SONIFICATION OF A REAL-TIME PHYSICS SIMULATION WITHIN A VIRTUAL ENVIRONMENT

Rhys Perkins

Anglia Ruskin University,
Department of Music and Performing Arts,
Cambridge, CB1 1PT, UK
rhysperkins@rhysperkins.com

ABSTRACT

There has been an increasing amount of research utilising 3D virtual environments as a core component of interactive sonifications. While showing considerable potential for their ability in producing both real-time visualisation and sound, they often come with constraints as a result of their design decision processes. This paper presents developments of a prototype that has arisen out of my attempts to address some of the issues involved in bringing sonification to a wider audience through a universal metaphor. These new additions allow for an intuitive, elementary introduction into to the world of auditory display, while providing a more flexible and immersive environment for composition and sound design.

1. INTRODUCTION

The emergence of improving technology has provided an opportunity to overcome the limitations of previous work [1] where computational requirements would produce significant latency between the audio and the visual, inhibiting real-time interaction. Dedicated hardware, such as the graphics processing unit (GPU), allow for the sharing of resources and have recently shifted their focus towards more general purpose computing [2] including accelerating physics simulations. Freely distributed 3D physics engines such as Bullet[1] and Open Dynamics Engine,[2] common middleware solutions for modern game engines, are readily available thanks in part to an increasing demand for realism amongst gamers. Their cross-platform approach sees widespread use of long established solutions such as the desktop computer but also extends to mobile devices. As a result this presents an opportunity to revisit previous compositional techniques [1] and reach a wider audience in the process.

Nguyen's approach to TONAL DisCo [3] acknowledges the benefits of using a game engine when combining dynamic visuals with audio but chooses to use a pre-processed sample library over real-time sound generation. The prototype laid out in this paper addresses that limitation by utilising a messaging system that provides a link between the visual and the sound synthesis. Pre-processed sample libraries alongside dynamic visuals have started to emerge in commercial software with applications like PhysSynth[3] for the iPad. However, I would consider this application to be compromised by only simulating basic particles with simple collision detection and response, comprising interactions between points and segments and restricting the process to two dimensions. By sonifying 3D physics engines we can surpass these restrictions to cover more interesting and complex interactions, communicating and exploring a wealth of new dynamic ideas within an auditory display whose interface with the user resembles the physical world.

Sturm [1] laid out some of the benefits that a sonification of particle physics simulation would bring to science, including new ways of understanding physical phenomena, and refers to several artistic merits such as bending of those scientific laws to suit a composer's taste. Pedagogical advantages that Western music composition students might gain from using an audiovisual simulation are also discussed. In particular, he suggests there is an advantage to combining the audio and visual modalities in order to present musical ideas, likening it to listening to a piece whilst reading a score, rather than partaking in one or the other activity separately. Metaphorical correlations between particle physics and sound synthesis have been explored [1][4], serving as a means of providing a bridge between the two aforementioned fields whilst highlighting cultural differences and the problems that might arise from them. For instance Sturm [1] states that composers must possess skills in physics to begin with and only the audience need not be versed. The system described in this paper was designed to cater for all levels, from the well-versed user who comprehends and wishes to explore and extend the open-source scientific algorithms to those that would like to immediately compose and play.

RedUniverse provided a toolkit for sonifications of dynamic systems [5] with the aim of producing a playground for compositional ideas. These systems were also limited to two dimensions and lacked accessible interactivity, requiring a good knowledge of the SuperCollider programming language in order to take full advantage of their potential customisation. What I present here will allow for immediate use and configuration via standard input, such as a keyboard and mouse, but will also cater for further inputs using standard protocols such as MIDI and OSC.

Interactive compositional tool, VR-RoBoser [6] makes a case against predetermined, repetitive soundscapes in a virtual environment by using a context dependent sonification. They present the idea of a user-controlled or autonomous avatar that continuously reacts to its unchanging surroundings in order to overcome this issue. I would argue that the dynamic nature of physics would help create a less static environment. Continuous user interaction would stimulate audible results as the simulated objects react accordingly. Automated movement can be

[1]http://bulletphysics.org/wordpress/
[2]http://www.ode.org/
[3]http://www.physynth.com/

accomplished through inter-object logic allowing the user to pay attention to other aspects of this system, such as the proposed camera control.

