
**Jonas Braasch,* Nils Peters,†
and Daniel L. Valente***

*CA³RL, School of Architecture
Rensselaer Polytechnic Institute
110 8th Street

Troy, New York 12180 USA
{braasj, valend2}@rpi.edu

†CIRMMT, Schulich School of Music
McGill University
555 Sherbrooke Street West
Montreal, QC H3A 1E3 Canada
nils.peters@mcgill.ca

This article describes a new sound-projection system for multichannel loudspeaker setups that has been developed by the authors. The system, called *Virtual Microphone Control* (ViMiC), is based on the simulation of microphone techniques and acoustic enclosures. In auditory virtual environments (AVEs), it is often required to position an anechoic point source in three-dimensional space. When sources in such applications are to be displayed using multichannel loudspeaker reproduction systems, the processing is typically based upon simple amplitude-panning laws. With an adequate loudspeaker setup, this approach allows relatively accurate positioning of spatial images in the horizontal plane, but it lacks the flexibility many composers of computer music would like to have. This article describes an alternative approach based on an array of virtual microphones. In the newly designed environment, the microphones, with adjustable directivity patterns and axis orientations, can be spatially placed as desired. Each virtual microphone signal is then fed to a separate (real) loudspeaker for sound projection. The system architecture was designed for a maximum flexibility in the creation of spatial imagery.

Despite its flexibility, the system is intuitive to use because it is based on the geometrical and physical principles of microphone techniques. It is also consistent with the expectations of audio engineers to create sound imagery similar to that associated with standard sound-recording practice, but it goes beyond the original concept by allowing strategic violations of physically possible parameters; namely, new supernatural microphone directivity patterns can be implemented into the ViMiC system.

This article begins with a review of various microphone techniques on which the ViMiC system

A Loudspeaker-Based Projection Technique for Spatial Music Applications Using Virtual Microphone Control

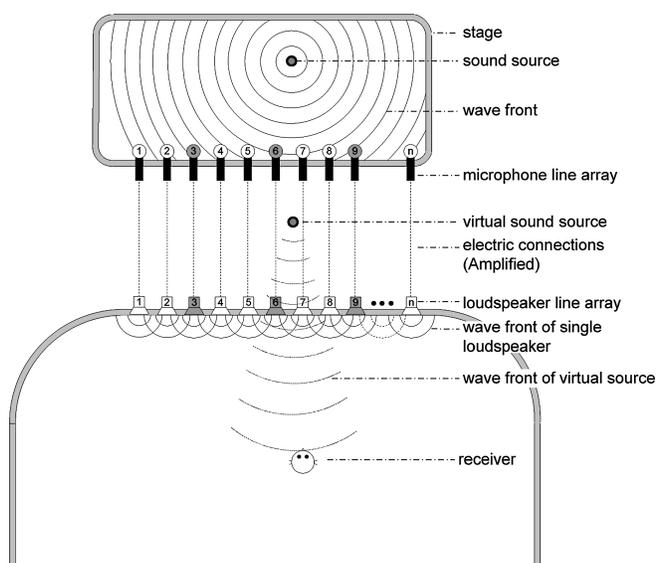
relies and alternative sound-projection techniques. Next, the fundamental physical concepts on which the ViMiC system is based are described. In the following section, software implementation of the system is outlined with a focus on strategies to keep processor load and system latency low. The article concludes with a description of several projects that involved the ViMiC system.

Historical Background

In the 20th century, electroacoustic and electromechanical devices were invented to spatialize sounds dynamically, before computers became advanced enough to fulfill this task with real-time audio-processing algorithms. First, the introduction of sonic spatialization techniques based on microphone arrays should be mentioned. In 1931, Alan Blumlein filed a patent on a newly developed two-channel recording scheme based on two bidirectional microphones with coincident diaphragms angled at 90° (Blumlein 1931). The family of stereo recording techniques with coincident microphone diaphragms were later referred to as *XY techniques*. In all XY techniques, the directionally dependent microphone sensitivities are used to encode the azimuth angle of the recorded sound source as level differences between both channels. Mr. Blumlein's invention, which was made practical through the introduction of the ribbon microphone by Siemens (Weiss 1993), was the birth of stereo recording.

Almost at the same time, Steinberg and Snow (1934) introduced another microphone-based recording technique that was used in 1933 for a telepresentation with Leopold Stokowski and the Philadelphia Orchestra. The basic idea was to capture a wavefront with several microphones, as shown in Figure 1. Their setup consisted of only two

Figure 1. Steinberg and Snow's (1934) recording setup.



or three transducers, owing to technical limitations, but this number was shown in a psychoacoustic experiment to be sufficient for spatial coding purposes. In their recording scheme, the directions of the sound sources are coded almost solely through inter-channel time differences and to a lesser extent through inter-channel level differences. This type of method, which is based on spaced microphone placements with two receivers, is now well known as the *AB technique* (or *spaced technique*) in the USA and parts of Europe.

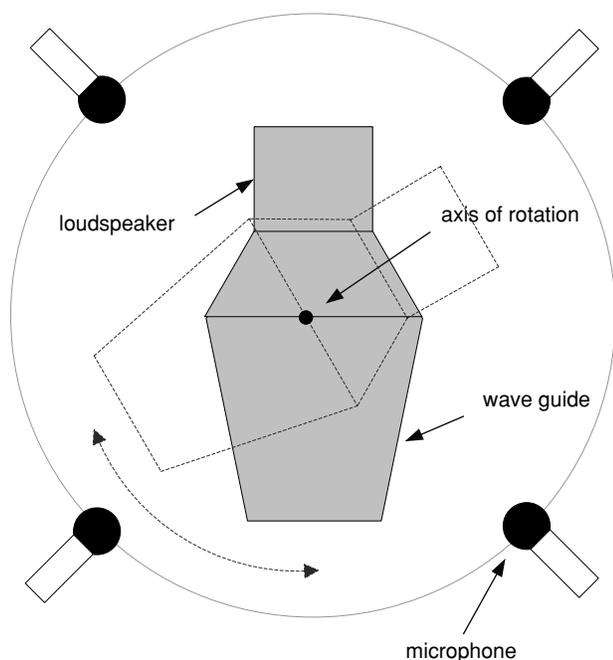
It took many years until a full “curtain of microphones” could be processed simultaneously to address a large loudspeaker array using the concept of wave field synthesis (WFS) (Berkhout 1988; Berkhout, Vogel, and de Vries 1993), which is based on Huygens’s Principle and the Kirchhoff–Helmholtz Integral. The present WFS technology shares two important features with the original work of Steinberg and Snow: Both systems only capture and process sound pressure and disregard the wave field’s particle velocities, and both systems claim to lead to a superior depth perception. In fact, based on their psychoacoustic findings, Steinberg and Snow recommended the use of three audio channels instead of two to allow for a better coding of depth and to achieve a more stable center image. Given these facts, one can claim that both the

AB technique and WFS have the same origin in the pioneer work of Steinberg and Snow at Bell Laboratories.

Ambisonics is an alternative means of accurately encoding and decoding a sound field. The underlying theory focuses on describing the sound field at a singular point using the Fourier-Bessel Series. For a long time, only the zero- and first-order Bessel functions were considered; the zero order represents the monopole component, and the three first-order components represent the dipoles in the x -, y -, and z -directions. Theoretically, the four channels can be recorded using an omnidirectional microphone for the monopole component and three bidirectional microphones for the dipole components. However, the only commercially available ambisonic microphone, the Soundfield Microphone, operates with four cardioid microphone capsules that are mounted on a tetrahedron. The first- and second-order signals can be easily derived from the four cardioid signals. Gerzon (1973, 1985) developed the mathematical framework for ambisonics, which also includes procedures to convert the signals among various formats and to generate loudspeaker signals for different configurations to reproduce a recorded soundfield. Recently, higher-order (second-order and above) ambisonic systems have received great attention (Malham 1999; Daniel 2000; Daniel, Nicol, and Moreau 2003). Because physical microphones are still restricted to the first order, their soundfields must be computer-generated, or higher-order microphone signals must be simulated using complex microphone arrays (e.g., Laborie et al. 2003). Although theoretically, higher-order ambisonics systems only consider the soundfield reproduction of a single point in space, practically, soundfields with large “sweet-spot” areas can be generated similar in size to those created by WFS.

