Audio Engineering Society

Convention Paper 10622

Presented at the 153rd Convention
2022 October

Comparisons between VBAP and WFS using Spatial Sound Synthesis

Timothy Schmele[1,2] and Jose J. Lopez[3]

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

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

ABSTRACT

This paper presents analysis methods that make use of synthesis concepts by means of movements of sounds in virtual 3D auditory space at audio rate. The main focuses lies on the modulating effect caused by moving a sound around a listener, commonly referred to as *panning*. Here, we will demonstrate a comparison between two common panning approaches: Vector-Base Amplitude Panning (VBAP) and Wavefield Synthesis (WFS). Both apply some type of secondary source selection process, reducing the number of active loudspeakers at any moment in time. Under specific conditions it therefore becomes possible to compare VBAP and WFS directly. Using a technique originally developed for 3D sound synthesis over multichannel loudspeaker arrays at audio rate, both VBAP and WFS are adapted to be emulated by this technique. By using this sonification process, the differences between panning with VBAP or WFS can then not only be visualised but also heard. This makes spatial sound synthesis not just a creative tool for sound creation, but also an analytic tool that can help illustrate characteristics of spatialisation approaches.

1 Introduction

Most discussions related to spatial sound center around the correct reconstruction of a sound scene or, in the case of post-production, the correct placement of an audio stream as a *phantom source* via cartesian or spherical parameters in 3D virtual audio space. For as long as sound is thought of as an object in space, represented by a single audio channel with metadata, placing sound points in auditory 3D space will remain the basic mode of thinking in virtual spatial audio. Even though Ambisonics [1] is capable of offering an alternative to the aforementioned approach by utilizing a spherical harmonic decomposition, two other common approaches to spatial audio that rely entirely on post-production techniques to produce a spatial sound scene are Vector-Base Amplitude Panning (VBAP) [2] and Wavefield Synthesis (WFS) [3, 4].

1.1 Vector-base Amplitude Panning

VBAP is generally considered a psycho-acoustic panning method and is mainly concerned with choosing the minimum amount of *active loudspeakers* necessary for each sound source [2]. In two dimensions, this is no more than 2 per source. Between those 2 loudspeakers, VBAP then uses the vectors to each active loudspeaker to derive the respective gain factors. The vectors are used as a vector-base to reproduce the phantom sound source by regarding the scalars as equivalent to the required gain factors for psycho-acoustically effective reproduction. Variants of VBAP also exist, such as distance based amplitude panning (DBAP), where the gain coefficients are determined as a function of distance to each loudspeaker instead [5].

1.2 Wavefield Synthesis

On the other hand, WFS is a physical approach to reconstructing the sound field by providing a solution the Kirchhoff-Helmholtz integral [3, 4]. In theory, all secondary sources along the boundary of the listening area contribute to this reconstruction by means of the Huygens principle. However, the accuracy and perceived spatial audio quality is somewhat diminished due to several theoretical and practical constraints of WFS systems, such as spatial aliasing [6, 7]. In terms of perception, one can improve the effective listening space for general, non-focused sources by applying an active secondary source selection criterion [8, 7]. As such, WFS also aims to reduce the amount of active loudspeakers and, under certain conditions, both WFS and VBAP thus become directly comparable.

1.3 Spatial Sound Synthesis

Spatial sound synthesis refers to synthesis approaches that intertwine the creation of sound spectra and sound spatialisation. In most cases, spatial sound synthesis allows a more holistic view of spatial audio instead of point sources in space; a spatially cohesive sound can have shape [9, 10, 11] and texture [12] in space. While most approaches apply spatialisation to individual components at an early stage in the sound synthesis process, some aim to create sound synthesis by means of the spatialisation itself.

In the former category we can find nearly all common synthesis methods, both temporal or spectral [13]. Spatial granular synthesis, for example, treats each generated grain as an individual sound source and applies spatial parameters to each [14, 15, 16]. Spectral methods are those that usually split the frequency content of a source into several frequency bands and spatialise each band individually [17, 18, 19].