In previous work [7] I laid out the design process behind a musical tool employing a 3D space that could be populated with audiovisual objects acting under the laws of physics. Each object exposed fundamental data dimensions that could then be mapped to sound dimensions via the OSC protocol, providing an audible insight into their behaviour within the current environment. A modular approach to the object design meant that a user would construct logic through object interaction and association without the need for coding keywords, operators and the understanding of basic programming paradigms.

The tool comprised an environment where polygonal models, visually representing the underlying simulation data, were introduced alongside a graphical user interface (GUI) that offered the opportunity to control fundamental properties, sonify and compose in real-time. Since designing the initial prototype I have found the need for improvements in a few key areas. This paper will discuss some of the fields I have considered when refining my approach to the sonification of an interactive physics engine. In the next section I will explore the idea of human interaction and how the process came to shape the design of the basic objects. The objects are then discussed along with some thoughts on their potential behaviour and how this can affect the overall output. This is followed by the mapping section which explains why I believe that the same simulated objects are inherently easier to map due to their physical grounding, along with several theories that underpin the conception of the mapping function tool. The camera section describes how the user will view and traverse the environment, how it accentuates the user's experience and why they should be presented with the option for both automatic and manual control configurations. This then leads into the messaging system where the aim is to make the same experience more flexible and personal. The paper concludes by stating some of the advantages the proposed system has over another comparatively close environment [8] along with my thoughts about potential future work and emerging areas of interest.

2. KEY COMPONENTS

2.1. Human Interaction

Investigation [9] has shown that applying human interaction in real-world contexts to sonification can help improve interface design. In this paper, the researchers state that humans are adapted for interaction within their physical environment and making continuous use of all their senses. When we perform an action on an object we expect some kind of reaction and our perception of objects builds up over time through this interactive process. The objects found in this system adhere to the unchanging laws of physics that our neural hardware has been effectively programmed to deal with over many years of evolution. This enables the user to utilise their acquired skills in order to manipulate high-dimensional data via objects with familiar behaviour and response. The authors also argue that one of the main problems in the domain of data exploration is that the data often inhibits a high-dimensional data space that is different from the 3D space we are familiar with. In this prototype the simulation data emulates rigid bodies, and their behaviour in our natural environment, providing familiar grounds for both exploration and interaction.

Interaction with the objects has a direct effect on the procedurally generated simulation in a similar manner to model-based sonification [10] where the user supplies the initial excitation. By grabbing, moving and throwing objects it is feasible to perform a wide range of actions from striking, to more delicate procedures such as plucking; each of which results in changes to the data dimensions. This direct process introduces information manipulation to the average user at a more accessible level when compared to other similar applications [11] since no coding knowledge is required. For example, saved scenarios containing preconfigured entities can be loaded, ensuring that user interaction yields instant audible results.

If the user wishes to create and save their own scenario, or edit existing ones, some basic GUI control knowledge is required. The controls can be toggled at any time and aim to present a more traditional and precise means of modifying the attributes that influence each aspect of the prototype. Presentation of the data in this manner brings its own set of problems in that the potential to overwhelm the viewer with information is increased. When considering high-dimensional spaces one study [12] argues for a mental model simpler than brute-force awareness of every detail in order to avoid cognitive overload. The authors suggest that parameters should be cross-coupled so that the performer naturally thinks of certain parameters as varying together in predefined patterns. The high-dimensional data, encapsulated visually by each model, allows us to intrinsically understand how the parameters vary together. Throwing an object would imply a change in velocity that would be influenced by the mass of the object. Spherical objects are more naturally inclined to roll, providing smoother changes in angular velocity as opposed to the sudden, erratic changes of their square shaped equivalents.

