MYtrOmbone: exploring gesture controlled live electronics in solo trombone performance

James Dooley The Open University james.dooley@open.ac.uk

ABSTRACT

MYtrOmbone is an interactive system that allows a trombonist to manipulate live electronic audio processing by analysis of the slide position engaged. The system is designed and developed at Integra Lab using an IMU prosthetic (Thalmic Labs Myo armband) worn on the right forearm as a control device. MyoMapper software translates armband data to OSC messages, and a Neural Network machine learning system detects trombone positions from the incoming OSC data. These positions are then mapped to audio signal processing in the Integra Live software that is controlled by the trombonist.

This paper presents the musical context that has led to the development of the system, outlining the way that the system works and its application within the first piece composed for it, *146 Lucina*. Finally, we suggest further work to develop and refine the system in the future.

1. INTRODUCTION

For the instrumentalist, performing a work that involves live electronics can be a daunting experience and one that can present a number of issues. Bullock et al [1] hypothesise that a common scenario involves a highly complex system, where the "performer and technology are 'divergent", resulting in a "disruptive effect on the musical experience". Often the instrumentalist may rely on an additional performer to operate the electronic part, transferring control of this away from the musician. The development of Augmented Musical Instruments (AMIs) can be seen as a way to transfer control back to the performer, allowing autonomy of audio processing through an extended interface. In the MYtrOmbone system, and resulting demonstration composition 146 Lucina for bass trombone, live electronics and gestural controller, we examine a way to transfer control of the live electronics back to the performer. Through the use of a commercially available gestural controller coupled with machine learning, we explore the development of an interface that is mindful of the musician's existing trombone technique and avoids the requirement to learn a complex set of new performance gestures. Consequently, in this paper we restrict ourselves to works for trombone and live electronics.

2. PROBLEM AND LEGACY SOLUTIONS

Whilst various commercial manifestations of wind controllers such as the AKAI EVI, AKAI EWI and Yamaha WX ranges

Copyright: © 2020 First author et al. This is an open-access article distributed under the terms of the <u>Creative Commons Attribution License</u> <u>3.0 Unported</u>, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Simon Hall Royal Birmingham Conservatoire simon.hall@bcu.ac.uk

have been in existence since the 1980s, there has been minimal development to date towards the creation of a commercial musical controller based around the trombone. In spite of this, the trombone has still presented itself as an interesting source with potential for the development of electronic augmentation for musical applications. Repertoire for solo trombone with both "tape" accompaniment, as well as electronic processing is intertwined with the European modernist traditions and American experimental traditions. Solo works that push the boundaries of the instrument have been developed by performer-protagonists such as Vinko Globokar, Benny Sluchin, Christian Lindberg, John Kenny, David Purser, Abbie Conant, George Lewis, James Fulkerson and Hilary Jeffrey, and this often has led to collaborations between trombone players and composer-technologists to realise a range of different incarnations of an electronically-augmented trombone.

In terms of musical HCI, audio signal analysis is an obvious method to use to analyse a performer's musical gestures, to programme actions and reactions, and to apply audio DSP-based processing. There are many examples of works that implement this approach to great and diverse musical effect. For example, trombonist and improviser George E. Lewis's Voyager software uses feature extraction to "guide the musical behaviour" of a computer "orchestra" that generates music in response to his playing style [2]. Similarly, Jamie Bullock's *Undercurrent* [3] uses the libxtract audio feature extraction library to analyse the trombonist's playing and control live electronic manipulations. The result produces an "obvious and immediate correlation between performance gesture–both musical and physical and the change in the electronics part" [3].

It is, however, notable that with the nature of the natural harmonic series, brass instruments have a particular facility to be able to produce certain pitches in multiple ways utilising different lengths of tubing, and in the case of the slide trombone, slide positions. For example, F3 may be produced in either first or sixth position; F4 in first, fourth or sixth. This can potentially add an additional creative dimension if one is able to obtain data from a performer about the slide position that is being utilised at any particular point. This is not easily obtainable purely by audio signal analysis.

In *I will not kiss your f.ing flag*, Lemouta, Stroppa and Sluchin [4] utilise a wireless, laser-based system to create an augmented trombone. Viewing the trombone's slide as a continuous controller and using IRCAM's WiSe box as intermediary hardware and MaxMSP as a host for the calibration system, all seven positions of the slide are tracked

and mapped to the control parameters of a range of audio processing that changes with score tracking throughout the piece. The system, and the piece, tries to create an intimate musical relationship by embedding the control elements of the musical work into the instrument itself, as opposed to it being a part of the software. However, there are a number of issues that surround the technology, including sensor reliability and accurate position tracking, as well as unsatisfactory battery life.

