ViMiC (Virtual Microphone Control) is a new toolbox for real-time synthesis of spatial sounds, particularly for concert situations and sound installations, and especially for larger or non-centralized audiences. Based on the concept of virtual microphones positioned within a virtual 3-D room, ViMiC supports loudspeaker reproduction up to 24 discrete loudspeaker channels whereby the loudspeakers do not necessarily have to be placed uniformly and equidistant around the audience. Through the integrated Open Sound Control protocol (OSC), ViMiC is easily accessed and manipulated.

2. SPATIALIZATION WITH MAX/MSP

This section briefly overviews available loudspeaker spatialization techniques for Max/MSP. 

**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 loudspeakers. 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.

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

**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 certain order. To date, Max/MSP externals up to the 3rd order for horizontal-only or periphonic speaker arrays have been independently presented in [11] and [13].

**Space Unit Generator**, also called room-within-the-room model, dates back to [8]. Four loudspeakers representing "open windows" are positioning around the listener and creates an "inner room", embedded in an "outer room" with virtual sound sources. The created audio signals contain ICTD and ICLD. Specular Early reflections are created according to the size of the outer room. A Max/MSP implementation was presented in [16].

**Spatialisateur**, in development at IRCAM and Espaces Nouveaux since 1991, is a library of spatialization techniques, including VBAP, 1st Order Ambisonics and algorithms to simulate stereo recording techniques (XY, 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, originally developed for Pure Data \[2\], is an integral part of the network music project SoundWIRE\[^{1}\] and the Tintinnabulate Ensemble\[^{2}\] directed by Pauline Oliveros. ViMiC was recently rewritten and extended for the Max/MSP composer and sonic artist community, and will be applied in the MusiMarch Festival 2008\[^{3}\].

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, the axis orientations of their microphone directivity patterns. Besides the direct sound component, a virtual microphone signal can also include early reflections and an adequate reverberation tail. Upon both 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 greatly modified by changing the exponent \(q\). Thus, the effect of distance attenuation can be boosted or softened. The minimum distance to a loudspeaker is limited to 1 meter in order to avoid amplification.

\[
g_i = \frac{1}{d_i^q} \quad d \geq 1
\]  

Further attenuation happens through the chosen microphone characteristic and source directivity (see Fig. \[2\]). The directivity for a certain angle of incident \(\delta\) of all common microphone characteristics can be imitated by calculating Eq. \(2\) and applying a set of microphone coefficients of Table 3.1.1. By increasing the exponent \(w\) to a value bigger than one will produce artificially sharper directivity pattern. Unlike actual microphone characteristics, which vary with frequency, microphones in ViMiC are designed to apply the concept of microphone directivity and not to simulate undesirable frequency dependencies.

\[
\Gamma = (a + b \cdot \cos \delta)^w \quad 0 \leq a, b \leq 1
\]  

Directional properties of sound sources are 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 in 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 \[^{4}\]. 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 3-D through the well-known image method \[^{5}\]. 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 reflection pattern accordingly. The spectral influence of the wall properties are simulated through high-mid-low shelf-filters. Because larger propagation path 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 rays 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 demanding.

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

\[^{1}\]http://ccrma.stanford.edu/groups/soundwire/
\[^{2}\]http://www.myspace.com/tintinnabulate
\[^{3}\]http://www.festivalmnm.ca/mnm.e
\[^{4}\]http://www.festivalmnm.ca/mnm.e
\[^{5}\]http://www.festivalmnm.ca/mnm.e

![Figure 1](image-url)
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. ViMiC flowchart](image)

### 4. MOVING SOURCES

In Fig. [3] 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 characteristic, 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 remain static until one of the paths has been changed by more than a certain 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 strong audible phase modulations.

![Figure 3. Geometric example 1](image)

![Figure 4. Geometric example 2](image)

### 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 a 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. [14]) 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 edited or defined through cartesian and spherical coordinates (Fig.5).

5.2. Controllability
ViMiC was structured as higher level modules using the Jamoma framework for Max/MSP [9] to increase easy and flexible controllability. Jamoma offers a clear advantage in its standardization of presets and parameter handling. Each ViMiC parameter has been sorted into three primary namespaces: source, microphone and room; and has a specified data range. ViMiC is fully controlled through a GUI and through external OSC-messages [15].

/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 develope an object which handles multiple sound source and handles multiple sound sources are under discussion. ViMiC for Max/MSP is available in the SVN Repository of Jamoma.

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).

References

[16] S. Yadegari, F. R. Moore, H. Castle, A. Burr, and

Figure 5. Interface to position microphones

---

4 The [Ambimonitor] by ICST is used to display microphones [11].
5 https://jamoma.svn.sourceforge.net/svnroot/jamoma/branches/active