A SPATIAL AUDITORY DISPLAY FOR TELEMATIC MUSIC PERFORMANCES

J. BRAASCH

School of Architecture, Rensselaer Polytechnic Institute, Troy, NY 12180, USA, www.rpi.edu, E-mail: braasj@rpi.edu

N. PETERS

CIRMMT, Schulich School of Music, McGill University, Montreal, Quebec H3A 1E3, Canada

P. OLIVEROS

Arts Department, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

D. VAN NORT

School of Architecture & Arts Department, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

C. CHAFE

CCRMA, Department of Music, Stanford University, Stanford, CA 94305, USA

This paper describes a system which is used to project musicians from two or more co-located venues into a shared virtual acoustic space. The sound of the musicians is captured using near-field microphones and a microphone array to localize the sounds. Afterwards, the near-field microphone signals are projected at the remote ends using spatialization software based on Virtual Microphone Control (ViMiC) and an array of loudspeakers. In order to simulate the same virtual room at all co-located sites, the ViMiC systems communicate using the OpenSound Control (OSC) protocol to exchange room parameters and the room coordinates of the musicians. Using OSC they also receive localization data from the microphone arrays.

Keywords: Virtual Auditory Display, Telepresence, Telematic Music.

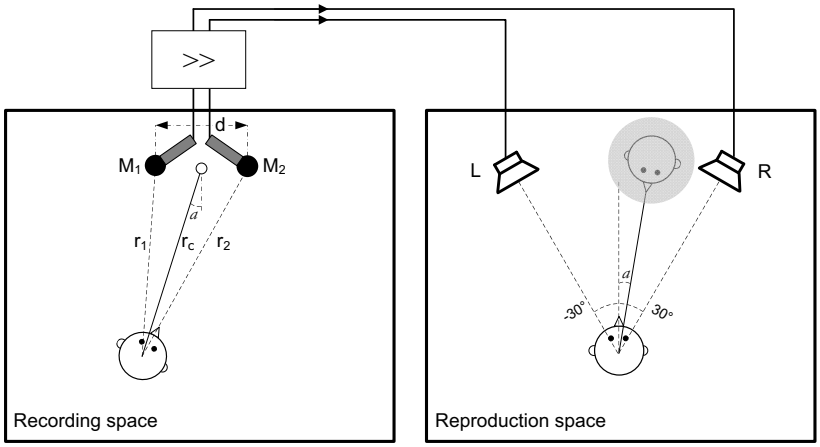

Fig. 1. Sketch of a microphone-based recording and reproduction set-up.

1. Introduction

Live networked music performances have gained in popularity over the last few years. In these concerts, musicians are distributed over at least two remote venues and connected via the internet. Some of the challenging technical requirements that these projects have imposed on the underlying research have been addressed in previous work.[1–6] One of the problems that has not been solved to the full extent is the accurate spatial reproduction of the broadcasted sound field at the remote end. Especially with the introduction of High-Definition (HD) video, the need for accurate spatial sound reproduction has become more eminent. In this chapter, a system for accurate spatial sound reproduction in telematic music performances is described.

The idea of this research goes back to the 1930s, when Steinberg and Snow described a system that enabled the world-renowned conductor Leopold Stokovski and the Philadelphia Orchestra to broadcast music live from Philadelphia to Washington, D.C. The authors used the then newly invented main-microphone techniques to produce a stereophonic image from the recorded sound. Figure 1 shows a general diagram of how microphones and loudspeakers have to be set-up and the signals have to be routed for stereophonic imagery. The spatial positions of sound sources in the recording space are encoded by placing and orienting two or more

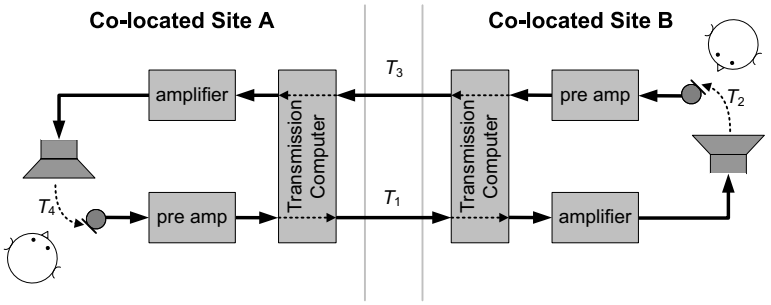

Fig. 2. Feedback loop in a telematic transmission.