Research into improving sonification tools [4][9] has questioned how information should be distributed to different modalities in order to maintain the best usability. As stated previously, our everyday interactions with physical objects providing a base level for our conceptual understanding of the data dimensions found within. With this in mind I highlighted what I believed were the important elements of the underlying data, choosing to expose those that had a direct impact on the representative model's behaviour. Given the longstanding synergy between humans and physical entities I would suggest that less mental bandwidth is required to comprehend the visual events. Instead, the attentive capacity of the user can focus on the audio, and its governing mapping process, encouraging sonic exploration and creativity.

2.2. Objects

Objects provide a modular approach to the way the user experiences the underlying data. Depending on the object's configurable physical parameters the program will automatically simulate subsequent interactions as the object reacts to its current environment and user intervention. However, it can be argued that there are parameters that have no direct effect on the simulation which are just as important for the user to exploit. These properties can enhance a user's experience, and encourage them to learn, by creating associations through further visual abstractions that can be audibly reinforced. As one example, the object colour could be changed in order to present object information in a new manner. According to research in Gestalt laws of grouping [13] there is a stronger tendency to group local elements by common colour than by similarity of shape. This would imply that, in some cases, our brains are more receptive to the material that encompasses each shape, rather than the shape itself. Therefore, by involuntarily grouping similar coloured entities, the audience's attention would be drawn to a single contrastingly coloured object, perceiving it as being outside of the group. The performer could then take advantage of this visual phenomenon by using it to introduce a solo theme or demonstrating object-specific sonic behaviour.

In a typical physics simulation most objects will likely come to rest until excitation provides the impetus for movement. If frequent changes in data are desired then further logic can be introduced via object specific context menus (Figure 1). For example the user could define a point where rigid bodies can spawn at regular intervals. Each body created would have a lifespan where the associated object would be automatically removed from the scene after such time had elapsed. Inter-object logic can be extended further by defining connecting mechanical joints and introducing external forces such as gravitational fields. Automated mechanical contraptions would be a logical step in complexity, allowing for the creation of visual algorithms. With the basic building blocks, it should be possible to conceive and construct contraptions in the style of Heath Robinson[1] or Rube Goldberg[2], providing unfamiliarity through the extraordinary.

2.3. Mapping

In this prototype the parametric mapping process grants an insight into the composer's conceptual understanding of the data dimensions. It has been suggested that metaphors help create more intuitive mappings [14] and is well suited to parameter mapping sonification [4]. Whereas the universal laws of physics can represent a predictable visual behaviour by employing a metaphor that fits our everyday observations, the sound representation is more subjective. The mapping of the objects serves to reflect the experiences of the user, making it difficult to produce general metaphors that are valid in any context. What may be coherent and intuitive to one mindset could be judged differently by those from another cultural background. There have been attempts to create online databases [15][16] suggesting mappings based on experimental evidence although it is widely accepted that an affective mapping can't always be predicted [4][14]. A heuristic approach to this area should be adopted to allow for a compositional process that encourages experimentation in order to express creativity where the audience can reflect on the implications of a musician's cultural and physical experiences.

When interviewing scientists, Vogt and Holdrich [4] discovered that strong metaphors emerge from their professional experience. They found that more mapping associations were suggested for the well-known particles and fewer for the rarer proposing that perhaps this arose from fewer encounters, lack of interaction, and therefore less prominent in the mind. They also discovered that everyday properties such as mass were cited more often than abstract ones. This would imply that an object visually described through a recognisable metaphor, encapsulating everyday properties, can be easier to map.

Our experience with physical objects allows us to inherently determine complex data relations. Properties are implied by a rigid body's response to collisions with its surroundings. By referring to the visual behaviour during this event we perhaps reduce the need to refer to the linking of parameters for interpretation. This can be illustrated by focusing on two dimensions, such as mass and velocity, where one could map them to pitch and envelope time, respectively. We understand that an object of greater mass would provide an object of less mass with a higher velocity upon collision. A spectrum of sound can be obtained afterwards where we could assume that objects that have travelled further will differ in pitch, and duration, to those that travelled less distance over the same period of time.

