapping.

2.3. Audio-lighting engine

The audio engine uses Pd patches from Integra Live's[3] Pitch Shifter and Reverb modules to control pitch shift and reverb respectively. MyoSpat's spatialiser uses the HOA library[4] for distributing the direct and reverberated sound through

[3] http://integra.io/integralive/
[4] http://www.mshparisnord.fr/hoalibrary/en/

the four speakers. A low-pass filter is added to simulate air absorption, and a delay line for establishing the time of arrival of the sound from each loudspeaker. Spatialisation parameters are controlled by mapping the yaw value through transfer functions.

The Pd patch controlling the lights maps and converts values into serial data. Data is sent to an Arduino with Tinkerkit DMX Master Shield, which then connects directly to the DMX lights.

3. EVALUATION

A direct evaluation approach [21] was adopted for assessing the gesture recognition process. Usability and user experience have been evaluated through a user study conducted at Berklee College of Music, Valencia Campus and at Integra Lab, Birmingham Conservatoire.

3.1. Methodology

Demographic and background information was collected through a semi structured interview. The system was then demonstrated to the participant, after which they could practice with and learn about the system for a period of 10 minutes. Next, the usability test asked participants to perform a series of sound manipulation tasks with the system. An audio file containing the sound of water flowing was used during this part of the study, with participants manipulating it using the system's gesture set. During the usability test, the gesture recognition accuracy was evaluated by monitoring when the gesture recognition algorithm successfully detected each user interaction; participants were interviewed through an informal, semi-structured interview to gather qualitative data about their perception of the audiovisual feedback. Lastly, participants were asked to fill a UEQ[5] to quantitatively assess their experience using the system.

3.2. Results

The user study was attended by nine participants (seven musicians and two non-musicians). Results from the UEQ reported that participants found MyoSpat 'highly attractive' and 'stimulating', finding it effective during the execution of specific task as well as free exploration. Most participants had prior experience with interactive music systems, potentially presenting a bias in our results. Results reported in Fig. 3 shows a high accuracy of the gesture recognition system. However, due to bugs in the implementation of the *crumpling* and *throwing* gestures mapping strategy, the associated sound manipulations resulted harder to perceive. Results from the UEQ (Fig. 4) report that MyoSpat is a highly attractive and stimulating system, and was found to be effective during the execution of specific tasks (pragmatic quality) as well as free exploration (hedonic quality).

[5] http://www.ueq-online.org/

During the user study, participants perceived a strong link between sound and gesture. They described light projections as enhancing the level of immersiveness and perception of the auditory feedback and its relation with hand gestures. Musically skilled participants described MyoSpat as being easy to use and incorporate into their performance practice. One of the participants (professional dancer) highlighted MyoSpat's potential application in dance performance. All participants considered the interaction with the audio file containing the sound of flowing water as natural and embodied. Interestingly, the lower gesture made participants interact with the sound as if they were submerging their hand in a tub filled with water. Other gestures included splashing and swirling water (see video[6]), demonstrating MyoSpat's potential to explore embodied interaction with sonic objects, in line with similar research [22].

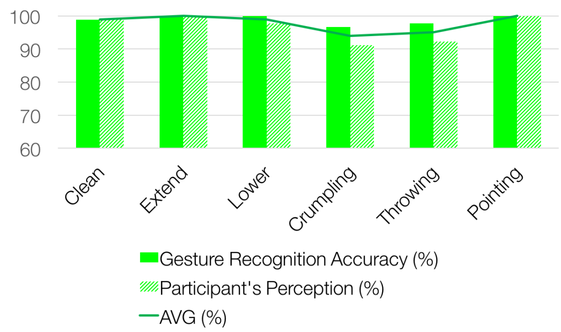

Figure 3: Usability study outcome.

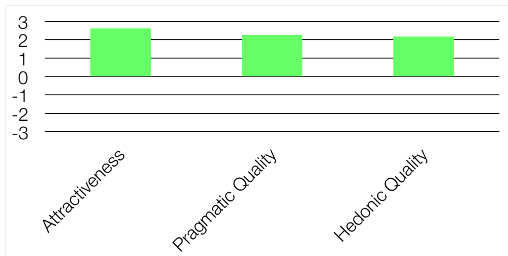

Figure 4: User Experience Questionnaire outcome.

4. CONCLUSIONS

We have presented MyoSpat, an interactive hand-gesture controlled system for creative audio manipulation in musical performance. Machine learning and mapping functions were successfully implemented to recognise a number of physical gestures, enabling audiovisual manipulations to be mapped to each one of them. The current and previous user studies demonstrate that the system can support musical improvisation and composition such as The Wood and

[6] https://vimeo.com/221800824

The Water by Eleanor Turner[7] and empowering users to explore a novel range of embodied interactions during the music making process. Results also demonstrate that the Myo armband does not restrict user movements, and that MyoSpat has the potential to be employed in different fields such as dance and Virtual Reality.

