Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York, Oct. 17-20, 1999

UNIFORM SPREADING OF AMPLITUDE PANNED VIRTUAL SOURCES

Ville Pulkki

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

ABSTRACT

The perceived spatial spread of amplitude panned virtual sources is dependent on the number of loudspeakers that are used to produce them. When pair-wise or triplet-wise panning is applied, the number of active loudspeakers varies as a function of the panning direction. This may cause unwanted changes in spatial spread and coloration of a virtual source if it is moved in the sound stage. In this paper a method is presented to make the directional spread of amplitude panned virtual sources independent of their panning direction. This is accomplished by panning the sound signal to multiple directions near each other simultaneously. This forms a single virtual source with constant directional spread as a function of direction.

1. INTRODUCTION

In many theaters and auditoriums there exists loudspeaker systems that include a large number of loudspeakers. Audio systems with multiple loudspeakers are also becoming more common in domestic use. In domestic use the speakers are most often placed to a horizontal plane around the listener (two-dimensional speaker setups). In theaters there are also systems in which also elevated or descended loudspeakers exist (three-dimensional setups).

Virtual sources can be positioned to such loudspeaker systems using various methods. In two-dimensional setups the panning is most often performed using pair-wise panning methods [1]. Pair-wise panning can be generalized to triplet-wise panning for three-dimensional loudspeaker setups [2]. Pair-wise panning and triplet-wise panning yields an acceptable virtual source quality in relatively large listening area. However, the virtual source quality is dependent on panning direction because the number of loudspeakers emanating the same sound signal varies in different directions. This may be perceived inconvenient when moving virtual sound sources are applied.

In matrixing sound reproduction techniques, like Ambisonics [3], sound signals are encoded analogically to few audio channels. In the decoding stage loudspeaker signals are decoded using some matrixing operations. In these techniques the same sound signal is existent in all loudspeakers, which may degrade the virtual source quality. However, when moving sound sources are applied, the number of loudspeakers producing a virtual source is not dependent on panning direction, thus the directional spread does not vary prominently.

Basically the matrixing systems and pair-wise or triplet-wise panning methods are quite similar techniques. In both methods a sound signal is applied to a number of loudspeakers with different amplitudes. The goal of this study is to form methods for virtual source positioning that would be something in between the matrixing techniques and pair-wise or triplet-wise panning.

This paper is organized as follows: At first some basics of spatial hearing and amplitude panning are reviewed. The method for spreading amplitude-panned virtual sources is then presented. The directional spread of pair-wise panned virtual sources and spreaded virtual sources are computed using a binaural model.

2. SPATIAL HEARING

Spatial and directional hearing has been studied intensively; for overviews, see e.g. [4]. The duplex theory of sound localization states that the two main cues of sound source localization are the *interaural time difference* ITD and the *interaural level difference* ILD which are due to wave propagation time difference (primarily below 1.5 kHz) and the shadowing effect by the head (primarily above 1.5 kHz), respectively. Both cues depend on frequency.

In the median plane where the distances to both ears are equal, the ITD and ILD values are close to zero. Other effects, such as spectral cues and head movements, are considered to carry elevation and front-back information in the median plane. Spatial discrimination is difficult also in so called cones of confusion where ILD and ITD vary only slightly due to the unsymmetry of the head. A cone of confusion can be approximated with a cone which has symmetry axis in the line passing through listener's ears.

If the cues of an auditory object are conflicting, the listener may perceive the object in multiple directions simultaneously. The directional spread denotes how well the auditory object is situated in one direction: what smaller directional spread, that more point-like auditory object.

3. AMPLITUDE PANNING

Amplitude panning [5] is most often applied to two loudspeakers which are in a standard stereophonic listening configuration, as depicted in Fig. 1. Loudspeakers 1 and 2 are placed in front of the listener with aperture of 60°. A signal is applied to each loudspeaker with different amplitudes, this can be formulated as

$$s_i(t) = g_i s(t), \quad i = 1, 2, \tag{1}$$

where $s_i(t)$ is the signal to be applied to loudspeaker i, g_i is the gain factor of the corresponding channel, and t is the time parameter.

