Doctoral Colloquium: The Application of Established Gestural Languages in the Control Mappings of Free-hand Gestural Musical Instruments

Dom Brown

University of the West of England
Bristol, UK.

dom.brown@uwe.ac.uk

Abstract. The "mapping problem" is a long standing issue in the development of digital musical instruments, and occurs when an instrument's musical response fails to reflect the performer's gestural actions. This paper describes ongoing doctoral research that seeks to address this issue by studying the existing gestural languages of Soundpainting and American Sign Language in order to influence the control mappings of free-hand gestural musical instruments. The research seeks to contribute a framework of mapping strategies influenced by these gestural languages, as well as developments in the field of gesture recognition algorithms and novel evaluation methods for free-hand gestural musical instruments.

Keywords: Doctoral Colloquium, Digital Musical Instruments, Gesture Recognition, Gesture Communication, Sign Language, Soundpainting, Mapping.

Introduction

While traditional musical instruments generate a musical response that is entirely dependent on the interaction between the musician, the instrument and the laws of physics, digital musical instruments are entirely freed from these restraints, and can instead map any conceivable input to any conceivable sonic output. However, without limitations these instruments lack cohesion and definition, which, once their initial novelty has waned, can disengage and frustrate audiences and performers alike. This doctoral research seeks to define a framework for free-hand gestural instruments by studying and learning from the sign languages American Sign Language and Soundpainting.

The research will expand on the latest research on gestural interactive music systems by studying and learning from long-established gestural languages for music improvisation and performance. The work will unite these gestural languages designed for human-human communication with novel developments in machine learning to enable fluid and transparent human-computer interaction for music. The research will contribute to the rapidly emerging and exciting field of computational gestural interaction with an overarching goal to develop a comprehensive system for musical interaction that incorporates an effective, nuanced gestural vocabulary that is inspired by pre-existing gestural languages.

The work will focus on asking: What can be learnt from drawing influence from gestural sign languages in the design of mapping strategies for gestural digital musical instruments?

This question can be further expanded as:

- What can be drawn from gestural language systems to achieve sophisticated musical expression and virtuosic performance?
- How can sign languages influence the encoding of musical meaning in gestural control?
- What methods of evaluating approaches and mappings of gestural interaction should be used, and what are the human factors that need to be considered?

The research will work to use gestural language systems that have been subject to decades of development and refinement to explore the issue of mapping freehand gestures to musical output, a long-standing issue in the development of novel digital musical instruments, and otherwise known as the "mapping problem". The research will take an empirical approach, with novel methods of analysis determining the effectiveness of the application of gestural languages to music. The research will focus on the languages of Soundpainting and American Sign Language, and will work towards developing a framework for effective gestural mapping for digital musical instruments.

Background

Gestural human-computer interaction is becoming increasingly prominent, with the development of low cost sensor technologies alongside the abundance of computational technology that enables the real-time execution of sophisticated gesture recognition algorithms. This has enabled the development of Virtual Musical Instruments (VMIs) (Mulder 1994) or Digital Musical Instruments (DMIs) (Miranda and Wanderley 2006), allowing musicians to interact with computers in novel ways, creating entirely new methods of composing and performing music.

Digital musical instruments provide the valuable opportunity to separate the site of physical interaction from the sound producing parts of a musical instrument (Winkler 1995). This process allows a musician using such a device to be able to harness a greater and more varied degree of sonic expression than that of acoustic instruments, as sound production relies not on physical but on virtual constraints. However, this decoupling of a musician's input and a musical instrument's output is unnatural for both musicians and their audiences, and can leave both parties feeling frustrated if the computer's musical output does not appear to relate to the musician's gestural input, resulting in what has been called perceptual disconnection (Nakra 2002) or "the mapping problem" (Wessel and Wright 2002).

Digital Musical Instrument mapping strategies are usually categorised via the relationship between input and output parameters, such as "one to one", where individual inputs are mapped to individual outputs, with no inter-parameter relationships; "one to many", where one input will be used to control multiple outputs; "many to one", in which multiple inputs will have varying effects on a single output; and "many to many", in which multiple input parameters will have varying degrees of control on a variety of outputs (Rovan et al. 1997) (Hunt and Kirk 2000). Studies have been conducted (Hunt, Wanderley, and Paradis 2003) (Hunt, Wanderley, and Kirk 2000) that show that multi-parametric interfaces with "many to many" mapping strategies provide the most engaging experiences for users. The key to intuitive expression in a gestural interface lies in its mapping (Dobrian and Koppelman 2006) and creating complex, nuanced controls that allow a musician to interact with a computer with the same level of detail and sophistication as they would interact with an acoustic instrument. However, complex mappings have the potential to fall into the mapping issues previously discussed.