5. REFERENCES

[1] B. Di Donato, J. Dooley, J. Hockman, J. Bullock, and S. Hall, "MyoSpat: A hand-gesture controlled system for sound and light projections manipulation," in *International Computer Music Conference*, Oct. 2017, forthcoming.

[2] E. McNutt, "Performing electroacoustic music: a wider view of interactivity," *Organised Sound*, vol. 8, pp. 297–304, Apr. 2004.

[3] J. Bullock, L. Coccioli, J. Dooley, and T. Michailidis, "Live Electronics in Practice: Approaches to training professional performers," *Organised Sound*, vol. 18, pp. 170–177, July 2013.

[4] C. Lippe, "Real-Time Interactive Digital Signal Processing: A View of Computer Music," *Computer Music Journal*, vol. 20, p. 21, Dec. 1996.

[5] A. Cox, "Embodying Music: Principles of the Mimetic Hypothesis," *Society for Music Theory*, vol. 12, pp. 1–24, July 2011.

[6] L. Vainio, R. Ellis, and M. Tucker, "The role of visual attention in action priming," *The Quarterly Journal of Experimental Psychology*, vol. 60, pp. 241–261, Feb. 2007.

[7] A. I. Goller, L. J. Otten, and J. Ward, "Seeing Sounds and Hearing Colors: An Event-related Potential Study of Auditory–Visual Synesthesia," *Journal of Cognitive Neuroscience*, vol. 21, pp. 1869–1881, Oct. 2009.

[8] K. Nymoen, M. R. Haugen, and A. R. Jensenius, "MuMYO — Evaluating and Exploring the MYO Armband for Musical Interaction," in *International Conference on New Interfaces for Musical Expression*, June 2015.

[9] B. Caramiaux, M. Donnarumma, and A. Tanaka, "Understanding Gesture Expressivity through Muscle Sensing," *ACM Transactions on Computer-Human Interaction*, vol. 21, pp. 1–26, Jan. 2015.

[10] G. Dubost and A. Tanaka, "A Wireless, Network-based Biosensor Interface for Music," *International Computer Music Conference (ICMC)*, Mar. 2002.

[7] https://vimeo.com/204371221

[11] A. Tanaka and R. B. Knapp, "Multimodal Interaction in Music Using the Electromyogram and Relative Position Sensing," in *International Computer Music Conference*, pp. 1–6, Apr. 2002.

[12] C. Benson, B. Manaris, S. Stoudenmier, and T. Ward, "SoundMorpheus: A Myoelectric-Sensor Based Interface for Sound Spatialization and Shaping," in *International Conference on New Interface for Musical Expression*, pp. 332–337, July 2016.

[13] M. Weber and M. Kuhn, "Kontraktion," in *Audio Mostly*, pp. 132–138, ACM Press, Oct. 2016.

[14] P. Dourish, *Where the Action is: The Foundations of Embodied Interaction*. Bradford books, MIT Press, 2004.

[15] R. W. Gibbs, "Artistic understanding as embodied simulation," *Behavioral and Brain Sciences*, vol. 36, pp. 143–144, Mar. 2013.

[16] N. Schnell and F. Bevilacqua, "Engaging with Recorded Sound Materials Through Metaphorical Actions," *Contemporary Music Review*, vol. 35, pp. 379–401, Jan. 2017.

[17] B. Leibe, T. Starner, W. Ribarsky, Z. Wartell, D. Krum, B. Singletary, and L. Hodges, "The Perceptive Workbench: toward spontaneous and natural interaction in semi-immersive virtual environments," in *IEEE Virtual Reality 2000*, pp. 13–20, IEEE Comput. Soc, Mar. 2000.

[18] J. Streeck, C. Goodwin, and C. LeBaron, "Embodied Interaction in the Material World: An Introduction," in *Embodied Interaction, Language and Body in the Material World* (J. Streeck, C. Goodwin, and C. LeBaron, eds.), pp. 1–10, Cambridge, United Kingdom: Cambridge University Press, Nov. 2011.

[19] J. Bullock and B. Di Donato, "Approaches to Visualizing the Spatial Position of 'Sound-objects'," *Electronic Visualisation and the Arts*, pp. 15–22, Jan. 2016.

[20] P. Wigham and C. Boehm, "Exploiting Mimetic Theory for Instrument Design," in *International Computer Music Conference*, pp. 1–4, May 2016.

[21] R. Fiebrink, P. R. Cook, and D. Trueman, "Human model evaluation in interactive supervised learning," in *Conference on Human Factors in Computing Systems*, pp. 147–10, ACM Press, May 2011.

[22] E. O. Boyer, L. Vandervoorde, F. Bevilacqua, and S. Hanneton, "Touching Sounds: Audio Virtual Surfaces," in *IEEE nd VR Workshop on Sonic Interactions for Virtual Environments*, pp. 1–5, Mar. 2015.