There are other techniques, though, that apply spatialisation first and achieve newly synthesized spectra as a consequence. This is often achieved by means of accelerating the spatial movement of the phantom sources to be able to modulate their spatial position at audio rate. For example, rapidly modulating a phantom source in it's distance causes a constant change in the Doppler shift, which results in a synthesised sound spectrum comparable to FM synthesis [20]. Audio rate panning approaches have also been described for pair-wise panning [21], similar to 2D VBAP, as well as DBAP [11] and successfully applied in documented musical compositions [22, 23].

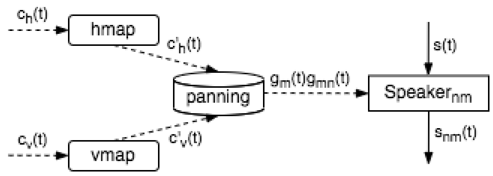

Fig. 1: *The specific schema of the audio path for a loudspeaker in RPMS [21].*

1.4 Rapid Panning Modulation Synthesis

One method for panning audio sources that originally uses equally spaced virtual loudspeakers as an abstraction layer is called *Rapid Panning Modulation Synthesis* (RPMS) [21, 22]. The main goal of RPMS is to modulate sounds and create rich overtones merely by using spatialization algorithms. In order to achieve this, the azimuth and elevation parameters of a source cannot be controlled via regular *messages* at control-rate. Instead, a respective audio-rate control signal for each azimuth and elevation are used.

Its main workings rely on a pair-wise panning algorithm, based on a panning curve stored in a table. In order to achieve both horizontal and vertical panning, a loudspeaker setup for this approach is required to be designed in m regularly spaced rings, with each ring having n_m regularly spaced loudspeakers. For each spherical dimension, a respective input audio signal c_h and c_v is mapped to retrieve an interpolated gain value from the panning curve table (see Fig. 1). Since each loudspeaker has a unique position in the array, each specific mapping operation needs to be shifted to match the position of the respective loudspeaker and ring.

The result of this operation is a modulation effect akin to amplitude modulation (AM). In the example shown in Fig. 2, a $12kHz$ input sound is panned horizontally

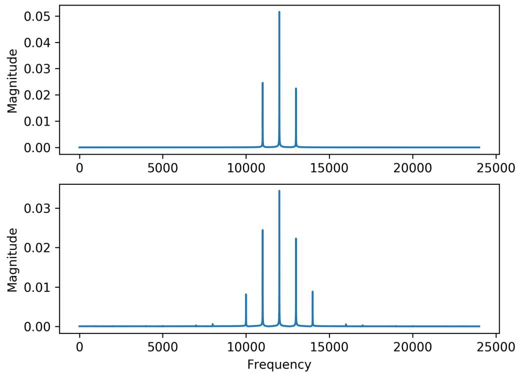

Fig. 2: *Spectral comparison of 2 (above) vs 3 (below) channel RPMS with a* $12kHz$ *input signal being rotated at a angular velocity of* $1kHz$ *using a sine curve panning function.*

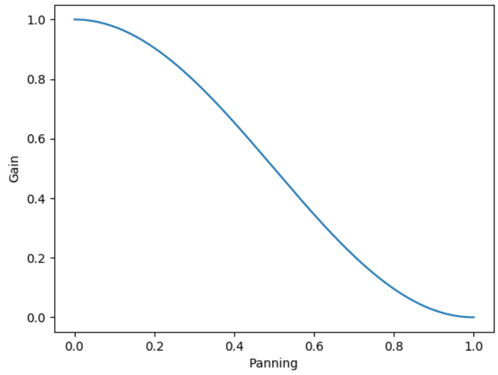

Fig. 3: *A panning function derived from a sine curve.*

over a simple 2 or 3 channel loudspeaker array. The control signal c_h used is a sawtooth waveform tuned to $1kHz$. The sawtooth signal is used to achieve a linear increment in angular distance traveled horizontally. Thus, the input signal is effectively rotated through the loudspeaker system 1000 times per second. Note that the spectrum shown in Fig. 2 is from a single loudspeaker only. However, because all loudspeakers are regularly spaced around the circle, the spectra in other loudspeakers are merely a phase shifted copy of the first. Therefore, it is sufficient to look at the spectrum of a single loudspeaker to understand the distortions produced.

