PROTOCOL AND MODULARITY IN NEW DIFFUSION SYSTEMS

Bridget Johnson
New Zealand School of Music
Victoria University of Wellington

Michael Norris
New Zealand School of Music
Victoria University of Wellington

Ajay Kapur
School of Engineering and Computer Science
Victoria University of Wellington

ABSTRACT

This paper proposes a move to further modularity in spatialisation systems. The focus is specifically on systems that are driven by the design of new diffusion user interfaces. The authors recognise a need for system specificity based on the communication needs of different user interface designs. As examples the paper presents two case studies of new user interfaces and how they are each integrated into fully modular real-time spatialisation systems. The paper also recognises a need for intuitive and shared protocol formats in order to achieve modularity, some examples of this are proposed.

1. INTRODUCTION

In the past 15 years a wide range of new spatialisation systems have emerged, many of which exhibit the potential to be segmented. Full spatialisation systems are made up of many parts from the user interface through the spatialisation algorithms and finally the audio drivers and speaker systems. One common trend amongst new systems is to have them be made up of a series of modules that allow integration between modules of varying systems. In order for this to be achieved all systems must implement the same communication protocols. This paper will discuss two options for implemented modularity and protocols that are currently under development at the author's institution. The solutions proposed in this paper are aimed at systems that place a strong importance of user interface design, and the incorporation of physical user interfaces into wider spatialisation systems.

The paper first discusses movements and systems in the paradigm of diffusion performance that have led to a move towards modularity in spatialisation systems. It then goes on to discuss new practice driven research that is currently being conducted by the authors, and how the desire for modularity has influenced the design of two new spatialisation projects; tactile.motion and Chronus_2.0, in different ways. The fourth section continues to use these two new systems as examples for discussion about the need for universally accepted protocol in order to ensure the modular aspects of diffusion systems are viable for the future. The paper is concluded with a look to the future of spatialisation systems for performance practice.

2. RELATED WORK

For the most part developments in spatialisation performance (other than algorithm developments) have remained specific to the institutions developing them. This is largely due to the fact that each spatialisation performance system depends on the hardware available.

A number of the first speaker orchestras including the GRM Acousmonium (Desantos, Roads, and Bayle 1997), BEAST (Harrison 1999), and the Gmebaphone (Clozier 2001) were able to travel and therefore had to be easily configurable for varying performance venues. Traveling with these systems involved moving the entire system, from computers to amplifiers and mixing desks as well as any number of loud speakers. The physicality and cost of moving such immense amounts of hardware may have limited the potential distances these orchestras were able to travel. With increasing technological advancements in the 1990s, artists began to work on the development of customized software that would allow advanced spatial trajectories to be performed, as well as decreasing the amount of hardware necessary.

In an attempt to counter some of the issues with traveling, the M2 diffusion system (Moore, Moore, and Mooney 2004) was developed in the mid 2000s with a focus on the capability of being compactable and robust enough to take on a plane. The Motu24i/o interface and relatively small fader console can be packed into a flight case for traveling. The flight case then opens up at one side allowing all required connections to be made without removing the hardware from the flight case.

One of the first significant attempts at making certain parts of a diffusion system more easily compatible with others, was to release the software that drives a system. As mentioned for the most part spatialisation systems, including their software, are custom-built for an institution. Once the systems were being driven by custom-built software it became very easy to release and share the tools they had created with the wider international diffusion community. Birmingham released BEASTMulch, a system which allows control of multi-channel sound diffusion with a range of spatialisation algorithms and an ability to communicate with standard MIDI controllers as well as custom OSC fader boards (Harrison and Wilson 2010). This system and others like it meant that institutions were able to incorporate the software into use with their own speaker configurations.

Significant developments in spatialisation algorithms such as Vector Base Amplitude Panning (Pulkki 1997), Ambisonics (Gerzon 1977) and Wave Field Synthesis have also encouraged an increased modularity in diffusion systems. These techniques for spatialisation are featured in the more traditional diffusion systems previously mentioned, as well as range of new diffusion systems including the SoundScape Renderer (Bredies et al. 2006), ReSound (Mooney and Moore 2008), and The Allosphere (Hollerer, Kuchera-Morin, and Amatriain 2007).