As an alternative to XY and AB techniques, near-coincident techniques such as ORTF (named after the French national broadcasting agency *Office de Radiodiffusion et de Télévision Française* where it was first introduced) and NOS (named after the public Dutch broadcaster *Nederlandse Omroep Stichting*) were developed, in which the spatial coding process is derived from a unique combination of inter-channel time differences (ICTDs)

Figure 2. Functional sketch of Stockhausen's rotational table.



and inter-channel level differences (ICLDs) that gives each technique its individual sound. The general difference between near-coincident and AB techniques is the amount of space between both microphones' diaphragms, which is typically less than 20 cm for near-coincident techniques and above this value for the AB technique. Omnidirectional microphones are commonly used in the AB technique, whereas the microphones used in near-coincident techniques typically have a cardioid, sub-cardioid, or hyper-cardioid pattern.

Spatialization of sounds in early electroacoustic music pieces was also frequently achieved using microphones. An interesting example for such an approach is the rotational table (*Rotationstisch*) of Karlheinz Stockhausen. The device's fundamental principle is depicted in Figure 2. The spatial orientation of the loudspeaker was recorded with four microphones, which encoded both ICTDs and ICLDs. The inter-channel time differences created audible artifacts that are related to the Doppler shift while the loudspeaker is in angular motion. In particular, all channels become slightly detuned in pitch relative to each other, because the sound source always approaches a subset of the

microphones while moving away from others. This effect cannot be observed for coincident techniques. Although inter-channel time differences create the perception of plasticity, they also lead to artifacts and an aesthetic that could be termed "surrealistic." It is questionable whether the artifacts induced by the Doppler shift were intended by Mr. Stockhausen or whether he was originally interested in a realistic-sounding method.

In modern electroacoustic compositions, individual sonic elements are generally spatialized using computer algorithms. The first system of this kind was introduced by Chowning (1971, 1977), a quadraphonic system based on amplitude panning using the tangent law and Schroeder's (1962) reverberation algorithm.

Another approach to spatializing sounds is based on virtual microphones. Here, virtual omnidirectional microphones are employed in auditory virtual environments to calculate the gains and delays between the virtual sound sources and the virtual microphones that can be positioned freely in a computer-generated space (Corey et al. 2001; Mouchtaris, Narayanan, and Kyriakakis 2003). The classic approach of Moore (1983) initiated the idea of virtual microphones, although Moore's concept was to create a room within a room for sound reproduction with a quadraphonic loudspeaker system. His virtual sensors were spaced as omnidirectional microphones in rectangular configuration. The crosstalk among channels was reduced through fully absorbent virtual walls along the rectangular basis pattern with windows at the microphone positions.

Publicly available real-time audio software systems such as Max/MSP and Pure Data have since revolutionized the way many artists work. Most of these applications are loudspeaker-based display systems similar to the one introduced by Mr. Chowning, where sound sources can be virtually placed in between speakers by using amplitude-panning laws to calculate the gain factor for each speaker for a given sound-source position—e.g., the tangent law (Jot 1992) and Vector-Based Amplitude Panning (VBAP), described by Pulkki and Karjalainen (2001) and Pulkki (2001). By addressing only the two or three speakers that are nearest to the virtual sound source, the apparent source width (the perceived

spatial extent of the sound-source image) can be kept fairly narrow, although it will typically be heard as wider than an actual point source in space. One can show that pure inter-channel amplitude differences between two frontal speakers transform well into the natural combinations of inter-aural time differences and inter-aural level differences as they are found for natural sound sources at the ear entrances of a human observer (Braasch 2005).

Alternatively, panning laws can also be derived from virtual directivity patterns of two or more receivers, which generally allows more flexibility than implementations based on the tangent law. Martin et al. (2001) tested human localization for an amplitude-panning technique that had been derived from cardioid microphone directivity patterns. Later, Studer developed a digital mixing console in which the panning laws for standard surround formats are derived from classic microphone techniques (Horbach et al. 2000). TC Electronic's multichannel processing System 6000 also incorporates virtual directivity patterns for sound sources and receivers (Nielsen 2001). Both the Studer and TC Electronic systems process early reflections in addition to the direct sounds using the same spatialization techniques.

The use of panning laws for 3-D sound projection seems to be the optimal choice for many applications that require realistic reproduction of a recorded event, for example applications in architectural acoustics. However, in the case of music reproduction, experienced listeners often prefer the artificially designed spaces associated with microphone-based audio engineering practices over more realistic sounding virtual rendering techniques. In electroacoustic music, an even more abstract and surrealistic environment is often desired.

Motivation

One of the reasons why many electroacoustic composers and performers use tools that were designed for lifelike sound reproduction is obviously the ready availability and minimal learning curve, rather than the unique sound they produce. One

of the keys for the success of a new auralization tool is to allow the manipulation of many spatial parameters while maintaining the intuitive control of present spatialization techniques.

Consequently, a new architecture for a virtual environment was developed to meet the needs of experimental spatial music. Instead of using panning laws to address loudspeakers, the system is based on an array of virtual microphones. In addition to the approaches of Corey et al. and Moore, the virtual microphones in the investigation reported here have been equipped with virtual directivity patterns. The axial orientation of these patterns can be freely adjusted in 3-D space, and the directivity patterns can be varied between the classic patterns: omnidirectional, cardioid, hyper-cardioid, sub-cardioid, or figure-eight characteristics that are found in actual microphones. This improvement allows a new way to control the spatial images in AVEs, a feature that characterizes Virtual Microphone Control.

In most AVEs, the physical parameters of the simulated environments (e.g., room dimensions, wall and floor materials) are adjusted to achieve the desired sound, but in a typical sound recording situation, the variation of room parameters is usually limited to the use of curtains and baffles. For this reason, sound engineers typically adjust the positions of the microphones relative to each other to capture the musical event in an optimal way. The proposed ViMiC system preserves these procedures to control the sound image in virtual environments. The goal was also to develop a technique that can be used with great flexibility. To serve the needs for electroacoustic composers, the architecture was designed to allow for dynamic changes of all parameters in real time, thus going beyond the concept of simply positioning sources in space using preprocessed concepts. As will be demonstrated later, ViMiC incorporates many previous approaches into one design, enabling the dynamic and interchangeable use of different techniques in parallel.

The basic concept of ViMiC is described in the following section. An interesting alternative within the ViMiC environment is the use of artificial directivity patterns for microphones. If desired, directivity patterns that represent various classic

panning laws can be easily implemented within the new system. Consequently, systems based on amplitude panning laws become a subset of systems made available by ViMiC. As an example, the implementation of the tangent panning law is demonstrated.

