Microphone Array[1]

Tutorial 4

In the array signal processing, typically a spreading wave is measured with the help of group of sensors in different places (spatial sampling) in order to recognize certain signal characteristics through appropriate processing of the sensor signals.
The direction to the sound source could be identified with the help of an array with only 2 microphones with the distance Δx, see fig.1. Here the algorithm Delay-and-Sum, which is very famous, can be learned. In this context one speaks sometimes about beamforming, because the microphones are interconnected in such a way, that they represent a beam, which can be swiveled electronically.

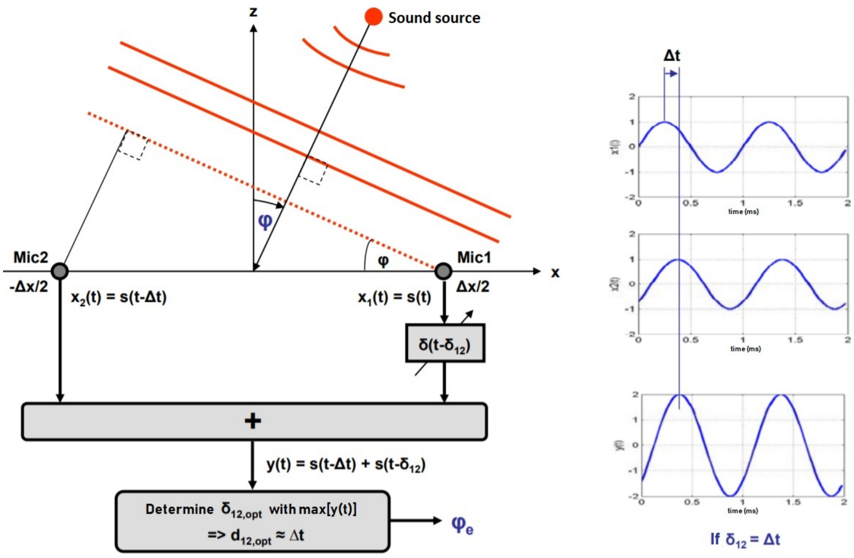

Figure 1: Sound source localization with the help of an array with 2 microphones

Look at the fig.1. If the sound source is located far enough from the microphones, in the microphone array meet "flat" acoustic waves. The angle of incidence φ can be identified indirectly through the estimation of the time delay Δt of the acoustic wave to the both microphones ($c = 343m/s$).
Time delay Δt should be in turn approximated with the help of Delay-and-Sum algorithm. The delay δ_{12} is step by step moved for a certain distance and each time the maximum value of y(t) is determined. The optimal delay value $\delta_{12,opt}$ with the maximum value of y(t) usually matches the time delay Δt, as the amplitude of y(t) becomes maximum when the both microphone signals $x_1(t)$ and $x_2(t)$ are added together temporally coherent, in other words, cophasal.
Based on the trigonometrical equations, the time delay Δt equals:

$$\Delta t = \Delta x \cdot sin(\varphi)/c \tag{1}$$

The distance between microphones Δx should be actually as large as possible in order to make the time delay Δt bigger, especially for the angle of incidence φ close to 0°.

[1]original: `https://home.zhaw.ch/~rumc/dsv2/praktikum/dsv2p6arraydsp/dsv2p6arraydsp.pdf`

For a given distance between the microphones Δx the time delay is maximum when the angle of incidence is $\varphi = \pm 90°$. The sound source is located then on the x-axes.
Time delay should not be however larger, than the half of the period T_{min} of the highest in the sound presented frequency f_{max}, this means:

$$\Delta t < T_{min}/2 = 1/(2 \cdot f_{max}), \tag{2}$$

because otherwise the time delay and with this the angle of incidence cannot be clearly identified.

Example

If in the fig.2 the time delay $\Delta t = T_{min}/2$ is reached, it cannot be clearly determined if the signal $x_2(t)$ against $x_1(t)$ comes ahead or behind, in other words if the angle of incidence φ_0 is positive or negative.

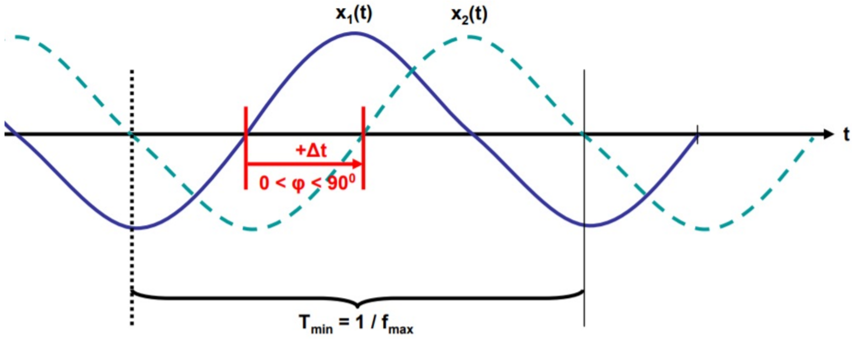

Figure 2: Microphone signals $x_1(t)$ and $x_2(t)$ (lagging)

It follows from the equations 1 and 2, that for the highest clearly located sound frequency f_{max} in the angle range $|\varphi| \leq \varphi_{max} \leq 90°$ it holds:

$$f_{max} < c/(2 \cdot \Delta x \cdot sin(\varphi_{max})), \tag{3}$$

So, the smaller the distance between the microphones Δx, the higher sounds can be determined.