The signal c_h is ultimately used as an index to determine the correct read position from the panning function buffer. The panning function used is shown in Fig. 3 and is derived from a sine curve. This means that the fade in and fade out of the $12kHz$ input within each loudspeaker is smooth and the harmonic distortion produced in Fig. 2 is minimal, particularly in the 2 channel case. Therefore, in the specific example of 2 channel RPMS shown in Fig. 2, the result is equivalent to traditional AM synthesis.

However, similar to using a complex modulator with AM synthesis, one can populate the panning table with a different panning curve and further distort the modulation of the input signal, resulting in a more complex and rich set of overtones. Moreover, RPMS is built with the intention to use more than just two secondary sources. Applying pair-wise panning to a signal through a circle of 3 loudspeakers, for example, will always result in silence in at least one loudspeaker at any given time, while the other two work to reproduce the phantom source at the desired location. This adds temporal deviations from the otherwise harmonic panning function and leads to further harmonic distortions in the resulting spectrum. The lower graph in Fig. 2 illustrates this increase in spectral richness using the $12kHz$ input tone panned by a $1kHz$ sawtooth control signal using the same sine curve panning function (Fig. 3) over a three-channel loudspeaker array.

Even though the remainder of this paper will focus on the 2D case, i.e. panning sounds in a circle in the horizontal plane, it should be noted that the vertical panning function through the circles of the loudspeaker array works similarly to the aforementioned horizontal technique. The original implementation of RPMS [21] simply regards each entire circle of loudspeakers as if it were a single virtual loudspeaker in its own right, and pans the audio vertically, up and down, by pairwise panning between each pair of adjacent circles. The interested reader may be referred to the references provided in Schmele [21], and Schmele and Gómez [22].

2 VBAP Equivalent Panning using RPMS

The flexibility of RPMS allows the use of any kind of panning curve, depending on the desired goal. Some panning curves may achieve a specific target spectrum (e.g. for creative, musical purposes) while others allow for greater panning accuracy. It is therefore possible to come close to the panning accuracy of 2D VBAP using 2D RPMS given the condition that a ring of n regularly spaced loudspeakers is used.

For 2D VBAP the panning curve is defined by two conditions. The first being the tangent law[2]:

$$\frac{\tan\phi}{\tan\phi_0} = \frac{g_1 - g_2}{g_1 + g_2}, \tag{1}$$

for phantom source angle ϕ, loudspeaker angle ϕ_0 and gain coefficients g_1 and g_2 for loudspeakers 1 and 2 at $\pm\phi_0$. In the appendix of Pulkki [2], it is shown that the vector base formulation:

$$\boldsymbol{g} = \boldsymbol{p}^T \boldsymbol{L}_{12}^{-1}, \tag{2}$$

for gain vector $\boldsymbol{g}$, phantom source position vector $\boldsymbol{p}$ and active loudspeaker position matrix $\boldsymbol{L}_{12}^{-1} = \left(\begin{smallmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{smallmatrix}\right)$ for loudspeakers 1 and 2 satisfy Eq. 1. In other words, the vector formulation is equivalent to the tangent law of Eq. 1 and we can use the trigonometric function to derive the panning curve from here on.

Furthermore, VBAP also defines a second condition, whereby the power output using both gain coefficients g_1 and g_2 should always equal some constant C that can be regarded as a global gain control [2]:

$$g_1^2 + g_2^2 = C. \tag{3}$$

Without loss of generality, we can consider $C = 1$. We can then proceed use Eq. 1 and Eq. 3 to receive the function for the VBAP panning curve (see Appendix A):

$$\sqrt{\frac{\tan^2\phi_0 + 2\tan\phi_0\tan\phi + \tan^2\phi}{2(\tan^2\phi_0 + \tan^2\phi)}} = g_1. \tag{4}$$

The effect of ϕ_0 on the shape of the VBAP panning curve is illustrated in Fig. 4. The graphs show the gain factor for a loudspeaker located at 0° as the phantom source would be panned away from (or towards) it. Each curve is drawn in relation to neighbouring loudspeakers located at angular distances from 10° to 170°, in steps of 10°. The graphs somewhat illustrate the relative complexity of the panning curve in relation to different neighboring loudspeakers at different distances.

