Digital Speech and Audio Processing

Vrije Universiteit Brussel

DSAP Theory course

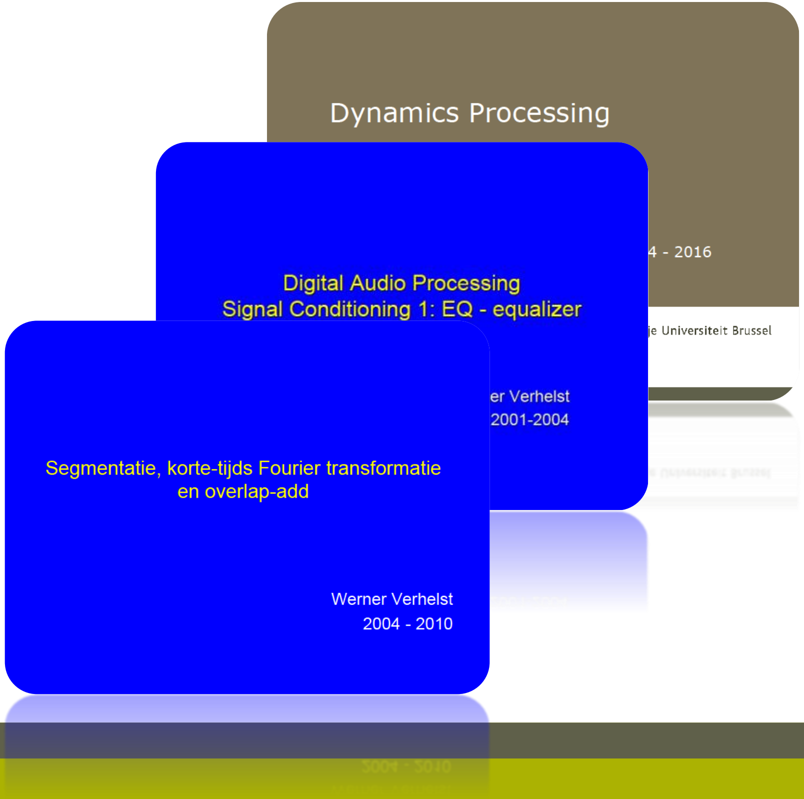
First…

A practical message from prof. Verhelst…

DSAP Theory course

Make sure you look into the theory course slides and notes BEFORE the exams!

- There’s a lot of information
- Possibility to ask questions



Digital Speech and Audio Processing

Introduction to project 2

Vrije Universiteit Brussel

Who are we?

- **Henk Brouckxon**

E-mail: hbrouckx@etro.vub.ac.be

Room: Ke.2.13, located in building Ke
Tel: +32 2 629 2939

Practical arrangements

30h practical exercises

- Week 22-23, 25 - 28: Wednesday 10:00-12:00 @ K.3.125
- Week 31 – 36: Wednesday 09:00-12:00 @ K.3,115

! Work in groups of 4 persons !

Practical arrangements

3 projects

- Project 1: Stereo to surround sound upmixing (week **23**, 25, 26)
- Project 2: surround sound reconstruction (week **27**, 28, 31, 32)
- Project 3: TBD (week **33**, 34, 35)
- Evaluation: week **36**

Week 23,27,28,32

- Explanation of the projects
- Attendance obligatory

Other weeks

- PC rooms available (attendance not obligatory)
- Assistance: available on request (send mail to make appointment)

Evaluation

- Send code + files (email) on May 19th (week 35)

- Evaluation May 24th (week 36)
    - 15 min presentation
    - 5 min demo
    - 10 min questions

Project 1:
Stereo to surround sound upmixing

Vrije Universiteit Brussel

Project 1: Some remarks

- Equation (8) for the window is wrong

$$\xcancel{w(n)=\begin{cases}\sin\left(\dfrac{\pi(n+\frac{1}{2})}{2(N-M)}\right), & 0\le n\le N-M-1\\ 1, & N-M\le n\le M-1\\ \sin\left(\dfrac{\pi(N-n-\frac{1}{2})}{2(N-M)}\right), & M\le n\le N-1\end{cases}}\qquad (8)$$

➔ replace by a suitable window (constant overlap-add requirement for synthesis !!)

