USING MULTI-CHANNEL SPATIALIZATION IN SONIFICATION: A CASE STUDY WITH METEOROLOGICAL DATA

Edward Childs, Ph.D.

Bregman Electronic Music Studio
Dartmouth College,
6242 Hallgarten Hall,
Hanover, NH 03755, USA
Edward.Childs@dartmouth.edu

Ville Pulkki, Ph.D.

Helsinki University of Technology
Acoustics Laboratory
P.O. Box 3000
FIN-02015 HUT, Finland
Ville.Pulkki@hut.fi

ABSTRACT

Spatialization can be used in sonification as an additional mapping dimension. Immersive spatialization also transmits more auditory data to a subject, compared to monophonic or stereophonic presentation. In this study the VBAP spatialization method is applied to the sonification of meteorological data gathered from a large geographical area in two different listening conditions. Parameter-driven FM synthesis, using JSyn (a Java-based sound API) [1] was used to reinforce the spatialization with timbral coloring. The sonification, designed to display a year of data in about twenty minutes, afforded a sense of storm movement, North-South seasonal variation, and comparison between years.

1. INTRODUCTION

The physical location of data to be sonified is frequently as important as the value of the data itself. A significant aspect of the "mapping" problem then becomes how to represent different spatial locations in auditory space. Mapping "space to space" by applying the techniques of 3D sound spatialization to data sonification seems a natural solution to this problem. The direction of the sound object itself thus contains information in a straightforward manner, e.g. the location of a hail storm may be mapped to the direction of a virtual source.

Unlike visual displays, which are restricted to the frontal section, sounds may be spatialized to all directions in spatial auditory displays. When the listener hears sounds from different directions, it is easier to decode multiple simultaneous sounds (the cocktail party effect). When multiple sounds are presented only monophonically or monaurally they mask each other, and the listener cannot distiguish between them. So, more information may be transmitted via spatialized sound objects.

There are many technologies to spatialize audio. They are based either on the use of headphones or loudspeakers. Loudspeakers can be used with two or more channels. The most desirable system for sonification would possess a relatively stable spatialization quality with relatively low numbers of loudspeakers in a large listening area, so that multiple listeners would perceive as similar a 3-D soundscape as possible. The VBAP method was chosen, since it fulfills these requirements quite well, and because of its applicability to arbitrary numbers and locations of speakers.

As a case study, a sonification of historical meteorological data sets was performed, in which the occurrence and location of all hail storms in the continental United States from 1955 to 1995 was provided [2].

The sonification of historical data at a "playback" speed faster than the occurrence of the original data has been used to discover larger scale trends or patterns in financial [3], geological [4], [5] meteorological [6], and other data, which were not apparent from other (sometimes more tedious) visual or analytical methods. Rapid scanning of large multi-variable data sets has also been used to "search", e.g. oil well logs [7], for anomalies or significance – another fruitful area in which to use sound spatialization techniques.

Choosing "natural" and "pleasant" sounds is an important consideration in designing sonifications for use in practical situations. One aspect of "naturalness" is pleasing timbre or tone quality. Chowning and others have observed that the time evolution of the spectral components in a sound determine, to a large degree, its timbre [8] and "liveliness". It is also easier to detect the location of spectrally rich sounds in a spatial sonification [9].

Systematic control of timbre has been cited as a desirable feature of any sonification system [10]. In this study, the implementation of an FM formant instrument was performed using the JSyn API, allowing data-driven manipulation of timbre in addition to other musical parameters such as pitch, envelope, duration, loudness, tempo, etc.

2. SPATIALIZATION METHODS

In this study, the intent is to synthesize spatial impressions; not to recreate spatial sound that existed and was recorded on some occasion. We are primarily concerned with the production of immersive 3-D sound stages, that would be audible by several people in a large listening area with about 10 loudspeakers. Therefore, headphone methods such as HRTF processing were not considered. Wave field synthesis requires a much larger number of loudspeakers and was not used.

2.1. Amplitude Panning

Amplitude panning is the most frequently used panning technique. In it a sound signal is applied to loudspeakers with different amplitudes. The amplitudes are controlled with gain factors denoted with g. The listener perceives a virtual source the direction of which is dependent on the gain factors. When the loudspeakers are in the horizontal plane, a pair-wise panning paradigm is often used, in which sound is applied to two adjacent loudspeakers [12]. If elevated or descended loudspeakers also exist, triplet-wise panning can be utilized, in which a sound is applied to three loudspeakers at one time forming a triangle from the listener's view point.