Basic Concepts

Virtual Microphone Control

To simplify the discussion, ViMiC will be first introduced for the anechoic situation; the generation of reflections and reverberation will be addressed later. In a typical recording situation, the transfer function between the sound source (e.g., a musical instrument treated as a one-dimensional signal in time $x(t)$) and a receiving microphone with signal $y(t)$ is determined by the distance and the orientation between the microphone's directivity pattern and the instrument. The distance determines the delay τ between the radiated sound at the source and its arrival at the microphone diaphragm:

$$\tau(r) = \frac{r}{c_s}, \quad (1)$$

with the distance r in meters and the speed of sound c_s . The latter can be approximated as 344 m/sec at room temperature (20°C). According to the $1/r$ law, the sound-pressure level radiated by a sound source will decrease by 6 dB with each doubling of the distance r :

$$p(r) = \frac{p_0 \cdot r_0}{r}, \quad (2)$$

with the sound pressure p_0 of the sound source at a reference distance r_0 . In addition, it should be considered that the sensitivity of a microphone varies with the angle of incidence according to its directivity pattern. In theory, only omnidirectional microphones are equally sensitive toward all directions, and in practice even this type of microphone is more sensitive toward the front for high frequencies, due to the physical limitations of the diaphragm size as well as the planar, directional nature of higher frequencies. The circumstance that real

Table 1. Values of a and b for Various Microphone Polar Responses

a	b	
1	0	omnidirectional
0.75	0.25	sub-cardioid
0.5	0.5	cardioid (unidirectional)
0.25	0.75	hyper-cardioid
0	1	figure-eight (bidirectional)

microphones generally have rotational directivity patterns simplifies their implementation into the AVE, because these types of directivity patterns $\Gamma(\alpha)$ can be written in a simple general form:

$$\Gamma(\alpha) = a + b \cos(\alpha), \quad (3)$$

where the variable α is the incoming angle of the sound source in relation to the microphone axis. Typically, the maximum sensitivity $a + b$ is normalized to one ($b = 1 - a$), and the different available microphones can be classified using different combinations of a and b , as shown in Table 1.

The overall gain g between the sound-source signal $x(t)$ (treated here as an omnidirectional point source) and a virtual microphone signal $y(t)$ can be determined as follows:

$$g = g_d \cdot \Gamma(\alpha), \quad (4)$$

with the distance-dependent gain $g_d = r_0/r$. For other sound sources, this equation can be generalized to

$$g = g_d \cdot \Gamma(\alpha) \cdot \Gamma(\beta), \quad (5)$$

assuming that the sound source has a rotational radiation pattern $\Gamma(\beta)$. The transfer function between the sound source and the microphone signal can be determined by two parameters only: the gain g and the delay τ , if the microphone and source directivity patterns are considered to be independent of frequency. Note that the directivity patterns of most real microphones are not fully independent of frequency, although this is often a design goal. The relationship between the sound radiated from an omnidirectional point source $x(t)$ and the

Figure 3. Microphone placement for (a) the Blumlein XY technique with two bidirectional microphones, and (b) the ORTF technique with two unidirectional microphones.

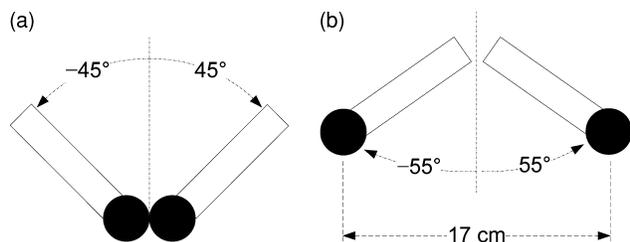


Figure 3

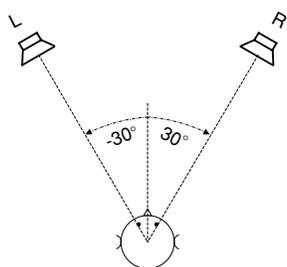


Figure 4

microphone signal $y(t)$ is found to be

$$y(t, r, \alpha) = g \cdot x(t - \tau) = g_d(r) \cdot \Gamma(\alpha) \cdot x\left(t - \frac{r}{c_s}\right). \quad (6)$$

By simulating this physical relationship within the AVE, the sound sources can be panned in the virtual space according to standard sound recording practices.

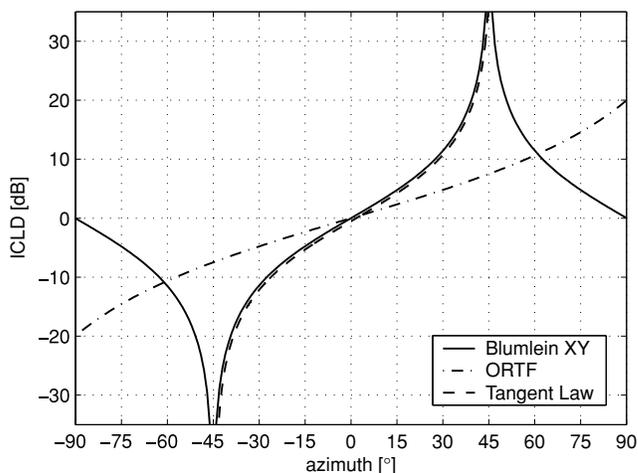
The Blumlein XY Technique

A good example to demonstrate ViMiC is the classic XY microphone technique mentioned in the introduction. Here, two bidirectional microphones are arranged at an angle of 90° in the horizontal plane as shown in Figure 3a. Theoretically, both microphone diaphragms are at the same location in space, which is not possible in a real setup as shown in the figure. The ratio between the signal amplitude at the sound source $x(t)$ and microphone signal amplitudes for the left and right channels $y_1(t)$ and $y_2(t)$ vary with the angle of incidence according

Figure 4. Standard stereo loudspeaker setup.

Figure 5. Inter-channel level differences as a function of azimuth for different recording and panning techniques. The Blumlein XY technique is based on two bidirectional

microphones with coincident placement; the ORTF technique uses two unidirectional microphones with near-coincident placement.



to Equations 3 and 6:

$$y_1(t) = g_d \cdot \cos(\alpha + 45^\circ) \cdot x(t - \tau), \quad (7)$$

$$y_2(t) = g_d \cdot \cos(\alpha - 45^\circ) \cdot x(t - \tau). \quad (8)$$

In general, both amplitude and time differences between the microphone channels determine the position of the spatial image that a listener will perceive when both microphone signals are amplified and played through two loudspeakers in standard stereo configuration (see Figure 4). When a virtual sound source is encircling the microphone setup in the frontal horizontal plane at a distance of 3 m ($-90^\circ < \alpha < +90^\circ$), the ICLD ρ as shown in Figure 5 can be calculated as follows:

$$\rho(\alpha) = 20 \cdot \log_{10} \left(\frac{y_2(t)}{y_1(t)} \right) \quad (9)$$

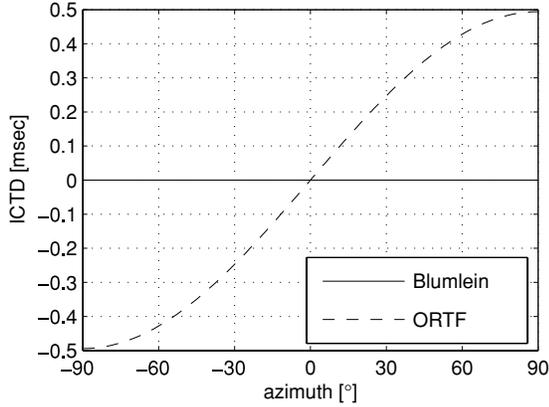
$$= 20 \cdot \log_{10} \left(\frac{g_d \cdot \cos(\alpha - 45^\circ)}{g_d \cdot \cos(\alpha + 45^\circ)} \right) \quad (10)$$

$$= 20 \cdot \log_{10}(\tan(\alpha + 45^\circ)). \quad (11)$$

ICTDs do not occur here, because both microphone diaphragms coincide (see Figure 6, solid line). This has been frequently criticized, apparently because the ICTDs are often confused with the inter-aural time differences (ITDs) that occur between the listener's eardrums—even though the underlying theory has been previously published. In fact, the binaural cues (inter-aural time and level differences)

Figure 6. Inter-channel time differences as a function of azimuth for different recording techniques. The Blumlein XY technique is based on two bidirectional

microphones with coincident placement; the ORTF technique uses two unidirectional microphones with near-coincident placement.



are fairly accurately encoded in the Blumlein recording technique, and the ITDs are generated within the transmission paths from the loudspeakers to the listener's eardrums (Braasch 2005). In fact, Blumlein's XY technique creates ICLDs that matches the tangent law for two loudspeakers that are placed at -45° and $+45^\circ$ (cf. Figure 5).

