THE DESIGN AND EVALUATION OF TWO-DIMENSIONAL AMBISONIC SYSTEM

Michal Trzos
Doctoral Degree Programme(1), FEEC BUT
E-mail: xtrzos00@stud.feec.vutbr.cz

Supervised by: Jiří Schimmel
E-mail: schimmel@feec.vutbr.cz

ABSTRACT

The paper deals with the design and implementation of a third-order two-dimensional ambisonic system. The design is focused on optimization based on psychoacoustic models. The paper also presents results of performed listening test, providing a proof of functionality of the designed system.

1. INTRODUCTION

Ambisonics is a method of recording information about a sound field and reproducing it over an arbitrary number of speakers based on a decomposition of harmonic functions of the sound field. The main difference between ambisonics and other methods of surround sound recording and reproduction is the separation of the number of channels needed for recording, and the number of speakers needed for reproduction. The ambisonics is a hierarchical method in which accuracy of the localization, number of channels needed for transmission, and number of speakers needed for reproduction increases with the increasing order. Contrariwise, it is possible to lower the hardware demands just by ignoring the higher order ambisonic components, reducing the localization accuracy.

Firstly, this paper presents a brief recapitulation of ambisonic theory. Then the main issues of ambisonic system design and implementation are addressed. In the end, a listening test of the proposed ambisonic system is presented, followed by evaluation of the results.

2. AMBISONIC PANNER DESIGN

The design of an ambisonic system can be divided into two parts – ambisonic encoder and ambisonic decoder. The ambisonic encoder encodes the input signal, which could be any arbitrary monophonic signal, and its azimuth into ambisonic signals. The encoder could be substituted by an ambisonic microphone, which captures the sound field into ambisonic signals directly or with some post-processing. The ambisonic signals are then decoded by the ambisonic decoder for given speaker configuration. A block diagram of such system can be seen in Figure 1. The number of channels needed for two-dimensional system is determined by [1]

$$N = 2M + 1, \tag{1}$$

where N is the number of channels and M is the ambisonic order.

The number of speakers needed for stable two-dimensional sound field reproduction is defined by [1]

$$S = 2M + 2, \tag{2}$$

where S is the number of speakers.

2.1. AMBISONIC ENCODER

The implementation of ambisonic encoder is straightforward. The input signal S is multiplied by the matrix of cylindrical ambisonic weighting functions giving the equation [2]

$$\mathbf{B}(\theta) = S\left[\frac{1}{\sqrt{2}}, \cos\theta, \sin\theta, \cos 2\theta, \sin 2\theta, ... \cos M\theta, \sin M\theta\right]^{\mathrm{T}}, \tag{3}$$

where $\mathbf{B}$ is the vector of horizontal ambisonic components of the M^{th} order, matrix A^{T} is the transpose of matrix A, and θ is the azimuth of the virtual source.

2.2. AMBISONIC DECODER

The ambisonic decoder can be realized using the re-encoding principle. It can be written as [1]

$$\mathbf{B} = \mathbf{CS}, \tag{4}$$

where $\mathbf{C}$ is the re-encoding matrix, which is formed by the speaker azimuths and $\mathbf{S}$ is a vector of signals reproduced by speakers. It can be interpreted as how much do the signals $\mathbf{S}$ reproduced by speakers contribute to the reconstruction of ambisonic components at the listening point. To get the speaker signals, equation (4) can be rewritten as

$$\mathbf{S} = \mathbf{C}^{+}\mathbf{B} = \mathbf{DB}, \tag{5}$$

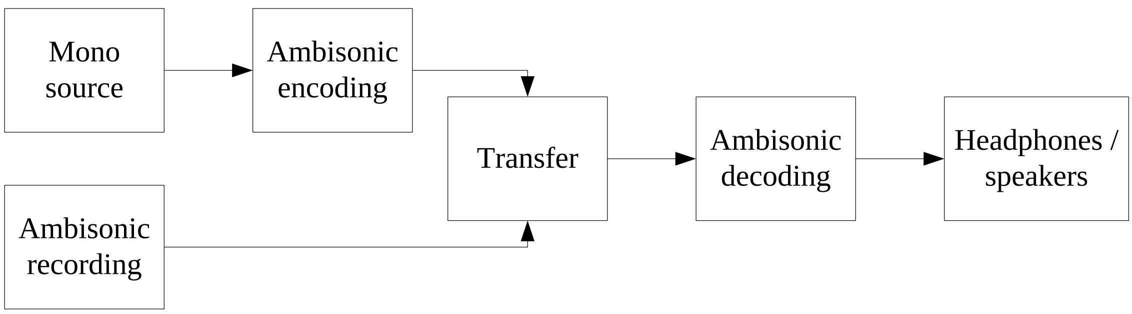

Figure 1: Block diagram of a typical ambisonic system.

where **D** is the decoding matrix and $\mathbf{C}^{+}$ is the pseudoinverse of **C**. The solution provided by (5) provides exact solution only if the speaker topology is regular [1]. The matrix pseudoinverse can be computed by the means of singular value decomposition [3].