2.2. Ambisonics

Ambisonics is basically a microphoning technique, however, it can also be used to synthesize spatial audio as an amplitude panning method [13]. The main problem with ambisonics is that the same sound signal appears generally in all loudspeakers. The directional quality degrades quickly outside the best listening area, since the virtual source appears to emanate from the nearest loudspeaker relative to the listener due to the precedence effect [9]. Also, even in the best listening postion, the directional quality is not optimal, since the directional cues produced with ambisonics may differ significantly from the real source cues [14].

2.3. Vector Base Amplitude Panning

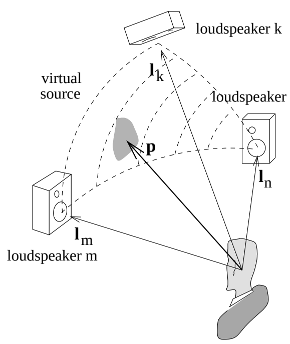

Figure 1: *A loudspeaker triplet forming a triangle formulated for three-dimensional vector base amplitude panning (VBAP).*

Vector base amplitude panning (VBAP) is a method to calculate gain factors for pair-wise or triplet-wise amplitude panning [15]. In VBAP the listening configuration is formulated with vectors; a Cartesian unit vector $\mathbf{l}_n$ points to the direction of loudspeaker n, from the listening position. In triplet-wise panning unit vectors $\mathbf{l}_n$, $\mathbf{l}_m$, and $\mathbf{l}_k$ then define the directions of loudspeakers n, m, and k, respectively. The panning direction of a virtual source is defined as a 3-D unit vector $\mathbf{p} = [p_n \ p_m \ p_k]^T$. A sample configuration is presented in Fig. 1.

The panning direction vector $\mathbf{p}$ is expressed as a linear combination of three loudspeaker vectors $\mathbf{l}_n, \mathbf{l}_m$, and $\mathbf{l}_k$, in matrix form:

$$\mathbf{p}^T = \mathbf{g}\mathbf{L}_{nmk}. \tag{1}$$

Here g_n, g_m and g_k are gain factors, $\mathbf{g} = [g_n \ g_m \ g_k]$ and $\mathbf{L}_{nmk} = [\mathbf{l}_n \ \mathbf{l}_m \ \mathbf{l}_k]^T$. Vector $\mathbf{g}$ can be solved

$$\mathbf{g} = \mathbf{p}^T\mathbf{L}_{nmk}^{-1} \tag{2}$$

if $\mathbf{L}_{nmk}^{-1}$ exists, which is true if the vector base defined by $\mathbf{L}_{nmk}$ spans a 3-D space. Eq. 2 calculates barycentric coordinates of vector $\mathbf{p}$ in a vector base defined by $\mathbf{L}_{nmk}$. The components of vector $\mathbf{g}$ can be used as gain factors; a scaling of them may be desired. When more than three loudspeakers are present, the loudspeaker system is triangulated, and one triplet is used at one time for panning.

The perceptual quality of virtual sources produced with VBAP, or more generally with amplitude panning has been studied in [17]. It was found that VBAP accurately predicts the azimuthal angle between the median plane and virtual source direction. Perception of the elevation angle was found to be more dependent on the individual. However, with triplet-wise panning, the perceived direction of virtual source is generally inside the loudspeaker triangle. This holds even if the listener is not located in the sweet spot.

3. SONIFICATION SCHEME

3.1. Data

The meteorological data files chosen for sonification provide (chronologically, one line for each event) the date, time (to the nearest minute) and location of every hail storm in the continental United States recorded at a National Weather Station (NWS), together with the diameter of the hail, and the number of injuries. The location information is by State, County, NWS office (a three letter code), latitude and longitude. A few lines from a file of 1995 data are shown as an example (columns not used in the sonification have been omitted).

