

## LOCALIZATION USING DIFFERENT AMPLITUDE-PANNING METHODS IN THE FRONTAL HORIZONTAL PLANE

Matthias Frank

Institute of Electronic Music and Acoustics,  
University of Music and Performing Arts Graz  
Graz, Austria  
frank@iem.at

### ABSTRACT

Amplitude panning is the simplest method to create phantom sources in the horizontal plane. The most commonly employed amplitude-panning methods are Vector-Base Amplitude Panning (VBAP), Multiple-Direction Amplitude Panning (MDAP), and Ambisonics. This article investigates the localization of frontal phantom sources created by VBAP, MDAP, and Ambisonics (with and without  $\max_E$  weighting) at the central listening position in a listening experiment. The experiment was conducted under typical non-anechoic studio listening conditions and utilized pink noise and a regular array of 8 loudspeakers for all methods. The experimental results are compared to different predictors: a binaural localization model using measured binaural room impulse responses, the direction of the measured sound intensity vector, and the directions of the simpler velocity and energy vectors. The article hereby addresses the questions of how close the actually localized directions of the different panning methods are compared to the desired directions, and how good the predictors match the experimental results.

### 1. INTRODUCTION

Amplitude-panning methods use simple level differences between the loudspeakers to evoke auditory objects between the loudspeakers, so-called phantom sources [1]. Although the computational effort is similar, the methods differ in their theoretical basis and the number of loudspeakers they use for each phantom source. This contribution examines perceptual differences between the most commonly used amplitude-panning methods, in particular the phantom source localization. This is done by employing existing experimental results from the thesis of the author [2] that used vector-base amplitude panning, multiple-direction amplitude panning, and Ambisonics with different weightings on the same loudspeaker arrangement. Although there are some studies about the localization of Ambisonics [3, 4, 5], comparisons to the other panning methods are rare [6]. The experiment focuses on frontal directions (with a maximum displacement from the median plane of  $45^\circ$ ) in order to compare the panning methods within the angular range where human sound source localization works most accurately [7].

The experimental results are compared to the directions of the simple velocity and energy vector. Although the suitability of these measures for the prediction of phantom source localization has not been proven yet, they are often applied in practice for Ambisonics decoder design [8]. This contribution examines their suitability and compares them to a state-of-the-art binaural localization model and measurements of the sound intensity vector.

The second section introduces the employed amplitude-panning methods and their theoretical basis. The experimental localization results for the panning methods are presented in the third section. In order to examine the controllability of the phantom source location, subjective variation is discussed and the results are compared to the desired panning direction. Section four presents localization predictors that are based on dummy head or microphone array measurements at the listening position within the actual sound field, as well as simpler predictors incorporating solely the loudspeaker gains and positions. Finally, their predictions are compared to the experimental results.

Throughout this contribution, the directions of  $L$  loudspeakers, as well as the panning directions are expressed as unit vectors  $\theta = [\cos(\phi), \sin(\phi)]^T$  depending on the azimuth angle  $\phi$  in the  $x$ - $y$  plane, cf. Figure 1. The scalar weight  $g_l$  of each loudspeaker  $l \in \{1 \dots L\}$  denotes its adjustable gain.

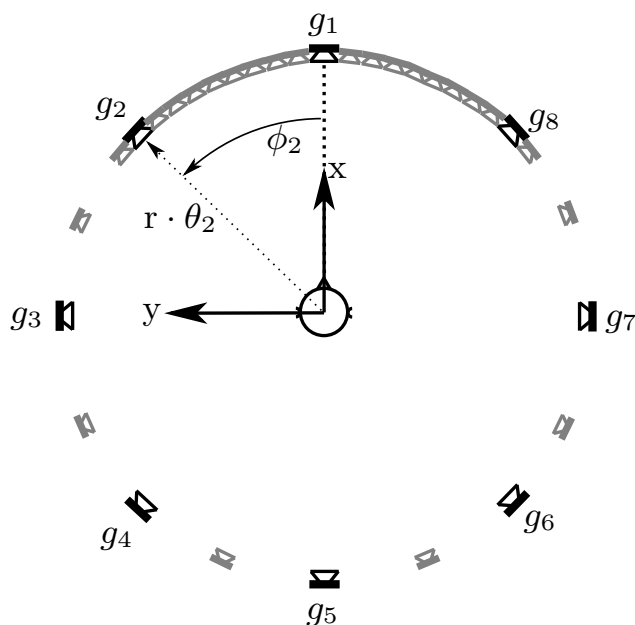


Figure 1: Experimental setup in the reference coordinate system. The smaller, gray loudspeakers are visible but inactive.

