1st EAA – EuroRegio 2010
Congress on Sound and Vibration
15 - 18 September 2010, Ljubljana, Slovenia
With Summer School for Young Researchers from 13 - 15 September 2010

Alps Adria Acoustics Association

Slovenian Acoustical Society

The Virtual T-Design Ambisonics-Rig Using VBAP

Franz ZOTTER[1], Matthias FRANK[1] and Alois SONTACCHI[1]

[1]Institute of Electronic Music and Acoustics, , University of Music and Performing Arts, Inffeldgasse 10/III, 8010 Graz, AUSTRIA, e-mail: {zotter,frank,sontacchi}@iem.at

ABSTRACT

Ambisonics can be regarded as a holophonic sound field rendering technique that decodes spherical harmonic encoded source-signals to discrete loudspeakers arranged on a sphere. The aim is the re-synthesis of sound sources perceivable from certain spatial directions, either by reproducing dedicated Ambisonics microphone recordings or synthetic signals. However, a truly holophonic decoding is frequently numerically ill-posed if the loudspeaker arrangement is nonuniformly covering the sphere. This contribution considers a new method that circumvents numerical difficulties by using a hybrid stereophonic-ambisonic approach.
A loudspeaker arrangement that corresponds to a spherical *t*-design provides an orthogonal sampling of the spherical harmonics. Therefore assuming such an arrangement, decoding to the loudspeakers is entirely unproblematic and only a matter of angular sampling the spherical harmonics expanded source-signal. However, setting up loudspeakers at positions of the considered *t*-design might be impossible due to practical limitations. To accomplish playback on any real loudspeaker arrangement, we employ vector-base amplitude-panning (VBAP) that decodes a virtual *t*-design. The paper concludes with an evaluation based on two objective characteristics of the hereby obtained decoder for arbitrary playback layouts.

[1]This project is funded by the Austrian Research Promotion Agency (FFG), the Styrian Government and the Styrian Business Promotion Agency (SFG) under the COMET programme.

1. INTRODUCTION

Stereophonic playback of sound uses two loudspeakers to create the impression of phantom sources that can be adjusted be perceived from locations between the loudspeakers. The evidence and restrictions of phantom sources have been clearly shown in the literature on psychoacoustics [2, 3] and have been generalized to triangles of loudspeakers (and combinations thereof) in the work of Pulkki [4] on *vector-base amplitude-panning* (VBAP).
Besides this experimentally well-proved method, an alternative way of acoustic imaging has evolved under the name *Ambisonics* [8, 9]. In particular, it is based on a mathematical decomposition of a virtual sound-field into finite-order discrete spherical harmonics. After its introduction using a low resolution, Ambisonics has been extended to a higher, adjustable but limited resolution [5, 14, 12]. Ambisonics only supports a limited flexibility of the loudspeaker arrangement and tends to pose mathematical obstacles otherwise. Its psychoacoustic evaluation is not as simple as for stereophony, however works exist [11, 10, 15].
In general, both panning methods are capable of spatializing sounds using an arrangement of loudspeakers, or in the language of audio technicians: using a *loudspeaker rig*. However, both have individual strengths and weaknesses. The following paragraphs discuss these aspects. In contrast to the *multiple-direction amplitude-panning* (MDAP [4]) approach, our contribution sketches a hybrid Ambi-VBAP approach to rule out remaining shortcomings. The idea to cascade VBAP with Ambisonics is not new and has been recently described in the papers of Batke and Keiler, e.g. [13]. This paper reversely cascades Ambisonics with VBAP, which seems to yield a highly generic approach applicable to arbitrary spherical loudspeaker arrangements.