microphones—the main microphone array—strategically, capturing spatial information by utilizing time and level differences between the different microphone channels. Each channel is then transmitted separately, amplified and fed to the matching loudspeaker of an array of at least two speakers, for example the classic stereo set-up shown in Fig. 1. The box in this figure that connects the microphones to the loudspeakers can either be an amplifier, a broadcasting unit or a sound-recording/reproduction system. Steinberg and Snow used two to three parallel telephone lines to transmit the spatially encoded sound from Philadelphia to Washington, D.C.[7]

While we now experience music broadcast via radio, satellite, and the internet in our daily life, music collaborations in which ensemble members are distributed over long distances are still in the experimental stage, because of technical difficulties that two-way or multicast connections bring along. A major challenge is the susceptibility of bidirectional set-ups to feedback loops. The latter can easily lead to audible colorations and echoes. Figure 2 demonstrates the general problem: the microphone signal recorded at Site A is broadcasted through a loudspeaker at the other Site B, where it is being picked up by a second microphone. This microphone signal is then broadcasted back to the original Site A, where it is re-captured by the first microphone. Due to the transmission latency, the feedback becomes audible as echo at much lower gains compared to the feedback situation known from local public address systems. Many audio/videoconferencing systems such as *iChat* or *Skype* use echo-cancellation systems to suppress feedbacks. In speech communication echo-cancellation systems work well, since the back-and-forth nature of spoken dialogue usually allows to suppress the transmission channel temporarily in one direction. In simultaneous music communication, however, this procedure tends to cut-off part of the perfor-

mance. Spectral alterations are a common side effect if the echo-cancellation system operates with a filter bank.

For the given reasons, the authors suggest to avoid using echo-cancellation systems completely. Instead, it is proposed to capture all instruments from a close distance (e.g., lavalier microphones) to minimize the gain and therefore the risk of feedback loops. Unfortunately, the exclusive use of near-field microphones contradicts the original idea of Steinberg and Snow, since the main microphones have to be placed at a further distance to capture the sound field stereophonically. To resolve this conflict, this paper describes an alternative approach to simulate main microphone signals from closely captured microphone signals and geometric data. The system—called Virtual Microphone Control (ViMiC)—includes a room simulation software to (re-)construct a multichannel audio signal from a dry recording as if it had been recorded in a particular room.[8–10] The position data of the sound sources, which is needed to compute the main microphone signals, are estimated using a microphone array. The array, which is optimized to locate multiple sound sources, is installed at each co-located venue to track the positions of the sound sources. The recorded position data is transmitted to the remote venue(s) along with the acoustic signals that were recorded in the near-field of the instruments. At the remote end, the sound can then be projected spatially correct using the ViMiC system. A sketch of the transmission system, which also includes video broadcast, is shown in Fig. 3. The low-latency audio transmission software *Jacktrip*[11,12] and the *Ultravideo Conferencing*[4] system are used for telecommunication.

2. Sound Spatialization using Virtual Microphone Control

2.1. *Basic Concept*

The following section deals with the fundamental principles of ViMiC. The system basically simulates a multichannel main microphone signal from the near-field recordings, descriptors about the room size and wall-absorption coefficients, and the sound-source positioning data. The latter is provided by a sound-localization microphone array as described in Section 3. To auralize the signals, each virtual microphone signal is then fed to a separate (real) loudspeaker.

The core concept of ViMiC involves an array of virtual microphones with simulated directivity patterns. The axial orientation of these patterns can be freely adjusted in 3D space, and the directivity patterns can be varied between the classic patterns that are found in real microphones: omnidirec-