The issues with devising a sophisticated mapping that retains its sense of embodiment has been previously addressed; Fels et al. (Fels, Gadd, and Mulder 2002) refer to this quality as the "transparency" of the instrument's mapping, and they describe how deriving the mapping of a new instrument from that of an existing instrument, or more succinctly, using a metaphor of an existing instrument, can increase the transparency of the instrument's mapping, enabling both the performer and the audience to understand the expressive qualities of the instrument. Wessel and Wright (Wessel and Wright 2002) also explore the application of metaphor in designing mapping strategies, drawing on the linguistic work of Lakoff and Johnson (Lakoff and Johnson 1980) and common human-computer interaction metaphors such as Drag and Drop.

However, those that wish to explore beyond the limitations of existing instruments are faced with new problems, especially systems that use spatial gestures to control musical output. With no standardised guidelines for designing mappings, the connection between gestural actions and musical responses often represent entirely arbitrary choices on the part of the designer (Nakra 2002). While this may result in a perfectly usable instrument, it requires players to learn the specific gestures for that system, and the designer is often unaware as to whether they have made suitable choices. Also, audiences that have no prior experience with a new DMI can struggle to perceive the connection between the performer's actions and the resulting sonic response. An audience's perception of a musical instrument is of as vital importance as that of the performer, and their understanding of intention and error in performance is a crucial factor for assessing a performer's skill and success (Fyans, Gurevich, and Stapleton 2010).

The field of recognising gesture in computing systems is dominated by machine learning techniques, most notably Hidden Markov Models (Yamato, Ohya, and Ishii 1992) and Neural Networks (Modler 2000). These techniques have been adapted and improved upon for recognising gestures, notably with the application of Dynamic Time Warping, which analyses a gesture as a series of successive states, capturing the repeatability and variability of a given gesture. This technique then "time-collapses" a prototype curve so that the velocity of a gesture becomes irrelevant in its recognition (Bobick and Wilson 1997).

The recognition of a gesture usually results in a discrete result: which gesture was performed and whether it was performed successfully. While this is fine for computer systems that only require specific triggering of events, it does not often suit the continuous nature of music performance. Existing music systems, such as SoundGrasp (Mitchell and Heap 2011), use this discrete technique to trigger one-shot events while tracking motion to control continuous parameters.

Research

This research seeks to address the issue of perceptually disconnected control mapping by learning and drawing influence from the established gestural language of American Sign Language, and apply these strategies in a gesturally controlled musical instruments using computational gesture recognition techniques.

Soundpainting (Thompson 2006) is intended to be used by conductors to control the musical performances of an ensemble, yet the languages instructional gestures share many similarities with how one would control the parameters of a synthesizer or sequencer, including directions and gestures for controlling pitch, dynamics, timbre and note triggering.

As well as using language systems that are designed specifically for musical interaction, musical contexts can be drawn from more generic sign languages by using the meaning that already exists within them (for example, signs for big and small in American Sign Language could be used to manipulate amplitude). These gesture languages have been developed over many decades, with (in some cases) many generations of speakers refining and shaping the languages into meaningful, ergonomic systems designed to convey their meaning effectively and efficiently. Using and deriving from these languages and systems could provide a much more effective means of gestural control than other more arbitrary methods employed in digital musical instruments.

Both Soundpainting and ASL raise concerns about expressive control. Both languages use gestures to convey discrete meanings, of which variation plays little part. However, there are exceptions; in ASL, certain adjectives are intensified with the speed of performance. This is highly apparent in the gesture "slowly", in which the dominant hand is slid from the palm to the elbow of the non-dominant arm. If the sign is performed quickly, the sign, slightly paradoxically, becomes "very slowly".

Instead of using the exact meaning encoded in ASL, the methods used to encode that meaning can be used to design mapping strategies. For example, meaning can be encoded in ASL's gestural signs either arbitrarily or iconically (Taub 2001). In semiotics, an arbitrary sign is one that has no perceived connection to the concept it signifies, while an iconic sign resembles the concept it represents (Chandler 2007). ASL signs that encode meaning through iconicity have a physical resemblance to their concepts, and can often be recognised by non-signers (Lieberth and Gamble 1991). The language also uses iconicity to encode conceptual metaphors (Taub 2001), and this research intends to apply this technique to encode musical conceptual metaphors, such as the spatial metaphor MORE IS UP, which is often associated with pitch (Wong 2011).

This research will make use of continuous real time gesture following, building on recent work (Bevilacqua et al. 2010) (Caramiaux et al. 2014) that uses probabilistic methods to estimate which gesture a user is exercising as it is being performed. Using continuous gesture following as opposed to motion tracking allows computer systems to differentiate between when a performer is using a performance gesture and when they are not, and refrain from interpreting trivial movements (such as itching one's nose) as musical interaction.

