A Microphone Array Tutorial

Iain McCowan

August 20, 2004
A microphone array is an array of microphones:
- Multiple microphones combined to act as a single device.

A microphone array can be used to:
- Discriminate between sounds based on direction.
  e.g. Input, Speaker 1, Speaker 2.
- Locate sound sources.

A microphone array provides hands-free/distant acquisition.
- Less constraining on users.
- Can be used for surveillance.
The Wave Equation

\[ \nabla^2 s(t, r) = \frac{1}{c^2} \delta^2 s(t, r) \]

where:

- \( \nabla^2 \) is the Laplacian operator
  (For Cartesian coordinates, \( \nabla^2 f = \frac{\delta^2 f}{\delta x^2} + \frac{\delta^2 f}{\delta y^2} + \frac{\delta^2 f}{\delta z^2} \))
- \( c \) is the speed of propagation, which depends on the type and temperature of the fluid.
- \( r \) is the position vector, \( r = [x \ y \ z]^T \).
- \( s \) is the amplitude of the wave (e.g. sound pressure level).
Sound propagates through a fluid (e.g. air) as a longitudinal pressure wave, with speed $c \approx 340 \text{ms}^{-1}$ in air at 20$^\circ$C.

**Plane Waves**

For a plane wave, the solution to the wave equation takes the form:

$$s(f, r) = s(f)e^{-jk \cdot r}$$

where:

$$k = \frac{2\pi f}{c} \begin{bmatrix} \sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta \end{bmatrix}$$

$(\theta, \phi)$ is the direction of propagation, and $r$ is a position vector relative to the sound source location.
An aperture is a spatial region that transmits or receives propagating waves, e.g. an antenna for EM waves, a hole in an opaque screen for optics, a microphone for acoustics.

The aperture sensitivity function $w(f, r)$ gives the response as a function of position on the aperture.
Aperture Directivity

- The received signal at a given point on the aperture is:
  \[ s(f)w(f, r)e^{-jk \cdot r} \]

Aperture Response

- The total response of the aperture to signal \( s(f) \) is thus:
  \[ D(f, k) = \int_{V} w(f, r)e^{-jk \cdot r} dr \]

- The aperture response is also known as the directivity function, as it gives the response as a function of the direction of arrival of the plane wave (recalling that \( k = g(f, \theta, \phi) \)).
Fourier Transform Relationship

From Time to Frequency Domain

The Fourier Transform operation transforms domain from $t \rightarrow \omega$:

$$X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt,$$

commonly denoted as:

$$X(\omega) = \mathcal{F}_t\{x(t)\}.$$

note $\omega = 2\pi f$
Fourier Transform Relationship

From Time to Frequency Domain

The Fourier Transform operation transforms domain from $t \rightarrow \omega$:

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt,$$

commonly denoted as:

$$X(\omega) = \mathcal{F}_t\{x(t)\}.$$ 

From Sensitivity to Directivity

At a fixed frequency, the directivity is:

$$D(k) = \int_V w(r)e^{-jkr} dr$$

which can be denoted:

$$D(k) = \mathcal{F}_r\{w(r)\}.$$ 

note $\omega = 2\pi f$
For a linear aperture $r = x$ of length $L$, having uniform sensitivity, $w(f, r) = \text{rect}(x/L) \Rightarrow D(f, r) = L \text{sinc}(k_x L)$.
This directivity can be approximated by a discrete aperture, which spatially samples the continuous aperture.

\[ \text{s}(f) \]

\[ w_n(f, r_n) \]

\[ L \]

\[ |D(f, k)| \]
A microphone array is a discrete receiving aperture.
A plot of the directivity function over different angles of arrival is known as the directivity pattern.

**Directivity of Continuous Aperture**

\[ D(f, k) = \int_V w(f, r) e^{-jk \cdot r} \, dr \]
Microphone Array Directivity Pattern

- A microphone array is a **discrete receiving aperture**.
- A plot of the directivity function over different angles of arrival is known as the **directivity pattern**.

