A VIRTUAL ACOUSTIC ENVIRONMENT AS AUDITORY DISPLAY FRONT-END FOR SONIFICATION

César D. Salvador

ISONAR, Universidad de San Martín de Porres
Av. Tomás Marsano 151 Surquillo, Lima 34, Perú
csalvador@comunicaciones.usmp.edu.pe

ABSTRACT

This paper describes work-in-progress on an immersive auditory display front-end that aims to collaborate with the development of a theoretical framework for spatial sonification in a soundscape context and an interactive sonification toolkit. Spatial sonification requirements are reviewed and related to the possibilities of the wave field synthesis spatial sound reproduction technique. The proposed real-time multichannel rendering system looks for immerse the user in a virtual acoustic environment, providing the capability of sound focusing and multiple channels of auditory information. A sonification project for signaling in sound art exhibition galleries is also introduced.

1. INTRODUCTION

Virtual acoustic environments look for immerse a listener in an almost real acoustic environment, synthesizing wave fronts with physical methods and rendering them through loudspeaker arrays. On the other hand, auditory displays in a sonification framework look for insight into data under analysis, while rendering sound in an organized and well-structured way [1]. As virtual acoustic environments provide multiple channels of auditory information, they emerge as a more realistic method for spatial auditory display.

In order to collaborate with the development of: 1) a theoretical framework for spatial sonification in a soundscape context and 2) an interactive sonification toolkit, the development of a real-time multichannel rendering system is reported here as a first stage of this project. Here, spatial sonification is mainly treated from the listener point of view; while spatial sound is mainly treated from the sound source point of view; consequently, psychoacoustics and acoustics approaches are used respectively, which can be unified later in a soundscape context.

This paper is organized as follows. Section 2 briefly reviews some results on spatial sonification, pointing out the general requirements of sound perception for ongoing research. Section 3 briefly introduces the wave field synthesis method, an acoustics-based spatial sound technique which has been chosen for the virtual acoustic display. Section 4 reviews a soundscape theoretical framework, highlighting the effects of the environment in sound perception. Section 5 proposes to link the requirements of spatial sonification with the possibilities of the wave field synthesis technique as a spatial auditory display, in a soundscape context. Section 6 introduces an application arisen in sound art exhibitions that aims to signaling and give information about the quantity of people that is being visiting a gallery, without interfering with the concept of the piece of art. Section 7 discuses the possibilities and limitations of the proposed virtual acoustic display in a sonification framework. Finally, the conclusions are given in section 8.

2. SPATIAL SONIFICATION

This section points out the results reviewed by Nasir and Roberts on sonification of spatial and non-spatial data based on spatial and non-spatial perception sound [2]. Spatial data is defined as any dataset that contains a location component along with other dependent variables. On the other hand, spatial perception of sound is affected by: Interaural Time Difference (ITD), Interaural Intensity Difference (ITD), Doppler and time-based effects, and the environment where the sound is displayed. Therefore, spatial sound mappings based on these facts enable the user to locate the origin of the sound. However, it is important to point out that users are unable to accurately locate the position of a sound source as accurately as they could locate the information in an equivalent graphical visualization. Hence, results about localization on spatial sonification experiments conducted by various researchers agree that:

a) The accuracy of spatial sound perception depends on the radial location of the sound source. Error metrics such as the Minimum Audible Angle should be referenced to create appropriate mappings and effective evaluations.
b) The maximum potential of spatial sonification could be on echo location and other factors such as Doppler effect, reverberation and spatial occlusion.
c) More accurate models such as HRTF's should be used to create accurate positional mappings.
d) Spatial sound is certainly not the only way to sonify spatial data and, reciprocally, non-spatial variables could be spatially sonified to maximize the perception of the information.

It can be seen from the previous list that the listener point of view has been remarked here, and it is important to highlight the effect of the sound environment. Next section now treats the sound source point of view, introducing the sound spatialization method chosen for this project, i.e., the Wave Field Synthesis technique. Then section 4 will return to the environment effects.

3. SPATIAL SOUND

The techniques of spatial sound reproduction can be classified into those mainly based on psicoacoustics and those mainly based on acoustics. Among perceptual methods there exist those that vary the intensity of sound such as: Quadraphonic and the generalized panning VBAP; and also those that introduce further delays between audio signals arriving to speakers, such as binaural spatialization. Alternatively, most current methods appeal to correct physical modeling, among which there are higher order ambisonics and wave field synthesis. This last is described below.