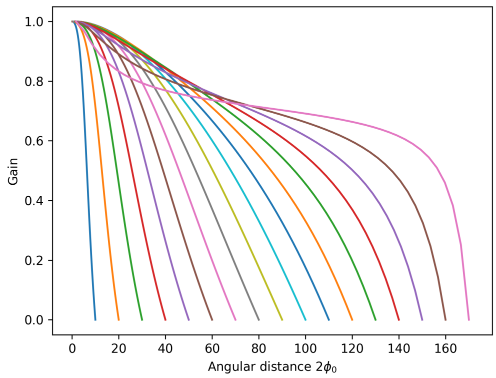

Fig. 4: *Illustration of the VBAP panning curve for different values of angular distance $2\phi_0$ to the next loudspeaker in steps of* 10°.

With respect to 2D RPMS, this angle ϕ_0 is directly correlated to the number of loudspeakers n used in a ring, given by the equation:

$$\pi/n = \phi_0. \tag{5}$$

Thus, the possible angles for ϕ_0 that can be used as a distance $2\phi_0$ between two neighbouring loudspeakers for RPMS changes in discrete steps, as n increases:

n	3	4	5	6	7	8	...
ϕ_0	60°	45°	36°	30°	25.71°	22.5°	...

As a consequence, we can produce characteristic spectra that can represent, by means of sonification, how the VBAP panning curve behaves in these particular instances. The spectra produced for values of $n \in [3, 8]$ using a $12kHz$ input source panned by a $1kHz$ sawtooth wave and cubic interpolation of the panning buffer are shown in Fig. 5.

Two particular aspects can be observed in the spectra of Fig. 5: first, an increase in magnitude of harmonics located further away from the fundamental with increasing number of loudspeakers in the circle n. This can be attributed to a longer period of silence within a single loudspeaker, since it's active range is reduced in angular width as more loudspeakers are present in the circle. Considering a constant angular velocity of the phantom source, the time that the phantom source is reproduced by any single loudspeaker is effectively

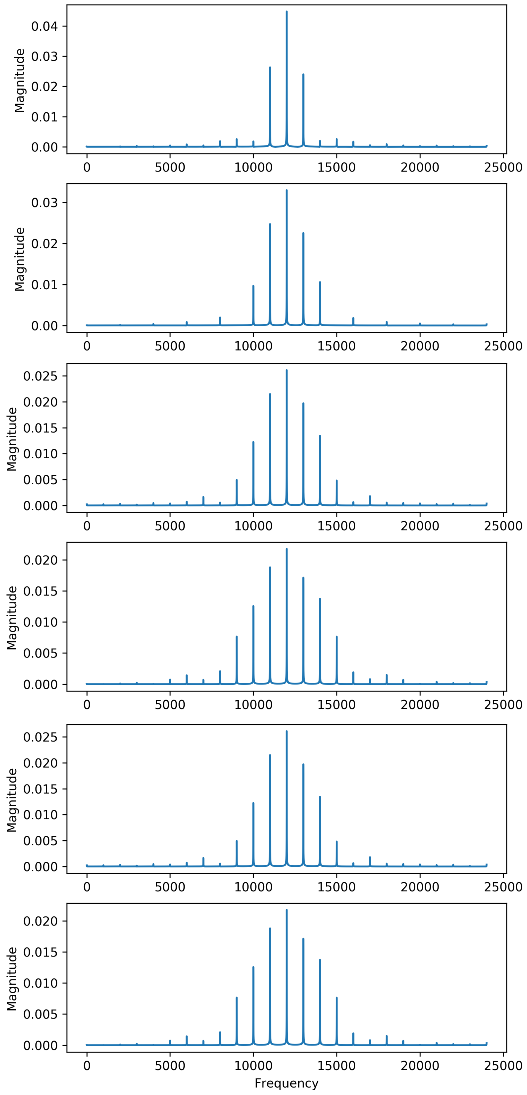

Fig. 5: *Spectra produced by RPMS using the VBAP panning curves for rings of loudspeaker arrays from $n = 3$ (top) to $n = 8$ (bottom).*

reduced, distorting the panning curve (gain envelope) further. Secondly, the changing shape of the VBAP panning function for different ϕ_0 is reflected in the varying relative gain distribution within the harmonics, contributing to differing timbres in each case. This characteristic is more pronounced for values $n > 7$, where the change in contortion of the gain envelope is less for increasing values of n.

