Simulations of second order microphones in audio coding

Simeon Delikaris-Manias

Research visit in the Department of Acoustics and Signal Processing, Espoo Finland: Simulations of second order microphones in audio coding

Delikaris-Manias Simeon

December 14, 2011

Contents

0.1 Higher order microphones

The method of array processing involves the use of multiple microphones to receive a signal carried by propagating waves. Microphone arrays have a variety of applications such as sonars, radars and acoustic tomography. The main objective of this work is the implementation of higher order microphone arrays in DirAC. A review on previous research followed by simulations of various microphone array designs has been performed. Prototype arrays will be constructed and evaluated within a DirAC system. Simulating the various microphone arrays revealed the advantages and possible drawbacks of each design. Various higher order designs have been implemented in the past and are used as guidelines.

First order microphones are well implemented into DirAC system by using the omnidirectional component to estimate pressure and the X,Y and Z to estimate the velocity. In principle DirAC input can be extended to higher order microphones. Any microphone array can be used for the estimation of the incoming direction of the wavefront and the diffuseness of the soundfield. One way of achieving that is by combining different microphone orders and create virtual microphones of higher order. There are various models that will be tested and applied in order to compute direction of the wavefront and diffuseness of the soundfield. Some of these are single and multiple source reverberant field models, cross correlation models or beamforming techniques. The use of these models has the advantage that the analysis and synthesis part will no longer be a separate process within the DirAC analysis but one single process using methods such as the cross-spectrum analysis.

0.2 Microphone array simulations

A microphone simulator is used to determine optimal design of microphone arrays and check their performance before construction. If the simulator is accurate it can be a very cost effective method. The initial design consisted of four microphone capsules placed at the edges of cube with inner distance of 0.5cm . By using this arrangement we can calculate components of 0^{th}, 1^st and 2^nd order. These components are extracted by simple additions and subtractions between the channel. The equation system below explain how these are calculated:

$$W = A_1(P_1 + P_2 + P_3 + P_4) \tag{1}$$

$$X = A_2(P_1 - P_3) \tag{2}$$

$$U = A_3(P_1 - P_2 + P_3 - P_4) \tag{3}$$

where A_1, A_2 and A_3 are normalization coefficients. The rotated components of 1^{st} and 2^{nd} order can be obtained as $Y = -X$ and $V = -U$. Figure 1 shows the calculated components if we use four omnidirectional capsules.

The same simulation is also performed by using cardioid capsules. Results are shown in Figure 1. The advantage in this case is the lack of boost in low frequencies especially in the 0^th order component. In addition to that the components X and U are closer to the ideal ones. According to this analysis it is possible to apply beamforming. Beamforming is a technique that combines a number of microphones in order to perform spatial filtering by creating a narrow directivity pattern. The purpose of a beamformer is to capture sound from a preferred location. There are various ways to implement beamformers. The cross-spectrum analysis described in this report can be also used a a beamforming technique.

0.3 Directional Audio Coding

Various spatial sound systems have been developed during the last 3 decades as a way for representing a sound field as accurately as possible. Some of them are based on the accurate reconstruction of the soundfield i.e. ambisonics, WFS while others are perception based. A good example of a latter systems is DirAC which is developed in the Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology. DirAC is based on the following assumptions and makes use that the human auditory system can process one directin of arrival for each time-frequency bin.

- Direction of arrival of sound will transform into interaural time difference (ITD), interaural level difference (ILD), and monaural localization cues. This has been proven and discussed extensively in Blauert's book, Spatial Hearing.

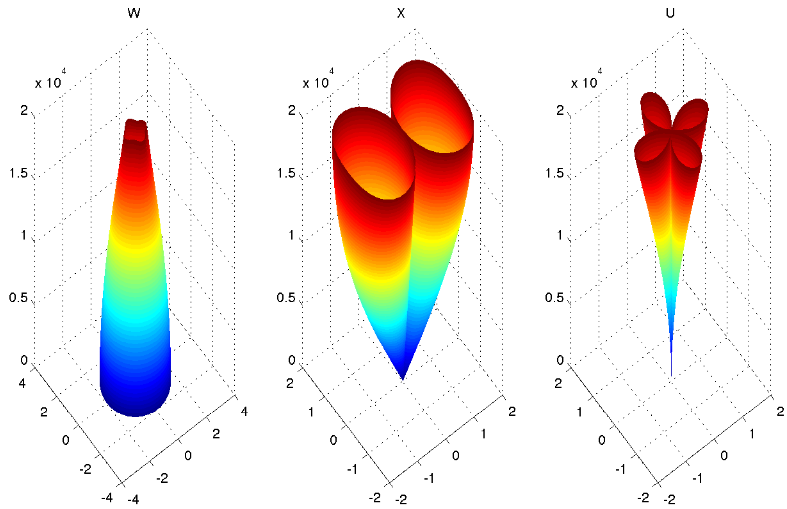

Figure 1: Unequalized components y using four omnidirectional microphones

- Diffuseness of sound will transform into interaural coherence cues [1].
- Timbre depends on the ITD, ILD, monaural spectrum and interaural coherence of sound [2].
- The direction of arrival, diffuseness, and spectrum of sound measured in a point with the temporal and spectral resolution of human hearing determines the auditory spatial image the listener is perceiving [3,4].