## 2. HORIZONTAL AMPLITUDE-PANNING METHODS

All presented amplitude-panning methods are also applicable to three-dimensional loudspeaker arrangements. However, this overview focuses on the two-dimensional case that is employed in this contribution.

### 2.1. Vector-Base Amplitude Panning (VBAP)

Vector-Base Amplitude Panning (VBAP) [9] can be seen as the generalization of the tangent law [10] for amplitude panning in two-channel stereophony [11]. The tangent law is based on a simple geometrical head model and is still the most popular panning law for pairwise panning. In order to create a phantom source within a loudspeaker pair located at  $\mathbf{L}_{ij} = [\boldsymbol{\theta}_i, \boldsymbol{\theta}_j]$ , VBAP calculates the weights  $\mathbf{g}_{ij} = [g_i, g_j]^T$  depending on the panning direction  $\boldsymbol{\theta}_s$ :

$$\mathbf{g}_{ij} = \mathbf{L}_{ij}^{-1} \boldsymbol{\theta}_s. \quad (1)$$

Typically, a subsequent normalization of the gains is necessary to keep the overall energy constant. The aperture angle of the loudspeaker pair should not exceed  $90^\circ$  [12] resulting in non-negative weights. To extend the panning range around one pair, more loudspeaker pairs can be attached [9].

Basically, the number of active loudspeakers is depending on the panning direction: two loudspeakers are active for directions between two loudspeakers and one is active for directions coinciding with a loudspeaker.

### 2.2. Multiple-Direction Amplitude Panning (MDAP)

For a more uniform panning, those cases of VBAP with only one active loudspeaker can be avoided. This is done by extending VBAP to Multiple-Direction Amplitude Panning (MDAP) [13].

MDAP superimposes the results of VBAP for  $B$  panning directions uniformly distributed around the desired panning direction  $\boldsymbol{\theta}_s$  within a spread of  $\pm\phi_{\text{MDAP}}$ . Typically, the maximum spread is related to the loudspeaker spacing  $\Delta\phi_L$  of a uniform loudspeaker arrangement. In this contribution, MDAP is always used with  $B = 10$  panning directions uniformly distributed within a spread of  $\phi_{\text{MDAP}} = 1/2\Delta\phi_L$ .

For this setting, three loudspeakers are active for most panning directions, even for directions on a loudspeaker. If the panning direction lies exactly in the middle between two loudspeakers, only these two loudspeakers are active. In these special cases, MDAP yields the same loudspeaker gains as VBAP.

### 2.3. Ambisonics

Ambisonics [14, 15, 16, 17] is a recording and reproduction method which is based on the representation of the sound field excitation as a superposition of orthogonal basis functions. In the horizontal case, these functions are the periodic trigonometric basis of the Fourier series, the so-called circular harmonics. Their maximum order  $N$  determines the spatial resolution and the number  $2N + 1$  of signals and minimum required loudspeakers.

For one source at a direction  $\boldsymbol{\theta}_s$ , the Ambisonic spectrum  $\mathbf{y}_N(\boldsymbol{\theta}_s)$  is calculated by evaluating the circular harmonics at  $\boldsymbol{\theta}_s$ . This calculation is frequency-independent and assumes that all sources and the loudspeakers lie on a circle of the same radius  $r$ .

The decoder derives the gains  $\mathbf{g} = \{g_1, \dots, g_L\}$  for the  $L$  loudspeakers of an arrangement from the Ambisonic spectrum  $\mathbf{y}_N(\boldsymbol{\theta}_s)$  by multiplication with the decoder matrix  $\mathbf{D}$ :

$$\mathbf{g} = \mathbf{D} \text{diag}\{\mathbf{a}_N\} \mathbf{y}_N(\boldsymbol{\theta}_s). \quad (2)$$

The matrix is derived from the circular harmonic spectra  $\mathbf{y}_N(\boldsymbol{\theta}_l)$  of each loudspeaker  $\mathbf{Y}_N = [\mathbf{y}_N(\boldsymbol{\theta}_1), \mathbf{y}_N(\boldsymbol{\theta}_2), \dots, \mathbf{y}_N(\boldsymbol{\theta}_L)]$ . It can be calculated by transposition or inversion of  $\mathbf{Y}_N$ , resulting in a sampling or mode-matching decoder [18], respectively. The energy-preserving decoder [19] uses more sophisticated techniques, such as singular value decomposition. See the appendix of [20] for an overview about different decoders. In this contribution, the regular arrangement of  $L$  loudspeakers (cf. Figure 1) ensures that the simple sampling decoder is also mode-matching and energy-preserving for all orders  $N \leq (L - 1)/2$ .