3 Comparisons between VBAP and WFS

The theory of WFS is based on the assumption that the sound field generated by sources outside of a volume V can be described by a set of secondary monopole sources located along the surface S of V. The general solution to this problem can be formulated using the Kirchoff-Helmholtz integral [3, 4]. In practice, these secondary sources are represented by loudspeakers. Therefore, the volume V becomes the space into which the loudspeakers project into and the loudspeaker perimeter defines the surface S. This means that the primary objective in WFS is to find the *driving functions*, which translate the phantom source signal into adequate driving signals for each loudspeaker n in the array. A simple and compact form for this function is given by the following expression:

$$s_n(t) = C(f)\frac{\cos\theta_n}{|\vec{r}_n|}e^{-jwr_n}s(t), \tag{6}$$

where, for each loudspeaker n, the phantom source signal $s(t)$ is both modified in amplitude by the geometric relation between the distance $|\vec{r}_n|$ of the phantom source to the loudspeaker n, the angle θ_n between $\vec{r}_n$ and the loudspeaker normal, as well as in phase by $e^{-j\omega r_n}$, while $C(f)$ is a frequency dependent constant.

From Eq. 6, it can be deduced that a phantom source close to the surface S, as defined by the loudspeaker array, approaches a singularity: as the phantom source comes close to a loudspeaker n, the distance between them, $|\vec{r}_n|$, tends towards 0 and the amplitude of the driving signal $s_n(t)$ becomes infinitely big. On the other hand, the angle θ_n between $\vec{r}_n$ and the normal of loudspeaker n will be 90° for a phantom source situated on top of the surface S. This creates an unwanted paradox where a phantom source directly panned onto the loudspeaker array would become inaudible.

The solution to this issue is to give a threshold, beyond which $|\vec{r}_n|$ cannot change anymore, as well as substitute the term θ_n with a term for the gain coefficient g_n. It has been shown, that this solution leads to conventional amplitude panning [7]. Given the regular distance Δx between each secondary source, one can define the immediate vicinity of a loudspeaker by a circle around it with radius Δx. Therefore, a phantom source that approaches the vicinity of two loudspeakers should only affect the gain coefficients of these two loudspeakers.

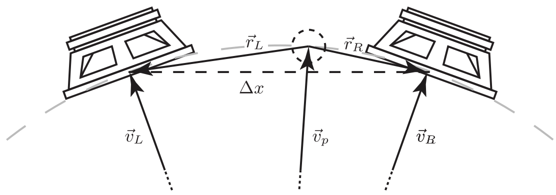

Fig. 6: *Depiction of the linear combination between the vectors $\vec{v}_p$, $\vec{v}_L$ and $\vec{v}_R$, and $\vec{r}_L$ and $\vec{r}_R$. Note that the vectors $\vec{v}_p$, $\vec{v}_L$ and $\vec{v}_R$ emerge from the circular array's center.*

This is expressed by the following equation [7]:

$$g_L = \frac{|\vec{r}_R|}{|\vec{r}_L| + |\vec{r}_R|}, \quad g_R = \frac{|\vec{r}_L|}{|\vec{r}_L| + |\vec{r}_R|}. \tag{7}$$

If the phantom source is directly on the line between two loudspeakers, Eq. 7 becomes:

$$g_L = \frac{|\vec{r}_R|}{\Delta x}, \quad g_R = \frac{|\vec{r}_L|}{\Delta x}. \tag{8}$$

Thus, we can now use these WFS panning functions to drive the RPMS algorithm. For this, we assume a circular 2D volume V. This means however that Eq. 8 does not exactly apply, as the phantom source will travel between loudspeakers along an arch and not on a straight line (see Fig. 6). Given the vectors to the phantom source position $\vec{v}_p$ and loudspeakers $\vec{v}_L$ and $\vec{v}_R$, the vectors $\vec{r}_L$ and $\vec{r}_R$ can be determined by:

$$\vec{r}_L = \vec{v}_L - \vec{v}_p, \quad \vec{r}_R = \vec{v}_R - \vec{v}_p, \tag{9}$$

where $|\vec{v}_p| = |\vec{v}_R| = |\vec{v}_L|$ is equal to the radius of the circular loudspeaker array. However, the ratio between $|\vec{r}_R|$ and $|\vec{r}_R| + |\vec{r}_L|$ is not affected with different array sizes, meaning that the radius of the array ultimately does not affect the shape of the panning curve.