Project 1: Some remarks

- Error in equation (9)

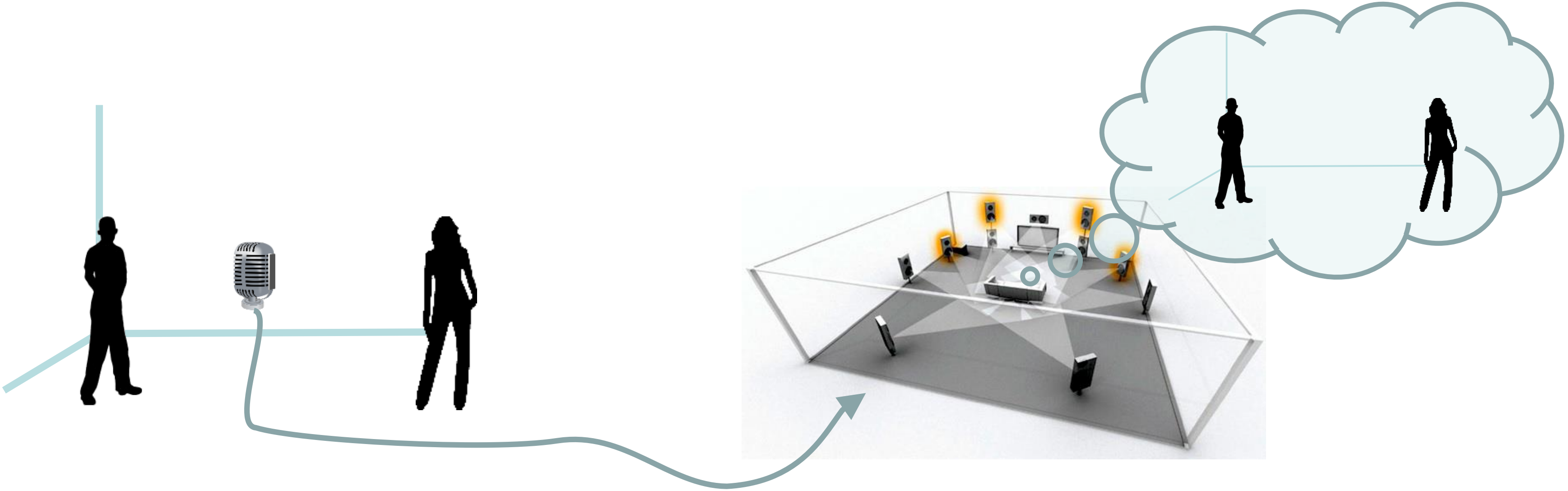
$$y(n)=\mathbf{w}^T(n)\mathbf{x}(n)=\cancel{\mathbf{w}(n)\mathbf{x}^T(n)}, \tag{9}$$