In order to control the main and side lobes emerging from the truncation of the circular harmonics, a weighting vector  $\mathbf{a}_N$  is applied in the harmonics domain [17]. The basic weighting uses a vector of ones  $\mathbf{a}_N = \mathbf{1}$ , whereas the max- $r_E$  weighting suppresses the side lobes at the cost of a wider main lobe by attenuating higher orders. This is done by an order-depend weight  $a(n) = \cos(\frac{n\pi}{2N+2})$ . Another weighting, called in-phase, yielded no convincing results in previous experiments [5] and is therefore not used here.

Basically, Ambisonics always uses all available loudspeakers for the creation of a single phantom source. However, the equiangular arrangement of  $L$  even-numbered loudspeakers yields an exception when using max- $r_E$  Ambisonics with an order of  $N = L/2 - 1$ : for panning directions exactly in the middle between two neighboring loudspeakers, only these two loudspeakers are active. In these cases, max- $r_E$  Ambisonics yields the same loudspeaker gains as VBAP and MDAP.

## 3. EXPERIMENT

The listening experiment evaluates the localization of phantom sources created by VBAP, MDAP, basic Ambisonics, and max- $r_E$  Ambisonics at the central listening position.

### 3.1. Setup and Conditions

All panning methods employ a regular ring of 8 Genelec 8020 loudspeakers at a radius of  $r = 2.5$  m. Figure 1 shows the experimental setup with additional inactive but visible loudspeakers placed in  $5^\circ$  steps in and around the angular range of the target directions. The height of all loudspeakers was set to 1.2 m which was also the ear height of the subjects. The experiment was performed in the IEM CUBE, a  $10.3 \text{ m} \times 12 \text{ m} \times 4.8 \text{ m}$  large room with a mean reverberation time of 470 ms that fulfills the recommendation for surround reproduction in ITU-R BS.1116-1 [2, 21]. The central listening position lies within the effective critical distance.

The control of the entire experiment and the creation of the loudspeaker signals used the open source software pure data<sup>1</sup> on a standard PC with RME audio interface and D/A converters. The perceived direction was assessed by a pointing method using a toy-gun that was captured by an infrared tracking system. Details about the pointing method can be found in [22].

<sup>1</sup>freely available on <http://puredata.info/downloads>

Both Ambisonics variants use a maximum order of 3 and MDAP is applied with  $B = 10$  panning directions uniformly distributed within a spread of  $\phi_{MDAP} = 22.5^\circ$ . The effect of the Ambisonics order has already been studied in [5] and is not part of this contribution. All panning methods were evaluated for 9 directions (with an even spacing of  $5.625^\circ$ ) between  $0^\circ$  and  $-45^\circ$  (to the right). Each of the  $36 = 9$  (directions)  $\times$  4 (panning methods) conditions was evaluated twice by each subject in random order. The stimulus consisted of 3 pink noise bursts, each with 100 ms fade-in, 200 ms at 65 dB(A), 100 ms fade-out, and 200 ms silence before the next fade-in. The stimulus playback could be repeated at will by the subjects.

There were 14 subjects participating in the experiment. All of them were part of a trained expert listening panel [23, 24].

### 3.2. Results

An analysis of variance (ANOVA) showed that the repetition was not a significant factor ( $p = 0.522$  for VBAP,  $p = 0.465$  for MDAP,  $p = 0.085$  for basic Ambisonics, and  $p = 0.91$  for max- $r_E$  Ambisonics). This confirms a high intra-rater reliability and thus no subjects were excluded from the results. On the other hand, the subjects were a highly significant factor ( $p \ll 0.001$ ) for all tested panning methods. This agrees with the inter-subjective localization found for lateral [2] and vertical phantom sources [25, 26]. Nevertheless, the following localization curves summarize all 28 answers from all subjects and repetitions for each condition.

Figure 2 shows the median values and the corresponding 95% confidence intervals for the 9 different panning angles using VBAP at the central listening position. Obviously, the panning angle is a significant factor ( $p \ll 0.001$ ). All neighboring conditions

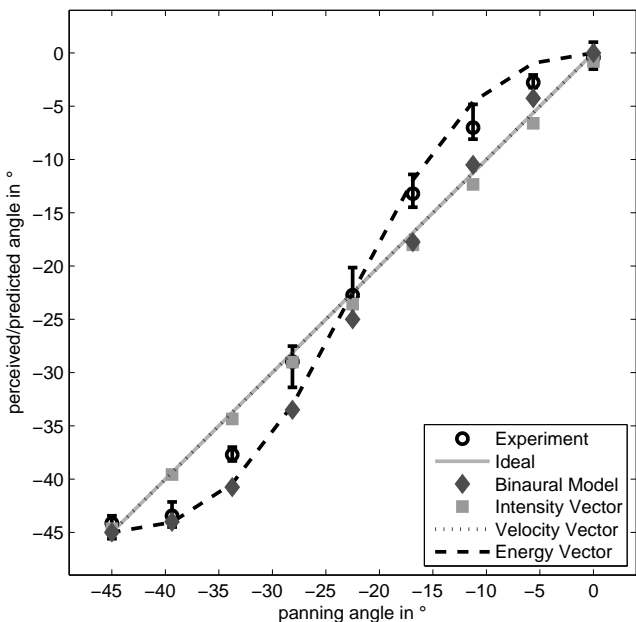


Figure 2: Localization curves for VBAP: experimental results (median values and 95% confidence intervals), ideal curve (perceived/predicted angle = panning angle), and predictions by binaural model, intensity vector, velocity vector, and energy vector.

