Audio Engineering Society

Convention Express Paper 119

Presented at the 155th Convention
2023 October 25–27, New York, USA

Emulating Vector Base Amplitude Panning Using Panningtable Synthesis

Timothy Schmele[1,2] and Niklas Reppel[3]

[1] *University of Music Karlsruhe, Institute for Musicology and Music Informatics, Karlsruhe, Germany*
[2] *Eurecat, Centre Tecnològic de Catalunya, Multimedia Technologies Group, Barcelona, Spain*
[3] *Toplab Barcelona, Barcelona, Spain*

Correspondence should be addressed to Timothy Schmele (`tim.schmele@eurecat.org`)

ABSTRACT

This paper presents Panningtable Synthesis (PTS) as an alternative approach to panning virtual sources in spatial audio that is both a generalization to and more efficient than Vector Base Amplitude Panning (VBAP). This new approach is inspired by a previous technique called Rapid Panning Modulation Synthesis (RPMS). RPMS however exhibits the limitation in that all secondary sources need to be regularly spaced across the circle and organized in equally spaced circles across the sphere. We demonstrate that PTS is not only able to overcome these restrictions, but that it is also fully compliant with VBAP, more computationally efficient and can be regarded as a generalization to the same. Furthermore, we demonstrate that PTS is also able to supersede RPMS both in its capacity to create and shape sound spectra, independently from the number of secondary sources used in the array. Considering creative spatial sound synthesis techniques, PTS can be compared to Wavetable or Wave-Terrain Synthesis, but with the added, inherent spatial characteristics. The flexibility of PTS allows any degree of trade-off between using perceptually correct panning curves and those that target specific sound spectra.

1 Introduction

The discussion around multi-channel spatial audio is mostly centered on the physical reconstruction of a sound field [1, 2, 3] or correct perceptual reproduction of the audible panorama [4, 5]. In relation to post-production, it often focuses on the correct placement of point sources using positional coordinates [6, 7, 8, 9] and descriptions of their widths [10, 11, 12]. For example, even though Ambisonics [1] represents the entire spatial scene in a holistic manner using spherical harmonics, for as long as an audible 3D scene is thought of as individual points in space, working with point sources and positional metadata will remain the basic mode of thinking in spatial audio productions. Furthermore, practical limitations and the cinema industry's focus on the action on screen would often steer common solutions to the front panorama [13].

Spatial Sound Synthesis methods represent a paradigm shift to this way of thinking. Here, we understand Spatial Sound Synthesis as methods that intend to tightly combine the creation of sound spectra with a sound's spatial appearance. Often, this leads to the sounding result to be distributed in space in a particular way, taking on a specific spatial shape alongside a specific sound spectrum. For example, time based approaches, such as granular synthesis, spatialize each grain of audio individually [14, 15, 16], while spectral approaches give each extracted frequency band an individual position in space [17, 18, 19].

Recently, it has been shown that the circular movement of a virtual source creates a spectral split in the velocity component of the acoustic field [20], which is comparable to amplitude modulation. Other methods that use a sound's movement to modulate the input signal include modulating the virtual source's distance in order to make use of the Doppler effect [21], which was compared to the results achieved with FM synthesis [22], as well as the expansion of the pair-wise panning technique, to be able to pan a virtual source at audio rate using panning curves stored in a buffer, coined Rapid Panning Modulation Synthesis (RPMS) [23]. It has also been proposed to use the scan lines of wave terrain synthesis [24] as trajectories in spatial audio, as a way to more tightly combine the sound synthesis technique with a sound's movement through space [17, 25, 26].

In this paper we present *Panningtable synthesis* (PTS), which is an expansion on the aforementioned RPMS method to utilize pair-wise panning as a starting point for creative sound synthesis. The primary intention behind RPMS is to move a virtual source at velocities far beyond the perceptual limit of rotation [27], which produces audible artifacts that can be shaped. While this causes spectral changes to the sound a virtual source emits into the listening space, it also transforms a point source to an immersive, all-encompassing sound. As such, PTS, like RPMS, can be regarded as a creative spatial sound synthesis technique, in which the specific sensation of immersion with which the sound engulfs the listener is tightly linked to the spectral modulation applied to the input audio. Similar to wavetable synthesis [24], PTS makes use of panningtables to control both the characteristic spectral artifacts applied to the virtual source, while determining the specific gain values in each loudspeaker for every spatial position on a circle or sphere.