The sound signals arrive to listener's both ears from both loudspeakers. If the wave propagation time difference is taken into account, but the shadowing effect of the head is neglected, we may presented the sine law [6] as

$$\frac{\sin\varphi}{\sin\varphi_0} = \frac{g_1 - g_2}{g_1 + g_2}, \tag{2}$$

where φ is the perceived angle of a virtual source and φ_0 is the loudspeaker base angle, as in Fig. 1. The equation is valid only when the frequency is below 600 Hz and when the listener's head is pointing directly forward. In the equation it is also assumed that the elevation is $0°$. The equation does not set limitations to φ, but in most cases it its value is set to satisfy $|\varphi| \leq \varphi_0$. If $|\varphi| > \varphi_0$ the amplitude panning will produce antiphase loudspeaker signals which may distort the virtual source [4].

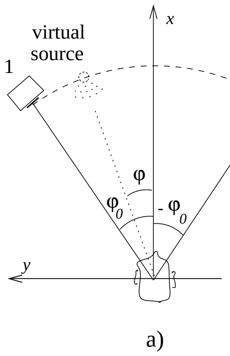

Figure 1: *A standard stereophonic listening configuration. The virtual source can be positioned between the loudspeakers.*

If the listener is facing towards the virtual source, the tangent law is more correct [7]

$$\frac{\tan\varphi}{\tan\varphi_0} = \frac{g_1 - g_2}{g_1 + g_2}. \tag{3}$$

This equation has the same limitations as the sine law.

The perception of the virtual source can be modeled by constructing the sound signal entering listener's ears using measured HRTFs and listening setup model [8]. The frequency-dependent ILD and ITD can be then calculated using a binaural auditory model. The virtual source qualities can then be monitored by comparing the ITDs and ILDs of real and virtual sources with each other. Further, the ITD and ILD can be expressed in direction angles, according to individual table lookup, which enables comparation of sound object direction perception between individuals. It has been shown that the model is able to predict many phenomena existing in spatial reproduction [11].

3.1. Pair-wise amplitude panning

When the number of loudspeakers exceeds two, pair-wise panning methods can be used. The sound signal is panned to the adjacent loudspeaker pair between which the virtual source direction lies. The virtual source is produced using one or two loudspeakers. If the virtual source is panned coincident with a loudspeaker, only that particular loudspeaker emanates the sound signal.

When the sound signal is panned to a single loudspeaker, it is actually not a virtual source, but a real source. This means that its directional spread is as low as it can be. When the virtual source is panned to two loudspeakers, the directional spread is higher, it can be localized erroneously and it may be colored. The directional spread of an auditory object may be estimated by simulating the listening condition and by calculating the main localization cues. The average directional spread q perceived by an individual listener may be estimated as

$$q = w_{\mathrm{ILD}}||\alpha - \mathrm{ILD}|| + w_{\mathrm{ITD}}||\alpha - \mathrm{ITD}||, \tag{4}$$

where w_{ILD} and w_{ITD} represent a frequency-dependent salience function of ILD and ITD, respectively. Variable α denotes the median value of ILD and ITD, which represents the center direction of the sound object around which the spectral spread appears. ITD and ILD are expressed as azimuth angles at 42 ERB (equivalent rectangular bandwidth) channels.

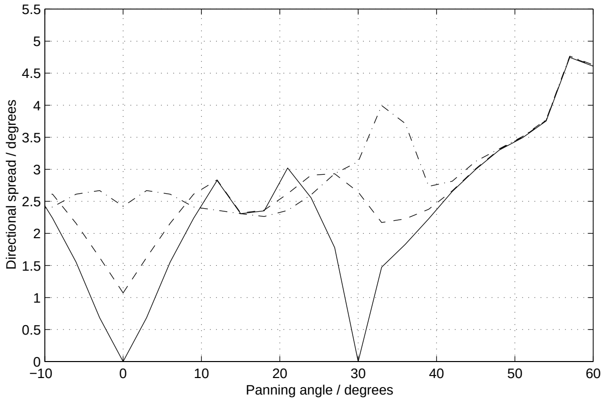

Figure 2: *Simulated directional spread of pair-wise amplitude panned and multiple-direction amplitude panned virtual sources. Solid: pair-wise panning. Two-direction panned virtual sources: dashed - spread angle* $20°$, *dash-dotted - spread angle* $30°$. *5 loudspeakers, directions* $-90°$, $-30°$, $0°$, $30°$, *and* $90°$ *were used. 2-D VBAP,* $||g|| = 1$