were perceived from significantly different directions ( $p \ll 0.001$ ), except for the angles  $-45^\circ$  and  $-39.375^\circ$  ( $p = 0.053$ ). In comparison to the ideal localization curve (perceived angle = panning angle), the perceived angles tend towards the loudspeakers. This tendency is known from [27] and is even more distinct for lateral directions.

Using MDAP, the panning angle is still a significant factor ( $p \ll 0.001$ ). This holds true for the direct comparison of neighboring panning angles. Compared to VBAP, the median perceived angles for MDAP are closer to the ideal panning curve and yield a reduced tendency towards the loudspeakers, cf. Figure 3.

The experimental results for basic Ambisonics show a stretched trend, i.e. a steeper slope than the ideal curve, cf. Figure 4. The discriminability of the panning angles is comparable to MDAP.

The significant discriminability of the panning angles holds true for max- $r_E$  Ambisonics. Figure 5 shows that the median experimental results are very close to the ideal localization curve for this panning method.

Table 1 compares the median deviation of the different panning methods from the ideal localization curve, i.e., how much the perceived angle deviated from the panning angle. The angles deviate most for VBAP, in fact more than two times as much as for max- $r_E$  Ambisonics. The angular match is best for max- $r_E$  Ambisonics, followed by MDAP, and basic Ambisonics.

Table 1: Average absolute deviation of median experimental results from ideal localization curve for different panning methods.

VBAP	MDAP	basic	max- $r_E$
2.35°	1.28°	1.58°	1.05°

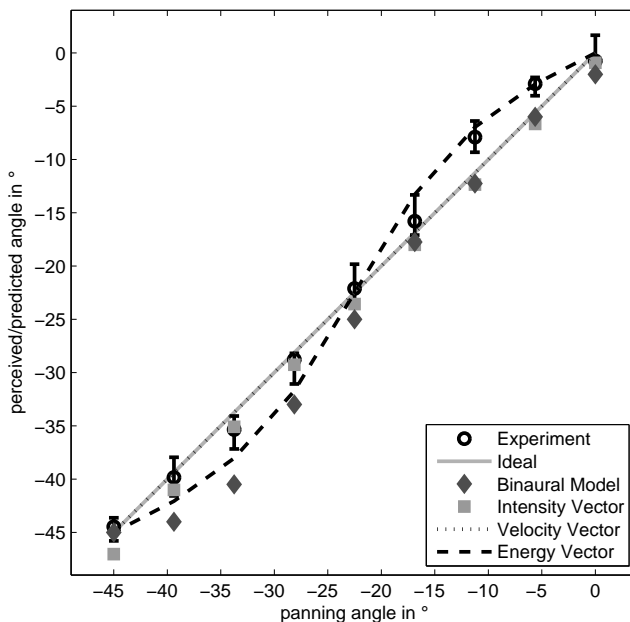


Figure 3: Localization curves for MDAP: experimental results (median values and 95% confidence intervals), ideal curve (perceived/predicted angle = panning angle), and predictions by binaural model, intensity vector, velocity vector, and energy vector.

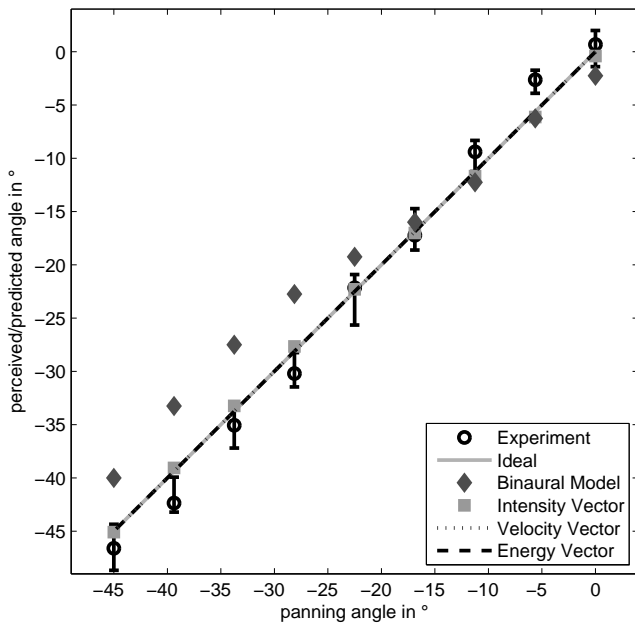


Figure 4: Localization curves for basic Ambisonics: experimental results (median values and 95% confidence intervals), ideal curve (perceived/predicted angle = panning angle), and predictions by binaural model, intensity vector, velocity vector, and energy vector.