In particular, this paper will primarily focus on demonstrating that PTS is a valid panning technique by emulating Vector Base Amplitude Panning (VBAP) [5] within precisional error. This initial step is necessary in order to claim that the spectral changes that occur in PTS are indeed caused solely by the spatial *movement* of a sound – as long as it is accepted that perceptual panning truly constitutes moving a virtual source in space. Moreover, it will be shown that PTS is able to outperform the standard VBAP implementation while also facilitating shaping the panning curve for different potential applications. Furthermore, we shall also demonstrate that PTS supersedes RPMS: not only is PTS able to produce the same results as RPMS, but PTS demonstrates greater flexibility and capacity to target specific spectra more easily. More importantly, due to its flexibility, PTS can target any sound spectra independently from the number of secondary sources used, which is a crucially limiting aspect of RPMS.

2 Background

2.1 Vector Base Amplitude Panning

VBAP is a perceptual panning technique and essentially expands traditional stereo panning to entire arrays over a sphere [5]. It is mainly concerned with choosing the *active* loudspeakers (secondary sources) for each virtual source and determines a panning curve based on the distance between the chosen ones. In a ring, no more than 2 secondary sources per virtual source are chosen at each time, while on the sphere this is no more than 3. Using a vector base defined by the secondary sources' positions, VBAP derives the gain factors of each secondary source as the equivalent to the scalars needed to obtain the vector to the virtual source.

2D VBAP formulates the condition, that the gain coefficients g_1 and g_2 of the active secondary sources should always adhere to $g_1^2 + g_2^2 = C$, which can be considered a global gain control [5]. Furthermore, without loss of generality, we can consider $C = 1$. Then the 2D VBAP panning curve in relation to the virtual source position ϕ and the angular distance between two secondary sources in a ring $2\phi_0$ can be described as [28]:

$$\sqrt{\frac{\tan^2\phi_0 + 2\tan\phi_0\tan\phi + \tan^2\phi}{2(\tan^2\phi_0 + \tan^2\phi)}} = g, \qquad (1)$$

if $-\phi_0 < \phi < \phi_0$, else $g = 0$. However, on a sphere, a trigonometric solution is not easily obtained. Here, the panning curve is best described by a rasterization of the gain factors using the above mentioned vector base approach.

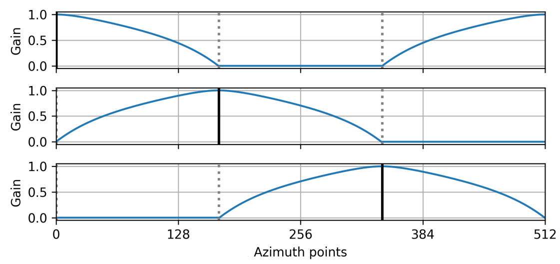


Fig. 1: Three panningtables containing VBAP gain changes across a circular array with three respective, equally spaced secondary sources (solid, vertical lines).

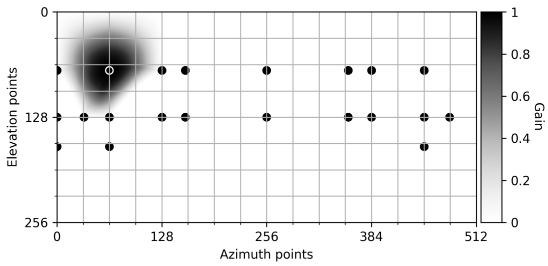


Fig. 2: A spherical panningtable for one secondary source (white circle) in a Hamasaki 22.2 layout. Other secondary sources are depicted as black dots for orientation.

2.2 Rapid Panning Modulation Synthesis

RPMS was designed to pan virtual sources using arbitrary panning curves at angular velocities beyond our limit to audibly perceive rotations [23, 27]. These high velocities produce spectral artifacts that can be used as a means to synthesize new sounds. Moreover, because of the pair-wise panning approach, irrespective of the velocity, the virtual source is reproduced by a limited number of secondary sources at any given moment. As a result, the spectral artifacts are highly decorrelated [29], creating additional spatial effects.