The ORTF Technique

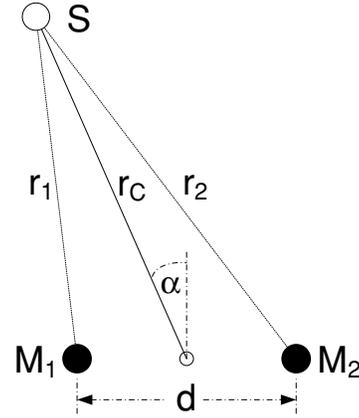
A stereo setup often uses two cardioid microphones to replace the bidirectional microphones. Owing to the broader width of the directivity lobe of the cardioid pattern compared to the lobe of the figure-eight pattern, the angle between both microphones is typically adjusted wider than was the case for the Blumlein technique (e.g., 110° instead of 90°). Again, the ratio between the signal amplitude at the sound source $x(t)$ and signal amplitude at the microphones $y(t)$ can be easily determined for both microphones:

$$y_1(t) = g_d \cdot 0.5 \cdot (1 + \cos(\alpha + 55^\circ)) \cdot x(t - \tau), \quad (12)$$

$$y_2(t) = g_d \cdot 0.5 \cdot (1 + \cos(\alpha - 55^\circ)) \cdot x(t - \tau). \quad (13)$$

The ICLD ρ can be calculated for this setup analogous to the Blumlein technique that was described in Equations 9–11. Figure 5 shows the ICLD as a function of the angle of incidence α . Apparently, the level difference between both microphones remains to be rather low for all angles when compared to the XY technique. However,

Figure 7. Physical relations in a two-channel, near-coincident microphone setup, M_1 and M_2 , to record a point source S .



increasing the angle between both microphones is rather problematic, as this would result in a very high sensitivity of the setup toward the sides. Instead, the diaphragms of both microphones are spaced 17 cm apart in the ORTF configuration (cf. Figure 3b). This way, ICTDs τ_Δ are generated in addition to the ICLDs. The ICTDs can be easily determined from the geometry of the setup (see Figure 7):

$$\tau_\Delta(\alpha) = \frac{(r_1 - r_2)}{c_s}, \quad (14)$$

where c_s is the speed of sound, and

$$r_{1,2} = \sqrt{r_c^2 + (d/2)^2 - r_c d \cos(90 \pm \alpha)}. \quad (15)$$

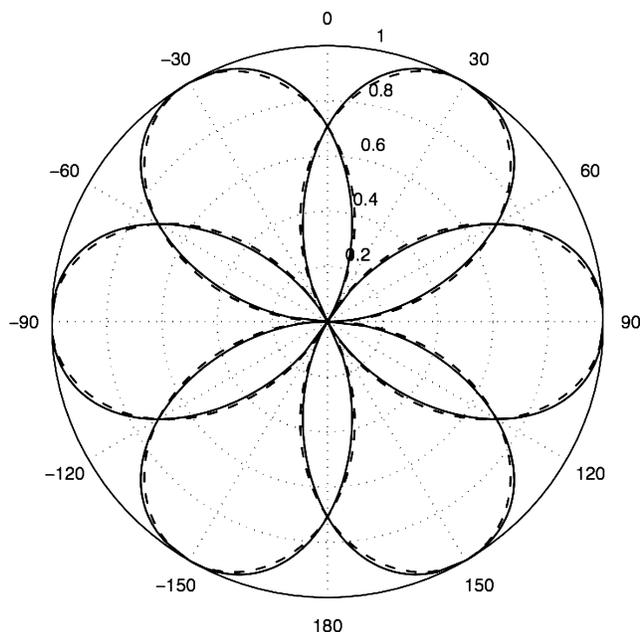
The variable d is the distance between both microphones in meters, and r_c is the point of symmetry between both microphone diaphragms. In most cases, it is sufficient to estimate the ICTD using the far-field approximation:

$$\tau_\Delta(\alpha) = \frac{d}{c_s} \sin(\alpha). \quad (16)$$

Note that even though the microphones are now spaced 17 cm apart, Equations 12 and 13 are still valid to determine the ICLDs for most cases. The $1/r$ term can be neglected in the ICLD calculation for the far-field condition, i.e., when the distance between the sound source and the center of the recording setup r is much larger than the distance d between both microphone diaphragms ($r \gg d$).

Figure 8. Polar plots of the sensitivity magnitudes of the directivity patterns that were derived using cosine functions according to Equation 18 (dashed lines) compared to those derived from the tangent panning law with constant power (solid lines). The plot shows an arrangement of six coincident (or near-coincident) microphones with a constant angle of 30° between two adjacent

microphone angles. This arrangement could be used to address a circular speaker array of six equally spaced loudspeakers.



In this case, which holds true for typical recording scenarios, the occurring ICLDs are still almost solely generated by the different orientations of the cardioid patterns of both microphones, because the term g_d has approximately the same value in both Equations 12 and 13.

To simulate spaced-microphone techniques, the ICLD ρ is determined by applying the $1/r$ law (Equation 2), as has been often described in previous literature (e.g., Mouchtaris, Narayanan, and Kyriakakis 2003):

$$\rho = 20 \log_{10} \left(\frac{r_2}{r_1} \right). \quad (17)$$

The ability to control the spatial image of sound sources by simulating microphone-based recording techniques using dynamically adjustable virtual microphones with various directivity patterns is a novel feature of the proposed AVE. In particular, the AVE allows audio engineers to intuitively apply time-based panning methods by simulating near-coincident microphone techniques. The AVE is, so to speak, a physical model of a recording/reproduction chain with virtual microphones and real loudspeakers. It can be assumed that our auditory system has

been accustomed to these kinds of artificial binaural cues, because many commercially available recordings have been established this way. This was an important factor for the design of the proposed AVE.

Implementation of Classic Panning Laws

The flexibility of an AVE to simulate classic microphone techniques is restricted by the relatively broad directivity patterns of physical microphones. In order to allow pure amplitude-panning techniques, artificial microphone directivity patterns are implemented into the system. A simple method to create a rotational directivity pattern for this case is to use the cosine function in the following way:

$$g_i = \begin{cases} \cos \left(\frac{\varphi}{\varphi_0} \cdot \frac{\pi}{2} \right) & : 0^\circ \leq \varphi \leq \varphi_0 \\ 0 & : \varphi_0 < \varphi \leq 180^\circ \end{cases} \quad (18)$$

where φ_0 is the angular distance to the next speaker. The approach is useful for equi-angular loudspeaker placements (e.g., 45° or 60°), which are very common in electroacoustic music. In this case, the sensitivity is unity if the source is facing the corresponding speaker and declines to zero when the source reaches an adjacent loudspeaker. An example for a loudspeaker arrangement with 60° inter-loudspeaker spacing is shown in Figure 8. Note that the differences between the directivity patterns produced by the tangent and cosine panning laws are very small, which explains why both algorithms are common in audio engineering practice.

For setups with variable angles between loudspeakers, such as the five-channel surround arrangement according to the ITU standard BS.775-1 (1994), asymmetrical directivity patterns can be created. In the following example, the directivity patterns are designed to meet the tangent panning law that is frequently used to pan a signal in between two adjacent loudspeakers. The tangent panning law is defined as

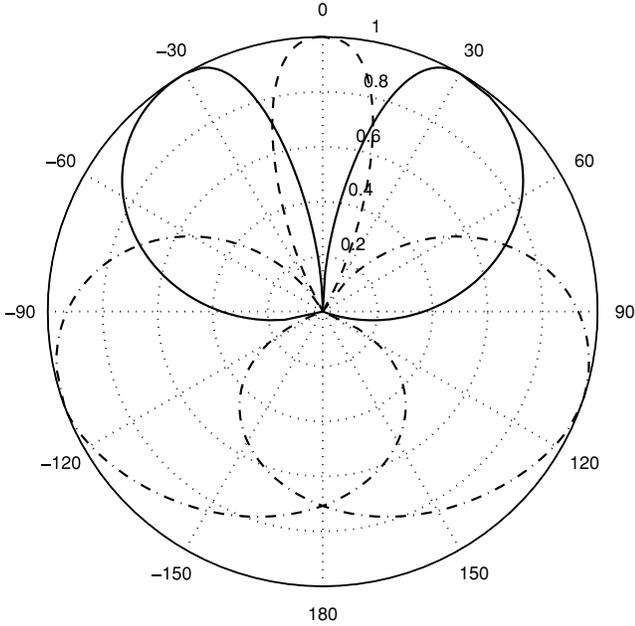
$$\frac{\tan \alpha}{\tan \alpha_0} = \frac{g_1 - g_2}{g_1 + g_2}. \quad (19)$$

where α_0 is half the angle between both loudspeakers, α is the angle of the virtual sound source