A number of new spatialisation systems have recognized a need for published open protocol across spatialisation systems. One common movement is towards the spatial scene descriptors. The SpatDIF (Spatial Sound Description Interchange Format) (Peters 2008) implemented in the HOLO-Edit system (Peters et al. 2009) proposes a format for scene description and also outlines a modular approach to design of a full spatialisation system.

The SoundScapeRenderer, an example of a user interface driven spatialisation system, also proposes a scene description format the Audio Scene Description Format (ASDF) (Geier et al. 2007).

The 2008 International Computer Music Conference held a panel discussion about the need for a recognized and implemented spatial scene description format proposed and lead by authors of the systems mentioned above (Kendall, Peters, and Geier 2008).

3. MODULARITY

The vast majority of the first author's research has been based around the design and development of new performance interfaces for diffusion practice. In designing new interfaces, one major consideration is how they will be integrated in to the larger diffusion system. One of the design goals for these interfaces is to have them not be system specific so that performers may choose to use the interfaces with any spatialisation systems they have access to. However, even so there is often a need for full systems to be developed for prototypes of new interfaces, observations made whilst developing these systems have influenced the majority of research presented in this paper.

Two user interfaces are currently under development, each of them require different support from the wider system and as such are not fully modular, however the intention is to make the two interfaces as interchangeable as possible whilst still adhering to their specific design goals.

3.1. Case Study 1: Tactile.motion

Figure 1. The tactile.motion graphical user interface. Each coloured circle represents a draggable audio object.

tactile.motion (shown in Figure 1), is an iPad application designed to replace the mixing desk as a user interface for diffusion performance. The original concept was to port the authors touch table application tactile.space (Johnson and Kapur 2013) to iPad, however the advantageous features of the iPad as a user interface have allowed significant development of new features. The feature most relevant to this paper is the incorporation of Apple's Zero Configuration Protocol (known as Bonjour). This allows a host (the computer) to broadcast itself as an available OSC service via Bonjour, and the iPad can present the user with a list of all available services so they only need select the appropriate host from the list.

The main advantage of this type of communication system is that inexperienced performers can connect to the network without any knowledge of the set up procedure. There is no need for input of an IP address or port number, the system handles the passing of data when it resolves the broadcasted service. This communication is outlined diagrammatically in Figure 2.

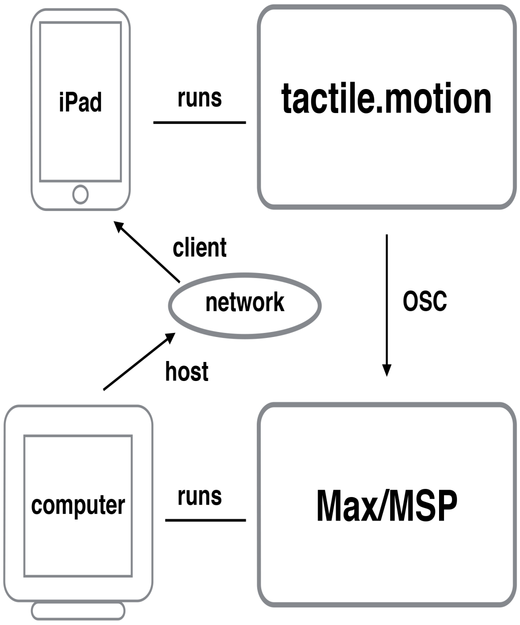

Figure 2. The tactile.motion networking and communication system overview.

On start-up the application links itself to the desired host therefore there is no permanent connection to any specific max Patch or audio software or even any specific computer. This means the application could be used by any spatialisation system.

Other advantages of the system are its stability and remarkably short set-up time. In a concert setting, the computer hosts an ad-hoc network and the performer can select said network from the system preferences on the iPad as they would any Wifi service. This action is very familiar to the majority of performers.