In a ring, RPMS takes two audio signals as input: the source audio and a control signal that determines the virtual source's position. The normalized control signal is scaled to 2π to describe the movement on the ring. A constant circular motion is therefore described by a positive-only, full-scale sawtooth wave. The gain factors are determined by a buffer containing the panning curve, that is read using the control signal. Per secondary source, the control signal is scaled, shifted and truncated within $[0,1]$ to only read from the buffer if the virtual source is approaching and receding from the respective secondary source in relation to it's immediate neighbors [23]. Because one buffer is used for all secondary sources, all secondary sources should ideally be placed regularly across the circle.

On a sphere, a second control signal is used to determine the vertical position. The secondary sources are organized in several rings, that are equally distributed across the sphere. For vertical panning, RPMS simply considers each ring as if it were a secondary source and applies the same panning curve to move the virtual source between rings. This means that virtual sources can be reproduced by up to four secondary sources (two per ring) on a sphere, making it incompatible with VBAP in the spherical case.

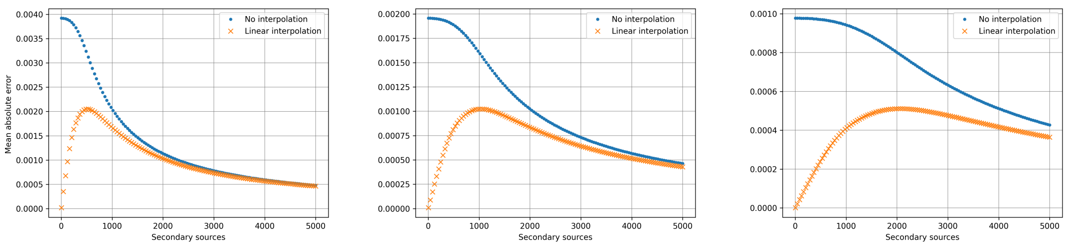
3 Panningtable Synthesis

As mentioned in section 2.2, RPMS suffers from a series of limitations that make it impossible to emulate VBAP properly, particularly across a sphere. Along a circle, adapting RPMS for irregular spaced secondary sources requires individual buffers for each secondary source. However, the asymmetry of the panning curve would complicate the scaling and shifting of the input control signal greatly. Thus, PTS was developed in an effort to simplify the approach taken by RPMS and be fully compatible with common panning techniques, such as VBAP.

3.1 Circular Panningtable Synthesis

Similarly to RPMS, PTS stores the panning curves in buffers, henceforth referred to as *panningtables*. However, one panningtable is allocated per secondary source. Furthermore, instead of recording only the panning curve itself, the gain changes for each secondary source across the entire circle are recorded in each respective panningtable (see Fig. 1), effectively rasterizing (1).

This simplifies mapping the input control signal to the panningtable read position, while also facilitating the use of asymmetric panning curves in the case of irregularly spaced arrays. The control signal is hence scaled to the size of the panningtable, directly reading a gain value for each audio sample of the input sound.



(a) 256 points **(b)** 512 points **(c)** 1024 points

Fig. 3: Changes in mean absolute error produced by PTS when compared to VBAP using a constant full scale input signal, rotated at 1HZ across the full circle over 1s with a 48kHz sampling rate. Figs. (a)–(c) depict the outcome using different panningtable sizes.

3.2 Spherical Panningtable Synthesis

On a sphere, the same rasterization principle is used to record a gain value per azimuth and elevation pair in a 2D panningtable (see Fig. 2). There are many ways to project the surface of a sphere to a flat surface. However, it has been found that the most straightforward way is to simply map each step in azimuth to the rows and in elevation to the columns. This allows for a simple and efficient conversion of the control signals to the panningtable position. In this approach, the top and bottom row of the 2D panningtable is populated with a constant gain value, because this reflects the ineffectiveness of changes in azimuth if the virtual source is panned to one of the two poles.

An up and down motion is simulated by scaling a triangular input control signal for elevation at half scale. In order to simulate a revolution using the elevation control signal, a modulated shift has to be applied to the azimuth control signal to invert the phase. Thus, as the virtual source is panned beyond the poles, the azimuth control signal is shifted by 180° (half a panningtable) and the elevation signal needs to be inverted so that the virtual source returns to the equator, on the other side of the sphere.

4 Emulating VBAP using PTS

Rendering the VBAP panning curve (1) using RPMS has been demonstrated in [28]. In particular, it was shown how the table needs to be modified according to (1) in relation to the secondary sources' distance $2\phi_0$ and how this affects the spectral distortions produced. Here, we show how the size of the panningtable and the number of secondary sources affects both the accuracy and efficiency, when comparing PTS to VBAP.