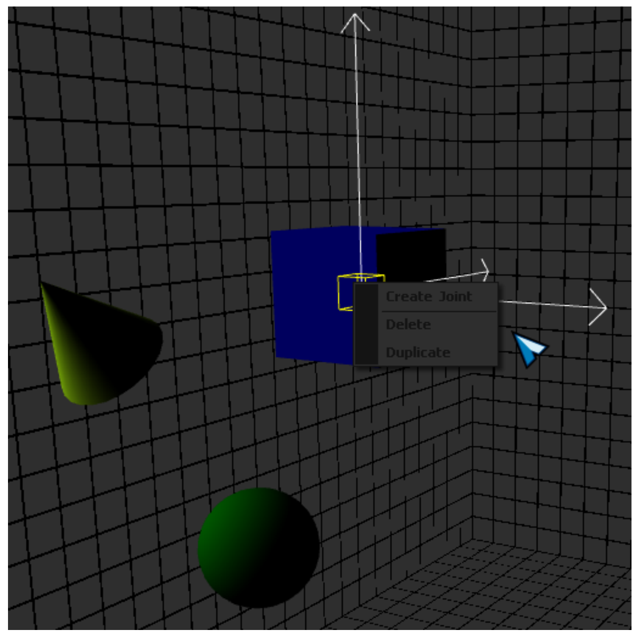

Figure 1: Three objects in the environment. The context menu is displayed over the selected object.

Understanding of the meaning in sonifications depends on the metaphors implied where the choices made during the process are crucial for how a design is understood by its listeners [17]. For instance, the coupling of coloured objects mentioned in section 2.2. Walker [18] states that in order to achieve an effective mapping choice, one must go beyond that of polarity and linear scaling functions while avoiding restrictions placed on the user through bad design [19]. The mapping window controls were devised to encourage flexibility by employing a messaging system, discussed later in this paper, to allow the user to map exposed parameters to potentially any input of a synthesiser. In conjunction with these GUI controls (Figure 2) I created a function editor that serves to display the relationship between the two dimensions. The editor itself contains two permanent breakpoints that define the input domain (x axis) and the output range (y axis). Further breakpoints can be added and removed in order to construct a bijective mapping curve or polyline. The curvature of the segments, found between each breakpoint, can also be configured in order to account for both linear and non-linear responses.

---

[1]http://heathrobinson.org/exhibition/index.htm

[2]http://www.rubegoldberg.com/

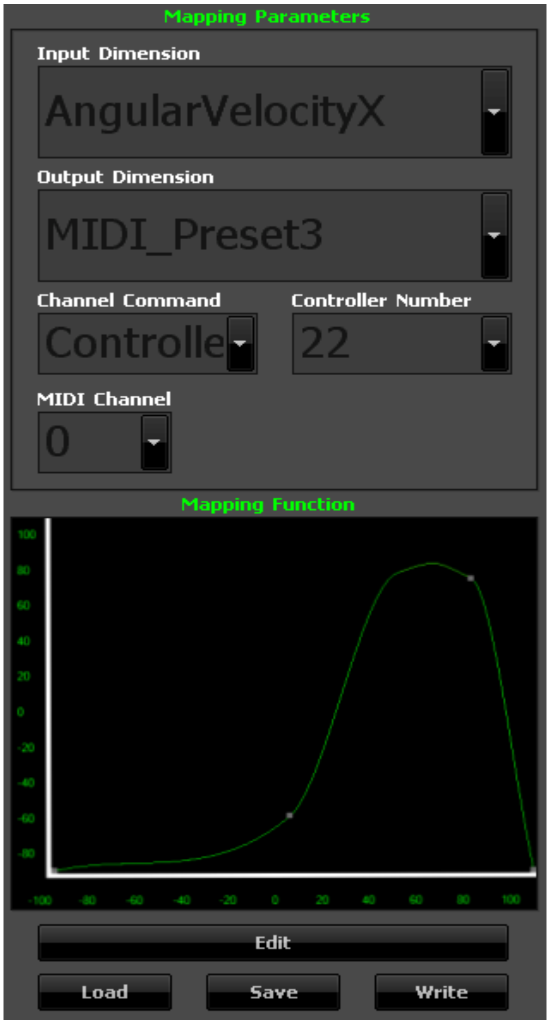

Figure 2: The mapping window along with the available controls for editing the relationship between the input and output dimensions.