Neal Farwell's uSlide [5] similarly demonstrates a bespoke approach to developing a gestural controller (along with the eMouth and eMute) for *Rouse*, a work commissioned by Hilary Jeffrey. The uSlide uses ultrasonic transducers attached to the slide to detect its position. The sensor is tethered to an audio interface capable of operating at a sample rate of 96kHz to transmit and receive the sensor signal, though the connecting lead "sometimes tangles when the instrument is first picked up". Despite this, the author states that the system is reliable in performance, but it has not been systematically tested.

La Kitchen's now-obsolete Kroonde Gamma system facilitated sensor data capture and transmission by way of wireless interface, and allowed any sensor capable of dealing with 0-5Volt analogue voltage range to be attached to the system. The system was used in Flectar [6], in conjunction with a flection sensor secured to the antecubital space of the trombonist's right arm using a sport's compression sleeve. As the trombone slide moves upwards from full extension of the arm, in trombone position seven, to first position, the amount of flex in the flexion sensor is increased. By applying a calibration process prior to performance, smoothing is applied to the incoming data, allowing the angle of flex and corresponding voltage to generate useful OSC values within a prescribed range. This is then mapped within the MaxMSP software to identify the slide position that the trombone player is occupying at any point. Titled Flectar, the Latin word "to bend", the musical work has a clear link between physical gesture and resulting sound, with the position of the slide analysed in combination with FFT analysis of incoming attacks of notes. This is mapped onto the triggering of sonic events (samples), transformational DSP processing and spatialisation of the trombone sound in real time.

More recently, Chet Udell's eMersion Gesture Controlled System [7] presents a modular solution that allows musicians to control audio processing parameters. Sensors (eMotes) transmit wirelessly to a receiver connected via USB to a computer. Data is received, processed and mapped with the Digital Data Workstation software. Using the MIDI, OSC or DMX protocols, the system can output and interface with "virtually any software or hardware device", making it a flexible solution. The system was available via a Kickstarter campaign [8], though the last updates to this webpage were in 2014.

Whilst computer power and interactive technologies have both moved forwards significantly, the enabling technologies cited above are now effectively obsolete. There has been little in the way of development to facilitate comparable creative interactions between trombonist and computer. The development of musician-centred HCI solutions within Royal Birmingham Conservatoire's Integra Lab utilising Thalmic Labs Myo gestural control armband suggest potential for the development of a wearable technology solution that may be applied to the trombone player. By analysis of the outputs of the various parameters of the armband, it is quite viable through machine learning to extract trombone positions from the data transmitted. From a player's point of view, the wearable technology solution, by application of a minor prosthetic or small exoskeleton, feels quite natural, and is not too invasive of either the player or the instrument.

3. MYtrOmbone

3.1 Introduction

In MYtrOmbone, we approach the problem of performer controlled live electronics by combining a commercial controller device with machine learning and open-source software. The concept to the system itself is quite simple: It is designed to recognise when the trombone player is in a particular slide position, and on positive identification of that position, will then wait for a preprogrammed amount of time before making a crossfade to the newly selected preset state that correlates with this new slide position. Rather than treating the slide as a continuous controller, we use slide positions as a method of activating different audio processing hosted in Integra Live.

3.2 System Overview

The system used in MYtrOmbone relies on a number of components for capture, transmission and processing of gestural data; Figure 1 gives an overview of the system. Thalmic Lab's Myo Armband is used to track the slide arm movements of the bass trombonist. Placed around the trombonist's right forearm, data from the Myo armband is received by the Myo Mapper software¹. Here yaw, pitch and roll data are collected and scaled values (0-1) sent to Pure Data (Pd) via the Open Sound Control (OSC) messaging protocol. Pd is used as a hub to receive, send and convert data types between the other software components of the system. Pd sends the pitch, yaw and roll data to Wekinator² where three artificial neural network machine learning algorithms are used to identify slide positions 1,4 and 6 on the bass trombone. A total of 712 data samples were used to train position 1's (slide position 1) neural network, 892 were used to train position 2 (slide position 4), and 847 to train position 3 (slide position 6). Data from the yaw, pitch and role position was used to train each of the three neural networks. Wekinator returns this gesture recognition data to Pd, where the OSC data is then converted to MIDI. Using macOS's in-built IAC MIDI driver, Pd sends this data to Integra Live where it is mapped to a number of audio processing modules.

A single large diaphragm condenser microphone on a boom microphone stand is used to capture audio from the trombone. The use of a fixed position microphone ensures that the trombonist remains in a consistent position relative to the armband once calibrated. The audio processing in Integra Live produced a stereo output for this first prototype.