For exemplary purposes, we shall consider the 72 loudspeaker array setup with $2m$ diameter that is present at the Universitat Politècnica de València (see Fig. 7). The angular distance between each loudspeaker is therefore $2\phi_0 = 5°$ in this specific case.

The resulting panning curves for such a loudspeaker array are plotted in Figure 8. It is immediately clear that the pair-wise panning Eq. 7 for WFS results in a nearly linear curve, while VBAP maintains some kind of curvature.

In Fig. 9 the different spectra produced by RPMS for the each the VBAP and WFS panning curve in this specific 72 loudspeaker array is shown. Additionally, the spectrum of the simple difference signal between the RPMS results using both VBAP and WFS is plotted in the bottom plot of Fig. 9.

In terms of spectral result of the RPMS method, one can appreciate a slightly larger main lobe for the much harsher WFS panning curve, versus stronger side lobes for the panning curve produced by VBAP. Looking at the magnitude of the simple difference signal between the two, it is clear that most differences lie in these side lobes. Particularly the distribution of the harmonic weights alternates between the two methods. Since using the WFS panning curve with RPMS produces a more spread out, but ultimately weaker main lobe, we can also appreciate the amplitude differences in the harmonics of the main lobe between the two approaches.

4 Conclusions

While VBAP and WFS appear to tackle the issue of sound spatialisation over a multi-channel loudspeaker array from two completely different approaches, they can, under specific circumstances, be directly comparable to one another. Comparisons between VBAP and WFS are usually limited to perceptual studies between the two [24, 25], or as a way to combine them in order to achieve vertical panning in WFS systems [26, 27].

Here, it was shown that in the case of a phantom source being panned along a regularly spaced 2D loudspeaker ring, we can extract the pair-wise panning curves applied in both VBAP and WFS respectively. For VBAP,

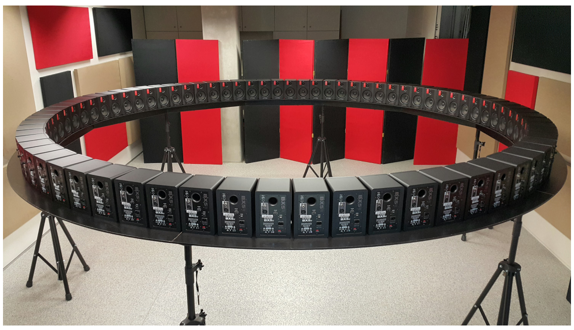

Fig. 7: *The circular WFS array of 72 loudspeakers at the Universitat Politècnica de València.*

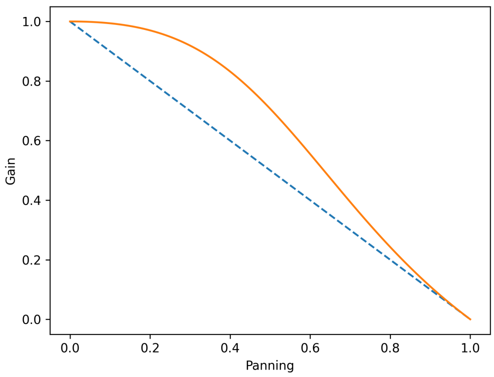

Fig. 8: *Panning curves for both VBAP (solid) and WFS (dashed) for a 72 loudspeaker array.*

this is relatively straightforward by obtaining the panning curve as a function of the angular distance between two loudspeakers $2\phi_0$. The equation for WFS, on the other hand, approaches a singularity if the phantom source is panned close to the line defined by the loudspeaker array. In this case, a limit needs to be defined to prevent the amplitude of the phantom source from exploding, or, paradoxically, disappearing at the same time.

Using a technique for spatial sound synthesis by pairwise panning a phantom source at audio rate around the listener (RPMS), it was shown that these two techniques, VBAP and WFS, can be directly compared by sonifying the specific artifacts caused by their respective panning functions. Originally, RPMS was intended as a tool for creative 3D sound synthesis. However, through the work done here, it became clear that RPMS can also be considered a tool for analysis through sonification.

