SPATIAL SOUND RENDERING IN MAX/MSP WITH VIMIC

Nils Peters[1], *Tristan Matthews*[1], *Jonas Braasch*[2], *Stephen McAdams*[1]
[1]Schulich School of Music - Music Technology Area, McGill University, Montréal, CA
[2]Rensselear Polytechnic Institute - School of Architecture, Troy, US
nils.peters@mcgill.ca
CIRMMT - Centre for Interdisciplinary Research in Music Media and Technology

ABSTRACT

Originally conceptualized [2] for the software Pure Data, ViMiC was recently refined and extended for release to the Max/MSP community. ViMiC (Virtual Microphone Control) is a tool for real-time spatialization synthesis, particularly for concert situations and site-specific immersive installations, and especially for larger or non-centralized audiences. Based on the concept of virtual microphones positioned within a virtual 3D room, ViMiC supports loudspeaker reproduction up to 24 discrete channels for which the loudspeakers do not necessarily have to be placed uniformly and equidistant around the audience. Also, through the integrated Open Sound Control protocol (OSC), ViMiC is easily accessed and manipulated.

1. INTRODUCTION

Besides the traditional concepts of pitch, timbre, and temporal structures, composers have long felt the desire to integrate a spatial dimension into their music. First, through static placement and separation of musicians in the concert space, and later, through dynamic modifications of the sound source position, effects of spatial sound segregation and fusion were discovered. In the 20th century, especially due to the invention and integration of microphones and loudspeakers in the musical performance, spatialization was popularized.

One of the earliest composers using the newly available electronic tools was Karlheinz Stockhausen. For his composition ”Kontakte” (1958-60) he developed a rotational table, mounting a directed loudspeaker surrounded by four stationary microphones that receive the loudspeaker signal. The recorded microphone signals were routed to different loudspeakers arranged around the audience. Due to the directivity and separation of the microphones, the recorded audio signals contained Inter-channel Time Differences (ICTDs) and Inter-channel Level Differences (ICLDs). Depending on the velocity of the speaker rotation, the change in ICTDs can create an audible Doppler effect. ViMiC follows somehow this Stockhausen's traditions by using the concept of spatially displaced microphones for the purpose of sound spatialization. Relation to pioneering works by Steinberg and Snow [13], Chowning [3], and Moore [8] also apply.

2. SPATIALIZATION WITH MAX/MSP

This section briefly overviews available loudspeaker spatialization techniques for Max/MSP. For further details, refer to the indicated references.

Vector Based Amplitude Panning (VBAP) is an efficient extension of stereophonic amplitude panning techniques, applied to multi-loudspeaker setups. In a horizontal plane around the listener, a virtual sound source at a certain position is created by applying the tangent panning law between the closest pair of loudspeaker. This principle was also extended to project sound sources onto a three dimensional sphere and assumes that the listener is located in the center of the equidistant speaker setup [11].

Distance Based Amplitude Panning (DBAP) also uses intensity panning applied to arbitrary loudspeaker configurations without assumptions as to the position of the listener. All loudspeakers radiate coherent signals, whereby the underlaying amplitude weighting is based on a distance attenuation model between the position of the virtual sound source and each loudspeaker [5].

Higher Order Ambisonics extends Blumlein's pioneering idea of coincident recording techniques. HOA aims to physically synthesize a soundfield based on its expansion into spherical harmonics up to a specified order. To date, Max/MSP externals up to the 3rd order for horizontal-only or periphonic speaker arrays have been presented in [12] and [14].

Space Unit Generator, also called room-within-the-room model, dates back to [8]. Four loudspeakers represented as ”open windows” are positioned around the listener and creates an ”inner room”, and embedded in an ”outer room” with virtual sound sources. Sound propagation of the virtual source rendered at the ”open windows” creates ICTDs and ICLDs. Some early reflections are calculated according to the size of the outer room. A Max/MSP implementation was presented in [17].

