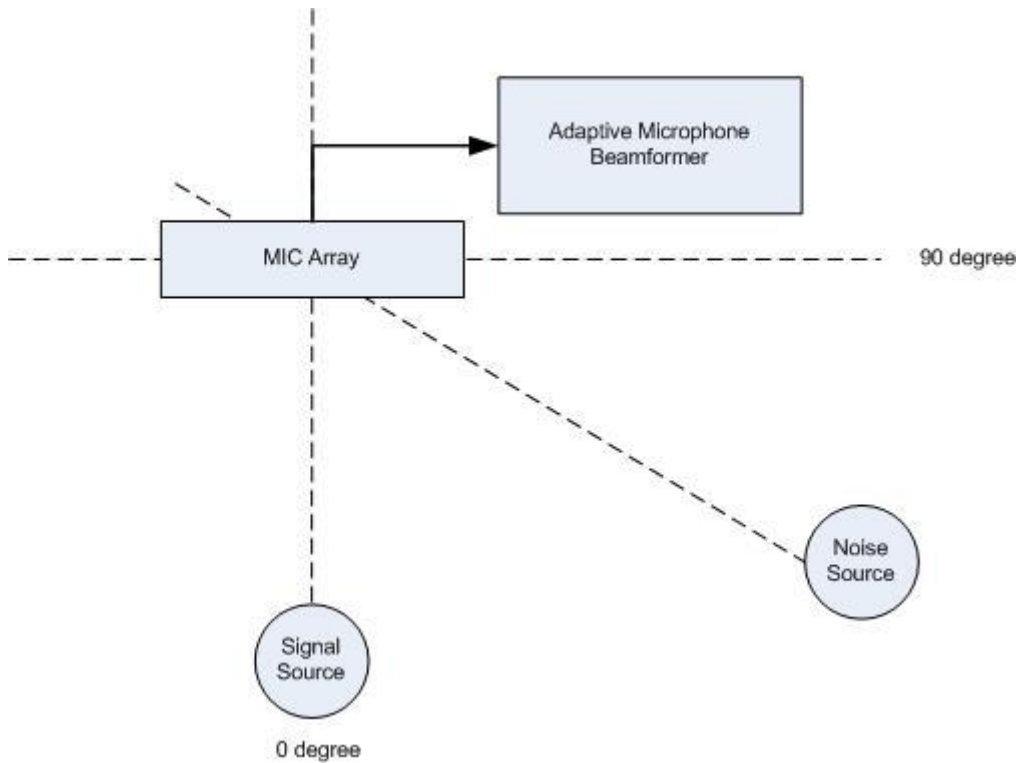# Microphone Beamforming (with Adaptive Sidelobe Canceller)

Overview

Microphone Beamforming is a technology used for discriminating sound sources by virtue of their position in space. It requires the use of an array of microphones to filter sounds coming from different directions even if they have overlapping spectral content. This can be used in the context of noise cancellation, when a desired sound source (speech) and interfering sound (noise) are originating from different positions in space.

The Microphone Beamforming component described here supports a 2 microphone linear array. It is compatible with both 8 kHz and 16 kHz sampling. The component uses adaptive sidelobe cancellation technique for cancelling interfering sound (noise) from a position different from the sound source (speech).

Configurable Properties

| Properties	Scaling	Default Value	Description
`mu`	1.x	0.1	The step size for adaptive sidelobe cancellation. Smaller value results in slow adaptation but convergence is stable. Larger value results in faster adaptation but may lead to unstable behaviour specially if adaptation happens during speech segments (due to misclassification of voice activity detector)
`INIT_PRD`	Float	.1 sec.	Initial Estimation Period (in secs) . This is the time for which initial estimation is done before the voice acitivity detector (VAD) or the sidelobe adaptation is activated.
`VAD_TH_DB`	Float	6 dB	VAD Threshold (in dB). The VAD threshold determines the margin between signal power and noise power. If the signal power is higher than the noise power by the given margin, then voice activity has been detected. The recommended range is 3 dB to 9 dB. Higher threshold may cause less voice detection in noisy background condition while lower threshold may cause false voice detection in non-stationary background noise.
`VAD_HANG_PRD`	Float	0.1 sec	VAD Hangover Period (in secs.). VAD will remain ON during the hangover period even after signal level gets lower than noise level by threshold margin. Longer hangover period will make VAD ON-OFF switching less frequent while VAD will remain ON longer after the end of speech. The recommended value is 0.1 sec to 0.5 sec.