950150101272330	342808752	HSV	00059	75C
950302801280029	320208853	MEI	00023	75C
950312801280030	322408839	MEI	00075	75C
950322801280100	314208906	MEI	00067	75C

The columns contain the following information:

1. The year, sequence number, state, date and time.
2. The latitude and longitude in degrees and minutes.
3. A three letter code for the name of the NWS at which the data was recorded.
4. The number of injuries, and the FIPS (Federal Information Processing Standards) code for the county.
5. The size of the hail in hundredths of an inch ($75 = 0.75$ inches).

As an example, the first entry depicts a storm (hail, diameter 0.75 inches) which took place on 27 January, 1995 at 2330, in the state of Alabama (FIPS code 1), at a latitude of 34 degrees 28 minutes, and a longitude of 87 degrees 52 minutes, recorded at the Huntsville (HSV) weather station, in Franklin county (FIPS code 59), in which no injuries occurred. The next 3 entries track, presumably, the same storm, occurring just past midnight the following day in Mississippi (FIPS code 28) in three counties Clarke (23), Lauderdale (75) and Jones (67), but all reported at the Key Field Airport in Meridian (MEI) (see Fig. 2).

3.2. Java VBAP Implementation and Spatial Mapping

JSyn is a Java API which provides an extensive toolkit of unit generators, filters, effects, mixers, etc. which may be encapsulated and customized by the user [1]. The low-level synthesis is carried out by native-compiled C subroutines. It is convenient to design sonifications in this environment because the necessary input/output, GUI design, and analytical calculations and sound generation may all be implemented in the same project. The existing VBAP C subroutines were accessed from Java via the Java Native Interface (JNI).

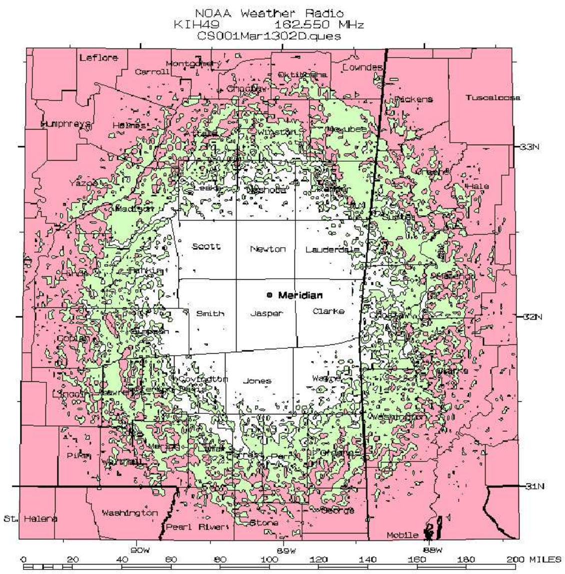

Figure 2: *Meridian, Mississippi Weather Station Area*

In the current VBAP implementation, an automatic method is used to divide the loudspeaker setup into pairs or triplets based on loudspeaker directions. An automatic triangulation was also adapted based on existing code [19]. It selects the triangles based on the greedy triangulation method. The method outputs the inverse matrices and other data needed in VBAP gain calculation. The system can be thus adapted to a new loudspeaker setup simply by defining the number and directions of loudspeakers in it.

In the Java implementation, a single VBAP object is instantiated using the loudspeaker location data (azimuth, elevation). When the sonification calculation is started, each hail storm location (in latitude and longitude) is read from a file and then converted to the azimuth and elevation of a single virtual source and passed to the VBAP object. The VBAP object then returns the gain factors for each loudspeaker, which are then used, via a standard JSyn SynthMixer object, to control the gain on the output channels.

The mapping of geographical space to auditory space was carried out by conceptually superimposing a map of the continental US on the hemispherical dome created by the 3D loudspeaker arrays. In this implementation, the geographical center of US (in the state of Kansas) is "mapped", approximately, to an elevation of 90° (ie. in a position directly above the listener). The distance r from this central location to any other location is calculated from the differences of the longitudes and latitudes, and the desired elevation ϕ (in degrees) of the sound source is then obtained from:

$$\phi = (1 - r^\star) \times 90 \tag{3}$$

where $r^\star$ is r/R, and R is the distance from the center of the US to the most distant point on the continent. As $r^\star \rightarrow 1$, $\phi \rightarrow 0^\circ$, and the virtual sound source elevation moves from 90° to 0°.