Methodology

Various frameworks exist for evaluating digital musical instruments (O'Modhrain 2011), most notably the method proposed by Wanderley and Orio (Wanderley and Orio 2002), which takes methodologies for evaluating human computer interaction and applies them to a musical context. Although the framework was designed to analyse devices and interfaces, their basic principles apply to many aspects of the digital musical instrument design process and can be adapted to evaluate freehand control mappings. They emphasise a set of contexts to aid in the analysis of different devices, five of which apply to hand-based gestural control of music: Note-level control (directly manipulating synthesiser parameters); Score-level control (controlling features of sequencers); Sound processing control (manipulating audio effects or spatialisation); Traditional HCI Contexts (drag and drop, scrubbing, navigation); and Interaction with Multimedia Installations (triggering audio without needing prior musical skill). Each context reflects a level of extraction between the user's actions and the computer's response. Wanderley and Orio stress that they do not represent fixed classifications, but it will be important to define the contexts in which the gestural mapping will be used when studying their effectiveness.

Wanderley and Orio also provide guidelines for developing musical tasks, which work as a benchmark to analyse the capabilities and usability of a controller in a musical context. The main features of these tasks are: "Learnability"; "Explorability"; Feature Controllability; and Timing Controllability.

Jordà also suggests criteria for the evaluation of digital musical instruments (Jordà 2004). He proposes that there are three areas of diversity (macro, mid and micro) that determine the nature of a newly developed instrument. His framework could be useful for analysing the characteristics that a gestural system would give to a digital musical instrument.

Their are also many evaluation methods intended to analyse DMIs from the perceptions of spectators. Barbosa et al. outline five aspects of comprehension that can be tested for: cause, effect, mapping, intention and error (Barbosa et al. 2012). Similarly, Fyans et al. (Fyans, Gurevich, and Stapleton 2009) explore a method for evaluating a spectator's perception of error in performance, highlighting five issues raised by Bellotti et al. in general human-computer interaction (Bellotti et al. 2002): address, attention, action, alignment and accident; and re-contextualising them for digital musical instruments.

As well as making use of these methods, this research seeks to develop new methods of evaluation to analyse the efficacy of the application of techniques derived from gestural language in music control. This will draw from a broad spectrum of methods, such as adapting traditional human-computer interaction evaluation methods (Wanderley and Orio 2002) and those designed for digital musical instrument analysis (O'Modhrain 2011). These will cover both audience and performer perspectives, and range from qualitative studies, such as discourse analysis (Stowell, Plumbley, and Bryan-Kinns 2008) and heuristic evaluation methods (Nielsen 1994), to more quantitative methods, such as using reproductive musical tasks to compare the accuracy of a user's performance over time (Hunt and Kirk 2000).

Contributions

The contributions this doctoral research intends to make includes a framework of mapping methods to be used in free-hand gestural digital musical instruments, as well as a series of rules and strategies for designing new mappings, influenced by what can be learnt from gestural languages. It also seeks to contribute further developments in the field of gestural recognition algorithms, as well as a new in depth analysis method for evaluating free-hand gestural digital musical instruments.

References

Barbosa, Jerônimo, Filipe Calegario, Veronica Teichrieb, Geber Ramalho, and Patrick McGlynn. 2012, May. "Considering Audience's View Towards an Evaluation Methodology for Digital Musical Instruments." *Proceedings of New Interfaces for Musical Expression (NIME)*. Ann Arbor, MI, USA.

Bellotti, Victoria, Maribeth Back, W Keith Edwards, Rebecca E Grinter, Austin Henderson, and Cristina Lopes. 2002, April. "Making Sense of Sensing Systems: Five Questions for Designers and Researchers." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Minneapolis, MN, USA, 415–422.

Bevilacqua, Frédéric, Bruno Zamborlin, Anthony Sypniewski, Norbert Schnell, Fabrice Guédy, and Nicolas Rasamimanana. 2010. "Continuous Realtime Gesture Following and Recognition." In *Gesture in Embodied Communication and Human-Computer Interaction*, 73–84. Springer Berlin Heidelberg.

Bobick, Aaron F, and Andrew D Wilson. 1997. "A State-Based Approach to the Representation and Recognition of Gesture." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (12): 1325–1337.

Caramiaux, Baptiste, Nicola Montecchio, Atau Tanaka, and Frédéric Bevilacqua. 2014. "Adaptive Gesture Recognition with Variation Estimation for Interactive Systems." *ACM Transactions on Interactive Intelligent Systems (TiiS)* 4 (4): 18.

Chandler, Daniel. 2007. *Semiotics: the basics*. Routledge.

Dobrian, Christopher, and Daniel Koppelman. 2006, June. "The 'E' in NIME: Musical Expression with New Computer Interfaces." *Proceedings of New Interfaces for Musical Expression (NIME)*. Paris, France.