3.1. Wave Field Synthesis

Wave field synthesis (WFS) is a sound reproduction technique whose theoretical framework was initially formulated by Berkhout et al., [3], [4]. WFS is actually emerging as an optimal format for spatialization of virtual auditory scenes that look for immerse a listener in an almost real acoustic environment, synthesizing wave fronts with physical methods and rendering them through loudspeaker arrays. WFS allows to synthesize virtual acoustical environments by rendering room impulse responses with plane wave models, as well as to synthesize virtual sources that appear to emanate from a defined position by rendering them with spherical wave models. Thus, it provides the listener with consistent spatial localization cues over large listening areas, but it utilizes a high number of loudspeakers [5].

In practical WFS applications, it is necessary to compute prefiltering, filtering, delaying and scaling operations on the audio signal to be spatialized before it drives each loudspeaker. The rendering of sound pressure fields is possible by using a set of loudspeakers uniformly distributed along contours, such as lines and circular arcs. The sound pressure field is computed by adding the effect of each loudspeaker, where the distance between two adjacent loudspeakers defines the maximum reproducible frequency [6].

The rendering of sources positioned behind the loudspeakers is possible with WFS, that is, the synthesis of plane and non-focused spherical wave fields. Furthermore, the rendering of sources positioned in between the loudspeakers and the listener, called focused sources, is also possible thanks to the time-reversal invariance of the wave equation: for each burst of sound diverging from a source, there exists a set of waves that retraces its paths and converges simultaneously at the original source site as if time were running backwards [7], [8], [9].

The following simulations have been done using MATLAB in order to illustrate the rendering of sound pressure fields with WFS [10]. Figure1 shows a plane wave field of 950Hz propagating in the -45° direction. Figure2 shows a focused spherical wave field of 1050Hz emanating from the position (0.0m, 0.5m). In both figures, the sound pressure fields were synthesized with a circular array of 24 loudspeakers, where the distance between two adjacent loudspeakers is 15.93cm, defining a maximum reproducible frequency is 1067Hz.

3.2. Possibilities of WFS for sonification

The following list highlights the possibilities of using the WFS reproduction technique as a spatial auditory display front-end for sonification experiments with spatial data:

a) In his WFS perceptual experiments, Sonke asked persons to discriminate between the two most different orientations of a plane wave field. He found that the most experienced listener could perceive differences in rotated versions of an 11-sided plane wave polygon. Most listeners lost their discrimination power about an 8-sided polygon, whereas for many persons a 4-sided polygon, i.e., a square of plane waves was enough to give a diffuse field of perception [11].

b) Moving sources can be easily done by rendering an audio signal with focused and non-focused spherical models with moving centers. Furthermore, the Doppler Effect has also been simulated using WFS by Ahrens et al. [12].

c) Although common research on WFS aims at spatial sound control for large audience, the projects at IEM are dedicated to a single person behind his/her computer, where beam formed and focused sources can be synthesized at the position of the listeners head as an attractive alternative to the use of headphones [13].

d) The ability of a listener to localize a sound is determined by frequencies up to about 1500 Hz [14]. When WFS is performed correctly up to frequencies of this order, correct source localization is warranted. The addition of non-localizable sound of higher frequencies often leads to the perception of an increasing "apparent source width": the source sound broader than its actual width [5].

4. SOUNDSCAPES

This section briefly resumes the theoretical approach of Valle et al., [15]. The term *soundscape* has been introduced by Schafer, who studied for the first time the relation between sounds, environment and cultures [16]. This concept plays an important role at the crossing of many sound-related fields, which includes multidimensional data sonification [17] and auditory displays using nature sounds [18]. Thus, the integration of soundscapes emerged as crucial in order to ensure a believable experience in human-computer interaction.

Soundscape studies have highlighted the relevance of different listening strategies in the perception of sonic environments: From a phenomenological perspective, [19], [20], it is possible to identify:

- an *indexical* listening, when sounds are brought back to their source,
- a *symbolic* listening, which maps a sound to its cultural specific meanings,
- an *iconic* listening, indicating the capabilities of creating new meanings from a certain sound material.

