**Directional Audio Coding Using Planar Microphone Arrays**

*Maneesh kumar meena, Sachin singh and Suresh Kumar Prajapat*

**Abstract**

Directional Audio Coding (DirAC) is a efficient approach to analyze spatial sound. It mainly focus on direction of arrival (DOA) and diffuseness of sound within frequency subbands. These two properties are used for spatial encoding of the sound. Three dimensional microphone array has some physical constrain, which makes this approach more acceptable. Planner microphone is used to encode sound and based on energetic sound field analysis, we are estimating DOA and diffuseness calculated.

**Introduction**

Spatial audio processing is becoming more important as the variety of possible applications for multichannel audio is constantly increasing. So, enhancing the audio quality is becoming more popular. Dirac is a more preferable reproduction approach, since, it doesn’t put any constrains on microphone and loudspeaker. Its core assumption is that interaural time difference(ITD) and interaural level difference (ILD) are perceived correctly when the DOA of sound field is correctly reproduced. Correct measure of interaural coherence (IC) leads to good reproduction of diffuseness. B-format signals are used for our purpose. In general 3D-array is used to produce B-format signal. Since, we are using 2D-microphone array, missing signals are replaced by an approximation.

**Directional Audio Coding: Analysis**

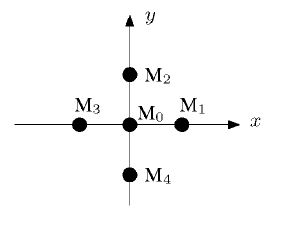
DOA of sound is determined using the active sound intensity vector and diffuseness is estimated by relating sound intensity vector to overall energy. B-format signals are made of an omnidirectional signal w(t) (represent sound pressure), and three signals x(t), y(t), and z(t) (these three represent particle velocity vector)correspond to the output of three dipoles aligned with the x-, y-, and z-direction.

n = frequency index, k = time block index of STFT. Where represent STFT of w(t) and represent STFT of active intensity vector. is mean air density and is speed of sound. DOA from is opposite of direction of . The frequency representation of the energy density

Where is the diffuseness. Here it is assumed that dipole signals are scaled by a factor of . is expectation operator it is implemented using temporal smoothing. Signal variance is little bit high which is reduced by smoothing directional information.

**Planner Microphone Array**

In this section, estimation of B-format signal will be explained. To do so, 5-omnidirectional microphones are placed in X-Y plane. M1 is placed at the origin.

**Fig1: Planner M-array**

The omnidirectional signal and the dipole signals corresponding to the -plane can be computed from the planar microphone array as follows:-

where denotes the sound pressure at the -th microphone . frequency-dependent normalization factor and is given by

where N is no of frequency bins, d is distance between 2 opposite microphones and is sampling frequency.

**3-D Sound Field Analysis**

If we consider simple trigonometry, using elevation angle p and azimuth angle s, than we can approximate as follows-

where denotes the phase of the omnidirectional signal The term corresponds to the frequency-domain magnitude of an auxiliary signal. Azimuth angle can be approximated from

Here elevation angle is calculated with an ambiguity between the upper and lower hemispheres. In common operation scenarios only positive p occur.

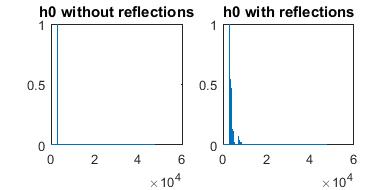
**2-D Sound Field Analysis**

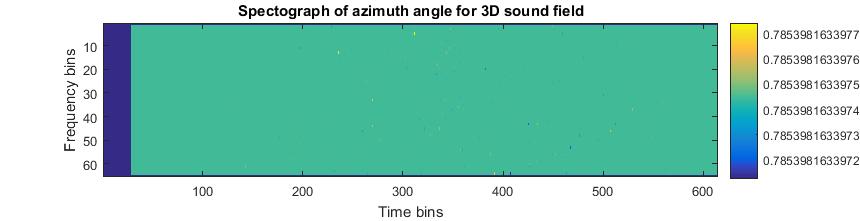
The application of this part is to calculate azimuth angle. Since, only X and Y directional signal are needed, dipole signal in X and Y are good enough to calculate azimuth angle. We simply discarded the Z-dipole signal. But, discarding the z-component of the intensity vector leads to overestimation of diffuseness. In order to calculate diffuseness replace by and by . Where

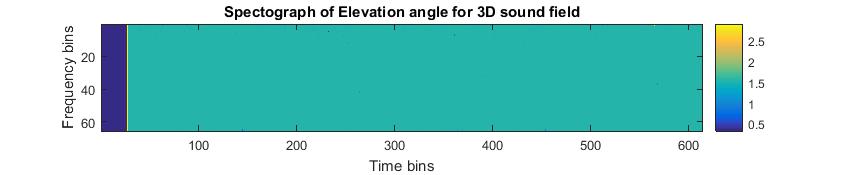
In this case, even if we increase elevation of the source, the diffuseness keeps a constant value.

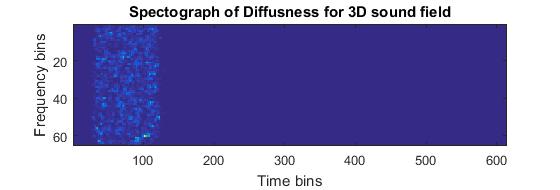
**Experimental Results**

For an experiment a source is placed at [15 15 15] in a room of [20 20 20]. M0 is placed at [2 2 2] and distance between M1-M3 and M2-M4 is 10cm. A white noise of 10db magnitude noise is coming from the source. In the diffusion diagram diffuse subband can be seen.

**Fig2: impluse response of microphone M0**

** Fig3: Azimuth angle in radian**

** Fig4: Elevation angle**

**Fig5: Diffuseness**

From the azimuth angle diagram, it can be that azimuth angle value is 0.7854 radian, which is true in our case. elevation come out to be 0.983 radian, which is closer to the actual value.

**Conclusion**

Low elevation turns out to be good for estimating azimuth angle. Dipole approximation in z-direction perform better at high elevation. Azimuth is calculated using planner waveform (neglecting z-part) which doesn’t hold for diffused sound, which is a bad thing for this method. In most of the cases Azimuth angle are predicted more correctly than elevation. We only considering the case of positive elevation. The accuracy of estimation increases for increasing elevation angle.

**References**

1. Fabian Kuech, Markus Kallinger, Richard Schultz-Amling, Giovanni Del Galdo, Jukka Ahonen and Ville Pulkki, “*Directional audio coding using planner microphone array*.”
2. V. Pulkki, “Directional audio coding: Filterbank and STFT-based design,” in 120th AES Convention, Paper 6658, Paris, May 2006.