Spatialisateur, in development at IRCAM and Espaces Nouveaux since 1991, is a library of spatialization algorithms, including VBAP, first-order Ambisonics and stereo techniques (XY, MS, ORTF) for up to 8 loudspeakers. It can also reproduce 3D sound for headphones (binaural) or 2/4 loudspeakers (transaural). A room model is included to create artificial reverberation controlled by a perceptual-based user interface.

3. VIRTUAL MICROPHONE CONTROL

ViMiC is a key part of the network music project SoundWIRE [1], and the Tintinnabulate Ensemble [2] directed by Pauline Oliveros. At the MusiMars Festival 2008 [3], *Ex Asperis*, a composition by Sean Ferguson, featured ViMiC for Max/MSP and integrated gestural controllers to manipulate ViMiC sound rendering processes in various ways.
ViMiC is a computer-generated virtual environment, where gains and delays between a virtual sound source and virtual microphones are calculated according to their distances, and the axis orientations of their microphone directivity patterns. Besides the direct sound component, a virtual microphone signal can also include early reflections and a late reverb tail, both dependent upon the sound absorbing and reflecting properties of the virtual surfaces.

3.1. ViMiC Principles

ViMiC is based on an array of virtual microphones with simulated directivity patterns placed in a virtual room.

3.1.1. Source - Microphone Relation

Sound sources and microphones can be placed and moved in 3D as desired. Figure 3 shows an example of one sound source recorded with three virtual microphones. A virtual microphone has five degrees of freedom: (X, Y, Z, yaw, pitch) and a sound source has four: (X, Y, Z, yaw). The propagation path between a sound source and each microphone is accordingly simulated. Depending on the speed-of-sound c and the distance d_i between a virtual sound source and the i-th microphone, time-of-arrival and attenuation due to distance are estimated. This attenuation function, seen in Eq. 1 can be greatly modified by changing the exponent q. Thus, the effect of distance attenuation can be boosted or softened. The minimum distance to a microphone is limited to 1 meter in order to avoid high amplification.

$$g_i = \frac{1}{d_i^q} \qquad d \geq 1 \tag{1}$$

Further attenuation happens through the chosen microphone characteristic and source directivity (see Fig. 2). For all common microphone characteristics, the directivity for a certain angle of incidence δ can be imitated by calculating Eq. 2 and applying a set of microphone coefficients from Table 3.1.1. By increasing the exponent w to a value greater than 1 will produce an artificially sharper directivity pattern. Unlike actual microphone characteristics, which vary with frequency, microphones in ViMiC are designed to apply the concept of microphone directivity without simulating undesirable frequency dependencies.

$$\Gamma = (a + b \cdot \cos\delta)^w \qquad 0 \leq a, b \leq 1 \tag{2}$$

[1] http://ccrma.stanford.edu/groups/soundwire/
[2] http://www.myspace.com/tintinnabulate
[3] http://www.music.mcgill.ca/musimars/

Characteristic	a	b	w
Omnidirectional	1	0	1
Subcardioid	0.7	0.3	1
Cardioid	0.5	0.5	1
Supercardioid	0.33	0.67	1
Hypercardioid	0.3	0.7	1
Figure-of-8	0	1	1

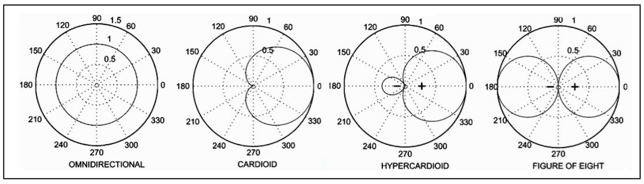

Figure 1. Common microphone characteristics

Source directivity is known to contribute to immersion and presence. Therefore ViMiC is also equipped with a source directivity model. For the sake of simplicity, in a graphical control window, the source directivity can be modeled through a frequency independent gain factor for each radiation angle to a 1° accuracy.