2. REAL-VALUED AMPLITUDE PANNING METHODS

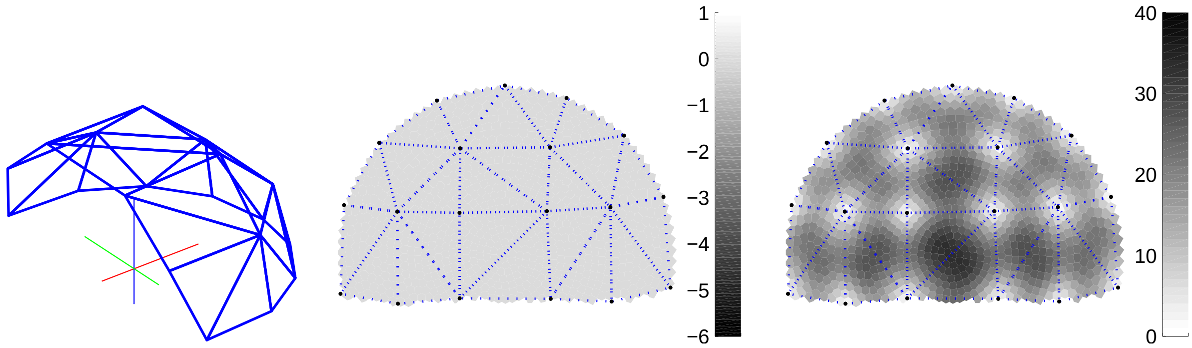
In order to discuss advantages and disadvantages of the basic approaches, it is necessary to find a notation and to introduce characteristic measures first. The directions associated with the loudspeakers $\{\boldsymbol{\theta}_l\}_{l=1,\dots,\mathrm{L}}$ and the panning direction $\boldsymbol{\theta}_s$ are vectors of unit length $\boldsymbol{\theta} = [\cos(\varphi)\sin(\vartheta),\ \sin(\varphi)\sin(\vartheta),\ \cos(\vartheta)]^{\mathrm{T}}$ that depend on the azimuth and zenith angle, φ and ϑ, respectively.
In general, amplitude-panning distributes the desired signal with the set of gains $\{g_l\}$ to a loudspeaker rig. We employ two technical measures estimating the *total signal power* $E = \sum_{l=1}^{\mathrm{L}} g_l^2$ and its *angular spread* $\sigma_{\mathrm{E}} = \arccos\left(\sum_{l=1}^{\mathrm{L}} < \boldsymbol{\theta}_l, \boldsymbol{\theta}_s > g_l^2/E\right)$ around the virtual source direction $\boldsymbol{\theta}_s$. These measures intend to characterize the loudness and the source-width.

2.1. Vector-Base Amplitude-Panning (VBAP)

VBAP uses a triangulation of the convex hull around the given loudspeaker rig. It is necessary to consider that the triangulation can have holes when the loudspeakers do not surround the listener. VBAP only considers virtual sources $\boldsymbol{\theta}_s$ inside a loudspeaker-triple $\{i, j, k\}$ of the triangulation, Fig. 1(a), aiming to evoke the impression of a phantom source. The corresponding loudspeaker-triple described by the vertices $\boldsymbol{\theta}_i$, $\boldsymbol{\theta}_j$, $\boldsymbol{\theta}_k$ uses strictly positive gains $\boldsymbol{g}_{ijk} = [g_i, g_j, g_k]^{\mathrm{T}}$

$$\begin{aligned} \boldsymbol{L} &= [\boldsymbol{\theta}_i, \boldsymbol{\theta}_j, \boldsymbol{\theta}_k], & \boldsymbol{L}\tilde{\boldsymbol{g}}_{ijk} &\stackrel{!}{=} \boldsymbol{\theta}_s, \\ \Rightarrow \tilde{\boldsymbol{g}}_{ijk} &= \boldsymbol{L}^{-1}\boldsymbol{\theta}_s, & \boldsymbol{g}_{ijk} &= \tilde{\boldsymbol{g}}_{ijk}/\|\tilde{\boldsymbol{g}}_{ijk}\|. \end{aligned} \tag{1}$$

VBAP has the attractive feature that for every line of the loudspeaker triples it reproduces panning laws for stereophonic phantom sources [4]; inside a loudspeaker triple it has been shown to work reasonably well. Nevertheless, the perceived source-width varies depending on the panning direction $\boldsymbol{\theta}_s$. For the above-mentioned technical measures see Tab. 1 and Figs. 1(b),1(c).



(a) Triangulation. (b) $E = 0$[dB]. (c) $\sigma_{\mathrm{E}} = 0 \ldots 33[^\circ]$.

Figure 1: VBAP: Triangulation, energy-measure $E(\boldsymbol{\theta}_s)$, and spread measure $\sigma_{\mathrm{E}}(\boldsymbol{\theta}_s)$ as a function of the virtual source direction $\boldsymbol{\theta}_s$ on the triangulation.