4.1 Accuracy

Using 2D PTS with a regular secondary source distribution, it can be demonstrated that the accuracy depends mainly on two factors: the size of the panningtable and the density of the secondary sources. The first factor is relatively trivial; the more points the panningtable contains, the finer the resolution of the rendered panning curve. The average error can be halved using larger panningtable sizes of $p = 2^n$ for each power n. While an acceptable accuracy can be achieved with a sufficiently large panningtable, it is also possible to improve the accuracy using an interpolation method. Using 3 secondary sources in a ring, linear interpolation can reduce the error against VBAP by roughly $1.12e^{-0.67p}$ compared to using no interpolation, given the number of points p in a panningtable.

The relationship between the accuracy of the rendered panning curve and the number of secondary sources in a ring is a bit more complex. Fig. 3 shows the mean absolute error PTS produces across all secondary sources when compared to VBAP while rendering a 1s long output of a constant signal at full scale being rotated at 1Hz and 48kHz sampling rate. Figs. 3a-3c demonstrate the error rate as the number of secondary sources is increased for 256, 512 and 1024 point panningtable sizes respectively, using both no and linear interpolation.

First, it should be pointed out that the overall accuracy increases with increasing panningtable sizes. Secondly, using no interpolation, increasing in the number of secondary sources decreases the error produced in general, but the effect is somewhat "delayed" for larger panningtable sizes. Having more secondary sources in a circle means that each secondary source effectively remains silent for a greater angular range of the virtual source's movement. This silence mostly exhibits a perfect overlap with the VBAP ground truth, minimizing the error on average.

Using linear interpolation the accuracy of the rendered output compared to VBAP first decreases as more secondary sources are added. The final portion of the gain curve (1) creates a discontinuity as the gain change is truncated at 0. Adding more secondary sources, effectively reducing the active angular area in each individually, steepens the slope at this truncated point of the VBAP gain curve, which increases the difference to the interpolated gain slope. Once the active angular area is sufficiently small due to so many secondary sources populating the circle, the difference between the linear interpolation and no interpolation becomes negligible as the two curves approximate each other.

4.2 Run Time Analysis

To compare the complexity, each approach was broken into its constant time operations and then quantified. The original VBAP implementation written for PureData in the C programming language was chosen as a benchmark for this comparison.[1] It is also assumed that the virtual source changes it's position at audio rate, meaning at every audio sample a new set of gain factors needs to be calculated. Also, since a table access operation belongs to $O(1)$, the complexity does not rise with increasing panningtable sizes for PTS, thus limiting this comparison against the number of secondary sources.

For circular VBAP, given the number of tuples t, a total of $13+8t$ additions, $23+8t$ multiplications, 7 divisions, 3 square roots, 1 modulo and 7 trigonometric operations are necessary. Considering that the number of non-overlapping tuples on a circle is equal to the number secondary sources themselves, this results in $54+16s$ operations for the number of secondary sources $s>2$.

For spherical VBAP there are $16+16t$ additions, $27+16t$ multiplications, 7 divisions, 3 square roots, 2 modulo and 7 trigonometric operations. The number of faces t on a triangulated hull is given by $t=2s-4$ using Eulers formula $s-e+t=2$ and stating that $2e=3t$, meaning that each face has 3 edges, while each edge of an outer hull triangulation belongs to exactly 2 faces. Thus, the total number of operations per sample yields $-82+72s$ for the number of secondary sources $s>3$.

Conversely, for both circular and spherical PTS, only $6+s$ additions, $7+s$ multiplications, 1 division and 2 modulo operations are necessary, amounting to $16+2s$ operations for the number of secondary sources s. Comparing the increase in complexity against VBAP, the ratios of the determined inclination factors amount to 5.3 and 24 for the circular and spherical case respectively.

The analysis shows that both algorithms belong to the class $O(n)$, with PTS having an initial advantage and being more efficient for increasing number of secondary sources. Because VBAP uses no more than a limited number of active secondary sources for any given panning location, these indices can be precomputed and stored. This can remove PTS' dependency on the number of secondary sources, moving it into the class $O(1)$. However, in favor of maintaining the flexibility of PTS to accommodate any panningtable, the run time analysis against VBAP is done using the $O(n)$ version only.