Microphones Array and Geometry

The SF algorithm is ultimately limited by the inter-microphone distance. Spatial aliasing refers to the phenomenon where sounds arriving from different angles can be misconstrued to be arriving from the same direction. To avoid spatial aliasing the inter-microphone distance is limited by the following relationship:

$$d_{max} \leq \frac{c}{2f_{max}}$$

Where f_{max} is the highest frequency of interest, c = speed of sound (typ. 340 m/sec.)

For example, if the highest frequency of interest was 4000 Hz, the maximum inter-microphone distance would be 4.3 cm, while at 8000 Hz, it would be 2.1 cm.
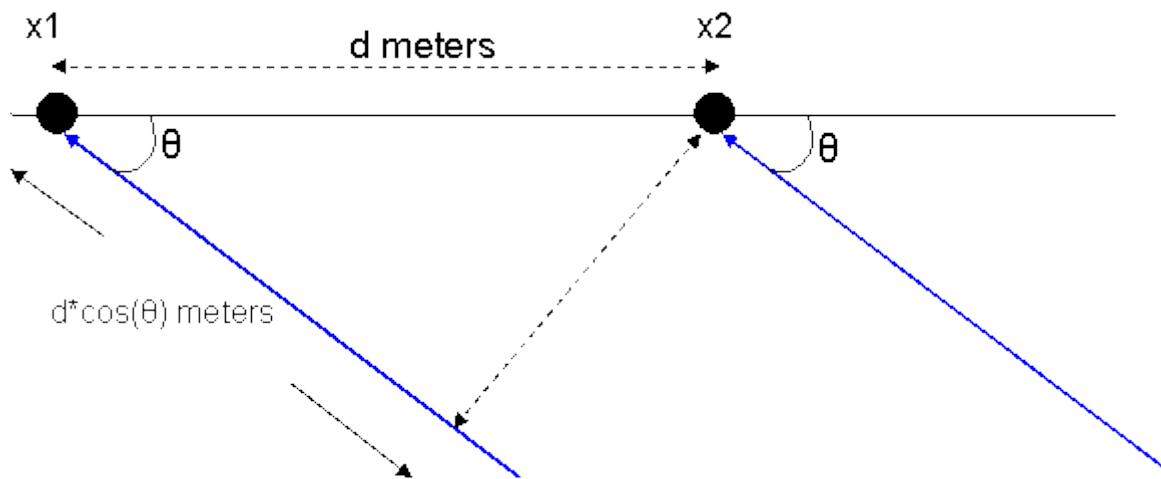
An illustration of a microphone array with 2 microphones and sound waves incident at an angle of θ.

Example Beampattern for d = 8 cm

In the following figure, we illustrate snapshots of the beampatterns at different frequencies. It can be observed that f = 2125 Hz has a well formed beam while any frequency above f = 2125 results in extra beams (sidelobes) due to spatial aliasing. The frequency 2125 is the critical frequency for d = 8 cm.