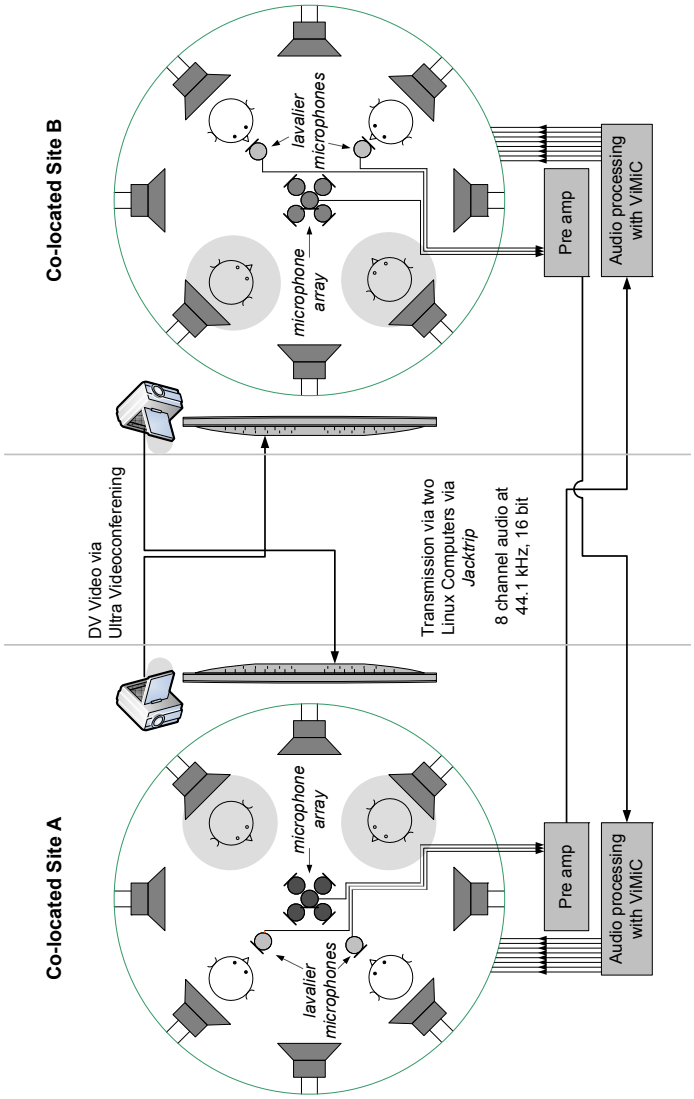

Fig. 3. Sketch of the internet-based telematic music system used by the authors.

tional, cardioid, hyper-cardioid, sub-cardioid, or figure-eight characteristics.

The transfer function between the sound source (e.g., a musical instrument which can be treated as a one-dimensional signal in time $x(t)$) and

a virtual microphone is then determined by the distance and the orientation between the microphone's directivity pattern and the recorded sound source (e.g., musical instrument). The distance determines the delay τ between the radiated sound at the sound-source position and the microphone signal:

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

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

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

with the sound pressure p_0 of the sound source at a reference distance r_0. In addition, the system considers that the sensitivity of a microphone varies with the angle of incidence according to its directivity pattern. In theory, only omni-directional microphones are equally sensitive towards all directions, and in practice even this type of microphone is more sensitive toward the front for high frequencies. The circumstance that real microphones generally have rotational directivity patterns simplifies their implementation in ViMiC, since these types of directivity patterns $\Gamma(\alpha)$ can be written in a simple general form:

$$\Gamma(\alpha) = a + b\cos(\alpha). \tag{3}$$

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, with omnidirectional: a=1, b=0; cardioid: a=0.5, b=0.5, and bi-directional: a=0, b=1.

The overall gain g between the sound source and the virtual microphone can be determined as follows:

$$g = g_d \cdot \Gamma(\alpha) \cdot \Gamma(\beta), \tag{4}$$

with the distance-dependent gain $g_d = r_0/r$, and the sound source's rotational radiation pattern $\Gamma(\beta)$.

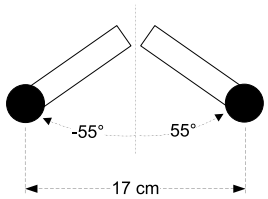

Fig. 4. Microphone placement for the ORTF technique with two cardioid microphones.

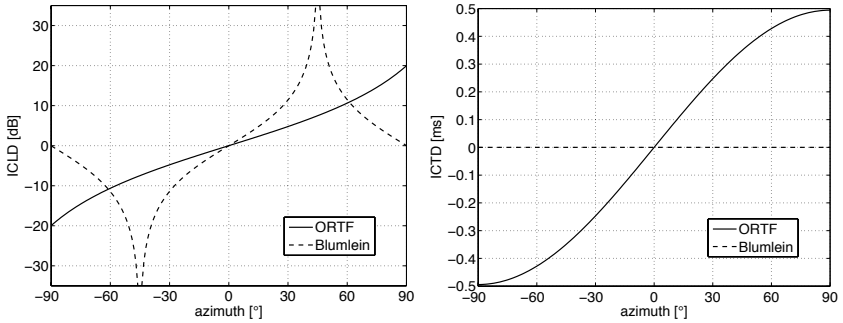

Fig. 5. The left graph shows calculated inter-channel level differences (ICLDs) for the ORTF technique as a function of azimuth in comparison to the Blumlein/XY technique. The Blumlein/XY technique is based on two bi-directional microphones with coincident placement at an angle of 90°. The right graph shows the results for inter-channel time differences (ICTDs).

