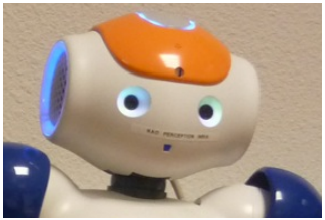
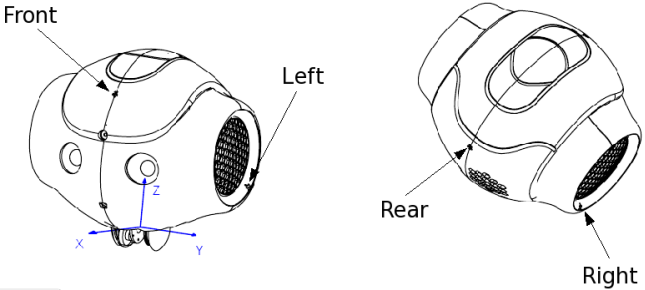
3. Sound-Source Localization

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

Microphones Embedded in a Robot Head





Sound Localization with the NAO Robot

NAO has four non-coplanar microphones

The microphones signals are corrupted by various sources of noise:

- environmental noise
- reverberations
- robot noise (motors, fans, etc.)

A robust localization method is needed.

Far-field Assumption

The sound source, ***S*** is parameterized by its direction with respect to the robot head:

- Euclidean coordinates: x, y, z.
- Spherical coordinates: $r \sin\alpha \sin\beta,\ r \sin\alpha \cos\beta,\ r \cos\alpha$
- We make a **far field** assumption: r is large with respect to the head size:

We only estimate α (azimuth) and β (elevation), the source direction.

TDOA Estimation Between Microphone Pairs

- The TDOA between two microphone pairs can be estimated using cross-correlation (CC), normalized cross-correlation (NCC), generalized cross-correlation (GCC), for example:

$$Q_{ij}(t') = \mathrm{NCC}(x_i(t), x_j(t - t'))$$

- For a **known** sound direction (α, β), one can estimate the corresponding TDOA:

$$\tau_{i,j} = \underset{t'}{\mathrm{argmax}}\, Q_{ij}(t')$$

A More General Sound Propagation Model

These $(\alpha, \beta) \leftrightarrow \tau$ input-output values do not necessarily verify the direct-path propagation model:

$$\tau_{i,j} \stackrel{?}{=} \frac{1}{\nu} \left(\|\boldsymbol{S}(\alpha, \beta) - \boldsymbol{M}_i\| - \|\boldsymbol{S}(\alpha, \beta) - \boldsymbol{M}_j\| \right)$$

More generally, we seek a function *f* that **verifies** such a relationship between known input and known output:

$$\tau = f(\alpha, \beta)$$

Session Summary

- Nao has four microphones
- The microphones capture noise coming form the electronics inside the head
- The head plays the role of an acoustic filter
- The propagation model is approximated by a linear regression function