Binaural Hearing for Robots

Sound-Source Localization

Radu Horaud

informatiques mathématiques Inria

Binaural Hearing for Robots

1. Introduction to Robot Hearing
2. Methodological Foundations
3. **Sound-Source Localization**
4. Machine Learning and Binaural Hearing
5. Fusion of Audio and Vision

Radu Horaud

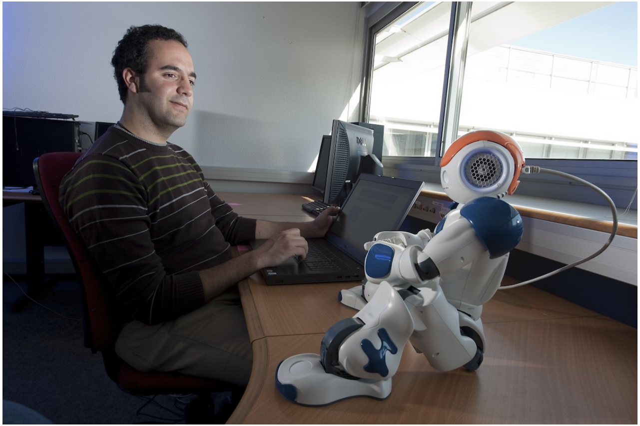
3. Sound-Source Localization

1. Time difference of arrival (TDOA)
2. Estimation of TDOA by cross-correlation
3. Estimation of TDOA in the spectral domain
4. The geometry of two microphones
5. Direction of arrival
6. Using more than two microphones
7. Embedding the microphones in a robot head
8. Learning a sound propagation model
9. Predicting direction of a sound with a robot head
10. Example of sound direction estimation

3. Sound-Source Localization

1. **Time difference of arrival (TDOA)**
2. Estimation of TDOA by cross-correlation
3. Estimation of TDOA in the spectral domain
4. The geometry of two microphones
5. Direction of arrival
6. Using more than two microphones
7. Embedding the microphones in a robot head
8. Learning a sound propagation model
9. Predicting direction of a sound with a robot head
10. Example of sound direction estimation

The Basic Setup



Notations and Definitions

We start by considering a simple situation: two microphones and one sound source:

- $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$: 3D **known** positions of microphones m_1 and m_2.
- $\boldsymbol{S}$: 3D **unknown** position of the sound source
- $s(t)$: **unknown** acoustic signal emitted by the source
- $x_1(t), x_2(t)$: **observed** acoustic signals recorded with the two microphones

Microphone Recordings

$$x_1(t) = s(t - \tau_1) + n_1(t)$$
$$x_2(t) = s(t - \tau_2) + n_2(t)$$
$$\tau = \tau_1 - \tau_2$$

- τ_1 and τ_2: times of arrival from the source to the microphone(s)
- $n_1(t)$ and $n_2(t)$: additive noise
- τ: time difference of arrival (TDOA)
  between the two microphones

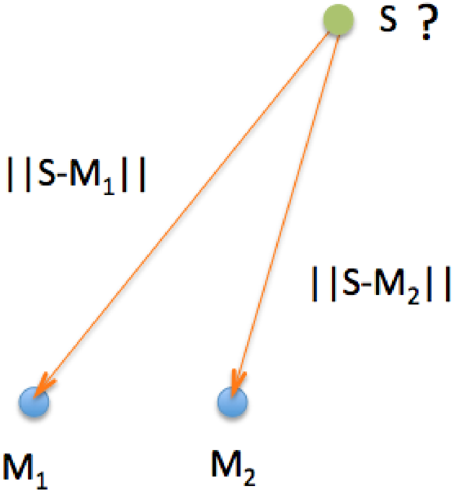
Direct Sound Propagation

We assume a direct propagation path:

$$\tau_i = \frac{\|\boldsymbol{S} - \boldsymbol{M}_i\|}{\nu}$$

- $\| \cdot \|$ is the Euclidean norm
- $\nu = 344$ m/s is the speed of sound

Geometry



Time Difference of Arrival

- The signal reaches the two microphones with a delay:

$$\begin{aligned}\tau &= \tau_1 - \tau_2 \\ &= \frac{\|\boldsymbol{S} - \boldsymbol{M}_1\|}{\nu} - \frac{\|\boldsymbol{S} - \boldsymbol{M}_2\|}{\nu}\end{aligned}$$

- τ is the:
- TDOA (**time difference of arrival**) or
- ITD (**interaural time difference**).

Session Summary

- Basic setup: two microphones and one source.
- Binaural recording with the direct propagation model.
- Geometric representation.
- Time difference of arrival (TDOA).