One of the main advantages of the system could also be seen as its disadvantage. The tactile.motion application only searches for services broadcast via as available to receive OSC over Bonjour, thus if one wanted to use the application with other software they would need to implement ZeroConf. For most custom built audio software the implementation of a broadcasting service is very simple and can be done with only minor adjustments to examples. In the Max Patch that drives the audio for tactile.motion we use the OSCBonjour externals and example patch[1]. For tactile.motion to communicate with another audio driver an understanding of the OSC protocol would also be necessary. This will be discussed in section 4.

3.2. Case Study 2: Chronus_2.0

The Chronus series has slightly different design specifications and as such requires modified communication protocols. The current version, Chronus_2.0 (shown in Figure 3) features a rotating disc with a potentiometer on top. The rotation of the disc provides an angle for phantom source positioning and the fader provides the distance. There are significant technical design features which allow this rotation whilst keeping the potentiometer oriented intuitively, they are outlined in (Johnson, Norris, and Kapur 2014).

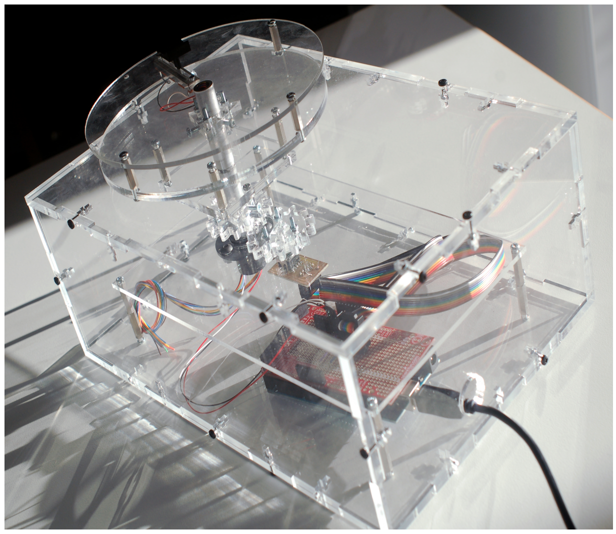

Figure 3. The Chronus_2.0 User Interface

The Arduino Uno microcontroller connects the Chronus_2.0 physical controller to the computer via a USB cable. Therefore there is no need for any wireless communication to occur, thus no broadcasting of Bonjour services. This represents the fundamental difference in communication requirements between the two interfaces. Once the computer receives the data there is an attempt for the interfaces communication streams to remain as similar as possible.

The microcontroller sends serial data with the radial and distance positions. This data is processed in a custom built Processing sketch, the main job of which is to convert the serial to OSC messages. At this point there are a few different options. The OSC messages can be read by a second Processing sketch implementing Pulkki's VBAP principals (Pulkki 1997), this sketch then sends the processed data as weighted gain factors to be unpacked and implemented by a very simple Max patch that drives the audio. This system is outlined in Figure 4.

[1] http://recherche.ircam.fr/equipes/temps-reel/movement/muller/index.php?entry=entry060616-173626

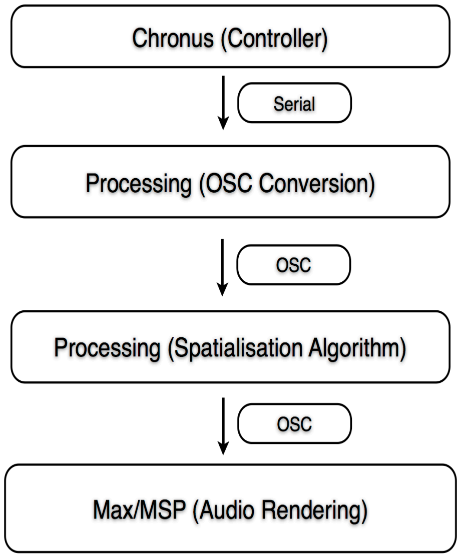

Figure 4. The Chronus spatialisation system separated into modules.

This process best outlines the intended modularity. The system is divided into a controller, a conversion process, an implementation of a real-time spatialisation algorithm and finally the audio rendering. There is a lot of flexibility within this system. For example it is very common for the spatialisation algorithm to be incorporated within the audio rendering application, this can easily be achieved by using a VBAP (or other algorithm) external for Max. Likewise at any point the OSC messages are not confined to Max, but can be received in by any application. The separation of each stage to implement a single modification allows the system to remain extremely modular and means that any single part could be easily incorporated into another spatialisation system. This flexibility is one advantage of such a modular system.