3.3. FM Formant Instrument and Sound Mapping

A three oscillator instrument consisting of one modulator and two carriers was chosen for the sonification, see Fig. 3. The inputs available for manipulation are labelled at the top of each unit. The timbre of the instrument may be adjusted via the manipulation of five parameters:

1. The index of modulation I.
2. The ratio of the frequency of Carrier 1 c_1 to the modulating frequency m (c_1/m).
3. The ratio of the frequency of Carrier 2 c_2 to that of Carrier 1 (c_2/c_1).
4. The ratio of the "depth" of Carrier 2 d_2 to that of Carrier 1 d_1 (d_2/d_1). If, for example, this ratio is less than one, the frequency deviation of Carrier 2 will be less than that of Carrier 1.
5. The ratio of the amplitude a_2 of Carrier 2 to that of Carrier 1 (a_2/a_1).

This instrument was originally proposed by Chowning [8] as a means of introducing a formant peak into the spectrum. It was implemented in JSyn by encapsulating three FMOperator objects into a SynthCircuit. Each FMOperator is controlled by an separate envelope, so that the time evolution of the modulation index, the formant peak and the fundamental tone (Carrier 1) can be independently controlled.

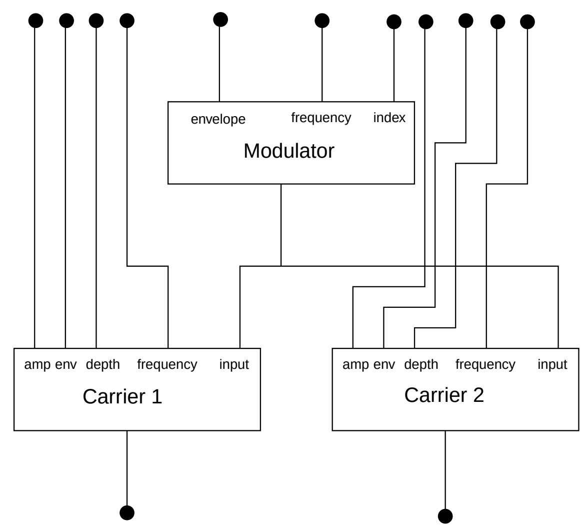

Figure 3: *FM Formant Instrument*

In the hail storm sonification, manipulation of the pitch of the fundamental tone, and of timbre via the above-mentioned five parameters, and the oscillator envelopes was primarily used to reinforce the spatial mapping described in Section 3.2, or to draw attention to a storm in which injuries had occurred (by increasing the index of modulation). The size of the hail was mapped to loudness and duration. The mapping and instrument afforded a variety of typical FM woodwind, percussion and bell sounds.

The spacing between notes was derived from the timings in the data file. The playback speed can be adjusted by the user, but a typical speed allows data for one year to be played in about 20 minutes.

3.4. User Interface

The hail storm data exploration program user interface allows the user to select the year to be explored, the "listening" location in the continental US, and the range from the listening area. For example, the user could simply elect to "listen" in the center of the US, to storms throughout the continent, or rather to listen at some location in Texas, to storms only in Texas. The sonification may be controlled in a standard fashion (start, stop, pause, rewind), and the "playback speed" may be adjusted as desired.

4. REALIZATION

The sonification was first realized in the Helsinki University of Technology Acoustics Lab Listening Room during April-May of 2002, using 5 speakers in the horizontal plane, at azimuthal positions of 30°, -30°, 90°, -90° and 180°, and 3 speakers at an elevation of 45°, at azimuthal positions of 40°, -40° and 180° and a subwoofer.

After the mapping scheme was modified, and a user interface was added, the sonification was realized a second time, during January 2003, in Bregman Electronic Music Studio listening room, using a 12 speaker system, comprising 8 loudspeakers on the horizontal plane at azimuthal positions of 15°, -15°, 45°, -45°, 90°, -90°, and 150°, -150°, and 4 loudspeakers an elevation of 45° at azimuthal positions of 0°, 90°, -90° and 180°.

5. RESULTS

The overall purposes of the sonification, to experiment with the concept of mapping physical space to auditory space, to observe both the local movements of hail storms and global seasonal changes in their location, to display a year of data sufficiently quickly that differences in patterns from year to year could be explored, and to use timbral differences to reinforce spatial location, were successfully realized.