To verify these findings, we ran each implementation on a Macbook Pro with a 3.1 Ghz Intel Core i7 processor and 16GB of DDR3 RAM. We carefully removed all dependencies to the PureData framework and converted the code to run as a script, keeping the core processes of the implementation as untouched as possible.[2] The processing time was measured using the builtin `time.h` library. The time measurement was inserted to only measure the time necessary to compute the gain factors inside the audio thread. The code was compiled using Apple clang version 14.0.0 using no optimizations. We measured the average time over 1000 trials, where each trial consisted of calculating a 1s output at 48kHz sampling rate, resulting in $48\cdot10^6$ measurements per number of secondary sources.

Fig. 4 shows the average computation time it took to retrieve the gain factors for one sample of audio as a function of the number of secondary sources contained in a circle and on the sphere respectively. It can be seen that the computational effort increases linearly with an increasing number of secondary sources. While the difference between VBAP and PTS is relatively small for small secondary source arrays, Fig. 4 shows how the

[1] `https://github.com/pd-externals/vbap`

[2] `https://github.com/multimedia-eurecat/pts`

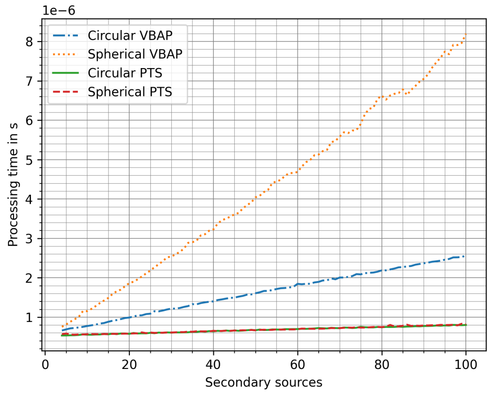


Fig. 4: Processing time necessary to compute gain factors for both VBAP and PTS.

approach taken by PTS outperforms the original VBAP implementation as the number of secondary sources increases. Furthermore, it confirms that there is no difference between running PTS on a circular or spherical layout. The approximate inclinations using linear regression measured in each case were $2.8 \cdot 10^{-9}$ for both versions of PTS, $1.97 \cdot 10^{-8}$ for circular VBAP and $7.29 \cdot 10^{-8}$ for spherical VBAP. The ratios between the inclinations of circular and spherical VBAP against PTS, 7 and 26 respectively, are similar to the theoretically determined ones, confirming the findings.

5 Comparing PTS to RPMS

As demonstrated above, PTS records the gain changes of any method over the entire circle or sphere in a panningtable for each secondary source respectively. Therefore, it is relatively trivial to see that any panning curve used with RPMS can be emulated with PTS as well. Conversely, trying to render any panningtable used in PTS with RPMS is not possible for several reasons. First, RPMS uses a single panning curve for all secondary sources, limiting its flexibility to adapt to complex situations. For example, as the distance between secondary sources varies, the shape of the VBAP gain curve, changes non-linearly [28]. Consequently, according to how RPMS is defined [23, 28], modifying a single "master" panning curve for each secondary source in a non-regularly distributed array using linear operations is not possible. Second, the RPMS "master" panning curve is only ever applied to the nearest secondary sources, meaning that a maximum of 2 secondary sources on a circle (or 4 secondary sources on a sphere) sound at any given moment. Conversely, PTS stores the gain coefficient for each secondary source across the entire circle or sphere, allowing any number of secondary sources in the array to sound at any time. This also means, that PTS, unlike RPMS, may also emulate other panning techniques, like Multiple-Direction Amplitude Panning (MDAP) [10].

PTS also supersedes RPMS with regard to the spectra that are theoretically possible, precisely because RPMS can only apply a panning curve to the virtual source's nearest secondary sources. Again, because PTS can rasterize and store any gain change produced by RPMS, PTS is able to emulate any RPMS panning behavior and, as a consequence, produce the same spectra. However, the opposite is not always true. Using a sine panning curve for RPMS [28], a virtual source emitting a single frequency, traveling in a circle of 4 regularly spaced secondary sources will produce a specific distortion pattern, due each secondary source being silent for half the time that the virtual source requires to traverse the circle. This is reflected in the first spectrum shown in Fig. 5. Using PTS we can populate the panningtable with a second copy of the sine panning curve in the region where the respective secondary source would otherwise be silent. By controlling the amplitude of this copy, we can transition between the distortion pattern and a spectral result akin to conventional amplitude panning (see Fig. 5).

