2. Methodological Foundations

1. Robot heads and acoustic laboratories
2. Binaural Processing Pipeline
3. Continuous-time Fourier transform
4. Continuous short-time Fourier transform
5. Discrete-time signals
6. Discrete short-time Fourier transform
7. Spectrogram of an acoustic signal
8. Cross-correlation
9. **Relative transfer function**
10. Binaural features

Using Two Microphones

- We now address the problem of defining some kind of **binaural features** that can be used to characterize sounds and to extract source parameters.
- Propagation model:

$$
\begin{aligned}
x_1(t) &= h_1(t) \star s(t) \\
x_2(t - t') &= h_2(t) \star s(t)
\end{aligned}
$$

- The signal x_2 is delayed by t' with respect to x_1.
- In the spectral domain we have:

$$
\begin{aligned}
X_1(f, l) &= H_1(f) S(f, l) \\
X_2(f, l) e^{\alpha(f, l, t')} &= H_2(f) S(f, l)
\end{aligned}
$$

Binaural Transfer Function

- The **relative transfer function** (RTF):

$$H_{\mathrm{rel}}(f) = \frac{H_2(f)}{H_1(f)}$$

- The RTF encodes binaural information:

$$H_{\mathrm{rel}}(f) \approx \frac{|X_2(f,l)|}{|X_1(f,l)|} e^{\alpha(f,l,t')}$$

Relative Transfer Function

- The RTF is a complex-valued function.
- The magnitude of the RTF encodes the ratio of the left and right magnitudes for each frequency f and frame l
- The argument of the RTF depends on the time-delay between the two microphone signals:

$$H_{\mathrm{rel}}(f, l) = \mathcal{H}_{\mathrm{rel}}(f, l) e^{\alpha(f,l,t')}$$

Estimation of the RTF

- We consider the power spectral density and cross-power spectral density of the two microphone signals:

$$\Phi_{x_1,x_1}(f) = \frac{1}{L} H_1(f) H_1^*(f) \sum_{l=1}^{L} |S(f,l)|^2$$

$$\Phi_{x_1,x_2}(f) = \frac{1}{L} H_1(f) H_2^*(f) \sum_{l=1}^{L} |S(f,l)|^2$$

- The RTF can be estimated with:

$$\hat{H}_{\mathrm{rel}}(f) \approx \frac{\Phi_{x_1,x_2}(f)}{\Phi_{x_1,x_1}(f)}$$

A Robot Head

Binaural head

Sound sources

The Head Related Transfer Function

- If there are no reverberations (anechoic environment) the impulse responses h_1 and h_2 model the filtering effects of the head and of the pinnae.
- In this case, the RTF is called the **head related transfer function** (HRTF):

$$H_{\text{head}}(f) = \frac{H_2(f)}{H_1(f)}$$

- The HRTF can be estimated using the same principle as the RTF, for each frame l:

$$\hat{H}_{\text{head}}(f, l) \approx \frac{\phi_{x_1,x_2}(f, l)}{\phi_{x_1,x_1}(f, l)}$$

Session Summary

- Relative transfer function
- Estimation of RTF
- Back to a robot head
- Head-related transfer function