2.4. Camera

Most software synthesisers use flat imagery in order to represent modifiable parameters making use of the two dimensions provided by a computer screen. Their interfaces often restrict the information to a window, presenting the intended audience with a multitude of controls that are difficult to decipher and engage with. The prototype's main interface attempts to address this problem, presenting the information in a more natural three dimensional form. This posed the question of how one might traverse the extra dimension in an immersive and inherent manner using the same display hardware.

Many first person perspective games attempt to provide immersive, high-level simulations of reality. In a typical setup, the player views, navigates and interacts with the game world by operating a camera. Recent studies have compared games in this category to sonification systems [20], where sounds are used to accentuate the player immersion by reacting to their behaviour or to provide sign posts for orienteering. I felt it was appropriate to adopt some of the ideas found in these proven systems by implementing a similar camera system for interactive traversal of the virtual environment.

Camera movement can be kinaesthetically controlled in real-time or automated along a user-defined pathway. Default manual control is that same as that of a typical PC first person shooter setup, utilising the W, A, S and D keys for camera translation, and a mouse for camera rotation. The camera rotation system required slightly different approaches to each mode of operation as manual rotation of a camera with six degrees of freedom would be disorientating with standard mouse control. If the camera pitch was allowed to be greater than 90 degrees in either direction, the mouse controls would be reversed along both the pitch and yaw axes. I therefore decided to emulate more natural head movement by restricting the camera pitch to a ±90 degree range in the same manner that a first person perspective camera does.

Node Property	**Description**
Position	Location of the node
Rotation/Orientation	Camera's orientation when reaching the node
Speed	The constant speed of the camera until the next node is reached *Time will be recalculated*
Time	The time (seconds) at which the camera arrives at the node *Speed will be recalculated*

Table 1: Configurable properties of an automated camera node.

Automated camera control frees the user from direct control giving them the opportunity to concentrate on other tasks, such as object interaction, and does not require the same restriction for rotation. Camera motion is defined by a series of nodes that comprise a Hermite spline-based path. The properties of each node (Table 1) allow for an increase in the accuracy and response time of the camera when compared to independent user control. Spherical linear interpolation is employed to ensure smooth changes in camera orientation when moving from one node to the next and prevents viewer disorientation. The timing of the camera can add to the overall sense of structure, guiding the viewer to focus on visual snapshots of the environment at designated points in time where precise values, for both speed and time, support various tempi.

By directing the camera, the user can create a sense of motion, guiding the audience through a visual soundscape. Choices made in constructing the camera's pathway become part of the creative process, enabling the viewer to observe through the cognitive lens of the composer. In this manner, attention can be drawn to specific areas of interest whilst providing an insight into the structures underlying the composition. Sturm [1] touched on this particular benefit of a camera system when he stated 'thus any sonification of a particle system is dependent on the state of the observer; each observer with a unique position and/or velocity will hear the system in a different way – a truly relativistic idea.'

2.5. Messaging System

The messaging system sends and receives OSC and MIDI based messages providing the user with an opportunity to customise the data flow both in and out of the software. Whereas traditional human interface devices, such as a mouse and keyboard, can be used without the need for this system, these two protocols provide a widely accepted standard for interface control, expanding upon the breadth of possible controllers and the levels of immersion they provide. For example, tactile feedback can be introduced by utilising this system. Specific data, sent when an object is grabbed, can be interpreted by the current controller in order to produce corresponding actions such as vibrating.

With the mapping system controlling the outgoing data, I separate the audio processing from the software. Implementing these protocols grants access for communication across a network, extending the reach to computers operating on different platforms, such as OS X or Linux, or mobile platforms such as Android or iOS. This would provide the opportunity to communicate with a vast range of audio applications and synthesis tools, exploiting their existing timbres and increasing potential sonic diversity when compared to an inbuilt synthesis engine.

I have produced a video [21] that demonstrates an example of my prototype connected to sixteen instances of the Alchemy VST plug-in, utilising all available MIDI channels on a loopback network. The table shown below the video contains the mapping details for each object and was automatically generated by the prototype.