2.3. PSYCHOACOUSTIC OPTIMIZATION

The optimization is based on the presumption that the sound localization is determined by the velocity vector [4] at lower frequencies and by the energy vector [4] at higher frequencies. However, it is not possible to match both velocity and energy vector at the same time. This can be solved by dividing the ambisonic signals to two frequency bands and decoding them separately. A diagram of such approach can be seen in Figure 2. The decoding matrix from (5) can be expanded to contain the weighting factors [2]

$$\mathbf{D} = \mathbf{C}^{+}\mathrm{diag}(\mathbf{w}), \tag{6}$$

with diag(**w**) being a diagonal vector of weighting factors computed according to table 1, where m is the actual ambisonic order, and M the highest ambisonic order.

The in-phase decoder is intended to be used when the listener is not in the centre of the listening area or with larger speaker setups such as in concert halls or cinemas.

Table 1: Weights for different types of ambisonic decoders [2].

decoder	velocity	max r_E	in-phase
weights $\mathbf{w}(m)$	1	$\cos\left(\frac{m\pi}{2M+2}\right)$	$\frac{M!^2}{(M+m)!(M-m)!}$

3. LISTENING TEST

Second order ambisonic system without psychoacoustic optimizations was used for the listening test. The system consisted of six speakers at azimuths 0°, ±60°, ±120°, and 180°. The test was performed in the acoustic laboratory with RT_{60} < 0.3 s. RT_{60} expresses the time it takes the given audio signal to decay 60 dB in a large room. Pink noise was used as the stimulus, as the human auditory system perceives it as having approximately equal

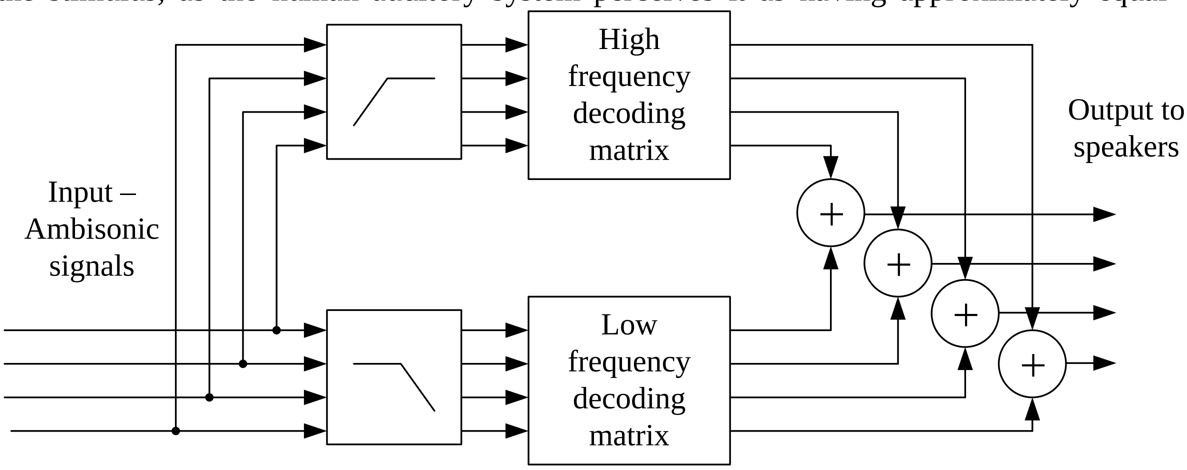

Figure 2: Block diagram of dual band ambisonic decoder [3].

magnitude on all frequencies. The stimulus is divided into four periods, each period has a rise time and fall time of 100 ms with 200 ms of unattenuated noise in between. The periods are separated by 100 ms of silence. Selection of the stimulus is based on [2].

Sixteen males ranging in age from 25 to 40 years without any hearing impairment participated in the test. The test consisted of 36 randomized measurements quantized to steps of 10°, which is consistent with the position of floor marks used to determine the azimuth of perceived stimuli, with the maximum error of this method being 5°. Each of the participants took part in half of the measurements, giving 18 results.

4. CONCLUSION

The results of the listening test are shown in Figure 3. The dots in the figure show azimuths, which were measured and the circles show the 25th quartile, 50th quartile, and 75th quartile respectively. The results, which had an error of about 180°, were ignored, as they suggest front-back confusion. It can be seen, that the most accurate localization was achieved in the front at ±45° and in the back, where the spread of the perceived azimuths is the lowest. The worst localization was achieved on the sides, between ±90° and ±120°. The accuracy of localization can be clearly seen in Figure 4, where the localization error is shown as a function of azimuth. While the predicted localization blur of a second order ambisonic system is 30° [5], the proposed system was able to perform with the maximal localization error of 25°.