### Directivity of Linear Array

\[ D(f, k) = \sum_{n=0}^{N-1} w_n(f) e^{-jk \cdot r_n}, \]

where there are \( N \) microphones and \( r_n \) is the location of the \( n^{th} \) microphone.
A microphone array is a discrete receiving aperture.

A plot of the directivity function over different angles of arrival is known as the directivity pattern.

Horizontal Directivity of Uniform Linear Array

\[ D(f, \phi) = \sum_{n=0}^{N-1} w_n(f) e^{-j \frac{2\pi f}{c} nd \cos \phi}, \]

where there are \( N \) microphones, \( d \) is the uniform inter-element spacing, and \( \theta = \pi/2 \).
Varying the Number of Microphones

- Varying the number of sensors, \( N \), for a given array length decreases the sidelobe level \((f=1 \text{ kHz}, L=0.5 \text{ m}, w_n(f) = \frac{1}{N})\):

Directivity Pattern for Varying \( N \)
Varying the Length of the Array

- Varying the array length, $L = Nd$, for fixed $N$ decreases the main lobe width ($f = 1$ kHz, $N = 5$, $w_n(f) = \frac{1}{N}$):

**Directivity Pattern for Varying $L$**

![Graph showing directivity pattern for varying array length](image)
Variation with Frequency

- Varying the frequency of interest $400 Hz \leq f \leq 3000 Hz$
  ($N=5, d=0.1 \text{ m}, w_n(f) = \frac{1}{N}$):

Directivity Pattern for Varying $f$
Spatial Aliasing: Analogous to the Nyquist frequency in temporal sampling, we have a restriction on minimum spatial sampling rate. Linear arrays require inter-element spacing $d < \frac{\lambda_{\text{min}}}{2}$ to avoid copies of the main lobe appearing in the directivity pattern, where $\lambda_{\text{min}}$ is the smallest wavelength of interest (corresponding to highest frequency).

Symmetry of Directivity Pattern: For a linear array, the directivity pattern is symmetrical about the array axis.
The term $w_n(f)$ represents a filter applied to microphone $n$.

The analysis so far has assumed uniform, frequency invariant, microphone filters $w_n(f) = \frac{1}{N}$.

In general, we can design filters to give a desired steering and shaping of the directivity pattern.

This is referred to as microphone array beamforming.
The term $w_n(f)$ represents a filter applied to microphone $n$.

The analysis so far has assumed uniform, frequency invariant, microphone filters $w_n(f) = \frac{1}{N}$.

In general, we can design filters to give a desired steering and shaping of the directivity pattern.

This is referred to as microphone array beamforming.
Delay-Sum Beamformer

- For the uniform linear array, using a simple time-delay filter:
  \[ w_n(f) = \frac{1}{N} e^{\frac{j2\pi fc}{c} nd \cos \phi_s} \]

  will lead to a horizontal directivity function:
  \[ D(f, \phi) = \frac{1}{N} \sum_{n=0}^{N-1} e^{-j\frac{2\pi fc}{c} nd(\cos \phi - \cos \phi_s)} \],

  steering the main lobe of the directivity pattern to the source direction \( \phi_s \).

- This is the well-known delay-sum beamformer.
Delay-sum is the simplest beamformer: it just ensures that the response maximum occurs for a given direction.

Many more sophisticated beamforming techniques exist, which differ in the design criteria for the beamforming filters \( w_n(f) \).

Examples include:

- the **superdirective** beamformer, which maximises the gain in the desired direction while minimising average gain over all other directions.
- **Adaptive** beamformers, which dynamically update filters to minimise the power from localised noise sources.
Delay-Sum vs Superdirective Beamforming

- For a circular array \((N=8, \text{ radius } = 10 \, \text{cm})\)

Delay-sum Beamformer

Superdirective Beamformer

For each frequency (250 Hz, 500 Hz, 1000 Hz, 2000 Hz), the diagrams illustrate the directivity pattern of both delay-sum and superdirective beamformers.