The transfer function between the sound source and the microphone signal can now be derived from two parameters only, the gain g and the delay τ, if the microphone and source directivity patterns are considered to be independent of frequency. It is noteworthy, that the directivity patters of most real microphones are not fully independent of frequency, although this is often a design goal. The relationship between the sound radiated from a point source $x(t)$ and the microphone signal $y(t)$ is found to be:

$$y(t, r, \alpha) = g \cdot x(t - \tau) = g_d(r) \cdot \Gamma(\alpha) \cdot \Gamma(\beta) \cdot x(t - \frac{r}{c_s}). \tag{5}$$

By simulating several of the virtual microphones as outlined above, the sound sources can be panned in virtual space according to standard sound recording practices.

2.2. *ORTF-technique implementation*

A good example to demonstrate ViMiC is the classic ORTF microphone technique, which is named after the French national broadcasting agency *Office de Radiodiffusion et de Télévision Française* where it was first introduced. The ORTF microphone placement is shown in Fig. 4. Due to the relative broad width of the directivity lobe of the cardioid pattern, the angle between both microphones is adjusted 110° wide. The ratio between the signal amplitude at the sound source x and microphone signal amplitudes for the left and right channels, y_1 and y_2, vary with the angle of incidence according to Eq. 5:

$$y_1(t) = g_{d_1} \cdot 0.5 \cdot (1 + \cos(\alpha + 55°)) \cdot x(t - \tau), \tag{6}$$

$$y_2(t) = g_{d_2} \cdot 0.5 \cdot (1 + \cos(\alpha - 55°)) \cdot x(t - \tau). \tag{7}$$

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 Fig. 1). When a virtual sound source is encircling the microphone set-up in the frontal horizontal plane at a distance of 3 m (α=−90° to 90°), the inter-channel level difference (ICLD) ρ as shown in Fig. 5 can be calculated as follows:

$$\rho(\alpha) = 20 \cdot \log_{10}\left(\frac{y_2(t)}{y_1(t)}\right) = 20 \cdot \log_{10}\left(\frac{g_{d_2} \cdot (1 + \cos(\alpha - 55°)}{g_{d_1} \cdot (1 + \cos(\alpha + 55°)}\right). \tag{8}$$

In the far-field—when the distance between the sound source and the center of the recording set-up r is much larger than the distance in between both microphone diaphragms d ($r \gg d$)—the $1/r$ term can be neglected in the ICLD calculation, and the occurring ICLDs are nearly solely generated by the different orientations of the cardioid patterns of both microphones. Figure 5 shows the ICLDs 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 coincidence techniques like the Blumlein/XY technique. However, increasing the angle between both microphones is rather problematic, as this would result in a very high sensitivity of the set-up toward the sides. Instead, the diaphragms of both microphones are spaced 17 cm apart in the ORTF configuration (compare Fig. 4. This way ICTDs τ_Δ are generated in addition to the ICLDs. The ICTDs, which are also shown in Fig. 5, can be easily determined from the geometry of the set-up (compare Fig. 1):

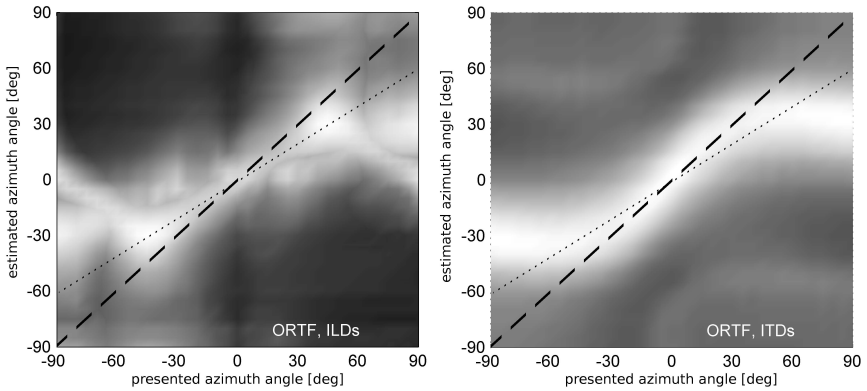

Fig. 6. Results for a binaural model to localize an ORTF en- and decoded sound source from various positions in the horizontal plane. The left graph shows the results of the ILD analysis, the one in the right the ITD-analysis results.