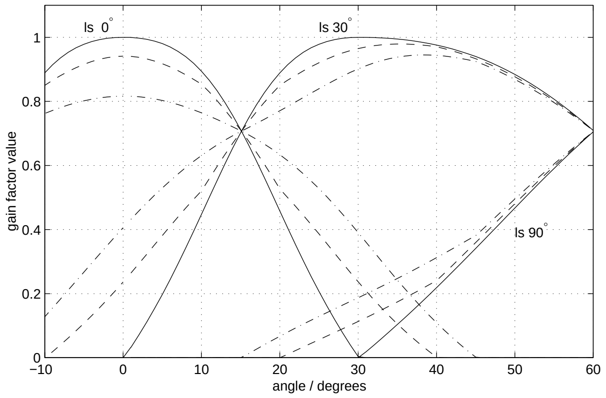

Figure 3: *The gain factors of loudspeakers at direction* $0°$, $30°$ *and* $90°$ *as function of panning angle. Solid: one panning direction. Two panning directions: dashed - spread angle* $20°$, *dash-dotted - spread angle* $30°$. *5 loudspeakers, directions* $-90°$, $-30°$, $0°$, $30°$, *and* $90°$ *were used. 2-D VBAP,* $||g|| = 1$.

In this study a rough approximation was used of the saliences of directional cues. The ITD salience function w_{ITD} was modeled with a function with value 1.0 at low frequencies and 0.1 at high frequencies. The ILD salience function w_{ILD} had value 0 at low frequencies and 0.4 at high frequencies. Between frequencies 400 Hz and 2000 Hz the saliences changed their values linearly.

A 5-loudspeaker system with speaker directions $0°$, $30°$, $90°$, $-90°$, and $-30°$ was simulated, and a pair-wise panned virtual source with pink noise was applied to it. The directional spread was computed between $-10°$ and $60°$, as shown in Fig. 2. The directional spread was calculated for 5 individuals, and the average value over individuals was set to be the directional spread estimate. The HRTFs of individuals have been measured by Wightman and Kistler [9].

In the figure it can be seen that the directional spread has local minima coincident with loudspeaker directions. Between the loudspeakers the virtual source is spread prominently. This simulation result corresponds well to known behavior amplitude-panned virtual sources. The corresponding gain factors are shown in Fig. 3.

3.2. Triplet-wise panning

In triplet-wise panning, the sound signal is applied to maximally three loudspeakers at time. The loudspeakers form a triangle when viewed from the listener's position, as in Fig. 4. This enables three-dimensional positioning of virtual sources. Three-dimensional vector base amplitude panning (VBAP) is a method for positioning virtual sources to arbitrary directions inside arbitrary loudspeaker triplets [2].

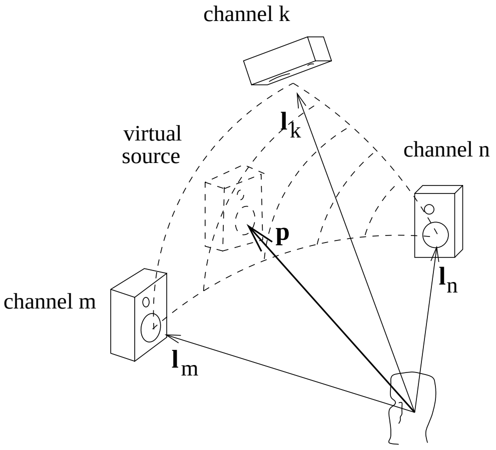

Figure 4: *Three-dimensional vector base amplitude panning.*

In three-dimensional VBAP a loudspeaker triplet is formulated with vectors. The unit-length vectors $\mathbf{l}_m, \mathbf{l}_n$ and $\mathbf{l}_k$ point from listening position to the loudspeakers. The direction of the virtual source is presented with unit-length vector $\mathbf{p}$. Vector $\mathbf{p}$ is expressed as a linear weighted sum of the loudspeaker vectors

$$\mathbf{p} = g_m \mathbf{l}_m + g_n \mathbf{l}_n + g_k \mathbf{l}_k. \tag{5}$$

Here g_m, g_n, and g_k are called gain factors of respective loudspeakers. The gain factors can be solved as $\mathbf{g} = \mathbf{p}^T \mathbf{L}_{mnk}^{-1}$, where $\mathbf{g} = [g_m \; g_n \; g_k]^T$ and $\mathbf{L}_{mnk} = [\mathbf{l}_m \; \mathbf{l}_n \; \mathbf{l}_k]$. The calculated factors are used in amplitude panning as gain factors of the signals applied to respective loudspeakers after suitable normalization, e.g. $||\mathbf{g}|| = 1$.