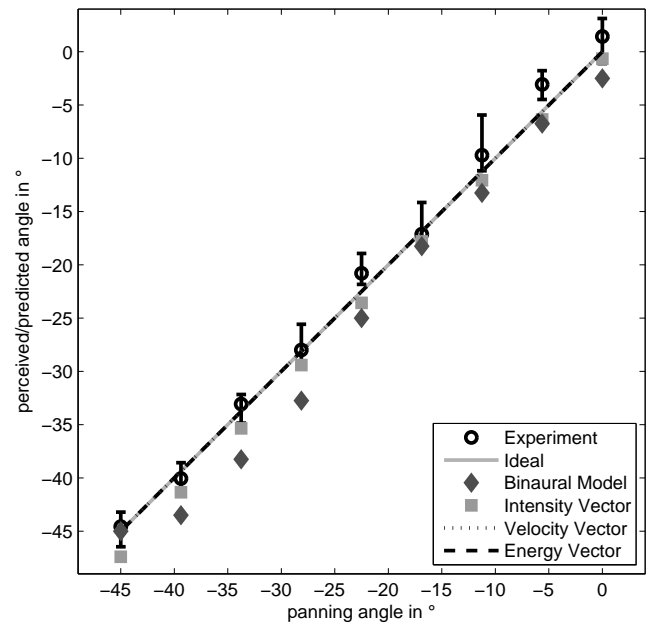


Figure 5: Localization curves for max-r<sub>E</sub> Ambisonics: experimental results (median values and 95% confidence intervals), ideal curve (perceived/predicted angle = panning angle), and predictions by binaural model, intensity vector, velocity vector, and energy vector.

Despite this ranking, VBAP yields the smallest standard deviations in the experimental results, cf. Table 2. This is mainly the case for panning angles close to the loudspeakers, which provide narrow and accurate localization in VBAP in comparison to phantom sources created by the other methods. Obviously, for higher number of active loudspeakers, the standard deviation increases. This holds true for the total standard deviation, as well as for the inter-subjective and the intra-subjective standard deviation. However, the inter-subjective standard deviation is greater than its intra-subjective counterpart for all panning methods, agreeing with the results from the ANOVA that showed the subjects to be a significant factor, but not the repetition.

Table 2: Mean total, inter-subjective, and intra-subjective standard deviations of experimental results for different panning methods.

	VBAP	MDAP	basic	max-r <sub>E</sub>
total	2.93°	3.32°	4.61°	4.07°
inter-subj.	2.52°	2.94°	3.94°	3.54°
intra-subj.	1.87°	1.90°	2.59°	2.28°

#### 4. PREDICTIONS

In order to save experiments in the future, it is desirable to find suitable predictors for the localization of phantom sources. This section presents a selection of predictors that differ in the measurement effort and it compares their predictions to the experimental results.

#### 4.1. Localization Predictors

##### Binaural Model

This contribution employs a binaural localization model after Lindemann [28, 29] which is part of the Auditory Modeling Toolbox<sup>2</sup>. It divides the binaural input signals into 36 frequency bands with a spacing of 1ERB (equivalent rectangular bandwidth) [30]. The auditory nerve is modeled by a half-wave rectifier and a low-pass filter at 800Hz. In each band, the inter-aural level-difference (ILD) is considered by monaural detectors and contra-lateral inhibition. The inter-aural time-difference (ITD) is then computed as the centroid of the inter-aural cross correlation function [31], which delivers one ITD value for each frequency band.

Within each frequency band, the ITD value of the phantom source is compared to the values of a single sound source in a lookup table. The best matching ITD is selected and the corresponding angle is regarded as the angle of the phantom source for the present frequency band. A single angle as prediction result is achieved by the median value of the angles for all frequency bands. The best fit to the median experimental results has been achieved when using 21 frequency bands covering the range from 164 Hz – 3558 Hz. On the one hand, the lower frequency limit seems reasonable as very low frequencies do not yield inter-aural differences because of the large wavelengths in comparison to the head diameter. On the other hand, the upper frequency limit underlines the dominant role of low-frequency ITDs, albeit it also supports the importance of ITDs at higher frequencies in comparison to the classical duplex theory [32].