Figure 9. Polar plots of the sensitivity magnitudes of the directivity patterns that were derived from the tangent panning law with constant power. The figure depicts a five-channel

microphone surround arrangement that can be used to address a five-channel surround sound loudspeaker system. The maxima of the lobes correspond to the

microphone/loudspeaker angles (solid lines, frontal left and right channels; dashed line, frontal center channel; dashed-dotted lines, surround channels).



measured from the midline between both loudspeakers, and g_1 and g_2 are gain factors for both speakers. Typically, the relationship between g_1 and g_2 is chosen such that their summed power, which is proportional to the summed squared gain factors, is always constant:

$$g_1^2 + g_2^2 = C. \quad (20)$$

Using the tangent panning law, the gain g_i of each speaker can be translated into a virtual microphone directivity pattern, assuming that the angles of the microphones correspond to the angles of the speakers:

$$g_i = \begin{cases} \sqrt{\frac{C \cdot (1 + a_l)^2}{2 \cdot (1 + a_l^2)}} & : \alpha_l \leq \alpha \leq \alpha_c \\ \sqrt{\frac{C \cdot (1 + a_r)^2}{2 \cdot (1 + a_r^2)}} & : \alpha_c \leq \alpha \leq \alpha_r \\ 0 & : \text{else} \end{cases} \quad (21)$$

where α_c is the horizontal angle of the specified speaker, $\alpha_l < \alpha_c$ is the angle of the closest speaker to the left, and $\alpha_r > \alpha_c$ the angle of the closest speaker to the right. The variables a_l and a_r are defined as

follows:

$$a_l = \frac{\tan(\alpha - 0.5 \cdot (\alpha_c + \alpha_l))}{\tan(0.5 \cdot (\alpha_c - \alpha_l))}, \quad (22)$$

$$a_r = \frac{\tan(\alpha - 0.5 \cdot (\alpha_c + \alpha_r))}{\tan(0.5 \cdot (\alpha_c - \alpha_r))}. \quad (23)$$

Figure 9 depicts the virtual directivity patterns for the five microphones of the standard surround setup. Note that the directivity patterns no longer show rotational symmetry as was the case the arrangement with cardioid directivity patterns shown in Figure 8.

Figure 9 shows a cross-section through a three-dimensional space. In this representation, a directivity pattern with rotational symmetry would be symmetric around its axis, which is spanned by the origin of the polar plot and the point of maximum sensitivity of the directivity pattern. In Figure 9, the left lobe of each pattern and the right lobe of the pattern left of it comply with the tangent panning law. Theoretically, a true three-dimensional microphone pattern can be synthesized, for example, based on the VBAP theory (Pulkki and Karjalainen 2001; Pulkki 2001). So far, a simpler method was pursued by decreasing the sensitivities of the virtual microphones with the factor $\cos(\vartheta)$ in the vertical dimension. (The variable ϑ is the elevation.)

In the current approach, the variable C in Equation 20 is decreased with the distance r between the sound source and microphone position according to the inverse-square law:

$$C(r) = \frac{C_0 \cdot r_0^2}{r^2}, \quad (24)$$

with C_0 set to 1 and r_0 adjusted to 0.1 m. To avoid clipping of the audio signal that is controlled with the directivity pattern, $C(r)$ is limited to values between 0 and 1.

Wave Field Synthesis

A wave-field synthesis (WFS) system can be easily simulated using ViMiC by following the original approach of Steinberg and Snow (1934). Instead of placing a curtain of real microphones in a

concert hall, an array of virtual microphones can be placed in the ViMiC environment. Ideally, the virtual microphone positions should correspond to the loudspeaker positions of the sound projection arrangement to capture the virtual wave front of a point source. The microphone signal of the n th omnidirectional microphone is determined in analogy to Equation 6:

$$y_n(t, r) = g_n \cdot x(t - \tau_n) = g_d(r_n) \cdot x\left(t - \frac{r_n}{c_s}\right), \quad (25)$$

where r_n is the distance between the n th microphone and the sound source. In principle, the achievable results with the ViMiC WFS approach are identical to traditional WFS implementations, and corrections for truncated and cornered arrays can be simulated through changes in the positions and the directional and frequency-dependent sensitivities of the virtual microphones. Because WFS can be integrated into the general framework of ViMiC, no additional software is needed to create a WFS (sub-)system.

Simulation of the Diffuse-Field

When simulating classic sound-recording techniques, one must take into account the fact that the ratio between the on-axis sound pickup of the microphone and the pickup of the diffuse sound, namely the late reverberation, varies among different directivity patterns. For each pattern, the relative power P_r of the recorded diffuse sound can be calculated from the intensity over a whole sphere:

$$P_r = \int_0^{2\pi} \int_0^\pi (\Gamma(\alpha))^2 \sin(\varphi) d\varphi d\vartheta. \quad (26)$$

The solution that covers all classic directivity patterns of the type $\Gamma(\alpha) = a + (1 - a)\cos(\alpha)$ (as defined in Equation 3) is

$$P_r(a) = 4\pi \left(a^2 + \frac{1}{3}(1 - a)^2 \right). \quad (27)$$

Important here is the ratio between each directivity pattern and the omnidirectional pattern. The latter shows the highest value of 4π . By definition, the power (and the sound pressure) of the direct and

diffuse sound should be equal at the *critical distance* in the stationary case. For an omnidirectional sound source and an omnidirectional microphone directivity pattern, the critical distance can be estimated as follows:

$$r_c = \sqrt{\frac{A\bar{\alpha}}{16\pi}}, \quad (28)$$

where A is the total area of the enclosing surfaces. The equivalent absorption coefficient $\bar{\alpha}$ is defined as

$$\bar{\alpha} = \frac{1}{A} \sum_{n=1}^N A_n \alpha_n. \quad (29)$$

The areas of the walls, ceiling, and floor A_n and the average absorption coefficient for each area α_n are determined by the simulated room characteristics. Using the critical distance, the stationary amplitude of the reverberation field y_r can be estimated for an omnidirectional microphone pattern:

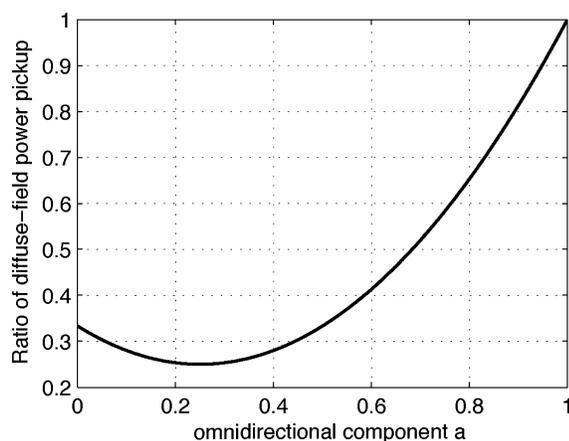
$$y_r = \frac{r_0}{r_c} \cdot y_s \quad (30)$$

The distance r_0 corresponds to the reference distance at which the signal gain was chosen to be one ($r_0 = 10$ cm in the current implementation). The parameter y_s represents the stationary amplitude of the sound source, as measured/simulated in the free field at the reference distance r_0 . Strictly speaking, the formula only applies for stationary signals y_s . For directivity patterns that are not omnidirectional, Equation 30 must be modified through multiplication by the square root of Equation 27 divided by 4π , which is the result for the omnidirectional case:

$$y_r = \frac{r_0}{r_c} \cdot \sqrt{a^2 + \frac{1}{3}(1 - a)^2} \cdot y_s. \quad (31)$$

The results for the power ratios between on-axis and diffuse sound pickup are shown in Figure 10 as relative values compared to the omnidirectional case.