3.1.2. Room model

ViMiC contains a shoe-box room model to generate time-accurate early reflections that increase the illusion of this virtual space and envelopment as described in the literature [9]. Early reflections are strong auditory cues in encoding the sound source distance. According to virtual room size and position of the microphones, adequate early reflections are rendered in 3D through the well-known image method [1]. Each image source is rendered according to the time of arrival, the distance attenuation, microphone characteristic and source directivity, as described in section 3.1.1. Virtual room dimensions (height, length, width) modified in real-time alter the refection pattern accordingly. The spectral influence of the wall properties are simulated through high-mid-low shelf-filters. Because larger propagation paths increase the audible effect of air absorption, early reflections in ViMiC are additionally filtered through a 2nd-order Butterworth lowpass filter with adjustable cut-off frequency.
Also, early reflections must be discretely rendered for each microphone, as propagation paths differ. For eight virtual microphones, 56 paths are rendered if the 1st-order reflections are considered (8 microphones · [6 early reflections + 1 direct sound path]). Although time delays are efficiently implemented through a shared multi-tap delay line, this processing can be computationally intensive.

3.2. Late Reverb

The late reverberant field of a room is often considered nearly diffuse without directional information. Thus, an efficient late reverb model, based on a feedback delay network [4] with 16 modulated delay lines diffused by a Hadamard mixing matrix, is used. By feeding the outputs of the room model into the late reverb a diffused reverb tail is synthesized (see Fig. 2), for which timbral and temporal character can be modified. This late reverb can be efficiently shared across several rendered sound sources.

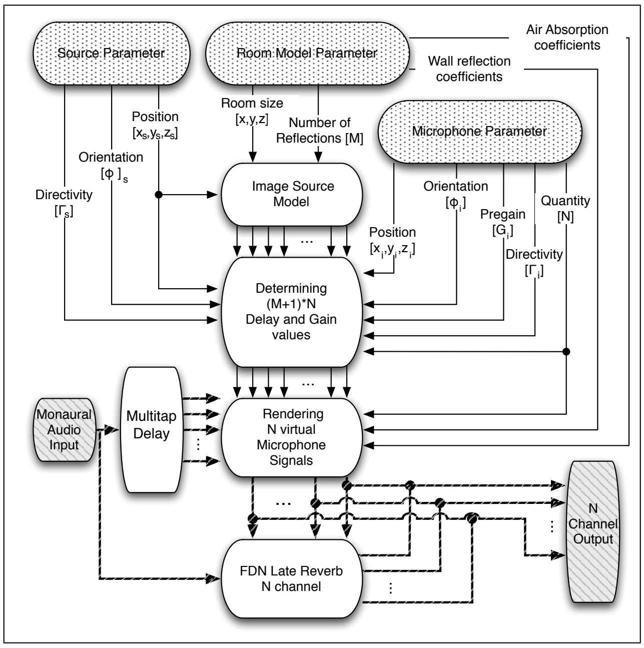

Figure 2. Flowchart of the Max/MSP processing

4. MOVING SOURCES

In Figure 4 the sound source moved from (x, y, z) to (x', y', z'), changing the propagation paths to all microphones, and also, the time delay and attenuation. A continuous change in time delay engenders a pitch change (Doppler effect) that creates a very realistic impression of a moving sound source. Doppler effect might not always be desired. ViMiC accommodates both scenarios.

4.1. Rendering with Doppler effect

For each changed sound path, the change in time delay is addressed through a 4-pole interpolated delay-line, the perceived quality of which is significantly better than with an economical linear interpolation. To save resources, interpolation is only applied when moving the virtual sound source, otherwise the time delay is being rounded to the next non-fractional delay value. At $fs = 44.1$ kHz and a speed-of-sound of $c = 344$ m/s, the roundoff error is approximately 4 mm. Some discrete reflections might not be perceptually important due to the applied distance law, microphone characteristics, and source directivity. To minimize processor load, an amplitude threshold can be set to prevent the algorithm from rendering these reflections.

4.2. Rendering without Doppler effect

This render method works without interpolation: the time delays of the rendered sound paths remains static until one of the paths has been changed by more than a specified time delay. In this case, the sound paths of the old and the new sound position are cross-faded within 50 ms, in order to avoid strongly audible phase modulations.