Thus, a soundscape can be defined as a temporal and typological organization of sound objects, related to a certain geo-cultural context, in relation to which a listener can apply a spatial and semiotic transformation.

5. SPATIAL SONIFICATION AND SPATIAL SOUND IN A SOUNDSCAPE CONTEXT

The next section proposes to unify the listener and the sound propagation approaches in the soundscape context, where an iconic listening of a sound object should be the appropriate phenomenological model for sonification experiments, where the data to be sonified becomes the new meaning of the sound object.

Annotations a), b) and c) about spatial sonification in section 2 can be directly related to agreements about spatial sonification results a), b) and c) in section 3.2. They clearly relate the radial location, the correct modeling of time effects, and the desired HRTFs, respectively.

The remaining agreement d) in section 2 and the annotation d) in section 3.2 will be exploited together in the next section, where an ongoing application on spatial sonification of non-spatial data using nature sounds is briefly introduced. Indeed, WFS with nature sounds in their spectral content below 1500Hz is enough to localize them.

6. APPLICATION TO SONIFICATION IN SOUND ART EXHIBITIONS

This section introduces an application that benefits from: 1) The sound focusing capability of WFS, and 2) the studies that have found that nature sounds are more easily recognized in an office environment than artificial tones [18]. This application has recently arisen in connection with activities at ISONAR Sound Research Group, devoted to soundscapes of Lima that includes urban and nature sounds. Since 2008 ISONAR is presenting its contents in the Lima Sonora sound art festival [21], which aims to encourage people to consciously perceive its sound environment, thus promoting acoustic ecology education to take further steps in reducing environmental noise.

Sound art exhibitions sometimes require darkness and a minimum of visual information to improve the reduced listening of sound objects in the sense of Schaeffer [19]. Signaling using sound for guiding the route to visitors and also to give information could be very useful in that context. In this application the data to be sonified is a flux of people, i.e., non-spatial data; the sonification method is mapping of sound; and the auditory display is based on an iconic listening of nature sounds.

We are working on the design of horizontal line arrays to be placed on the side walls of a corridor that divides two halls of a gallery, in order to point out if a hall is full or is available to be visited, using focused sound. As sound objects used for this purpose should not interfere with the piece of art and also should be recognized in an intuitive manner, the number of people is mapped into a stream of water. Furthermore, it could be possible to include extra intuitive information in the same sound object. Indeed, According to Chion [20], the following manipulations of a sound object would not affect its timbre: amplification, attenuation, echo and inverted echo.

Figure 5 resumes the proposed sonification project in a block diagram. Media examples of the selected sounds will be available at the festival [21]. The real-time multichannel renderer has been implemented at using Pure Data, an open-source real-time audio processing environment [21]. The graphical user interface, based on the room_sim_2d.pd function [23], allows the rendering of up to five virtual sources through 24 audio channels, giving the capability of including multiple channels of auditory information. This real system is being evaluated at ISONAR with the following equipment: one Mac Book Pro laptop, one M-audio Profire Lightbridge audio interface, three Behringer ADA8000 digital to analog converters, three QSC168X eight channel amplifiers, and 24 Behringer 1CBK loudspeakers.

7. DISCUSSION AND FURTHER WORK

The reproduction of sound using spherical and plane waves models are widely used in room acoustics. Spherical models are used for the rendering of audio tracks such as music and speech, while plane waves are used for the rendering of room impulse responses, reproducing in that way the acoustics of a different room [5]. Another less known application is done in the composition and recreation of soundscapes. Since a background sound are perceived as coming from a non-localized source and a foreground sound as coming from a localized source, their reproduction model can be done respectively with plane and spherical wave models. Both rendering modes could be useful in the rendering of information in sonification projects, however, the synthesis of directional sources [24], [25], moreover, the synthesis of directional and focused sources [26],

could reduce the interference between channels of information and improve the perception of sonified data.

The reproduction of moving virtual sources can also be done using a spherical wave models, however a more exact approach that takes into account spatial aliasing artifacts is being evaluated [12]. This would be useful for interactions such as the movement of the head to shift the focus, which can be achieved using head tracking video systems or ultrasonic position measurement based on trilateration. A critical point here is the measurement of head related transfer functions, which would be included at the final stage of rendering [27].