Spatialization of the hail storm data to 8 or 12 channels afforded a significantly enhanced perception of storm movement and location, compared to a stereo system. During the US winter, hail storms occur predominantly in the South. Intensity and frequency increases substantially during the spring. These variations in location and frequency were apparent in the sonification.

Following feedback from colleagues at the Helsinki University of Technology, modifications of the FM instrument parameter mapping were made to emphasize generally lower audio frequencies. Variable playback speed, so as to pause, rewind and replay areas of interest, and the ability to adjust of the location and extent of the "listening" area were added to the interface. After further tests at Dartmouth College, listeners commented that the spatialization and storm movement was clear. A time passage indicator, to announce the months of the year and a sonic "legend" to elucidate the effect of the different parameters on the sound, were requested.

6. ACKNOWLEDGMENTS

This work has received funding from the Academy of Finland (101339), the John Sloan Dickey Center for International Understanding and the Bregman Electronic Music Studio at Dartmouth College in Hanover, NH.

7. REFERENCES

[1] *JSyn Audio Synthesis*, http://www.softsynth.com

[2] *Storm Prediction Center*, http://www.spc.noaa.gov/archive/hail

[3] O. Ben-Tal, J. Berger, B. Cook, M. Daniels, G. Scavone and P. Cook, "SONART: The Sonification Application Research Toolbox", *Proceedings of the 2002 International Conference on Auditory Display, Kyoto, Japan, July 2 - 5, 2002*, p. 151.

[4] F. Dombois, "Using Audification in Planetary Seismology", *Proceedings of the 2001 International Conference on Auditory Display, Espoo, Finland, July 29 - August 1, 2001*, p. 227.

[5] F. Dombois, "Auditory Seismology: On Free Oscillations, Focal Mechanisms, Explosions and Synthetic Seismograms", *Proceedings of the 2002 International Conference on Auditory Display, Kyoto, Japan, July 2 - 5, 2002*, p. 27.

[6] M. Fernström, N. Griffith and S. Taylor, "Bliain Le Baisteach - Sonifying a Year With Rain", *Proceedings of the 2001 International Conference on Auditory Display, Espoo, Finland, July 29 - August 1, 2001*, p. 282

[7] S. Barrass and B. Zehner, "Responsive Sonfication of Well-logs", *Proceedings of the 2000 International Conference on Auditory Display, Georgia Institute of Technology, Atlanta, Georgia, April 2 - 5*

[8] John M. Chowning, "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation" *Journal of the Audio Engineering Society* 21(7), 1973.

[9] J. Blauert, *Spatial Hearing, Revised edition*. Cambridge, MA, USA: The MIT Press, 1997.

[10] C. Scaletti and A.B. Craig "Using sound to extract meaning from complex data", *SPIE Vol. 1459 Extracting Meaning from Complex Data: Processing, Display, Interaction II*, 1991

[12] J. Chowning. The simulation of moving sound sources. *J. Audio Eng. Soc.*, 19(1):2–6, 1971.

[13] D. G. Malham and A. Myatt, "3-d sound spatialization using ambisonic techniques," *Comp. Music J.*, vol. 19, no. 4, pp. 58–70, 1995.

[14] V. Pulkki. Microphone techniques and directional quality of sound reproduction. Convention Paper 5500 presented at the 112th Convention Munich, Germany, 2002.

[15] V. Pulkki. Virtual source positioning using vector base amplitude panning. *J. Audio Eng. Soc.*, 45(6):456–466, June 1997.

[16] J. C. Bennett, K. Barker, and F. O. Edeko, "A new approach to the assessment of stereophonic sound system performance," *J. Audio Eng. Soc.*, vol. 33, pp. 314–321, May 1985.

[17] V. Pulkki. *Spatial sound generation and perception by amplitude panning techniques*. PhD thesis, Helsinki University of Technology, 2001.

[18] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am*, vol. 93, May 1993.

[19] V. Pulkki and Tapio Lokki. Creating auditory displays to multiple loudspeakers using vbap: A case study with diva project. In *International Conference on Auditory Display*, Glasgow, Scotland, 1998. ICAD.