¹ http://www.balandinodidonato.com/myomapper/

² http://www.wekinator.org



Figure 1. MYtrOmbone system overview.

3.3 Interaction Design

Our approach to mapping gesture to audio processing aims to allow the performer to use a discrete number of slide positions to select between different audio processing presets. To create a system that is able to clearly differentiate between a number of different states, three of the seven slide positions are used that are harmonically related but physically quite different in terms of the pose of the right arm: position one (slide fully closed), position four (slide extended approximately midway, in line with the bell of the instrument), and position six (significant extension of the slide to near the stocking of the instrument). The trombone moves into a slide position and holds it for two or more seconds. The system then crossfades between the current selected processing chain to the new one over a period of one second. The system allows the position hold and fade times to be changed. Using this select-and-hold method of preset selection allows the performer to choose an effect chain, then continue playing divergent material whilst holding the selected effect in place.

Though Pure Data could have been used to host the audio processing, Integra Live was chosen for the first iteration of the system, primarily due to its ability to allow for rapid prototyping of audio interactions with ease. Its set of existing audio processing modules meant that, unlike in Pd, effects could be drag-and-dropped and parameters quickly mapped to incoming MIDI data. The audio processing comprises three harmonic resonator banks in parallel, tuned to the frequencies corresponding to the fundamental note associated with each slide position: Bb, G and F. Additional effects are then applied after each harmonic resonator, with each effects chain being activated when one of the three slide positions is detected. The position/effects chain relationships are as follows:

- **Position 1 Bb**: Resonant Harmonic Filter (117Hz) → Spectral Delay
- **Position 4 G**: Resonant Harmonic Filter (98Hz) → Pitch Shifter (down minor 3rd) → Spectral Delay
- Position 6 F: Resonant Harmonic Filter (87Hz) → Pitch Shifter (down perfect 4th) → Granular Delay

A video demonstrating the MYtrOmbone system can be viewed at the following address: https://youtu.be/e3y74R8Mfrw

3.4 146 Lucina: Composition and Performance The first piece to have been created for the system is *146 Lucina*, a relatively short (ca. 5minute) piece devised by Hall and Dooley to both test the first iteration of the system created, and to create a musical work that illustrates a quite direct correlation between the physical position of the trombone, the performed musical action, and the sonic output.

The processing chain outlined above is static in each of the states. As a position is detected, the harmonic fields related to the overtone content of each of the selected positions, the Myo effectively facilitating dynamic manipulation of the audio processing in real time. The musical intention is to allow the harmonic fields to come into focus and resonate as accompaniment to the material performed by the trombonist.

The title of the piece comes from the celestial body that resides in the main asteroid belt of our solar system, located between the orbits of Mars and Jupiter. The asteroid, and therefore the piece, is named after Lucina, the Roman goddess of childbirth who safeguarded the lives of women in labour, the name taken to mean "she who brings children to light".

The trombonist performs phoneme-based gestures taken from the name "Lucina" itself, and combines these with other pitched, semi-pitched (e.g. slap-tongue on the resonant tube) and unpitched gestures, that loosely act as a metaphor for themes of "mass", the movement from darkness to light, and the creation of life. The plunger mute acts as an additional filter, helping to move between phonetic and pitched gestures. As the piece progresses and crescendos, the ability to employ multiple positions to obtain a single pitch, (e.g. C3, F3, and F4) allows for the accompaniment to be triggered to move through various states whilst the melodic pitch movement becomes static.

4. DISCUSSION

The MYtrOmbone system was developed as a prototype/proof of concept of using a commercially available IMU device as a controller for musical HCI. This has demonstrated strong potential for further development as well as facilitating the creation of a compositional work. By employing machine learning, we were able to accurately detect the three selectable slide positions. Of the yaw, pitch and roll data, the yaw value exhibits the most significant change when the slide position is altered. However, yaw data drift in the Myo is a documented issue [9], and as such requires the trombonist to recalibrate the device at the beginning of every performance. Further work may include reconsidering the feature set used by the ANN to provide more accurate results, as well as considering the use of more accurate IMU devices.

The approach of using the slide position to activate a particular effects chain or "scene" proved successful and allowed the performer to "play" with differing sonic interactions. This did have its limitations in that once an effects chain was selected, no parameters of those effects could further be altered. Exploring different mapping strategies to find ways that combine the current "fixed" effects approach with "continuous" parameter changes could be further examined and would potentially allow the notion of virtuosity with the system to be explored. An example of this might be to map acceleration data to an effect parameter, enabling a specific scene to be selected, then utilising gestural activity to vary effect parameters within that scene.