$$\tau_\Delta(\alpha) = \frac{(r_1 - r_2)}{c_s}, \tag{9}$$

with the speed of sound c_s and the far-field approximation:

$$\tau_\Delta(\alpha) = \frac{d}{c_s}\sin(\alpha). \tag{10}$$

One of the core ideas of ViMiC is to be able to play with the spatial imagery of sound in a similar fashion to microphone-based sound recording practice. To illustrate this approach, the output of a binaural model is shown in Fig. 6. For this graph, the model was used to analyze the reproduced sound field of an ORTF recording via a dummy head. The figures show the estimated localization curves, the relationship between the azimuth of the original source position and the azimuth of the auditory event when listening to the reproduced signal. The left graph shows the analysis of interaural level differences (ILDs), and the estimated position of the auditory events are highlighted in white or light gray, the right graphs shows the same context but for interaural time differences. The figure shows, that within the range of interest ($-45°$ to $+45°$), the ILD cues project the sound source at a narrower angle compared to the natural condition, while the ITD cues suggest a wider angle. The mismatch between both cues leads to the perceptual widening of the auditory objects, which is often preferred and

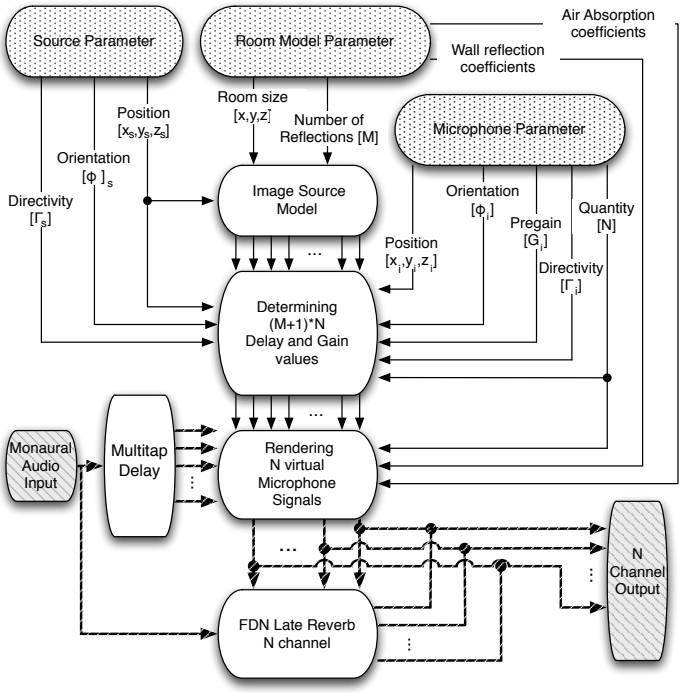

Fig. 7. Architecture of the auditory virtual environment based on Virtual Microphone Control (ViMiC).

makes the use classic microphone techniques so interesting. Further details about the model analysis can be found in Braasch (2005)[13] and Blauert and Braasch (2008).[14]

2.3. *Software Implementation*

Figure 7 shows the system architecture of the current ViMiC implementation, which is part of the Jamoma package.[15,16] The system has three larger signal processing components: an Image Source Model, a Multitap Delay Unit, and a Reverberation Unit. The Image Source Model determines the gains and delays between the sound source positions and the receiving virtual microphones. The algorithm considers the positions and orientations of both sources and receivers including their directivity characteristics. The model uses the mirror image technique[17] to calculate the positions and strengths of early room reflections for a rectangular enclosure

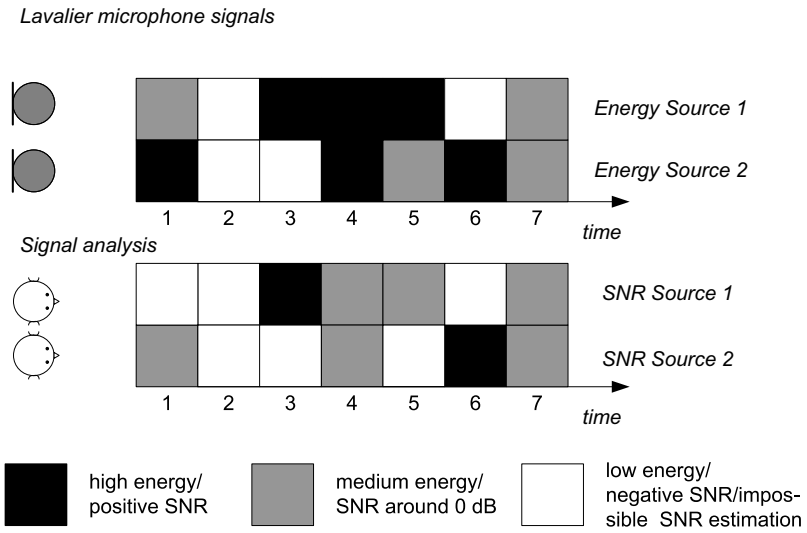

Fig. 8. Estimation of the signal-to-noise ratios for each sound source.

with adjustable dimension and wall-absorption characteristics.

