

Directional Audio Coding (DirAC) demonstrations for teleconferencing applications

Ville Pulkki, Jukka Ahonen
Laboratory of Acoustics and Audio Signal Processing
Helsinki University of Technology
Finland

12. May 2006

The use of Directional Audio Coding (DirAC) in telecommunication is demonstrated in these files. In high-quality DirAC reproduction, sound is divided into diffuse and non-diffuse parts, and then reproduced with different strategies. In telecommunication application, the natural reproduction of reverberation is unnecessary, and overall the data rate should be minimized. Thus, diffuseness analysis is disregarded, and only directional analysis and synthesis is performed.

The B-format microphone is composed of four miniature microphones in a square with the diameter of 4 cm. The omnidirectional microphone signal is a sum of all microphones, and X and Y dipole signals are composed by subtraction of two microphone signals. The transmitted data consists of a monophonic channel with bandwidth 300 Hz to 3400 Hz and a side band of azimuth values with rate 7 kbit/s. In the decoding part, pair-wise panning is used to decode the monophonic sound at each utilized frequency band. The loudspeaker setup is this time 5.0 setup, although any other setup can be utilized. The demonstration files are found in www.acoustics.hut.fi/demos/DirAC.

File extensions:

<code>DirAC.wav</code>	DirAC processed sound with loudspeakers in 5.0 setup
<code>mono.wav</code>	Transmitted monophonic audio track

The file format is wave-format-extensible, and it should be playable with different media players.

Demo1: Two simultaneous speech sources in moderately reverberant room (English male -45° , English female $+45^\circ$)

This demonstration has been recorded in reverberant space with four miniature microphones. One loudspeaker in about 1.5 m distance reproduced male speech in direction of -45 degrees, and the another reproduced female speech in direction of $+45$ degrees. Speech sources were simultaneous.

Filenames: `talk_pm45_*`

Demo2: Three simultaneous speech sources (English male -45° , Danish male 0° , English female $+45^\circ$)

Same as above but third loudspeaker reproduced Danish male speech from direction of zero degrees.

Filenames: talk_0pm45_*

Demo3: Speech with background noise (English male -45, white noise +45)

The demonstration has been recorded in an anechoic chamber. Loudspeaker with direction of -45 degrees reproduced English male speech. White noise was played by loudspeaker with direction of +45 degrees. The sound pressure level of the white noise increases linearly when signal-to-noise ratio is initially 0 dB and -25 dB at the end. The file `talk_noise_extract.wav` demonstrates how DirAC can be used as a directive microphone technique. Only the loudspeaker channel corresponding to speech source direction is reproduced, and the speech should be understandable with lower signal-to-noise ratios than in monophonic presentation.

Filenames: talk_noise_*

Reference

Directional Audio Coding: Filterbank and STFT-Based Design—*Ville Pulkki*,; *Christof Faller*. AES 120th Convention. *Convention Paper 6658* Paris, France.