Thalmic Labs, now rebranded as North, the parent company that developed the IP of the Myo IMU, as well as manufacturing and supporting the device itself and associated proprietary software, has recently undergone a significant change of corporate direction and strategy which has resulted in a cease of production of the Myo and limited continuing support. The impact of the vagaries of technology are not to be ignored, as can be seen in the examples cited earlier in this paper, with the Myo being locked behind proprietary licences. However, alternative hardware solutions are available, such as XIO Technologies' NGIMU, as well as the Arduino MKR WiFi 1010 paired with a MKR IMU Shield. Both could be explored as alternative IMU, with Arduino benefitting from being open, low-cost and having a large support community.

Finally, moving the software from a prototyping environment to a "production" environment would require reducing the number of components in the system to reduce the complexity of the message chain and operation. This could be as simple as using Pd for audio processing, thus removing the need for Integra Live, as well as exploring the Multilayer Perceptron implementation in Jamie Bullock and Ali Momeni's ml-lib machine learning externals [10]. Once again, Pd is open-source, backwards compatible software with a large community of users and this has its advantages when considering sustainability and longevity.

5. CONCLUSIONS AND FUTURE WORK

In many ways the system and piece act as a first proof-ofconcept of the MYtrOmbone, and within that scope may be claimed as a success. It is, in many respects, quite limited in this first iteration, and there is much that may be progressed further from both a technical and a musical point of view. It does, however, demonstrate that gestures based on existing right-hand trombone technique can be used to control audio processing reliably.

One most obvious development is to improve the accuracy of the position tracking with the machine learning. Being able to accurately identify all seven trombone positions, as well as accommodating the variations that occur quite naturally for a performer within the range of a particular position, would obviously open many additional musical possibilities. There is also scope for other performative gestures to be identified: once a position is selected, other movements that are made when in that position could change effects or other parameters. This may facilitate a greater theatricality in performance: for example, being able to identify motion in other planes, gestures with the left hand (utilising a second IMU device), or mimelike gestures without the instrument actually playing. Similarly, employing supplementary analysis as an addition to slide position identification, for example analysis of pitch tracking or other FFT-based audio feature extraction, may also facilitate greater subtlety of control and scope for musical development.

In 146 Lucina, the mapping of the positions to the processing employed makes for the creation of a particular soundworld that is quite constricted in terms of being able to maintain musical interest. It is a single scene that is explored, and this ultimately reaches a limit of creative application. Implementing changes to the processing through score following, development of multiple cues, or other ways of generating a creative structural progression is necessary if a more extended and substantial musical work is to be realised.

6. REFERENCES

- J. Bullock, L. Coccioli, J. Dooley and T. Michailidis, "Live Electronics in Practice: approaches to training professional performers," *Organised Sound*, vol. 18, no. 2, pp. 170-177, 2013.
- [2] G. Lewis, "Interacting with latter-day musical automata," *Contemporary Music Review*, vol. 18, no. 3, pp. 99-112, 1999.
- [3] J. Bullock, *Implementing audio feature extraction in live electronic music*, Birmingham: Doctoral dissertation, PhD thesis, Birmingham City University, 2008.
- [4] S. Lemouton, M. Stroppa and B. Sluchin, "Using the augmented trombone in I will not kiss your f. ing flag," in *Proceedings of the International Conference on New Interfaces for Musical Expression, 304–307*, IRCAM— Centre Pompidou, Paris., 2006.
- [5] N. Farwell, "Adapting the trombone: a suite of electroacoustic interventions for the piece Rouse.," in *New Interfaces for Musical Expression*, Paris, 2006.
- [6] L. Coccioli, "Flectar," 2005. [Online]. Available: https://www.lambertococcioli.com/flectar/. [Accessed 20th January 2018].
- [7] C. a. S. J. Udell, "eMersion Sensor-controlled Electronic Music Modules & Digital Data Workstation. In NIME (pp. 130-133).," in *New Interfaces for Musical Expression*, London, UK, 2014.
- [8] C. Udell, "eMersion Gesture Control System for Music Performance & More," 2014. [Online]. Available: https://www.kickstarter.com/projects/1865217699/emersi on-gesture-control-system-for-music-performa. [Accessed 2020].
- [9] K. Nyomen, M. Romarheim Haugen and A. R. Jensenius, "MuMYO — Evaluating and Exploring the MYO Armband for Musical Interaction," in *Proceedings of the International Conference on New Interfaces for Musical Expression, 215–218.*, Louisiana State University, Baton Rouge, Louisiana, USA, 2015.
- [10 J. Bullock and A. Momeni, "Ml. lib: robust, crossplatform, open-source machine learning for max and pure data.," in *Proceedings of the International conference on New Interfaces for Musical Expression (pp. 265-270).*, 2015.