Finally, the proposed immersive auditory display could also be included in interactive sonification toolkits based on Pure Data, such that the reported in [28].

8. CONCLUSIONS

Spatial sonification requirements have been briefly reviewed and related to the possibilities of the Wave Field Synthesis spatial sound reproduction technique in a soundscape context. A prototype version of an immersive auditory display front-end has been described. This system can be included in interactive sonification toolkits based on Pure Data. The proposed system is capable of immerse the user in a virtual acoustic environment, providing the capability of sound focusing. A sonification project for signaling and monitoring in sound art exhibition galleries has also been introduced.
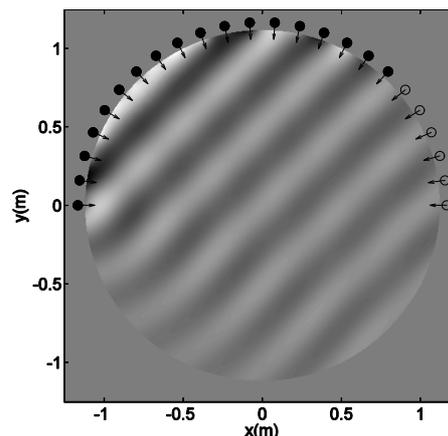
Figure 1. *A plane wave.*
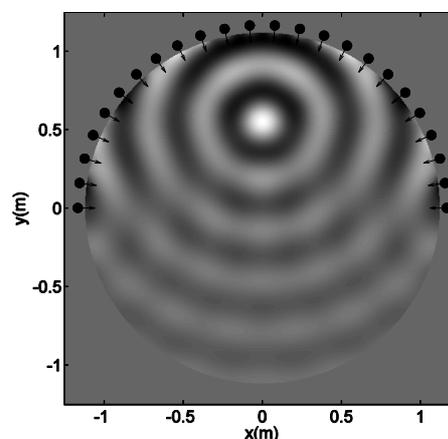
Figure 2. A *focused spherical wave.*
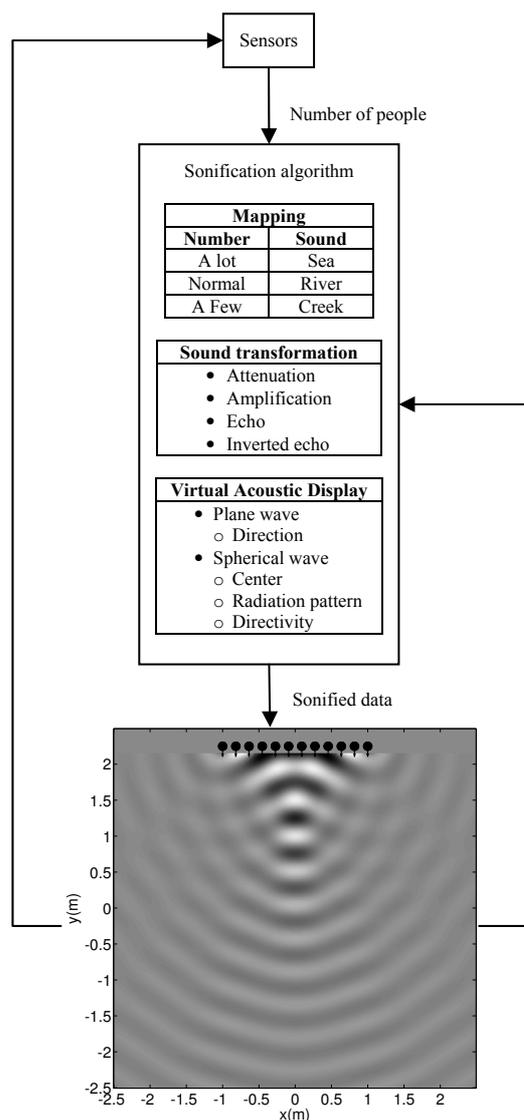
Figure 3. *Sonification project.*

9.   REFERENCES

[1]   T. Hermann, "Taxonomy and definitions for sonification and auditory display," in *Proceedings of the 14th International Conference on Auditory Display*, Paris, France, June 2008.