However, it must also be stressed that the work presented here should be considered exploratory and the techniques used for analysis are still limited. Further work will need to be done to refine the analysis methods, e.g. using different values for the source signal $s(t)$, as well as different control signal types for c_h, for a better, more targeted analysis. Particularly the case of expanding this analysis technique to 3D could reveal interesting results.

Considering this, this work also opens up the question if RPMS is capable of fully emulating VBAP in it's

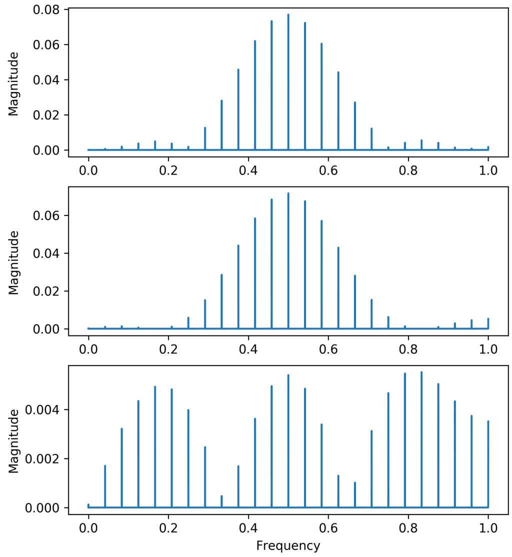

Fig. 9: *Spectra produced by RPMS using VBAP (above), WFS (middle) and the difference between VBAP and WFS (bottom) on a 72 speaker array.*

entirety. In essence, RPMS rasterizes the panning curve to be stored in a pre-computed buffer. In theory, this could be expanded to a multidimensional buffer that stores the specific panning curve for each loudspeaker considering it's particular neighbors.

Furthermore, the authors plan to verify the theoretical findings with practical measurements in the real loudspeaker array shown in Fig. 7. Here, considerations about further comparisons between VBAP and WFS arise. High density loudspeaker arrays in the magnitude of one loudspeaker per audio sample could hold the potential to bring VBAP and WFS even closer together. Therefore, one may further ask to what extend is VBAP actually capable of reconstructing a sound field and not merely a psychoacoustic panning method.

References

[1] Gerzon, M. A., "Periphony: With-height sound reproduction," *Journal of the audio engineering society*, 21(1), pp. 2–10, 1973.

[2] Pulkki, V., "Virtual sound source positioning using vector base amplitude panning," *Journal of*

the audio engineering society, 45(6), pp. 456–466, 1997.

[3] Berkhout, A. J., "A holographic approach to acoustic control," *Journal of the audio engineering society*, 36(12), pp. 977–995, 1988.

[4] Spors, S., Rabenstein, R., and Ahrens, J., "The theory of wave field synthesis revisited," in *Proceedings of the 124th AES convention*, pp. 17–20, Audio Engineering Society (AES), 2008.

[5] Lossius, T., Baltazar, P., and de la Hogue, T., "DBAP – distance-based amplitude panning," in *Processings of the International Computer Music Conference*, Montreal, QC, Canada, 2009.

[6] Theile, G. and Wittek, H., "Wave field synthesis: A promising spatial audio rendering concept," *Acoustical science and technology*, 25(6), pp. 393–399, 2004.

[7] Bleda Pérez, S., *Contribuciones a la implementación de sistemas Wave Field Synthesis*, Ph.D. thesis, Universitat Politècnica de València, 2009.

[8] Spors, S., "An Analytic Secondary Source Selection Criterion for Wavefield Synthesis," *Fortschritte der Akustik, DAGA 2007*, pp. 679–680, 2007.

[9] Smalley, D., "Spectromorphology: explaining sound-shapes," *Organised sound*, 2(2), pp. 107–126, 1997.

[10] Smalley, D., "Space-form and the acousmatic image," *Organised sound*, 12(1), pp. 35–58, 2007.

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

[12] Hagan, K. L., "Textural composition and its space," in *Proceedings of the 5th Sound and Music Computing Conference*, Berlin, Germany, 2008.

[13] von Coler, H., Lepa, S., and Weinzierl, S., "User-Defined Mappings for Spatial Sound Synthesis," in *International Conference on New Interfaces for Musical Expression*, Royal Birmingham Conservatoire, UK, 2020.