The model is fed with binaural room impulse responses recorded at the central listening position of the experimental setup with a

<sup>2</sup>freely available on [amtoolbox.sourceforge.net/](http://amtoolbox.sourceforge.net/)

B&K 4128C dummy head. It uses the first 80 ms of the impulse responses. As the model cannot distinguish between front and back, the ITD values in the lookup table were limited to the directions between  $\pm 45^\circ$ , where the conditions of the experiment lie.

### Intensity Vector

Sound intensity is a physical measure of the directional sound power flow and can thus be used to determine the direction where sound is coming from [33]. The intensity  $\mathbf{I}$  is computed from the scalar sound pressure  $p$  and the vectorial particle velocity  $\mathbf{v}$  as  $\mathbf{I} = p\mathbf{v}$  [34]. The sound pressure can be measured by an omni-directional microphone. The particle velocity is typically not measured directly but by the pressure gradients  $v_x$  and  $v_y$  using figure-of-eight microphones, each one aligned to the axis of the coordinate system. Here  $p$ ,  $v_x$ , and  $v_y$  are computed as the convolution of A-weighted pink noise with impulse responses measured with two Schoeps CCM 8 figure-of-eight microphones and an NTI MM2210 omni-directional microphone at the central listening position.

As a predictor for sound source directions, it is not suitable to compute the instantaneous direction of the intensity vector for each sample separately, i.e. 44100 times a second, but rather as a temporal average within a certain time window. The time window was set to 80 ms ( $S = 3528$  samples), which corresponds to the binaural localization model. The components  $\overline{I_x}$  and  $\overline{I_y}$  of the temporally averaged intensity vector  $\overline{\mathbf{I}} = [\overline{I_x} \ \overline{I_y}]$  are computed as

$$\overline{I_x} = \sum_{s=1}^S p(s)v_x(s) \quad \text{and} \quad \overline{I_y} = \sum_{s=1}^S p(s)v_y(s). \quad (3)$$

The direction of the intensity vector is calculated as  $\arctan(\overline{I_y}, \overline{I_x})$  and is equal to the direction of the velocity vector under free-field conditions.

### Velocity Vector

The direction of the velocity vector was proposed as a simple predictor for the localization of low frequencies ( $\leq 700$  Hz) [35, 36]. It is calculated as linear summation of the weighted loudspeaker directions:

$$\mathbf{r}_V = \frac{\sum_{l=1}^L g_l \boldsymbol{\theta}_l}{\sum_{l=1}^L g_l}. \quad (4)$$

As it is solely based on the loudspeaker directions and gains, it does not require any acoustical measurements. It assumes free-field conditions or at least a dominant direct sound. For two loudspeakers, the velocity vector points towards the same direction as intended by VBAP.

### Energy Vector

Following the idea of the velocity vector, the energy vector  $\mathbf{r}_E$  [35, 36] was defined as

$$\mathbf{r}_E = \frac{\sum_{l=1}^L g_l^2 \boldsymbol{\theta}_l}{\sum_{l=1}^L g_l^2}. \quad (5)$$

This model assumes an energetic superposition of the loudspeaker signals and is expected to model the localization direction for higher frequencies or broadband signals. The magnitude of the energy vector can also be used to describe spatial distribution of energy [17] and the perceived width of phantom sources [37].

## 4.2. Prediction of the Experimental Results

Along with the experimental results, Figures 2 to 5 show the different predictions. Obviously, the direction of the velocity vector is identical to the desired panning direction for all evaluated panning methods due to the regular loudspeaker arrangement. For both Ambisonics variants, it is also identical to the direction of the energy vector. The direction of the intensity vector is very close to the one of the velocity vector. This finding shows that the intensity vector is the measured counterpart of the velocity vector, even under non-free-field conditions.

Table 3: Average absolute deviation of predictions from median experimental results for different panning methods.

	VBAP	MDAP	basic	max-r <sub>E</sub>
Binaural Model	2.35°	3.07°	4.92°	3.37°
Intensity Vector	2.75°	2.12°	1.89°	2.41°
Velocity Vector	2.35°	1.28°	1.58°	1.05°
Energy Vector	1.60°	1.44°	1.58°	1.05°

Table 3 compares the deviation of the predictions from the median experimental results. The binaural model yields the worst prediction, especially for basic Ambisonics. Better results are achieved by the intensity vector. The vector models yield the smallest deviations from the experimental results. In detail, the energy vector predicts the VBAP localization better, as it includes the effect that the localization tends towards the loudspeakers.

The deviations of the predictions from the experimental results are similar to the standard deviations of the experimental results. Thus, all predictors seem to be suitable for the localization of frontal phantom sources at the central listening position. However, it is remarkable that the simplest models yield the best predictions at the same time and the most complex model the worst predictions.