0.4 Higher order systems

The main idea behind a higher order implementation in DirAC is to use higher order microphone directivities instead of the standard b-format. Higher order microphones are able to be narrower and hence provide more directional information for the arrival of sound and the calculation of diffuseness. The simulations that follow are based on the analysis of higher order ideal microphones and the use of a cross-spectrum function.

0.5 Theory

Spectral analysis shows how the energy is distributed in frequency. Autocorrelation is the correlation of a function with a shifted version of itself

$$R_{xx}[m] = \sum_{n} x^*[n]x[n+m] \tag{4}$$

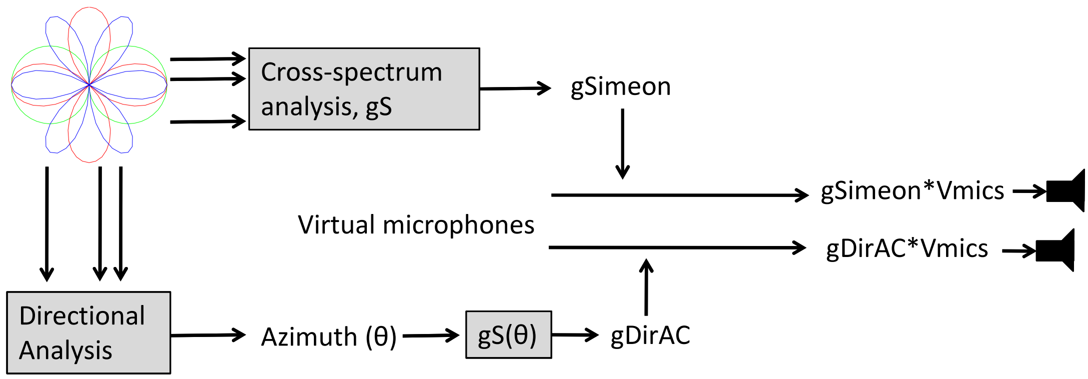

Figure 2: Analysis of Directional Audio Coding

Hence the auto-spectrum is the Fourier transform in a simplified way is:

$$G_{xx} = fft(x(t) * x(t)) \tag{5}$$

The cross-spectrum is similar but instead of using the autocorrelation, it makes use of the cross-correlation between two signals. In this work the cross-spectrum is described between two microphone inputs. For example:

$$g_{simeon} = \frac{\Re(M_x \times M_y)}{M_x(n)^2 + M_y(n)^2} \tag{6}$$

where M_x and M_y are the time-frequency analysis of two different microphone inputs. The quantity $M_x(n)^2 + M_y(n)^2$ is a normalization coefficient. $\Re$ indicates the real part and M_x^* the complex conjugate. The cross-spectrum describes the correlation between two inputs. It estimates the coherence between two input signals. If we take the real part of this quantity we get the in phase fluctuation between the inputs. A sample diagram of the process described in this chapter is shown in Figure 2

The input of this system is a microphone array from which higher order directivities are extracted by using either the analysis in the spherical harmonic domain or a simple addition and multiplication between the array's signals. In this diagram ideal microphones are shown. Specifically the microphones that are used in this example are up to 3_{rd}: an omnidirectional W, the dipoles X and Y and a quadrapole U. By taking the cross-spectrum of the X and U components a narrower directivity pattern can be obtained. The resulting directivity provided specific gain values for each angle. It is also possible to use the remaining components of 1^{st} 2^{nd} in order to construct directive microphones. A standard DirAC analysis is performed in order to calculate the active intensity and the azimuth. These quantities provide information on the direction and diffuseness of the soundfield by using the 0^{th} and 1^{st} order components. In order to synthesize this information for reproduction a loudspeakers gain needs to be computed: the g_{DirAC}. This is done by mapping the values of the azimuth to the values of calculated cross-spectrum for the specific angles. Hence we map

$$g_{XU}(Azimuth) \tag{7}$$

where g_{XY} is the cross-spectrum between the microphone inputs from X and U. Now the proposed analysis is performed where the cross spectrum between X and U is calculated as:

$$g_{simeon} = \frac{\Re(X \times U)}{X^2 + U^2} \tag{8}$$

Virtual microphones are created by using the calculated components $(X + U)$. It should be noted that any combination of components can be used in order to create virtual microphones. The combination of X and U provides a relative directional pattern which is usefull for this analysis. The calculated virtual microphones that contain the signal information are filtered through g_{DirAC} and g_{Simeon}. The analysis is based on the comparison between the existing DirAC system and proposed HODirAC by comparing:

$$g_{DirAC} \times (X + U), g_{Simeon} \times (X + U). \tag{9}$$

0.6 Summary

My research goals in the department of acoustics and signals processing in Aalto university can be summarized as follows:

1. A model to compare higher order with first order input in DirAC analysis is created. Coincident microphones were used and the theory was extended to microphone arrays.

2. A slight advantage is gained by using higher higher order microphones instead of first order.

3. Future work will support these findings with listening tests. Subjective evaluation will check the validity of a higher order microphone array system analysis in DirAC. A study on the subjective effects of the physical limitations of microphone arrays (i.e. spatial aliasing) will also be performed. The results will be compared with the original DirAC analysis to check the potential advantages of the implemented system.