3. CONCLUSION

This work sets out to add to the range of tools for experimentation and interaction of a data set using the combination of the visual and aural modalities. It intends to expand on the experience found in Versum [8] by incorporating objects that dynamically respond to real-time interaction, with a unique collision response determined by their configurable properties. Whereas the entities found in the aforementioned software lend themselves to more ambient sounds due to their continuous playback, the objects here also allow for more dynamic sound with full control of ADSR envelopes and sonic response. Furthermore, the flexible mapping system does not constrain the amplitude of objects based solely on their distance from the camera as other relationships can be explored via the messaging system.

Whilst much time has been spent creating the basic tools for both music creation and sonification, I feel that future work should be focused on musicality. Sonifying data into a systematic musical structure to understand patterns and trends in a more traditional sense would have the benefit of making the tools more widely understood by the potential audience.

Nguyen discusses an approach to musicality in the mapping process where he decides to lose resolution of the data in favour for an increase in musicality [3]. He argues that the use of musical structure in sonification has the potential to communicate compound relationships with an increase in clarity that might not be apparent with high resolution data. To integrate this I would suggest changes to the function editor that would accommodate user-defined bands of any width. These regions would then be displayed on top of the mapping transfer function (Figure 3), allowing the user to conceive the varying resolutions. For example, on the y axis, the output dimension of pitch can be constrained to a musical scale (Lydian, Chromatic, etc.) using standard frequency tuning. This idea can be extended into other areas such as rhythm where the triggering of sound can be quantised to match common subdivisions of a bar based on the global tempo assigned. The effect should be subtle as to not lose perceived concurrency between the audio and the visual events.

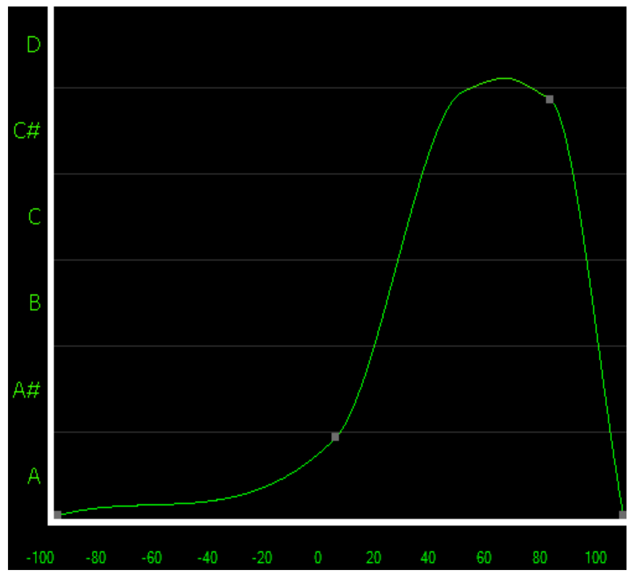

Figure 3: A function with granular regions denoted by the horizontal lines.

Recent developments in human computer interfaces, including the 'See Through 3D Desktop'[1] and the HoloDesk,[2] present more direct ways of interacting with virtual 3D objects. Model-based sonification has been shown to be intuitive by taking important dimensions of sound semantics into account and grounding them in physical sound generating processes in a natural and user-transparent way [22]. By combining projections of the simulated objects with interfaces that emulate a more innate way of interaction, we can extend the model-based method beyond its inherent physical constraints. This would benefit the interface building process as a variety of deformable, polygonal objects could be designed, created and saved in a portable format. It would also encourage the creation of more abstract and imaginative virtual controller shapes whose physical counterparts would be difficult or impossible to implement.

4. REFERENCES

[1] B. L. Sturm, "Synthesis and Algorithmic Composition Techniques Derived from Particle Physics," in *Proceedings*

[1]http://leejinha.com/See-Through-3D-Desktop
[2]http://research.microsoft.com/apps/video/dl.aspx?id=154571

of the Eighth Biennial Symposium on Technology and the Arts, Connecticut College, New London, Connecticut, 2001.