Fels, Sidney, Ashley Gadd, and Axel Mulder. 2002. "Mapping Transparency Through Metaphor: Towards More Expressive Musical Instruments." *Organised Sound* 7 (2): 109–126.

Fyans, A Cavan, Michael Gurevich, and Paul Stapleton. 2009, April. "Spectator Understanding of Error in Performance." *CHI'09 Extended Abstracts on Human Factors in Computing Systems*. ACM, Boston, MA, USA, 3955–3960.

———. 2010, June. "Examining the Spectator Experience." *Proceedings of New Interfaces for Musical Expression (NIME)*, Volume 10. Sydney, Australia, 1–4.

Hunt, Andy, and Ross Kirk. 2000. "Mapping Strategies for Musical Performance." *Trends in Gestural Control of Music* 21:231–258.

Hunt, Andy, Marcelo Wanderley, and Ross Kirk. 2000, September. "Towards a Model for Instrumental Mapping in Expert Musical Interaction." *Proceedings of the International Computer Music Conference*. Berlin, Germany, 209–212.

Hunt, Andy, Marcelo M Wanderley, and Matthew Paradis. 2003. "The Importance of Parameter Mapping in Electronic Instrument Design." *Journal of New Music Research* 32 (4): 429–440.

Jordà, Sergi. 2004, November. "Digital Instruments and Players: Part II–Diversity, Freedom and Control." *Proceedings of the International Computer Music Conference*. Miami, FL, USA, 706–710.

Lakoff, George, and Mark Johnson. 1980. *Metaphors we live by*. University of Chicago press.

Lieberth, Ann K, and Mary Ellen Bellile Gamble. 1991. "The Role of Iconicity in Sign Language Learning by Hearing Adults." *Journal of Communication Disorders* 24 (2): 89–99.

Miranda, Eduardo Reck, and Marcelo M Wanderley. 2006. *New digital musical instruments: control and interaction beyond the keyboard*. AR Editions, Inc.

Mitchell, Thomas, and Imogen Heap. 2011, June. "SoundGrasp: A Gestural Interface for the Performance of Live Music." *Proceedings of New Interfaces for Musical Expression (NIME)*. Oslo, Norway.

Modler, Paul. 2000. "Neural Networks for Mapping Hand Gestures to Sound Synthesis Parameters." *Trends in Gestural Control of Music*, vol. 18.

Mulder, Axel. 1994, August. "Virtual Musical Instruments: Accessing the Sound Synthesis Universe as a Performer." *Proceedings of the First Brazilian Symposium on Computer Music*. Minas Gerais, Brazil, 243–250.

Nakra, Teresa Marrin. 2002. "Synthesizing Expressive Music Through the Language of Conducting." *Journal of New Music Research* 31 (1): 11–26.

Nielsen, Jakob. 1994. "Heuristic Evaluation." *Usability Inspection Methods* 17 (1): 25–62.

O'Modhrain, Sile. 2011. "A Framework for the Evaluation of Digital Musical Instruments." *Computer Music Journal* 35 (1): 28–42.

Rovan, Joseph Butch, Marcelo M Wanderley, Shlomo Dubnov, and Philippe Depalle. 1997, October. "Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance." *Proceedings of Kansei - The Technology of Emotion Workshop*. Genoa, Italy, 3–4.

Stowell, Dan, Mark D Plumbley, and N. Bryan-Kinns. 2008, June. "Discourse Analysis Evaluation Method for Expressive Musical Interfaces." *Proceedings of New Interfaces for Musical Expression (NIME)*. Genova, Italy, 81–86.

Taub, Sarah F. 2001. *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge University Press.

Thompson, Walter. 2006. *Soundpainting: the art of live composition. Workbook I*. New York, NY, USA: Walter Thompson.

Wanderley, Marcelo Mortensen, and Nicola Orio. 2002. "Evaluation of Input Devices for Musical Expression: Borrowing Tools from HCI." *Computer Music Journal* 26 (3): 62–76.

Wessel, David, and Matthew Wright. 2002. "Problems and Prospects for Intimate Musical Control of Computers." *Computer Music Journal* 26 (3): 11–22.

Winkler, Todd. 1995. "Making Motion Musical: Gesture Mapping Strategies for Interactive Computer Music." *Proceedings of the International Computer Music Conference*. Banff Centre for the Arts, Canada, 261–264.

Wong, Peter Ho-Kin. 2011. "Conceptual Metaphor in the Practice of Computer Music." MSc, Mills College.

Yamato, Junji, Jun Ohya, and Kenichiro Ishii. 1992, June. "Recognizing Human Action in Time-Sequential Images Using Hidden Markov Model." *Proceedings of the 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 379–385.