Using the gain and delay data provided by the Image Source Model, the dry sound is processed using a multi-tap delay network for spatialization. Since typically a high number of delays have to be computed, for example, 42 delays have to be processed per primary sound source in a 6 channel surround system, if first-order reflections are considered (1 direct source plus 6 first-order reflections $\times$ 6 output channels). This number increases to 114 delays, if second-order reflections are simulated as well. Several measures have been taken to reduce the computational load. One of them is the automated shift between 4-point fractional delays for moving sound sources and non-fractional delays, which are activated once the sound source remains stationary. The late reverberant field is considered to be diffuse and simulated through a feedback delay network[18] with 16 modulated delay lines, which are diffused by a Hadamard mixing matrix. By feeding the outputs of the room model into the late reverb unit a diffused reverb tail is synthesized (see Fig. 7), for which timbral and temporal character can be modified. This late reverb can be efficiently shared across several rendered sound sources.

3. Sound Source Tracking System

So far, we have described the spatial decoding method using the ViMiC system, but it has not been discussed how the spatial positions can be captured at the remote site. For fixed instrument positions, as is often the case in classical music, a manual adjustment of the sound source positions is a viable option. However, this procedure can be cumbersome if the positions of the sound source vary over time.

The solution that was integrated into our telematic music system is based on a pyramidal five-microphone array, which has been described earlier.[10,19] The five omni-directional microphones are arranged in a square-based pyramid with 14-cm base side and 14-cm triangular side dimensions. Traditional microphone-array based systems work well to localize an isolated sound source by utilizing arrival time differences or amplitude differences of the sound source between the individual array microphones.[?,?] In multiple-sound-source scenarios (e.g., a music ensemble), however, determining the sound-source positions from the mixed signal and assigning them to the corresponding source is still a real challenge.

A solution for this problem is to use the near-field microphone signals in conjunction with a traditional microphone-array based localization system. The near-field microphone signals are then used to determine the signal-to-noise ratios (SNRs) between several sound sources, for example concurrent musicians, while still serving the main purpose of capturing the audio signals. The running SNR is calculated frequency-wise from the acoustic energy recorded in a certain time interval:

$$\mathrm{SNR}_{i,m} = 10 \log_{10} \left(\frac{1}{a} \int_{t_m}^{t_m+\Delta t} p_i^2 \cdot dt \right) \tag{11}$$

with:

$$a = \sum_{n=1}^{i-1} \int_{t_m}^{t_m+\Delta t} p_i^2 \cdot dt + \sum_{n=i+1}^{N} \int_{t_m}^{t_m+\Delta t} p_i^2 \cdot dt \tag{12}$$

and p_i the sound pressure captured with the i^{th} near-field microphone, t_m the beginning of the measured time interval m, Δt its duration and N, the number of near-field microphones.

Basically, the SNRs are measured for each time interval between each observed sound source and the remaining sound sources. The data can then be used to select and weight those time slots in which the sound source

dominates the scene, assuming that in this case the SNR is high enough for the microphone array to provide stable localization cues. Figure 8 depicts the core idea. In this example, a good time slot is found for the third time frame for Sound Source 1, which has a large amount of energy in this frame, because the recorded energy for Sound Source 2 is very low. Time Slot 6 depicts an example where a high SNR is found for the second sound source.

To improve the quality of the algorithm, all data are analyzed frequency-wise. For this purpose the signals are sent through an octave-band filter bank before the SNR is determined. Basically, the SNR is now a function of frequency f, time interval t, and the index of the sound source i: SNR=SNR(f,t,i).

The sound source position is determined for each time/frequency slot by analyzing the time delays between the microphone signals of the microphone array. The position of the sound source is estimated using the cross-correlation technique, which is used to determine the direction of arrival (DOA) from the measured internal delay (peak position of the maximum of the cross-correlation function) via this equation as shown by Würfel among others:[20]

$$\alpha = \arcsin\left(\frac{c}{f_s} \cdot \frac{\tau}{d}\right), \tag{13}$$

with the speed of sound c, the sampling frequency f_s, the internal delay τ, and the distance between both microphones d.

Since this technique cannot resolve two sound sources within one time-frequency bin, the estimated position is assigned to the sound source with the highest SNR. Alternatively, the information in each band can be weighted with the SNR in this band. To save computational cost, a minimum SNR threshold can be determined, below which the localization algorithm will not be activated for the corresponding time/frequency slot.