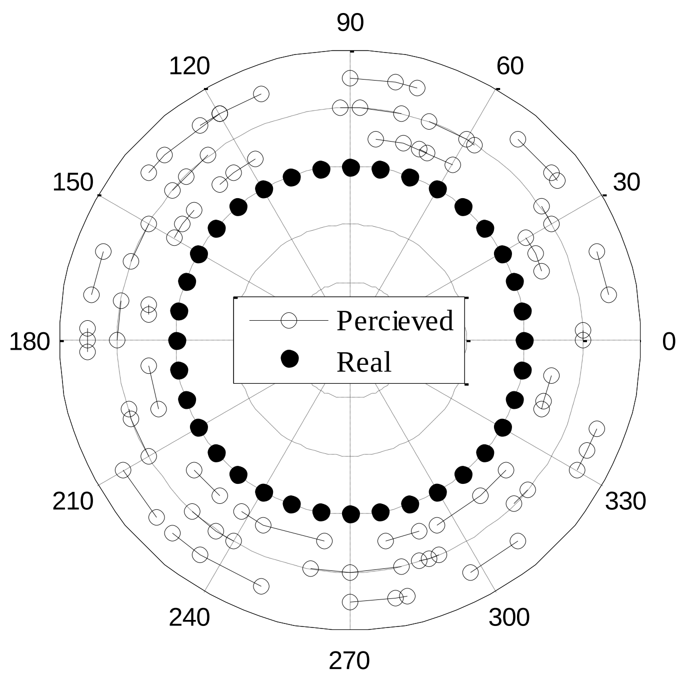

Figure 3: Real and perceived azimuths for second order ambisonic system; Symbols represent 25th quartile, median and 75th quartile [3].

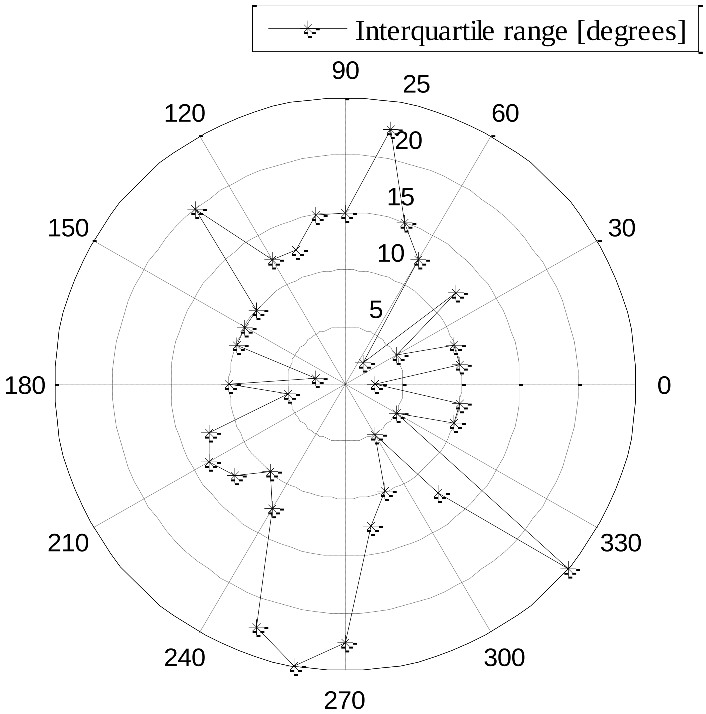

Figure 4: Interquartile range of perceived azimuths [3].

REFERENCES

[1] DANIEL, J. *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*. Paris, Francie : University of Paris, 2000. 319 p. Dissertation thesis.

[2] FRANK, M., ZOTTER, F., & SONTACCHI, A. *Localization Experiments Using Different 2D Ambisonic Decoders*. 25th Tonmeistertagung. Leipzig. 2008.

[3] TRZOS, M. *Popis a reprezentace dvourozměrných zvukových scén ve vícekanálových systémech reprodukce zvuku*. Brno: Brno University of Technology, Faculty of Electrical Engineering and Communication, 2009. 62 p. Supervisor Ing. Jiří Schimmel, Ph.D.

[4] BENJAMIN, E. M., LEE, R., HELLER, A. J. Localization in Horizontal-Only Ambisonic Systems. In *Preprints of the 121st Convention of the Audio Engineering Society*. San Francisco, 2007. p. 12.

[5] BERTET, S., DANIEL, J., PARIZET, E. LAETITA, G. WARUSFEL, O. Investigation of the Perceived Spatial Resolution of Higher Order Ambisonics Sound Fields: A Subjective Evaluation Involving Virtual and Real 3D Microphones. At *30th International Conference of the Audio Engineering* Society. Finland, Saariselkä, 2007.