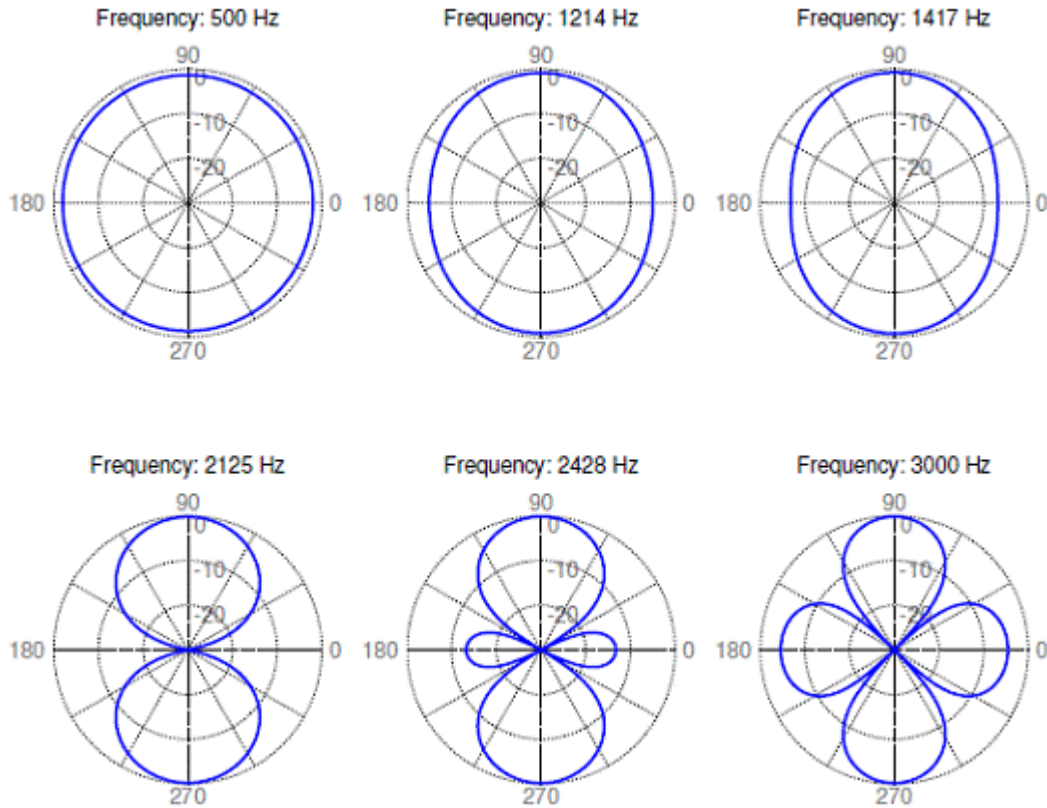Illustration of beampatterns at various frequencies

SNR performance of the 2-mic beamformer at fs = 16 KHz.

Speech = 85 dB SPL, 21 inches from the microphone array. Speech is at 0 degree relative to the microphone array

Noise = 70 dB SPL, 34 inches from the microphone array. Noise is varied from 20 degrees to 90 degrees relative to the microphone array. Noise Type = White Noise
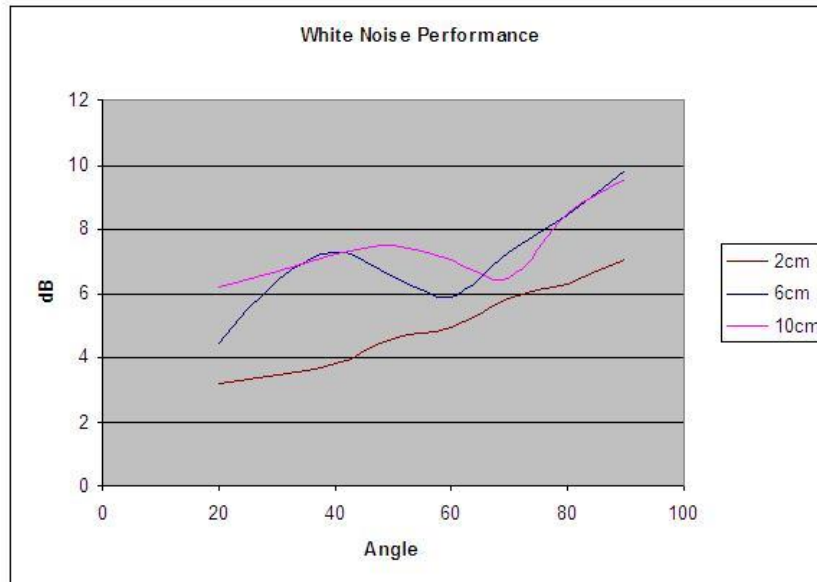
It is observed that the SNR improvement using the adaptive 2-microphone beamformer is around 7 – 9 dB.

When microphones are closely spaced:

- improvement in SNR is less (5 – 7 dB)

- however, there is no spatial aliasing (dips in the SNR plot)

- noise attenuation starts at a wider angle, so even if the signal is a bit off-centric it will not get heavily attenuated

When microphones are widely spaced:

- improvement in SNR is more (7 – 10 dB)

- however, there is spatial aliasing (dips in the SNR plot)

- noise attenuation starts at a narrower angle, so even if the signal is a bit off-centric it might get attenuated