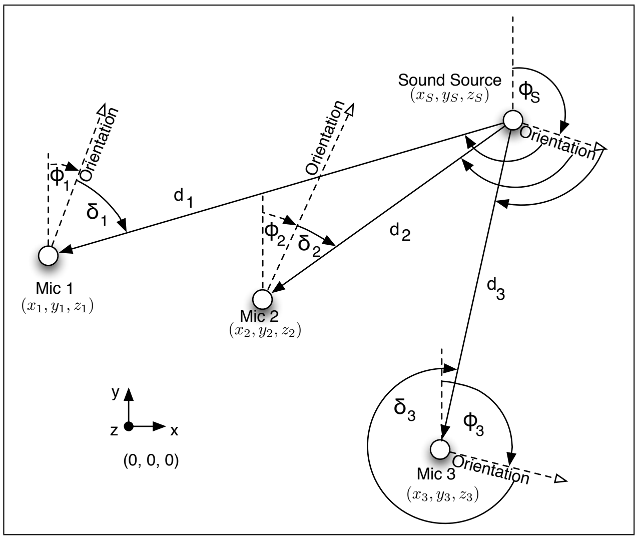

Figure 3. Geometric principles

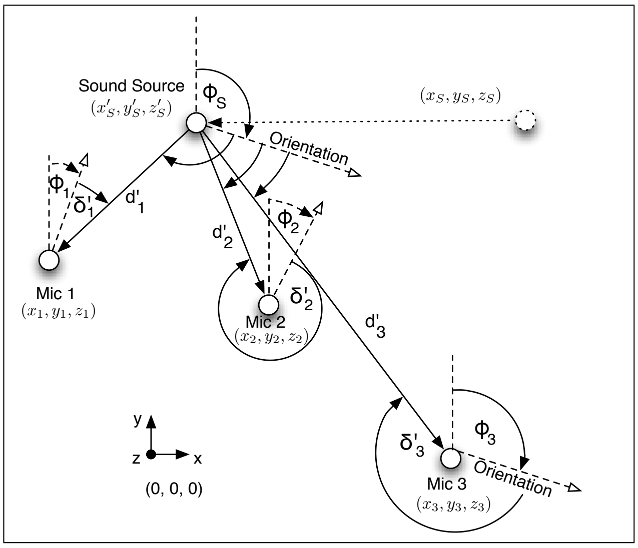

Figure 4. Moving source, geometric principles

5. PRACTICAL CONSIDERATIONS

5.1. How to set up the virtual microphones?

Typically, each virtual microphone is associated with one loudspeaker, and should be oriented at the same angle as the loudspeaker. The more spaced the microphones are, the bigger the ICTDs will be. The use of virtual microphones is especially interesting for arrays of speakers with different elevation angles, because the time-delay based panning possibilities help to project elevated sounds. Although ViMiC is a 24-channel system, for smaller loudspeaker setups the number of virtual microphones can be reduced. For surround recordings for the popular ITU

5.1 speaker configuration, Tonmeisters developed different microphone setups (e.g. [15]) applicable in ViMiC. To ease placing and modifying of the microphone positions, ViMiC provides an extra user interface where an array of microphones can either be graphically [4] edited or defined through cartesian and spherical coordinates (Fig. 5).

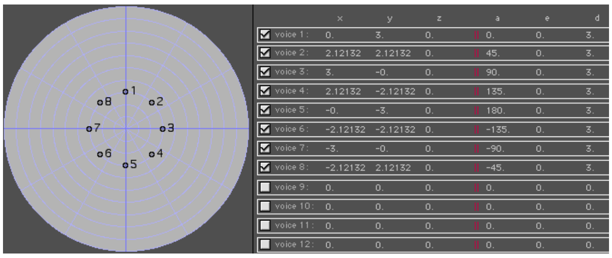

Figure 5. Interface to position microphones

5.2. Controllability

For easier and more flexible controllability, ViMiC was structured as high-level modules using the Jamoma framework for Max/MSP [10]. Jamoma offers a clear advantage in its standardization of presets and parameter handling. The ViMiC parameters have been sorted into three primary namespaces: *source*, *microphone* and *room*; and are specified with a data range. ViMiC is fully controlled through a GUI and through external OSC-messages [16]:

```
/source/orientation/azimuth/degree 45
/microphones/3/directivity/ratio 0.5
/room/size/xyz 10. 30. 7.
```

OSC enables the access and manipulation of ViMiC via different hardware controllers, tracking devices or user interfaces through OSC mapping applications (e.g. [6]) and allows gestural control of spatialization in real-time [7].