Figure 10. Power ratio for diffuse sound pickup between a directivity pattern with omnidirectional component a and the omnidirectional pattern.



Implementation

The ViMiC system has been implemented using Pure Data (PD), a software platform for real-time audio applications that was developed by Miller Puckette. In Pure Data, as in Mr. Puckette's earlier software environments Patcher and Max, new applications can be developed by graphically linking various objects. Owing to the complexity of the proposed system, several objects (`SoundFieldRenderer`, `MultiTapDelay~`, and `Visualizer`) were written as new C externals. The architecture of the system is shown in Figure 11. The ViMiC system has been successfully implemented onto all three major operating systems: Windows XP, Linux, and Mac OS X.

The SoundField Renderer

The `SoundFieldRenderer` calculates the gain and delay between each sound source and virtual microphone based on the concept described earlier. Strictly speaking, the soundfield is only calculated at the microphone positions. The three-dimensional Cartesian coordinates of the sound-source positions are specified in meters. The coordinates can be entered directly or alternatively by using a joystick that controls the velocity of the source position in the xy dimensions. The third option is to use the `Visualizer`. In this C external, all objects (sound sources and microphones) are graphically displayed

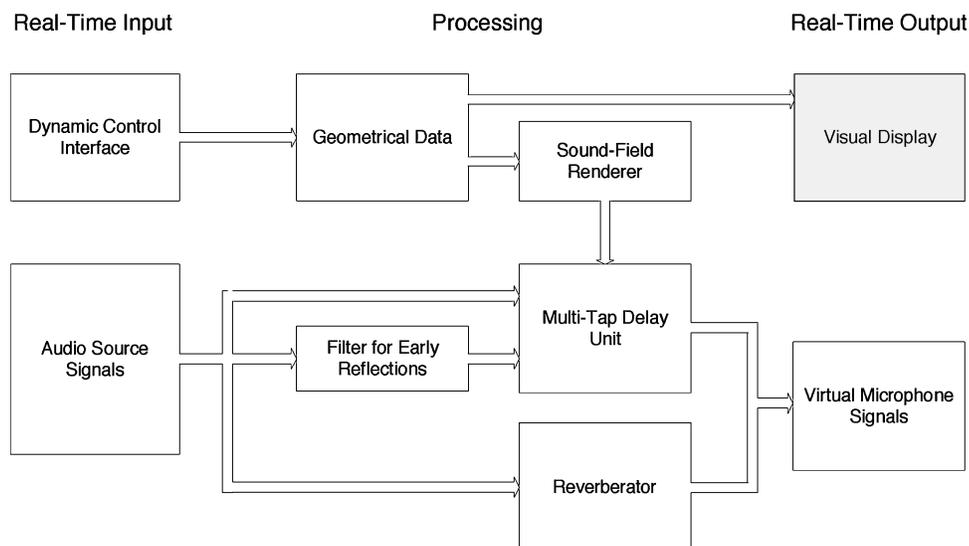
and can be moved in real-time using the mouse. The visualizer combines both the Dynamic Control Interface and Visual Display of Figure 11.

The positions of the microphones, their axes of orientation in both the horizontal and the vertical planes, and their directivity patterns can be entered separately for each individual microphone. A microphone array can also be adjusted globally according to the ITU standard for surround playback. In this mode, the angles for the front and surround microphones and the distances of the microphones from their common center can be adjusted. Afterward, the position of the microphones can be changed individually to improve the sound, a practice common in professional sound recording practice. At the present, all major directivity patterns (e.g., omnidirectional, cardioid, and figure-eight) can be modeled, as well as the artificial patterns for simulating the tangent panning law. Alternatively, directivity patterns can be read from files.

In addition to handling the direct sound sources, the `SoundFieldRenderer` module calculates the gains and delays between all first-order reflections and the microphones. Second-order reflections can be rendered as well if needed. To reduce the computational load, the second-order reflections can be kept at a static position rather than being updated with source movement. The coordinates of the reflections are calculated using the mirror-image technique (Allen and Berkley 1979). The directivity pattern of the mirror image is derived by mirroring the pattern of the direct sound source. The present algorithm is currently limited to rectangular rooms, because the Allen-Berkley algorithm only covers this type of room. Other techniques such as ray tracing could be implemented into the ViMiC system should it become necessary to simulate more complex room shapes. The three room dimensions (width, length, and height) can be set freely. The absorption coefficients of the walls, ceiling, and floor are simulated using a first-order low-pass filter or an FIR filter that can be fed with measured impulse responses of wall reflections. The simulation of diffuse acoustic wall reflections has not been implemented yet.

In the present implementation, one `SoundFieldRenderer` module determines the gains and

Figure 11. Architecture of the auditory virtual environment based on ViMiC.



delays for up to 24 microphone channels. Typically, the microphones are positioned at the same angles as the loudspeakers that reproduce the virtual microphone signals. For each new sound source, another `SoundFieldRenderer` module must be created.

A Multichannel Delay Unit

Using the data provided by the `SoundFieldRenderer` module, the dry sound is processed using a multi-tap delay network (in the module `MultiTapDelay`) for auralization of the specified sound field. The gain and delay for each output tap is taken from the output of the `SoundFieldRenderer` module. Even though PD provides several modules to write to and read from delay lines (`delwrite`, `delread`, `vd`), the multichannel delay unit was written as a C external for two reasons: to reduce the computational load and to optimize the interpolation algorithm.

Reduction of the Computational Load

Although the classes provided by PD are very efficient for most applications, one must keep in mind that the multi-tap delay network is a rather atypical

application, because the model requires a very large number of reflections. For the `SoundFieldRenderer` module, for example, 42 output taps must be processed if first-order reflections are considered (one direct source plus six first-order reflections, multiplied by six output channels), or 114 output taps if second-order reflections are simulated as well. Naturally, this number increases when more channels or sound sources are used. In order to minimize the computational load, several measures were taken. Firstly, both the delay-write and delay-read functions are hosted in the same module to avoid having to compensate an additional delay that occurs when the delay-read function is called by the CPU before the delay-write function. In the present system, the output taps outnumber the input taps to a great extent, because only three inputs are needed (one for the direct sound source, and two for the first- and second-order reflections that are fed with filtered signals). For this reason, the effort was made to write each delay-line element twice (the second time by adding the delay-line length to the pointer such that a copy of the delay line follows the first delay line). This way, only the pointer position for the delay line input must be observed and re-initialized when it reaches the end of the delay line. The pointer for the outputs can now be determined more easily by simply subtracting

the required number of delay taps from the second input-signal pointer.

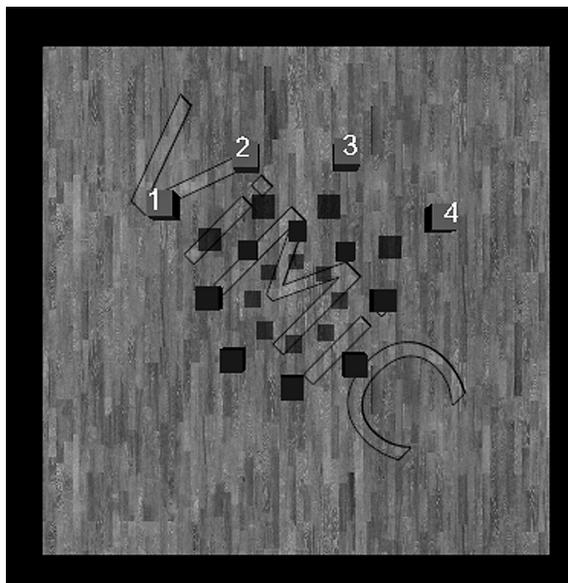
Interpolation Algorithms

Pure Data allows the use of fractional delays in the module `vd~`. Fractional delays have great advantages, for example when tuning a physical resonator model. In the case of the room model, the spatial resolution would be sufficient without interpolation (at a 48-kHz sampling frequency, one tap corresponds to a length of approximately 7 mm, when $c_s = 344$ m/sec), but the use of fractional delays has the advantage of increasing the “smoothness” for moving sound sources. To save resources, the multi-tap delay is programmed such that the four-point interpolation is only used when the sound source is being moved. Otherwise, the position of the sound source is rounded to the next tap, and the system switches back to using non-fractional delays. The multi-tap delay unit interpolates the input gain and delay values linearly at a rate that can be set in units of signal processing blocks. The clock rate of the `SoundFieldRenderer` module is set externally, e.g. by using the `PD metro` function. At the present, the `SoundFieldRenderer` operates at a frame rate of 10 msec.