[14] Kim-Boyle, D., "Spectral and Granular Spatialization with Boids," in *Proceedings of the International Computer Music Conference (ICMC)*, New Orleans, LA, USA, 2006.

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

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

[17] 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.

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

[19] 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)*, Perth, Australia, 2013.

[20] McGee, R. M., *Scanning Spaces: Paradigms for Spatial Sonification and Synthesis*, Ph.D. thesis, University of California, Santa Barbara, 2015.

[21] Schmele, T., "Exploring 3d audio as a new musical language," 2011, masters thesis, Universitat Pompeu Fabra, Barcelona, Spain.

[22] Schmele, T. and Gómez, I., "Three-dimensional sonification of fMRI brain data in the musical composition Neurospaces," in *Proceedings of the 9th Conference on Interdisciplinary Musicology – CIM14*, Berlin, Germany, 2014.

[23] Norilo, V. and Moreno, J., "Aural Weather Etude: Installing Atmosphere," in *Proceedings of the International Computer Music Conference (ICMC)*, International Computer Music Association, Santiago, Chile, 2021.

[24] Lopez, J. J., Gutierrez, P., Cobos, M., and Aguilera, E., "Sound distance perception comparison between wave field synthesis and vector base amplitude panning," in *Proceedings of the 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, pp. 165–168, IEEE, 2014.

[25] Camier, C. and Guastavino, C., "Perceptual evaluation of multichannel synthesis of moving sounds as a function of rendering strategies and velocity," *The Journal of the Acoustical Society of America*, 141(5), pp. 3511–3511, 2017.

[26] Chung, H., Chon, S. B., Yoo, J.-h., and Sung, K.-M., "Analysis of frontal localization in double layered loudspeaker array system," in *Proceedings of 20th International Congress on Acoustics*, Sydney, Australia, 2010.

[27] Rohr, L., Corteel, E., Nguyen, K.-V., and Lissek, H., "Vertical localization performance in a practical 3-D WFS formulation," *Journal of the Audio Engineering Society*, 61(12), pp. 1001–1014, 2013.

A Deriving the VBAP panning function

As stated in section 2, the VBAP panning function is derived from Eqs. 1 and 3. Substituting g_2 with Eq. 3 in Eq. 1, we receive:

$$\frac{\tan\phi}{\tan\phi_0} = \frac{g_1 - \sqrt{C - g_1^2}}{g_1 + \sqrt{C - g_1^2}} \tag{A1}$$

We then proceed to remove the fractions:

$$g_1 \tan\phi + \sqrt{C - g_1^2}\tan\phi = g_1 \tan\phi_0 - \sqrt{C - g_1^2}\tan\phi_0 \tag{A2}$$

By rearranging the terms, we can factor out the square root:

$$\sqrt{C - g_1^2}(\tan\phi_0 + \tan\phi) = g_1 \tan\phi_0 - g_1 \tan\phi \tag{A3}$$

Squaring both sides then removes the square root:

$$(C - g_1^2)(\tan\phi_0 + \tan\phi)^2 = (g_1 \tan\phi_0 - g_1 \tan\phi)^2 \tag{A4}$$

We then apply the binomial theorem:

$$\begin{aligned}(C - g_1^2)(\tan^2\phi_0 + 2\tan\phi_0\tan\phi + \tan^2\phi) = \\ g_1^2\tan^2\phi_0 - 2g_1^2\tan\phi_0\tan\phi + g_1^2\tan^2\phi\end{aligned} \tag{A5}$$

By resolving the brackets on the left side of the equation, we can proceed to get all terms with g_1^2 to the right:

$$\begin{aligned}C(\tan^2\phi_0 + 2\tan\phi_0\tan\phi + \tan^2\phi) = \\ 2g_1^2(\tan^2\phi_0 + \tan^2\phi)\end{aligned} \tag{A6}$$

Finally, we isolate g_1 on the right by dividing by $2(\tan^2\phi_0 + \tan^2\phi)$ and taking the square root:

$$\sqrt{\frac{C(\tan^2\phi_0 + 2\tan\phi_0\tan\phi + \tan^2\phi)}{2(\tan^2\phi_0 + \tan^2\phi)}} = g_1 \tag{A7}$$

Thus, if we consider $C = 1$, we receive Eq. 4.