Moreover, because RPMS applies the panning curve to only the nearest secondary sources, adding more secondary sources would further "squish" the sine panning curve, forcibly producing increased spectral distortion. Instead, PTS can maintain the shape of the panning curve, because it records the respective gain changes over the entire circle or sphere irrespective of the number of secondary sources used. Thus, PTS can give us greater control over the spectra created and their consistency across any reproduction array.

6 Conclusions

This paper introduced a new, flexible approach to spatial audio panning called PTS, that works using pannintables as pre-computed gain-coefficient buffers. We compared PTS against the original VBAP implementation written in C for PureData in both accuracy and performance. Testing our approach against all other implementations and interpretations of VBAP is out of the scope of this paper and must remain as part of

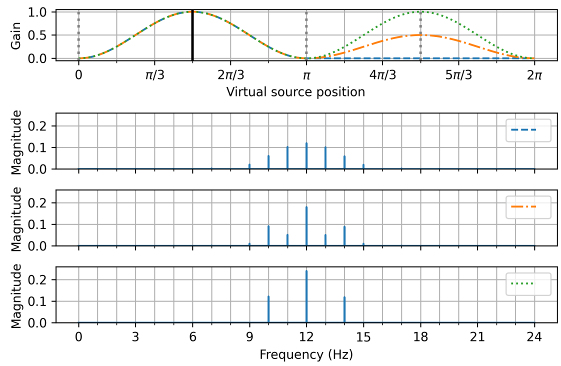


Fig. 5: Three different panning curves for the secondary source at $\pi/2$ and the resulting spectra for a virtual source emitting a 12kHz sine tone, rotating at 1kHz.

future work for now. Moreover, as this paper shows, PTS itself is a more efficient implementation compared to VBAP, with a minor trade-off to accuracy that can be mitigated with increased memory space. Most importantly, this confirms that PTS can be regarded as a valid spatial audio panning technique.

However, PTS is more than that. While VBAP would result in only those panning curves that aim for accuracte panning characteristics (unless modified in its mathematical formulation) PTS allows to directly modify the panningtable in a very intuitive way. Because PTS is controlled specifically using additional audio control signals, a new position is calculated on each audio sample, enabling virtual sources to be panned up to the Nyquist frequency. Reading a panningtable in such a way is akin to wavetable synthesis [24], turning each secondary source into an independent synthesizer. Using the VBAP characteristic gain changes produces a specific spectrum in each secondary source. But these curves can be modified individually to deviate from this approach in search of other, specific spectra or unorthodox spatial distributions. Additionally, if a trade-off between accurate panning and wavetable synthesis is struck, maintaining a degree of decorrelation between the secondary sources, a sense of spaciousness becomes inherent to the sound synthesis produced.

In the spherical case, a 2D panningtable can also be understood as a *wave terrain* in wave terrain synthesis [24]. The trajectory a virtual source takes is then the same as the *orbit* with which the wave terrain is scanned. In contrast to previous approaches to spatialise wave terrain synthesis [17, 25, 26], where a single spectrum is generated and then spatialised using the orbit as a basis for the trajectory, the concept is somewhat inverted and the trajectory is now the basis for the orbit, which describes the scan taken over each panningtable for each secondary source. Future work will include investigating the capabilities of such an approach.

References

[1] Gerzon, M. A., “Periphony: With-height sound reproduction,” *Journal of the Audio Engineering Society*, 21(1), pp. 2–10, 1973.

[2] Berkhout, A. J., de Vries, D., and Vogel, P., “Acoustic control by wave field synthesis,” *The Journal of the Acoustical Society of America*, 93(5), pp. 2764–2778, 1993.

[3] Daniel, J., *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*, Ph.D. thesis, University of Paris VI, France, 2000.

[4] Blumlein, A., “Improvements in and relating to sound-transmission, sound-recording and sound-reproducing systems,” 1931.

[5] Pulkki, V., “Virtual sound source positioning using vector base amplitude panning,” *Journal of the audio engineering society*, 45(6), pp. 456–466, 1997.

[6] Chowning, J. M., “The simulation of moving sound sources,” *Journal of the audio engineering society*, 19(1), pp. 2–6, 1971.

[7] Lossius, T., Baltazar, P., and de la Hogue, T., “DBAP–distance-based amplitude panning,” in *Proceedings of the 35th International Computer Music Conference*, Montreal, Quebec, Canada, 2009.