Reverberation

In the present implementation, the late reverberation is generated using a multichannel reverberation algorithm based on feedback loops as proposed by Miller Puckette. The algorithm is available in the `nSLAM` toolbox, available online at tot.sat.qc.ca/down/nslam/nSLAM-2.0_MSP.zip. Modifications were made to this algorithm to fulfill the particular needs for the proposed system. First, the algorithm that generates the first- and second-order reflections was removed, because these are processed dynamically and with greater resolution in the multi-tap delay unit. Secondly, the reverberation unit was extended from 4 to up to 24 channels by increasing the number of feedback loops with slightly altered delay times.

Figure 12. ViMiC visualization tool to display sound source positions (numbered boxes) and microphone positions (unnumbered boxes).



Control Interfaces

Control Data Protocol

The communication between the ViMiC environment and its user interfaces is established through OpenSound Control (OSC; Wright, Freed, and Momeni 2003). The OSC protocol allows users to address the ViMiC environment from another computer through a network connection. Currently, the following ViMiC parameters can be modified using OSC: the position of the sound sources (x, y, z); the ratio of direct, early, and late reflections; absorption properties of the reflecting surfaces; room size (x, y, z); microphone positions (x, y, z); microphone orientation angles (α, ϑ); microphone directivity patterns ($0 < a < 1$, as in Equation 3); and switching to classic amplitude panning laws (on/off).

Visualization Tool

A visualization tool was programmed in PD using the Graphics Environment for Multimedia (GEM; see gem.iem.at) extension to have visual control and feedback over the ViMiC environment. This tool, depicted in Figure 12, helps the user monitor the spatial scenery by displaying all source positions

Figure 13. ViMiC control as a VST plug-in. ViMiC control plug-ins are also available for the RTAS and Audio Units standards.



(numbered boxes) and microphone positions (unnumbered boxes) relative to the floor of the virtual room. All positions are updated in real time, and the user can switch between three different viewing angles.

Control Plug-In

A plug-in has been designed to control the ViMiC unit from most commercially available digital audio workstations (DAWs). The initial goal for designing the plug-in was to store ViMiC automation data together with pre-recorded audio material to create complex and dynamic soundscapes. The ViMiC plug-in is based on the Pluggo Runtime Environment for Max/MSP and runs on VST, RTAS, or Audio Units host applications (see Figure 13). A separate plug-in unit can be loaded for each audio track to be spatialized with ViMiC. DAW software automation can be used to control the values of all ViMiC parameters. The ViMiC control plug-in communicates with the dedicated auditory rendering system through OSC control messages

over a UDP network. The pre-recorded audio tracks can be streamed from the DAW to the ViMiC unit through a digital multichannel audio connection.

Live Control Interface

A second, standalone ViMiC graphical user interface (GUI) was created for live performances to control the ViMiC system with various input devices in real time. A number of presets can be stored, adapted, and recalled in the standalone interface. The parameter script is a text file that can be easily edited in an external text editor. In the text file, ramp times can be added and edited for every parameter to allow smooth transitions from one setting to another. The standalone GUI looks very similar to the VST plug-in shown in Figure 13.

Current ViMiC Projects

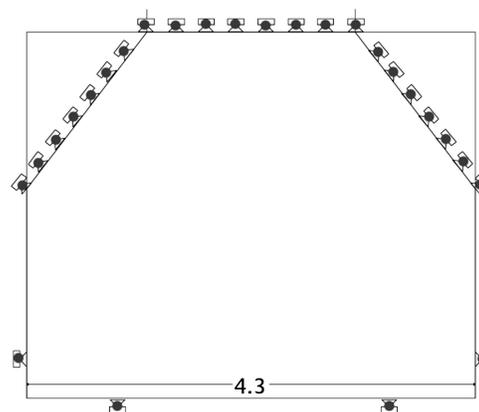
The ViMiC environment was originally designed for the 24-channel sound reproduction system of

the Immersive Presence Lab at McGill University. The system, which has been designed within the Valorisation-recherche Québec (VRQ) project “Real-Time Communication of High-Resolution Multi-Sensory Content via Broadband Networks,” is an augmented standard surround system with three vertically aligned rings of custom-built ribbon loudspeakers and six subwoofers (D-Box, Mini Mammoth; Woszczyk et al. 2005). The use of virtual microphones is ideal for arrays with speakers at different elevation angles, because the time-delay based panning possibilities help the user position sounds vertically. It is commonly known that summing localization, upon which amplitude panning techniques are based, does not work as well in the vertical compared to the horizontal (frontal) plane. Using the precedence effect through ICTDs offers new possibilities for improved elevation panning.

In the VRQ project, the ViMiC system was part of a newly built multimedia reproduction system for high-resolution, two-way audio-visual transmissions over the Internet. For this environment, the room acoustics were recreated from near-field microphone signals using ViMiC. This way, the use of room microphones—which show a strong tendency toward feedback problems in two-way transmissions—could be avoided.

By exploiting the advantages of WFS—which include a large listening area and the precise control over the resulting soundfield—and the flexibility of ViMiC, a hybrid system has been created to improve the spatial-image resolution of the virtual stage area (Valente and Braasch 2006). In this study, the three frontal speakers of a standard surround sound set-up were replaced with a linear WFS array as shown in Figure 14. The top 22 loudspeakers depict the linear WFS array to process the three frontal channels of a five-channel surround system. The two loudspeakers at the bottom left of the figure both project the unprocessed signal for the left surround channel, and the two loudspeakers shown at the right bottom both play back the unprocessed signal for the right surround channel. By maintaining the use of conventional surround speakers, both the space requirements and financial costs are within the limits of a commercially viable home-entertainment system. In addition to spatializing single anechoic

Figure 14. Top view sketch for the ViMiC/WFS hybrid system. The gray circles denote the virtual microphone positions. The width of the arrangement is 4.3 m, providing a large listening or performing area inside the array of loudspeakers.



sources in a virtual room, the system can be used to spatially upsample pre-existing recorded material by using virtual sound sources. In this application, the user can position, for example, all five channels of a surround sound mix. This mix can then be spatialized with the ViMiC system, and, by using WFS techniques, the “sweet spot” of a standard surround playback system can be greatly increased. This technique can be used with any pre-mixed or pre-recorded material.

The ViMiC system was also demonstrated in a commissioned piece by Sean Ferguson at the 2008 MusiMarch Festival in Montreal. In this piece, the ViMiC system was employed in connection with sensor-based human interfaces devices (HID), which were developed by Marshall et al. (2006). A gesture description interchange format (GDIF) has been designed to standardize the way gesture-related information is stored and shared in a networked computer setup.

The ViMiC system has also been used in conjunction with a 360° cylindrical digital video projection screen designed by Jeffrey Shaw at the University of New South Wales for the Wooster Group Installation *There is still time . . . brother*. The work, which was presented at the Zentrum für Kunst und Medientechnologie (ZKM) in Karlsruhe, Germany, has been commissioned by Rensselaer Polytechnic Institute’s Experimental Media and Performing Arts Center (EMPAC).