## 5. CONCLUSION

This contribution investigated frontal phantom source localization at the central listening position using VBAP, MDAP, basic, and max-r<sub>E</sub> Ambisonics on a circle of 8 loudspeakers. The match between the median experimental results and the desired panning direction was best for max-r<sub>E</sub> Ambisonics and worst for VBAP. However, the standard deviation of VBAP was the smallest of all panning methods. Obviously, the standard deviation increases with the number of active loudspeakers. This is expected to be even more relevant for off-center listening positions [2, 5, 38]. The standard deviation was found to be dominated by the inter-subjective standard deviation, i.e. the differences between the subjects.

The experimental results were compared to a binaural localization model, the measured intensity vector, and the velocity and energy vectors. All these predictors seemed to be suitable for the prediction of the experimental results. It is remarkable that the velocity and energy vectors as the simplest predictors yielded the best predictions at the same time. This finding justifies the use of these predictors in practice. Moreover, there exist first hints that they can be also applied to vertical or three-dimensional amplitude panning [39]. However, their applicability for lateral phantom sources or off-center listening positions is still under investigation.

## 6. ACKNOWLEDGMENTS

The author thanks all subjects for their participation and the reviewers for their helpful comments. This work was partly supported by the projects AAP and ASD, which are funded by Austrian ministries BMVIT, BMWFJ, the Styrian Business Promotion Agency (SFG), and the departments 3 and 14 of the Styrian Government. The Austrian Research Promotion Agency (FFG) conducted the funding under the Competence Centers for Excellent Technologies (COMET, K-Project), a program of the above-mentioned institutions.

## 7. REFERENCES

- [1] Klaus Wendt, *Das Richtungshören bei der Überlagerung zweier Schallfelder bei Intensitäts- und Laufzeitstereophonie*, Ph.D. thesis, RWTH Aachen, Germany, 1963.
- [2] Matthias Frank, *Phantom Sources using Multiple Loudspeakers in the Horizontal Plane*, Ph.D. thesis, University of Music and Performing Arts Graz, Austria, 2013.
- [3] Eric Benjamin, Aaron Heller, and Richard Lee, "Localization in Horizontal-Only Ambisonic Systems," in *Audio Engineering Society Convention 121*, 10 2006.
- [4] Stéphanie Bertet, Jérôme Daniel, Laëtitia Gros, Etienne Parizet, and Olivier Warusfel, "Investigation of the Perceived Spatial Resolution of Higher Order Ambisonics Sound Fields: A Subjective Evaluation Involving Virtual and Real 3D Microphones," in *Audio Engineering Society Conference: 30th International Conference: Intelligent Audio Environments*, 3 2007.
- [5] Matthias Frank and Franz Zotter, "Localization experiments using different 2D Ambisonics decoders," in *25. Tonmeis-tertagung, Leipzig*, 2008.
- [6] Gavin Kearney, Enda Bates, Frank Boland, and Dermot Furlong, "A comparative study of the performance of spatialization techniques for a distributed audience in a concert hall environment," in *Audio Engineering Society Conference: 31st International Conference: New Directions in High Resolution Audio*, 6 2007.
- [7] Jens Blauert, *Spatial Hearing*, MIT Press, 1983.
- [8] David Moore and Jonathan Wakefield, "A design tool to produce optimized ambisonic decoders," in *Audio Engineering Society Conference: 40th International Conference: Spatial Audio: Sense the Sound of Space*, 10 2010.
- [9] Ville Pulkki, "Virtual sound source positioning using vector base amplitude panning," *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456–466, 1997.
- [10] D. M. Leakey, "Some measurements on the effects of inter-channel intensity and time differences in two channel sound systems," *The Journal of the Acoustical Society of America*, vol. 31, no. 7, pp. 977–986, 1959.
- [11] Alan D. Blumlein, "British patent specification 394,325 (improvements in and relating to sound-transmission, sound-recording and sound-reproducing systems)," *J. Audio Eng. Soc.*, vol. 6, no. 2, pp. 91–98, 130, 1958.
- [12] Ville Pulkki, *Spatial Sound Generation and Perception by Amplitude Panning Techniques*, Ph.D. thesis, Helsinki University of Technology, 2001.
- [13] Ville Pulkki, "Uniform spreading of amplitude panned virtual sources," in *Applications of Signal Processing to Audio and Acoustics, 1999 IEEE Workshop on*, 1999, pp. 187–190.
- [14] Duane H. Cooper and Takeo Shiga, "Discrete-matrix multi-channel stereo," *Journal of the Audio Engineering Society*, vol. 20, no. 5, pp. 346–360, 1972.
- [15] Michael A. Gerzon, "With-height sound reproduction," *Journal of the Audio Engineering Society*, vol. 21, pp. 2–10, 1973.
- [16] David G. Malham and Anthony Myatt, "3D Sound Spatialization using Ambisonic Techniques," *Computer Music Journal*, vol. 19, no. 4, pp. 58–70, 1995.
- [17] Jérôme Daniel, *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*, Ph.D. thesis, Université Paris 6, 2001.
- [18] Mark A. Poletti, "A Unified Theory of Horizontal Holographic Sound Systems," *J. Audio Eng. Soc.*, vol. 48, no. 12, pp. 1155–1182, 2000.
- [19] Frank Zotter, Hannes Pomberger, and Markus Noisternig, "Energy-Preserving Ambisonic Decoding," *Acta Acustica united with Acustica*, vol. 98, no. 1, pp. 37–47, 2012.
- [20] Franz Zotter and Matthias Frank, "All-round ambisonic panning and decoding," *J. Audio Eng. Soc.*, vol. 60, no. 10, pp. 807–820, 2012.
- [21] ITU, "ITU-R BS.1116-1: Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems," 1997.
- [22] Matthias Frank, Ludwig Mohr, Alois Sontacchi, and Franz Zotter, "Flexible and Intuitive Pointing Method for 3-D Auditory Localization Experiments," in *Audio Engineering Society Conference: 38th International Conference: Sound Quality Evaluation*, 6 2010.
- [23] Alois Sontacchi, Hannes Pomberger, and Robert Höldrich, "Recruiting and evaluation process of an expert listening panel," in *Fortschritte der Akustik, NAG/DAGA*, Rotterdam, 2009.
- [24] Matthias Frank and Alois Sontacchi, "Performance review of an expert listening panel," in *Fortschritte der Akustik, DAGA*, Darmstadt, 2012.
- [25] Ville Pulkki, "Localization of amplitude-panned virtual sources II: Two- and three-dimensional panning," *J. Audio Eng. Soc.*, vol. 49, no. 9, pp. 753–767, 2001.
- [26] Florian Wendt, Matthias Frank, and Franz Zotter, "Application of localization models for vertical phantom sources," in *Fortschritte der Akustik, AIA-DAGA*, Meran, 2013.
- [27] Laurent S. R. Simon, Russell Mason, and Francis Rumsey, "Localization curves for a regularly-spaced octagon loudspeaker array," in *Audio Engineering Society Convention 127*, 10 2009.
- [28] Werner Lindemann, "Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals," *The Journal of the Acoustical Society of America*, vol. 80, no. 6, pp. 1608–1622, 1986.
- [29] Werner Lindemann, "Extension of a binaural cross-correlation model by contralateral inhibition. II. The law of the first wave front," *The Journal of the Acoustical Society of America*, vol. 80, no. 6, pp. 1623–1630, 1986.