6. FUTURE WORK

Currently, a ViMiC module renders one sound source. Plans to develop an object which handles multiple sound source and frequency dependent source directivity are under discussion. ViMiC for Max/MSP under MacOS is available in the Subversion repository of Jamoma [5].

7. ACKNOWLEDGMENT

This work has been funded by the Canadian Natural Sciences and Engineering Research Council (NSERC) and the Centre for Interdisciplinary Research in Music, Media and Technology (CIRMMT).

[4] The [Ambimonitor] by ICST is used to display microphones[12].

[5] `https://jamoma.svn.sourceforge.net/svnroot/jamoma/branches/active`

References

[1] J. B. Allen and D. A. Berkley. Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. Am.*, 65(4):943 – 950, 1979.

[2] J. Braasch. A loudspeaker-based 3D sound projection using Virtual Microphone Control (ViMiC). In *Convention of the AudioEng. Soc. 118, Preprint 6430*, Barcelona, Spain, 2005.

[3] J. M. Chowning. The simulation of moving sound sources. *JAES*, 19(1):2 – 6, 1971.

[4] J. Jot and A. Chaigne. Digital delay networks for designing artificial reverberators. In *90th AES Convention, Preprint 3030*, Paris, France, 1991.

[5] T. Lossius. *Sound Space Body: Reflections on Artistic Practice*. PhD thesis, Bergen National Academy of the Arts, 2007.

[6] J. Malloch, S. Sinclair, and M. M. Wanderley. From controller to sound: Tools for collaborative development of digital musical instruments. In *Proceedings of the International Computer Music Conference*, pages 65–72, Copenhagen, Denmark, 2007.

[7] M. Marshall, N. Peters, A. Jensenius, J. Boissinot, M. Wanderley, and J. Braasch. On the development of a system for gesture control of spatialization. *In Proceedings of the 2006 International Computer Music Conference*, pages 360–366, 2006.

[8] F. R. Moore. A General Model for Spatial Processing of Sounds. *Computer Music Journal*, 7(6):6 – 15, 1983.

[9] R. Pellegrini. Perception-based design of virtual rooms for sound reproduction. In *22nd AES International Conference, Preprint 000245*, 2002.

[10] T. Place and T. Lossius. Jamoma: A modular standard for structuring patches in Max. In *Proceedings of the International Computer Music Conference 2006*, New Orleans, US, 2006.

[11] V. Pulkki. Generic panning tools for MAX/MSP. *Proceedings of International Computer Music Conference*, pages 304–307, 2000.

[12] J. C. Schacher and P. Kocher. Ambisonics Spatialization Tools for Max/MSP. In *Proceedings of the 2006 International Computer Music Conference*, pages 274 – 277, New Orleans, US, 2006.

[13] J. Steinberg and W. Snow. Auditory Perspective-Physical Factors. *Electrical Engineering*, 53(1):12–15, 1934.

[14] G. Wakefield. Third-Order Ambisonic Extension for Max/MSP with Musical Applications. In *Proceedings of the 2006 International Computer Music Conference*, pages 123 – 126, New Orleans, US, 2006.

[15] M. Williams and G. Le Dû. The Quick Reference Guide to Multichannel Microphone Arrays, Part 2: using Supercardioid and Hypercardioid Microphones. In *116th AES Convention, Preprint 6059*, Berlin, Germany, May 8–11 2004.

[16] M. Wright and A. Freed. Open Sound Control: A New Protocol for Communicating with Sound Synthesizers. *Proceedings of the 1997 International Computer Music Conference*, pages 101–104, 1997.

[17] S. Yadegari, F. R. Moore, H. Castle, A. Burr, and T. Apel. Real-time implementation of a general model for spatial processing of sounds. In *Proceedings of the 2002 International Computer Music Conference*, pages 244 – 247, San Francisco, CA, USA, 2002.