2.2. Ambisonics with t-design loudspeaker-rig

Ambisonics is based on analytic panning functions with limited resolution; definitions can be found, e.g., in [6]. A basic panning function is similar to a periodic sinc-function with a limited order $n \leq \mathrm{N}$. Panning to the z-axis, its rotationally symmetric expression becomes (P_n is the Legendre-polynomial)

$$g(\vartheta) = \frac{1}{(\mathrm{N}+1)^2} \sum_{n=0}^{\mathrm{N}} (2n+1)\, P_n(\cos(\vartheta)). \tag{2}$$

Analytically, the angular spread σ_{E} depends on the order N and is independent of $\boldsymbol{\theta}_s$. Nevertheless, real playback employs discrete gains and loudspeakers. For general loudspeaker-rigs it might be challenging to find a suitable *Ambisonics decoder* computing discrete gains, cf. [7]. Specifically, spherical t-designs [1] (like Platonic solids) represent layouts $\{\boldsymbol{\theta}_l\}$ with superior properties. If the order is restricted to $\mathrm{N} \leq t/2$, their decoding is simplest: the analytic gain function $g(\boldsymbol{\theta})$ is sampled at the loudspeaker positions $\{\boldsymbol{\theta}_l\}$. For this purpose, the $\{\boldsymbol{\theta}_l\}$-sampled spherical harmonics $Y_n^m(\boldsymbol{\theta})$ of the orders $n \leq \mathrm{N}$ expressed as a matrix $\boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}}$ are weighted by the vector of the spherical harmonics $\boldsymbol{y}_{\mathrm{N}}(\boldsymbol{\theta}_s)$ representing the virtual source direction, cp. [6],

$$\boldsymbol{g} = \boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}}\, \boldsymbol{y}_{\mathrm{N}}(\boldsymbol{\theta}_s). \tag{3}$$

Unlike more general layouts and right-inverse decoders, t-design Ambisonics provides

$$E = \mathrm{const,\ for\ } t \geq 2\mathrm{N}, \qquad \mathrm{and\ } \sigma_{\mathrm{E}} = \mathrm{const,\ for\ } t \geq 2\mathrm{N}+1, \text{ see Tab. 1.} \tag{4}$$

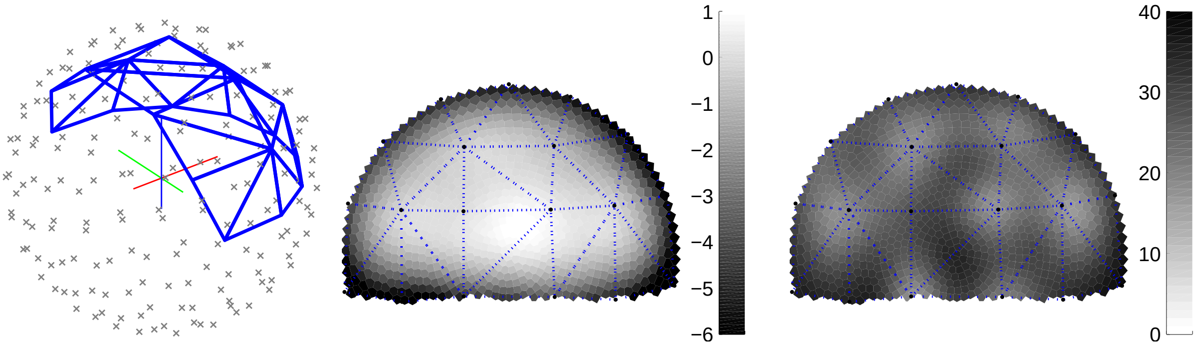
3. VIRTUAL T-DESIGN AMBISONICS VIA VBAP

Both above-mentioned approaches accomplish constant power levels for arbitrary panning angles $\boldsymbol{\theta}_s$. However, only VBAP supports arbitrary loudspeaker layouts $\{\boldsymbol{\theta}_l\}$, whereas only t-design Ambisonics supports a constant angular power-spread σ_{E}, see Tab. 1.