Again the advantage could be seen as the disadvantage. It could be argued that this separation over complicates the systems, and it is true that the same could be achieved with only two modules; the controller and the audio driver. One of the design goals for Chronus_2.0 was to build an interface that would encourage live spatialisation performance outside the diffusion paradigm. It is hoped that performers from the wider live electronic performance community will embrace the Chronus controller and integrate it into their established live set up. The current modularity of the system ensures that this integration could be a smooth one even for performers without experience with spatialisation algorithms. A simple Max for Live patch would also allow communication with Ableton Live as is a common choice of DAW for many live electronic performers.

4. PROTOCOL

The Open Sound Control (OSC) protocol (Wright, Freed, and Momeni 2003) was chosen because of its ease of implementation especially across different platforms and devices, and its flexibility and range. As mentioned, the configuration of the OSC messages sent from one module to the next are extremely important for ensuring the modularity of the system, and allowing for communication between systems. In designing this protocol there has also been an attempt to remain within the most standard input configurations for popular spatialisation algorithm externals so these may be easily incorporated. The messages sent from both interfaces are configured as polar coordinates oriented around the centre of the speaker array. Polar coordinates are a commonly required input data for spatialisation rendering tools including the custom built Processing sketch discussed in section 3.2 and the VBAP external used in the current Max/MSP patch for tactile.motion. As both interfaces are intended for multi-channel diffusion they also send an identification tag that directs messages to control the designated audio stem. An example of this protocol is shown in Figure 5.

Protocol
```
object i/f/f
object number/distance/theta
```
Example Message
```
object 3/3.25/2.91
```

Figure 5. Example of protocol output from spatialisation controllers.

Currently there is no 3-dimensional speaker array available for use with these spatialisation systems and thus has not been a focus of this research, however this may be implemented in the future and thus the potential need for a fourth data stream, the azimuth was also considered in designing the OSC protocol. The adaptation to the third dimension would only require that message to be added to the current data, rather than requiring any new protocol design.

In Chronus_2.0 the second module merely converts serial data messages into the OSC messages. If a separate spatialiatision algorithm module is used before the final audio driver, the most common form of messages for it to output is weighted gain factors. For earlier versions of tactile.motion this was the case. The OSC messages for weighted gain factors are configured with the stem (or object) number, the speakers number, and the gain as a floating point value. In a similar way to the positional coordinates the gain factors are formatted with the intention of being as universally accepted as possible. By using a custom built VBAP implementation the system can be incorporated into

DAW's without sophisticated multichannel implementations.

Implementing intuitive and universally recognized protocol is an important step to the modularization of spatialisation systems. All systems need to format messages in the same way in order to have modules be interchangeable between systems.

5. CONCLUSIONS AND FUTURE WORK

Both the systems presented in this paper have proved to be intuitive and have been used in performance making use of various modular configurations. The tactile.motion performance interface was featured in a concert at the Wellington City Gallery in September 2013. For this concert the system was configured as is mentioned for Chronus_2.0, without the need for the OSC conversion. The modularity incorporated into both systems meant that this transition was a fast and easy process to make. Interfaces from the Chronus family have also being used in performance with live audio input rather than standard diffusion models in a system configuration much more similar to the proposed tactile.motion set-up. Again the adaptability and modular separation built in to all aspects of the spatialisation system have proved easy to implement and afford a performative expressive range.

At the time of writing there are a number of concerts scheduled to take place featuring both the presented interfaces. A user study is also scheduled to evaluate the tactile.motion user interfaces expressivity as a new performance interfaces for more traditional diffusion concert settings.

Extensive new features are proposed for both interfaces presented and protocol specifications will be constantly reassessed, as new features are developed in order to assure the most intuitive protocol specifications are always in use. Adaptations to the system beyond the user interface are also intended in order to more closely sync the systems and encourage their use from a wider electronic performance community.