[8] Zotter, F. and Frank, M., “All-round ambisonic panning and decoding,” *Journal of the audio engineering society*, 60(10), pp. 807–820, 2012.

[9] Lopez, J. J., Gutierrez, P., Cobos, M., and Aguilera, E., “Sound distance perception comparison

between wave field synthesis and vector base amplitude panning," in *2014 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, pp. 165–168, IEEE, 2014.

[10] Pulkki, V., "Uniform spreading of amplitude panned virtual sources," in *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. WASPAA'99*, pp. 187–190, IEEE, 1999.

[11] Potard, G. and Burnett, I., "Decorrelation techniques for the rendering of apparent sound source width in 3D audio displays," in *Proceedings of the International Conference on Digital Audio Effects (DAFx'04)*, Naples, Italy, 2004.

[12] Frank, M., "Source width of frontal phantom sources: Perception, measurement, and modeling," *Archives of Acoustics*, 38(3), pp. 311–319, 2013.

[13] Torick, E., "Highlights in the history of multichannel sound," *Journal of the Audio Engineering Society*, 46(1/2), pp. 27–31, 1998.

[14] Kim-Boyle, D., "Sound spatialization with particle systems," in *Proceedings of the 8th International Conference on Digital Audio Effects (DAFX-05)*, pp. 65–68, Madrid, Spain, 2005.

[15] Wilson, S., "Spatial swarm granulation," in *Proceedings of the International Computer Music Conference (ICMC)*, pp. 1–4, Belfast, Northern Ireland, 2008.

[16] Bates, E., *The Composition and Performance of Spatial Music*, Ph.D. thesis, Trinity College Dublin, Ireland, 2009.

[17] James, S. and Hope, C., "2D and 3D Timbral Spatialisation: Spatial Motion, Immersiveness, and Notions Of space," in *Proceedings of the International Computer Music Conference (ICMC)*, pp. 77–84, Perth, Australia, 2013.

[18] Kim-Boyle, D., "Spectral spatialization – an Overview," in *Proceedings of the International Computer Music Conference (ICMC)*, pp. 1–7, Belfast, Northern Ireland, 2008.

[19] Torchia, R. H. and Lippe, C., "Techniques for multi-channel real-time spatial distribution using frequency-domain processing," in *Proceedings of the 2004 conference on New Interfaces for Musical Expression (NIME)*, pp. 116–119, Shizuoka University of Art and Culture, Hamamatsu, Japan, 2004.

[20] Schmele, T. and Garriga, A., "Sound spectrum modulation generated by circularly moving sound sources," in *Audio Engineering Society Convention 154*, Audio Engineering Society, Espoo, Finland, 2023.

[21] McGee, R., "Spatial modulation synthesis," in *Proceedings of the International Computer Music Conference (ICMC)*, pp. 246—-249, Denton, TX, USA, 2015.

[22] Chowning, J. M., "The synthesis of complex audio spectra by means of frequency modulation," *Journal of the audio engineering society*, 21(7), pp. 526–534, 1973.

[23] Schmele, T., *Exploring 3D audio as a new musical language*, Master's thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2011.

[24] Roads, C., *The computer music tutorial*, MIT press, Cambridge, MA, USA, 1996.

[25] Mills, A. and de Souza, R. C., "Gestural sounds by means of wave terrain synthesis," in *Proc. of the VI Brazilian Symp. on Computer Music*, volume 3, pp. 9–16, 1999.

[26] James, S., "Spectromorphology and Spatiomorphology of Sound Shapes: audio-rate AEP and DBAP panning of spectra," in *Proceedings of the 41st International Computer Music Conference*, pp. 278–285, Denton, TX, USA, 2015.

[27] Féron, F.-X., Frissen, I., Boissinot, J., and Guastavino, C., "Upper limits of auditory rotational motion perception," *The Journal of the Acoustical Society of America*, 128(6), pp. 3703–3714, 2010.

[28] Schmele, T. and Lopez, J. J., "Comparisons between VBAP and WFS using Spatial Sound Synthesis," in *Audio Engineering Society Convention 153*, Audio Engineering Society, 2022.

[29] Kendall, G. S., "The decorrelation of audio signals and its impact on spatial imagery," *Computer Music Journal*, 19(4), pp. 71–87, 1995.