4. Integrated system

Figure 9 depicts the whole transmission chain which includes the sonification system. At the recording site, the raw sound signals are captured through the near-field microphones which also feed the localization algorithm with information to calculate the instantaneous SNR. Both the audio data and the control data—which contains information on the estimated sound source position—is transmitted live to the co-located site(s). Here, the sound field is resynthesized from the near-field audio signals and the control data using rendering techniques such as ViMiC.

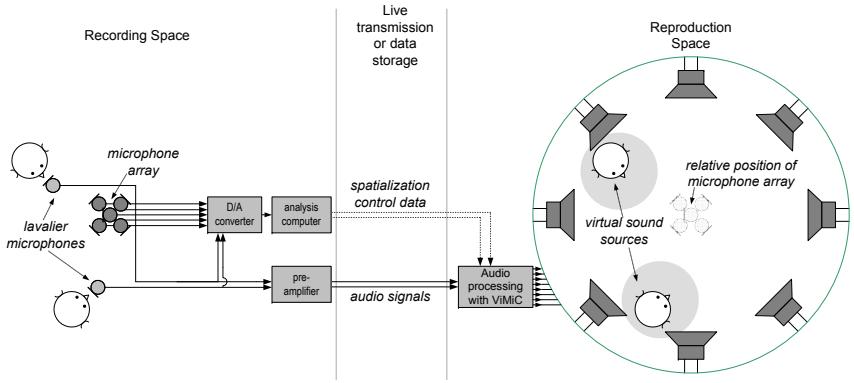

Fig. 9. Sketch of the spatial sound recording and reproduction set-up.

The sound source tracking unit is currently implemented in Matlab, which allows easier prototyping than an implementation in Max/MSP. The Matlab module runs in real-time using the *Data Acquisition Toolbox*. The module receives multichannel audio input and returns the calculated results (positions of individual sound sources) via OSC. Currently, we are also experimenting with an ambisonics-based microphone array (1st-order, B-Format) for sound localization.[21,22] Since, the spatial positions can be derived from amplitude differences, it requires less computational resources than the current pyramidal array, which localizes sounds through time delay analysis. The expected decrease in localization accuracy is acceptable for the given application and the described algorithm to analyze multiple sound sources can be applied equally well.

The ViMiC system has been used in several projects to spatialize telematically transmitted sound. The first commercial album using ViMiC in a telepresence scenario has been released with the Deep Listening Record Label in Kingston, New York.[23] The 5-channel Quicktime video is a recording of the Tintinnabulate and Soundwire ensembles performing live at the ICAD 2007 conference in Montreal, Canada (McGill University), RPI, Stanford University and KAIST, Seoul, South Korea. For a telematic concert at SIGGRAPH 2007, *Dynamic Spaces*,[24] we used ViMiC to create a dynamically changing acoustical space. In this piece, the room acoustics was altered in realtime using a joystick controller. The system was used to vary the acoustics in San Diego during a remote clarinet solo that was played by Bobby Gibbs at Rensselaer Polytechnic Institute. Reverberation time,

room size, sound pressure level of early reflections, and frequency response were among the parameters that were controlled. The project was a milestone in our current paradigm to explore the possibility of changing the acoustics of the concert space during the performance. This new possibility adds substantially to the way we perform and listen to music—creating a new awareness for the space surrounding us.

The project reported here has received support from the National Science Foundation (#0757454), the Canadian Natural Sciences and Engineering Research Council (NSERC, New Media Initiative), and a seed grant from Rensselaer Polytechnic Institute and the Experimental Media and Performing Arts Centers (EMPAC). We would also like to thank Johannes Goebel and Todd Vos from EMPAC for their support.