This research has exhibited a focus on the development of new performance interfaces for real – time spatialisation. In order to encourage wider use of new interfaces the authors have designed and implemented modularity throughout their spatialisation systems in order to make these user interfaces available to the wider diffusion performance community. It is hoped that the intuitive nature of the implemented OSC protocol will allow other performers to integrate the use of these new performance interfaces into their own spatialisation systems.

Acknowledgements

The authors would like to thank the New Zealand School of Music and the School of Engineering and Computer Science, Victoria University of Wellington for their on-going support of this research.

6. REFERENCES

Bredies, Katharina, Nick Alexander Mann, Jens Ahrens, Matthias Geier, Sascha Spors, and Michael Nischet. 2006. "The Multi-Touch SoundScape Renderer." *In Proceedings of the Working Conference on Advanced Visual Interfaces*. New York, USA.

Clozier, Christian. 2001. "The Gmebaphone Concept and Cybernephone Instrument." *Computer Music Journal MIT Press*. 25 (4). Winter: 81–90.

Desantos,, Sandra, Curtis Roads, and François Bayle. 1997. "Acousmatic Morphology: An Interview with François Bayle." *Computer Music Journal MIT Press* 21 (3): 11–19.

Geier, Matthias, Jens Ahrens, A Mohl, Sascha Spors, J Loh, and Katharina Bredies. 2007. "The SoundScape Renderer: A Versatile Software Framework for Spatial Audio Reproduction." *In Proceedings of WFS Workshop*. Ilmenau, Germany.

Gerzon, Michael. 1977. "Design of Ambisonic Decoders for Multispeaker Surround Sound". *Audio Engineering Society Convention Paper.* New York, USA.

Harrison, Jonty. 1999. "Sound, Space, Sculpture: Some Thoughts on the 'What', 'How' and 'Why' of Sound Diffusion." *Organised Sound, Cambridge University Press* 3 (2): 117–27.

Harrison, Jonty, and Scott Wilson. 2010. "Rethinking the BEAST: Recent Developments in Multichannel Composition at Birmingham ElectroAcoustic Sound Theatre." *Organised Sound, Cambridge University Press* 15 (3): 239–50.

Hollerer, T, J Kuchera-Morin, and X Amatriain. 2007. "The Allosphere: A Large-Scale Immersive Surround-View Instrument." *In Proceedings of the 2007 Workshop on Emerging Displays Technologies.* New York, United States of America: ACM Press.

Johnson, Bridget, and Ajay Kapur. 2013. "Multi-Touch Interfaces For Phantom Source Positioning In Live Sound Diffusion." *In Proceedings of New Interfaces For Musical Expression.* Kaaist, South Korea.

Johnson, Bridget, Michael Norris, and Ajay Kapur. 2014. "The Development of Physical Spatial Controllers." *In Proceedings of New Interfaces For Musical Expression.* London, England.

Kendall, Gary, Nils Peters, and Matthias Geier. 2008. "Towards An Interchange Format For Spatial Audio Scenes." *In Proceedings of International Computer Music Conference*. Belfast, Ireland.

Mooney, James, and D Moore. 2008. "Resound: Open-Source Live Sound Spatialisation." *In Proceedings of International Computer Music Conference*. Belfast, Ireland.

Moore, A, D Moore, and James Mooney. 2004. “M2 Diffusion - The Live Diffusion of Sound in Space.” *In Proceedings of International Computer Music Conference.* Miami, USA.

Peters, Nils. 2008. “Proposing SpatDif - The Spatial Sound Discription Interchange Format.” *In Proceedings of International Computer Music Conference.* Belfast, Ireland.

Peters, Nils, Trond Lossius, Jan Schacher, Charles Bascou, and Timothy Place. 2009. “A Stratified Approach For Sound Spatialization.” *In Proceedings of 6th Sound and Music Computing Conference*. Porto, Portugal.

Pulkki, Ville. 1997. “Virtual Source Positioning Using Vector Base Amplitude Panning.” *Journal of Audio Engineering Society* 45 (6): 456–66.

Wright, Matt, Adrian Freed, and Ali Momeni. 2003. “Open Sound Control: State of the Art 2003.” *In Proceedings of New Interfaces For Musical Expression,* 153–59. Montreal, Canada.