	VBAP	**t-design Ambi** $t \geq \mathrm{N}+1$	**virt. t-design Ambi**
$E = \mathrm{const.}$	+	+	+/-
$\vartheta_{\mathrm{E}} = \mathrm{const.}$	-	+	+/-
$\{\boldsymbol{\theta}_l\} = \mathrm{arbitrary}$	+	-	+

Table 1: Pros and Cons of presented panning approaches.

The idea of combining both approaches seems obvious to compensate for their weaknesses. Fig. 2(a) shows an 18-design Ambisonics-rig with 180 virtual loudspeakers used with N= 4. Fgs. 2(b)2(c) evaluate the result of 19 loudspeakers VBAP-rendering these virtual loudspeakers.



(a) Virtual 18-design. (b) $E = -6\ldots 1$[dB]. (c) $\sigma_{\mathrm{E}} = 25 \pm 10[^\circ]$.

Figure 2: Virtual 18-design Ambisonics: Measures E[dB] and $\vartheta_E\,[^\circ]$ for a virtual 18-design Ambisonics arrangement with $\mathrm{N} = 4$ using an $L = 180$ virtual loudspeaker rig.

4. DISCUSSION AND OUTLOOK

This paper roughly presented a novel idea to decode an optimal virtual Ambisonics-rig on arbitrary loudspeaker layouts with VBAP. This greatly simplifies Ambisonics-decoding on the one hand, and introduces a bottom limit of the angular spread into VBAP on the other hand. However, the diagram Fig. 2(b) shows that the signal power at the border of the panning domain drops by 6dB. We will show how to stabilize this shortcoming in a future publication going into more detail.

We thank our assisting colleagues at IEM, especially Thomas Musil and Hannes Pomberger.

REFERENCES

[1] N. J. A. Sloane and R. H. Hardin, *t*-designs: http://www2.research.att.com/~njas/sphdesigns/dim3/

[2] J. Blauert, “Räumliches Hören”, Hirzel, 1974.

[3] G. Theile, “Über die Lokalisation im überlagerten Schallfeld”, PhD Thesis, 1980.

[4] V. Pulkki, “Spatial Sound Generation and Perception by Amplitude Panning Techniques”, PhD Thesis, 2001.

[5] J. Daniel, “Représentation de Champs Acoustiques, Application à la Transmission et à la Reproduction de Scènes Sonores Complexes dans un Contexte Multimédia”, PhD Thesis, 2000.

[6] F. Zotter, H. Pomberger and M. Frank, “An Alternative Ambisonics Formulation: Modal Source Strength Matching and the Effect of Spatial Aliasing”, AES 126th Convention, Munich, May 2009.

[7] F. Zotter, H. Pomberger and M. Noisternig, “Ambisonic Decoding with and without Mode-Matching: A Case Study Using the Hemisphere”, 2nd Ambisonics Symposium, Paris, May 2010.

[8] P. Fellgett, “Ambisonics. Part one: General system description”, Studio Sound, 17, 1975.

[9] M. Gerzon, “Ambisonisc. Part two: Studio techniques”, Studio Sound, 17, 1975.

[10] M. Frank, F. Zotter and A. Sontacchi, “Localization Experiments using different 2D Ambisonics Decoders”, 25th Tonmeistertagung - VDT International Convention, Leipzig, Nov. 2008.

[11] S. Bertet, “Formats audio 3D hiérarchiques : caractérisation objective et perceptive des systèmes Ambionics d’ordres supérieurs”, PhD Thesis, 2009.

[12] D. B. Ward and T. D. Abhayapala, “Reproduction of a Plane-Wave Sound Field Using an Array of Loudspeakers”, IEEE Transactions on Speech ans Audio Processing, vol.9(6), Sept. 2001.

[13] J.-M. Batke and F. Keiler, “Using VBAP-derived panning functions for 3D ambisonics decoding”, 2nd Ambisonics Symposium, Paris, May, 2010.

[14] M. A. Poletti, “Three-Dimensional Surround Sound Systems Based on Spherical Harmonics”, Journal of the AES, Nov. 2005.

[15] M. Kratschmer and R. Rabenstein,“Implementing Ambisonics on a 48 Channel Circular Loudspeaker Array”, 1st Ambisonics Symposium, Graz, June 2009.