[2] M. Harris. "GPGPU: General-purpose computation on GPUs," in *Game Developers Conference*, San Francisco, California, 2005.

[3] V. Nguyen, "Tonal DisCo: Dissonance and Consonance in a Gaming Engine," in *Proceedings of the International Conference on Auditory Display*, Budapest, Hungary, 2011.

[4] K. Vogt and R. Holdrich, "A Metaphoric Sonification Method – Towards the Acoustic Standard Model of Particle Physics," in *Proceedings of the International Conference on Auditory Display*, Washington, USA, 2010.

[5] M. d'Inverno, F. Olofsson, "RedUniverse – a simple toolkit," in *Live Algorithms for Music Conference*, London, UK, 2006.

[6] S. Le Groux, J. Manzolli, P. Verschure, "VR-RoBoser: Real-Time Adaptive Sonification of Virtual Environments Based on Avatar Behavior," in *Proceedings of the Conference on New Interfaces for Musical Expression,* New York, USA, 2007.

[7] R. Perkins, "Mapping 3D Objects To Synthesised Sound Using A Simulated Physics System," in *Proceedings of the International Computer Music Conference*, Huddersfield, UK, 2011.

[8] T. Barri, "Versum: Audiovisual Composing in 3D," in *Proceedings of the International Conference on Auditory Display*, Copenhagen, Denmark, 2009.

[9] A. Hunt and T. Hermann, "The Importance of Interaction in Sonification," in *Proceedings of the International Conference on Auditory Display,* Sydney, Australia, 2004.

[10] T. Hermann, "Sonification for Exploratory Data Analysis," PhD thesis, Bielefeld University, Bielefeld, 2002.

[11] I. Bukvic and K. Ji-Sun, "µ Max-Unity3D interoperability toolkit," in *Proceedings of the International Computer Music Conference*, Montreal, Canada, 2009.

[12] G. Garnett and C. Goudeseune, "Performance Factors in Control of High-Dimensional Spaces," in *Proceedings of the International Computer Music Conference*, San Francisco, 1999.

[13] P. T. Quinlan and R. N. Wilton, "Grouping by proximity or similarity? Competition between the Gestalt principles in vision," in *Perception 27*, 1998, pp. 417–430.

[14] B. Walker and G. Kramer, "Mappings and Metaphors in Auditory Displays: An Experimental Assessment," in *ACM Transactions on Applied Perception*, vol. 2, issue 4, 2005.

[15] S. Barrass, "EarBenders: Using Stories About Listening to Design Auditory Interfaces," in *Proceedings of the First Asia-Pacific Conference on Human Computer Interaction*, Information Technology Institute, Singapore, 1996.

[16] B. Walker, and D. Lane, "Sonification Mappings Database on the Web," in *Proceedings of the International Conference on Auditory Display*, Espoo, Finland, 2001.

[17] A. de Campo, J. Rohrhuber, T. Bovermann and C. Frauenberger, "Sonification and Auditory Display in SuperCollider," in *The SuperCollider Book,* S. Wilson, D. Cottle, and N. Collins, eds., 2011, Cambridge, UK: The MIT Press, pp. 381-408.

[18] B. Walker, and D. Lane, "Psychophysical Scaling of Sonification Mappings: A Comparison of Visually Impaired and Sighted Listeners," in *Proceedings of the International Conference on Auditory Display*, Espoo, Finland, 2001.

[19] G. Kramer, et al. "Sonification Report: Status of the Field and Research Agenda," in *Report prepared for the National Science Foundation by members of the International Community for Auditory Display*, Santa Fe, NM, 1999.

[20] M. Grimshaw, "Sound and Immersion in the First-Person Shooter," in *Proceedings of the International Conference on Computer Games*, La Rochelle, France, 2007.

[21] R. Perkins, "Video demonstrating MIDI out capabilities of the prototype," Available at www.rhysperkins.com/Sonification/MIDI.htm, Accessed 8 February 2012.

[22] T. Hermann, H. Ritter, "Sound and Meaning in Auditory Data Display," in *Proceedings of the IEEE*, vol. 92, issue 4, 2004.