When the number of loudspeakers is greater than three, the loudspeaker setup is divided to triangles forming a triangle set. An automatic triangle forming algorithm has been presented in [10]. During the panning process a single triangle from the set is chosen to be used in panning. The selection can be made by calculating the gain factors in each loudspeaker triangle in the triangle set and selecting the triangle that produced non-negative factors. If the triangles in the set are non-overlapping, the selection is unambiguous. VBAP can be also formulated in two dimensions, in this formulation it can be applied also to pair-wise amplitude panning. 2-D VBAP is a reformulation of tangent law (Eq. 3) [2].

3-D VBAP has a similar feature with pair-wise amplitude panning had, the directional spread is dependent on the panning direction. The change of directional spread can be even higher because the number of loudspeakers emanating same sound signal varies from one to three.

4. UNIFORM SPREADING OF AMPLITUDE-PANNED VIRTUAL SOURCES

It has now be shown that the directional spread of pair-wise and triplet-wise virtual sources depends on panning direction. Methods for making the directional spread of a virtual source constant are now considered. The best way would be, of course, to make the virtual source as point-like as possible. There exists some attempts to this [11]. However, reducing the directional spread demands computational resources, and the spread can in very few cases removed totally.

Other way to solve the problem is to increase the directional spread in directions where the loudspeakers are in order to create a roughly constant directional spread. The directional spread can be increased by applying the same sound signal to more than one physical loudspeaker at each time. However, the virtual source quality on directions between the loudspeakers should not be degraded more.

This can be implemented with a method in which the sound signal is panned simultaneously to multiple panning directions near the wanted virtual source direction. Despite of the multiple panning directions, the listener will still perceive a single virtual source. The average direction of the panning directions is considered as the panning direction, where the virtual source is wanted to appear. The method is called multiple-direction amplitude panning (MDAP). The largest angle between the multiple panning directions is called the spread angle.

When the multiple panning directions are located inside the same loudspeaker set, MDAP is equivalent to traditional amplitude panning with a single virtual source, since the sound signals are panned to same loudspeakers. This implies that MDAP does not degrade further the virtual source quality in directions where the virtual source quality is at its worst in traditional amplitude panning.

The difference between traditional amplitude panning and MDAP appears when there exist loudspeakers between the panning directions in MDAP. The sound signals are then panned to different loudspeaker sets, which affects the gain factors and increases the amount of used loudspeakers. This spreads the virtual source. The amount of spreading can be selected by adjusting the spread angle of MDAP.

When the virtual source is desired to be spread in two-dimensional speaker setups, the effect can be achieved using two panning directions, as in Fig. 5. If the loudspeakers are not spread evenly, there might be situations in which there exist more than one speaker between panning directions of one virtual source, which may cause that some of the speakers between the panning directions may not emanate sound at all. This can be avoided by applying more panning directions to MDAP.

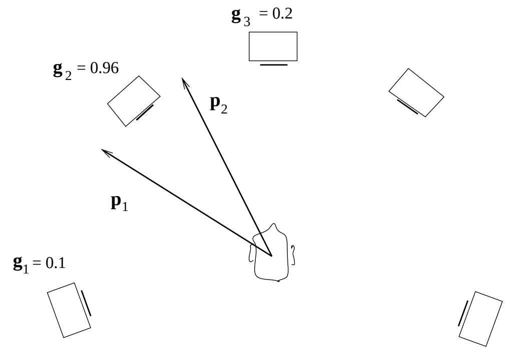

Figure 5: *Spreading the virtual source using two panning directions.*

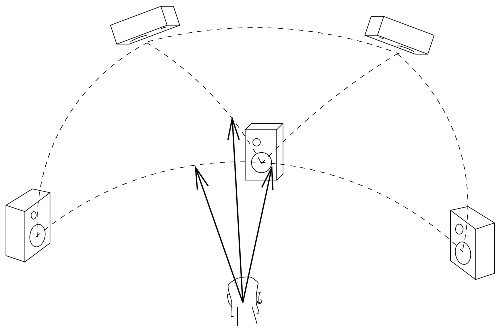

Figure 6: *Spreading the virtual source in 3-D loudspeaker setups using three panning directions*