References

1. P. Oliveros, J. Watanabe and B. Lonsway, *A collaborative Internet2 performance*, tech. rep., Offering Research In Music and Art, Orima Inc. Oakland, CA (2003).
2. E. Chew, A. Sawchuk, R. Zimmerman, V. Stoyanova, I. Tosheff, C. C. Kyriakakis, C. Papadopoulos, A. Franois and A. Volk, Distributed immersive performance, in *Proceedings of the 2004 Annual National Association of the Schools of Music (NASM) Meeting*, (San Diego, CA, 2004).
3. R. Rowe and N. Rolnick, The technophobe and the madman: an internet2 distributed musical, in *Proc. of the Int. Computer Music Conf. Miami*, (Florida, 2004).
4. J. Cooperstock, J. Roston and W. Woszczyk, Broadband networked audio: Entering the era of multisensory data distribution, in *18th International Congress on Acoustics*, (Kyoto, 2004).
5. F. Schroeder, A. Renaud, P. Rebelo and F. Gualdas, Addressing the network: Performative strategies for playing apart, in *Proc. of the 2007 International Computer Music Conference (ICMC 07)*, (Copenhagen, Denmark, 2007).
6. P. Oliveros, S. Weaver, M. Dresser, J. Pitcher, J. Braasch and C. Chafe, *Leonardo Music Journal* **19**, 95 (2009).
7. J. C. Steinberg and W. B. Snow, *Electrical Engineering* , 12(Jan 1934).
8. J. Braasch, A loudspeaker-based 3D sound projection using virtual microphone control (ViMiC), in *Proc. of the 118th Convention of the Audio Eng. Soc.*, (Barcelona, Spain, 2005). Paper Number 6430.
9. J. Braasch, T. Ryan and W. Woszczyk, An immersive audio environment with source positioning based on virtual microphone control (ViMiC), in *Proc. of the 119th Convention of the Audio Eng. Soc.*, (New York, NY, 2005). Paper Number 6546.
10. J. Braasch, N. Peters and D. Valente, *Computer Music Journal* **32**, 55 (2008).
11. J. Cáceres, R. Hamilton, D. Iyer, C. Chafe and G. Wang, To the edge with China: Explorations in network performance, in *ARTECH 2008: Proceedings*

of the 4th International Conference on Digital Arts, (Porto, Portugal, 2008).
12. J. Cáceres and C. Chafe, JackTrip: Under the hood of an engine for network audio, in *Proceedings of International Computer Music Conference*, (Montreal, QC, Canada, 2009).
13. J. Braasch, A binaural model to predict position and extension of spatial images created with standard sound recording techniques, in *Proc. of the Convention of the Audio Eng. Soc.*, (New York, NY, 2005). Paper Number 6610.
14. J. Blauert and J. Braasch, Räumliches Hören [Spatial hearing], in *Applications of digital signal processing to audio and acoustics*, ed. S. Weinzierl (Springer Verlag, Berlin-Heidelberg-New York, 1998) pp. 75–108.
15. T. Place and T. Lossius, Jamoma: A modular standard for structuring patches in Max, in *Proc. of the 2006 International Computer Music Conference (ICMC 06)*, (New Orleans, LA, 2006).
16. N. Peters, T. Matthews, J. Braasch and S. McAdams, ViMiC – A novel toolbox for spatial sound processing in Max/MSP, in *Proceedings of International Computer Music Conference*, (Belfast, Northern Ireland, 2008).
17. J. B. Allen and D. A. Berkley, *J. Acoust. Soc. Am.* **65**, 943 (1979).
18. J. Jot and A. Chaigne, Digital delay networks for designing artificial reverberators, in *Proc. of the Convention of the Audio Eng. Soc.*, (Paris, France, 1991). Paper Number 6610.
19. J. Braasch, D. Valente and N. Peters, An immersive audio environment with source positioning based on virtual microphone control (ViMiC), in *Proc. of the 123rd Convention of the Audio Eng. Soc.*, (New York, NY, 2007). Paper Number 7209.
20. A. Quazi, *IEEE Transactions on Acoustics, Speech and Signal Processing* **29**, 527(June 1981).
21. R. Hickling, W. Wei and R. Raspet, *JASA* **94**, 2408(Oct 1993).
22. W. Würfel, Passive akustische lokalisation [passive acoustical localization], Master's thesis, Technical University Graz (1997).
23. V. Pulkki, J. Merimaa and T. Lokki, Reproduction of reverberation with spatial impulse response rendering, in *Proc. of the Convention of the Audio Eng. Soc.*, (Berlin, Germany, 2004). Paper Number 6057.
24. V. Pulkki and C. Faller, Directional audio coding: Filterbank and STFT-based design, in *Preprint 120th Conv. Aud. Eng. Soc.*, (Paris, France, 2006). Paper Number 6658.
25. Tintinnabulate & Soundwire, J. Braasch, C. Chafe, P. Oliveros and B. Woodstrup, *Tele-Colonization* (Deep Listening Institute, Ltd., DL-TMS/DD-1, 2009).
26. P. Oliveros, C. Bahn, J. Braasch, C. Chafe, T. Hahn, Soundwire Ensemble, Tintinnabulate Ensemble, D. Valente and B. Woodstrup, Dynamic spaces(August 2007), SIGGRAPH 2007.