- [30] Brian C. J. Moore, Robert W. Peters, and Brian R. Glasberg, "Auditory filter shapes at low center frequencies," *The Journal of the Acoustical Society of America*, vol. 88, no. 1, pp. 132–140, 1990.
- [31] Lloyd A. Jeffress, "A place theory of sound localization," *Journal of comparative and physiological psychology*, vol. 41, no. 1, pp. 35–39, 1948.
- [32] Frederic L. Wightman and Doris J. Kistler, "The dominant role of low-frequency interaural time differences in sound localization," *The Journal of the Acoustical Society of America*, vol. 91, no. 3, pp. 1648–1661, 1992.
- [33] Juha Merimaa, "Energetic sound field analysis of stereo and multichannel loudspeaker reproduction," in *Audio Engineering Society Convention 123*, 10 2007.
- [34] Richard C. Heyser, "Instantaneous intensity," in *Audio Engineering Society Convention 81*, 11 1986.
- [35] Y. Makita, "On the directional localization of sound in the stereophonic sound field," Tech. Rep., EBU, Review Part A, 73, 1962.
- [36] Michael A. Gerzon, "General metatheory of auditory localisation," in *Audio Engineering Society Convention 92*, 3 1992.
- [37] Matthias Frank, "Source width of frontal phantom sources: Perception, measurement, and modeling," *Archives of Acoustics*, vol. 38, no. 3, pp. 311–319, September 2013.
- [38] Peter Stitt, Stéphanie Bertet, and Maarten van Walstijn, "Perceptual investigation of image placement with ambisonics for non-centred listeners," in *DAFx-13*, Maynooth, Ireland, September 2013.
- [39] Florian Wendt, Matthias Frank, and Franz Zotter, "Amplitude panning with height on 2, 3, and 4 loudspeakers," in *International Conference on Spatial Audio*, Erlangen, Germany, 2014.