In 3-D loudspeaker setups, if two panning directions are used in MDAP, the sound signal is applied at least to two loudspeakers at time. If the sound is desired to be applied at least to three loudspeakers at time, three panning directions are needed as is illustrated in Fig. 6.

The directional spread of virtual sources created with MDAP with 2 panning directions with spread angles 20° or 30° are shown in Fig. 2. The virtual sources created with MDAP do not have local minima on directions coincident with loudspeakers, as pair-wise-panned virtual sources have. It can also be seen that the amount of directional spread increases when the spread angle of MDAP is increased. The directional spread is also roughly constant over panning direction with virtual sources created with MDAP. The result corresponds well to informal listening tests in which it was found that the directional spread of virtual sources created with MDAP did not vanish on directions coincident with loudspeakers. The corresponding gain factors are shown in Fig. 3

In matrixing systems the reproduced sound signal is applied to virtually all loudspeakers at one time. This can be simulated with MDAP by applying the sound signal to needed amount of directions and by gaining the input signal properly in different directions.

5. CONCLUSIONS

The pair-wise and triplet-wise amplitude panned virtual sources can be spread by panning the same sound signal to multiple directions. This techniques is called multiple-direction amplitude panning (MDAP). When applying MDAP, the sound signal never emanates from only one loudspeaker. The directional spread is increased on directions coincident with loudspeakers, but the directional spread between loudspeakers remains on same value as in pair-wise or triplet-wise panning. The amount of directional spread of the virtual source can be controlled by adjusting the panning directions. The panning direction sets can be composed in such a manner that the spread of virtual sources is not dependent on panning direction. The directional spread was simulated using a binaural model, and examined with informal listening tests.

6. ACKNOWLEDGMENT

The work of Mr. Pulkki has been supported by the Graduate School in Electronics, Telecommunications and Automation (GETA). The author would like to thank Dr. Fred Wightman and Dr. Doris Kistler at University of Wisconsin, Madison for making the HRTF measurements available to scientific community.

7. REFERENCES

[1] J. Chowning, “The simulation of moving sound sources,” *J. Audio Eng. Soc.*, vol. 19, no. 1, pp. 2–6, 1971.

[2] V. Pulkki, “Virtual source positioning using vector base amplitude panning,” *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456–466, 1997.

[3] M. A. Gerzon, “Periphony: With-height sound reproduction,” *J. Audio Eng. Soc.*, vol. 21, no. 1, pp. 2–10, 1972.

[4] J. Blauert, *Spatial Hearing, Revised edition*, The MIT Press, Cambridge, Massachusetts, 1997.

[5] A. D. Blumlein, ,” U.K. Patent 394,325, 1931, Reprinted in Stereophonic Techniques, Audio Eng. Soc., NY, 1986.

[6] B. B. Bauer, “Phasor analysis of some stereophonic phenomena,” *J. Acoust. Soc. Am.*, vol. 33, no. 11, pp. 1536–1539, November 1961.

[7] B. Bernfeld, “Attempts for better understanding of the directional stereophonic listening mechanism,” in *44th Convention of the Audio Engineering Society*, Rotterdam, The Netherlands, February 1973.

[8] V. Pulkki, M. Karjalainen, and J. Huopaniemi, “Analyzing virtual source attributes using a binaural model,” *J. Audio Eng. Soc.*, vol. 47, no. 4, 1999.

[9] F. L. Wightman and D. J. Kistler, “HRTF measurements,” http://www.waisman.wisc.edu/hdrl/, 1998.

[10] 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, England, 1998, ICAD.

[11] V. Pulkki, M. Karjalainen, and V. Välimäki, “Localization, coloration, and enhancement of amplitude panned virtual sources,” in *The Proceedings of the AES 16th International Conference*, 1999, pp. 257–278.