[2]   T. Nasir and J. Roberts, "Sonification of spatial data," in *Proceedings of the 13th International Conference on Auditory Display*, Montréal, Canada, June 2007.

[3]   A. Berkhout, "A holographic approach to acoustic control," in *Journal of the Audio Engineering Society*, vol. 36, no. 12, december 1988, pp. 977-995.

[4]   A. Berkhout, D. de Vries and P. Vogel, "Acoustic control by wave field synthesis," in *Journal of the Acoustic Society of America*, vol. 93, no. 5, may 1993, pp. 2664-2778.

[5]   D. de Vries, *Wave Field Synthesis*, Audio Engineering Society Monograph, April 2009.

[6]   S. Spors, J. Ahrens and R. Rabenstein, "The theory of wave field synthesis revisited," in *124th Audio Engineering Society Convention*, Amsterdam, The Netherlands, May 2008.

[7]   S. Yon, M. Tanter and M. Fink, "Sound focusing in rooms: The time-reversal approach," in *Journal of the Acoustic Society of America*, vol. 113, no. 3, March 2003, pp. 1533-1543.

[8]   S. Yon, M. Tanter and M. Fink, "Sound focusing in rooms. II. The spatio-temporal inverse filter," in *Journal of the Acoustic Society of America*, vol. 114, no. 6, December 2003, pp. 3044-3052.

[9]   S. Spors, H. Wierstorf, M. Geier, and J. Ahrens, "Physical and perceptual properties of focused sources in wave field synthesis," in *127th Audio Engineering Society Convention*, New York, USA, October 2009.

[10] C. Salvador, "Wave field synthesis using fractional order systems and fractional delays," accepted for presentation in the *128th Audio Engineering Society Convention*, London, United Kingdom, May 2010.

[11] J. Sonke, *Variable Acoustics by Wave Field Synthesis*, Delft University of Technology, Delf, The Netherlands, 2000.

[12] J. Ahrens and S. Spors, "Reproduction of moving virtual sound sources with special attention to the doppler effect," in *124th Audio Engineering Society Convention*, Amsterdam, The Netherlands, May 2008.

[13] M. Strauss, "A spatial audio interface for desktop applications," in *24th International AES Conference*, Banff, Canada, 2003.

[14] J. Blauert, *Spatial hearing*, MIT Press, Boston, USA, 1983.

[15] A. Valle, V. Lombardo and M. Schirosa, "A graph-based system for the dynamic generation of soundscapes," in *Proceedings of the 15th International Conference on Auditory Display*, Copenhagen, Denmark, May 2009.

[16] M. Schafer, *The tuning of the world*. Knopf, New York, USA, 1977.

[17] T. Hermann, P. Meinicke and H. Ritter, "Principal curve sonification," in *Proceedings of the International Conference on Auditory Display,* 2000.

[18] B. Mauney and B. Walker, "Creating functional and livable soundscapes for peripheral monitoring of dynamic data," in *Proceedings of the 10th International Conference on Auditory Display*, Sydney, Australia, July 2004.

[19] P. Schaeffer, *Traité des objets musicaux*. Seuil, Paris, France, 1988.

[20] M. Chion, *El sonido: música, cine, literatura.* Paidós, Barcelona, España, 1999.

[21] Lima Sonora: http://www.limasonora.com

[22] Pure Data: http://www.pure-data.org

[23] T. Musil, IEMLIB: http://iem.at/pd/

[24] E. Corteel, "Synthesis of directional sources using wave field synthesis, possibilities, and limitations," in *EURASIP Journal on advances in signal processing*, Hindawi Publishing Corporation, 2007.

[25] J. Ahrens and S. Spors, "Implementation of directional sources in wave field synthesis," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, USA, October 2007.

[26] J. Ahrens and S. Spors, "Notes on rendering of focused directional sound sources in wave field synthesis," in *34rd German Annual Conference on Acoustics*, Dresden, Germany, March 2008.

[27] S. Takane, Y. Suzuki, T. Miyajima and T. Sone, "A new theory for high definition virtual acoustic display named ADVISE," in *Acoustical Science and Technology*, vol. 24, no. 5, 2003, pp.276-283.

[28] S. Pauletto and A. Hunt, "A toolkit for interactive sonification," in *Proceedings of the 10th International Conference on Auditory Display*, Sydney, Australia, July 2004.