In another project at Rensselaer Polytechnic Institute, the ViMiC system is being used as a

spatialization tool for telepresence music improvisations between RPI (Pauline Oliveros' Tintinnabulate Ensemble) and CCRMA/Stanford University (Chris Chafe's Soundwire Ensemble). In this project, the ViMiC system is integrated with Pauline Oliveros' Expanded Instrument System (Gamper and Oliveros 1998), CCRMA's low-latency JackTrip software (Chafe 2003; Caceres n.d.) and Jeremy Cooperstock's Ultra-Video Conferencing System (Cooperstock, Roston, and Woszczyk 2004). Two telepresence concerts demonstrating this configuration were conducted during the 2007 International Conference for Auditory Display (ICAD 2007) in Montreal and the 34th International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH 2007) in San Diego. A two-channel auralization using the ViMiC technology can be heard in Braasch (2006).

Currently, a number of extensions for the ViMiC system are being planned. To improve the naturalness of the reverberant field, diffuse filters for the early reflections and the modulation of early reflections and reverberation will be considered. In addition, psychophysical auditory visual results will be used to improve the ViMiC system's accuracy in applications that include a video component. A future version will also include the implementation of frequency-dependent directivity patterns.

References

- Allen, J. B., and D. A. Berkley. 1979. "Image Method for Efficiently Simulating Small-Room Acoustics." *Journal of the Acoustical Society of America* 65(4):943–950.
- Berkhout, A. J. 1988. "A Holographic Approach to Acoustic Control." *Journal of the Audio Engineering Society* 36(12):977–995.
- Berkhout, A. J., P. Vogel, and D. de Vries. 1993. "Acoustic Control by Wave Field Synthesis." *Journal of the Acoustical Society of America* 93(5):2764–2779.
- Blumlein, A. D. 1931. "Improvements in and Relating to Sound-Transmission, Sound-Recording and Sound-Reproducing Systems." British Patent 394,325.
- Braasch, J. 2005. "A Binaural Model to Predict Position and Extension of Spatial Images Created with Standard Sound Recording Techniques." *119th Convention of the Audio Engineering Society*. New York: Audio Engineering Society, Preprint 6610.
- Braasch, J. 2006. *Global Reflections*. Audio compact disc. Kingston, New York: Deep Listening DL 34-2006.
- Caceres, J.-C. n.d. "JackTrip – Multimachine Jam Sessions Over the Internet2." Stanford: CCRMA, Stanford University. Available online at ccrma.stanford.edu/groups/soundwire/software.
- Chafe, C. 2003. "Distributed Internet Reverberation for Audio Collaboration." *Proceeding of the 24th International Conference of the Audio Engineering Society*. New York: Audio Engineering Society. Available online at www.ccrma.stanford.edu/~cc/soundwire/dirac.pdf.
- Chowning, J. M. 1971. "The Simulation of Moving Sound Sources." *Journal of the Audio Engineering Society* 19(1):2–6.
- Chowning, J. M. 1977. "The Simulation of Moving Sound Sources." *Computer Music Journal* 1(3):48–52.
- Cooperstock, J. R., J. Roston, and W. Woszczyk. 2004. "Broadband Networked Audio: Entering the Era of Multisensory Data Distribution." Paper presented at the 18th International Congress on Acoustics, Paris, 7 April.
- Corey, J., et al. 2001. "Enhancements of Room Simulation with Dynamic Cues Related to Source Position." *Proceedings of the 19th International Conference of the Audio Engineering Society*. New York: Audio Engineering Society. Available online at www.tonmeister.ca/research/pubs/corey01b.pdf.
- Daniel, J. 2000. "Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimedia." PhD thesis, Université Paris 6.
- Daniel, J., R. Nicol, and S. Moreau. 2003. "Further Investigations of High-Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging." *Proceedings of the 114th Convention of the Audio Engineering Society*. New York: Audio Engineering Society, Preprint 5788.
- Gamper, D., and P. Oliveros. 1998. "A Performer-Controlled Live Sound-Processing System: New Developments and Implementations of the Expanded Instrument System." *Leonardo Music Journal* 8(1):33–38.
- Gerzon, M. A. 1973. "Periphony: With-Height Sound Reproduction." *Journal of the Audio Engineering Society* 21(1):2–10.
- Gerzon, M. A. 1985. "Ambisonics in Multichannel Broadcasting and Video." *Journal of the Audio Engineering Society* 33(11):859–871.
- Horbach, U., et al. 2000. "Implementation of an Auralization Scheme in a Digital Mixing Console Using Perceptual Parameters." *Proceedings of the 108th Convention of the Audio Engineering Society*. New York: Audio Engineering Society, Preprint 5099.

-
- ITU. 1994. "Multichannel Stereophonic Sound System with and without Accompanying Picture." Standard BS.775-1, International Telecommunication Union.
- Jot, J.-M. 1992. "Étude et réalisation d'un spatialisateur de sons par modèles physiques et perceptifs." PhD thesis, Télécom Paris.
- Laborie, A., et al. 2003. "A New Comprehensive Approach of Surround Sound Recording." *Proceedings of the 114th Convention of the Audio Engineering Society*. New York: Audio Engineering Society, Preprint 5717.
- Malham, D. G. 1999. "Higher Order Ambisonic Systems for the Spatialisation of Sound." *Proceedings of the 1999 International Computer Music Conference*. San Francisco, California: International Computer Music Association, pp. 484–487.
- Marshall, M. T., et al. 2006. "On the Development of a System for Gesture Control of Spatialization." *Proceedings of the 2006 International Computer Music Conference*. San Francisco, California: International Computer Music Association, pp. 360–366.
- Martin, G., et al. 2001. "Controlling Phantom Image Focus in a Multichannel Reproduction System." *Proceedings of the 107th Convention of the Audio Engineering Society*. New York: Audio Engineering Society, Preprint 4996.
- Moore, F. R. 1983. "A General Model for Spatial Processing of Sounds." *Computer Music Journal* 7(3):6–15.
- Mouchtaris, A., S. Narayanan, and C. Kyriakakis C. 2003. "Virtual Microphones for Multichannel Audio Resynthesis." *EURASIP Journal on Applied Signal Processing* 2003(1):968–979.
- Nielsen, S. H. 2001. "Methods of Processing a Signal." International patent WO0188901.
- Pulkki, V. 2001. "Localization of Amplitude-Panned Virtual Sources, Part 2: Two- and Three-Dimensional Panning." *Journal of the Audio Engineering Society* 49(9):753–767.
- Pulkki, V., and M. Karjalainen. 2001. "Localization of Amplitude-Panned Virtual Sources, Part 1: Stereophonic Panning." *Journal of the Audio Engineering Society* 49(9):739–752.
- Schroeder, M. R. 1962. "Natural Sounding Artificial Reverberation." *Journal of the Audio Engineering Society* 10(3):219–223.
- Steinberg, J. C., and W. B. Snow. 1934. "Auditory Perspective—Physical Factors." *Electrical Engineering* January:12–17.
- Valente, D., and J. Braasch. 2006. "A Hybrid Approach to Rendering Multi-Channel Audio, Using Wave Field Synthesis and Stereophonic Techniques, with Virtual Microphone Control." *Journal of the Acoustical Society of America* 119(5):3281(A).
- Weiss, E. 1993. "Audiotechnologie in Berlin bis 1943: Mikrophone." In H. K. Thiele, ed. *50 Jahre Stereo-Magnetbandtechnik: Die Entwicklungen der Audio Technologie in Berlin und den USA von den Anfängen bis 1943*. Oberursel: Berlebach, pp. 39–54.
- Woszczyk, W., et al. 2005. "Shake, Rattle, and Roll: Getting Immersed in Multisensory, Interactive Music via Broadband Networks." *Journal of the Audio Engineering Society* 53(4):336–344.
- Wright, M., A. Freed, and A. Momeni. 2003. "OpenSound Control: State of the Art 2003." *Proceedings of the 2003 Conference on New Interfaces for Musical Expression*. New York: Association for Computing Machinery, pp. 154–159.