Project 2:
Surround sound scene reconstruction

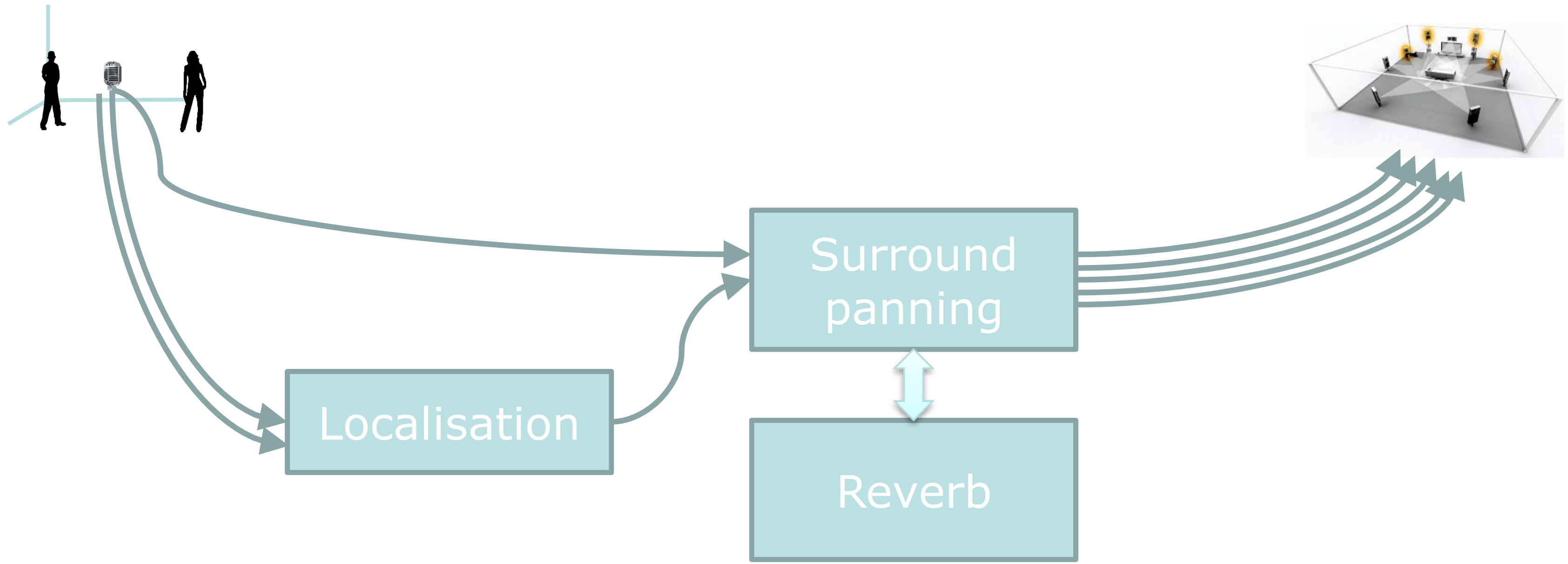
Vrije Universiteit Brussel

Surround sound scene reconstruction

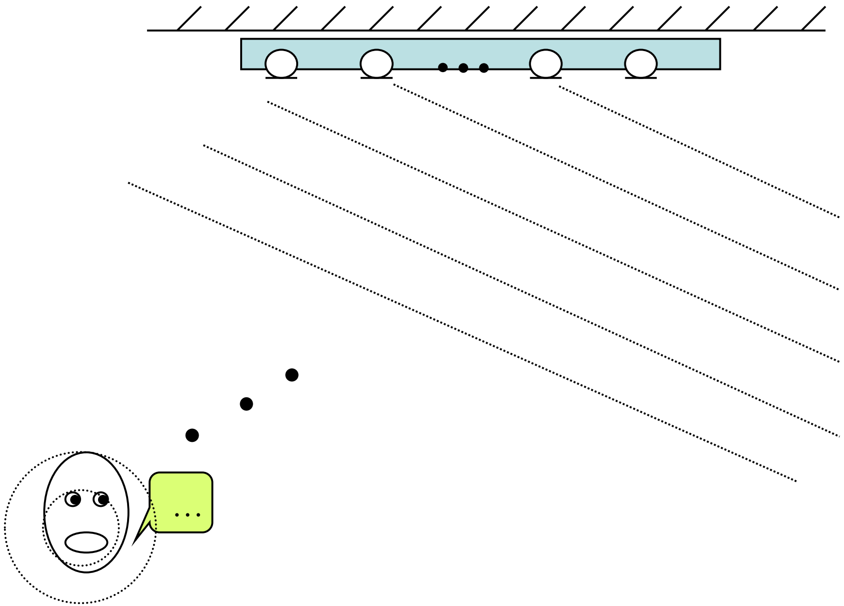
Goal: record and reconstruct a surround sound scene



Surround sound scene reconstruction



Sound source localisation



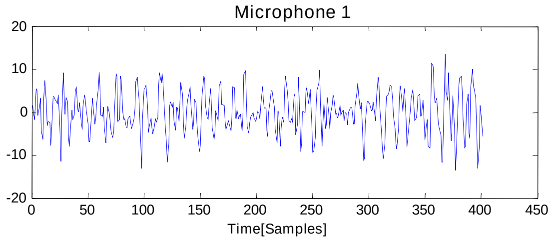
Signal, picked up by microphone i at location $\vec{r}_i$:

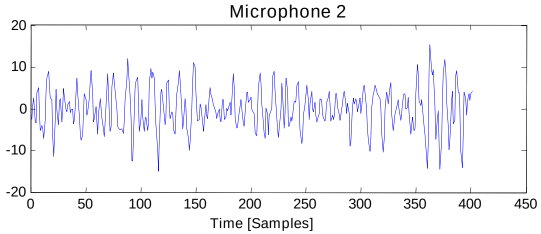
$$x_i(t) = A_i(\vec{r}_s - \vec{r}_i).s(t - \Delta_t(\vec{r}_s - \vec{r}_i))$$

By measuring the time difference on arrival (TDOA) of the sound at different microphones, a microphone array can localise sound sources

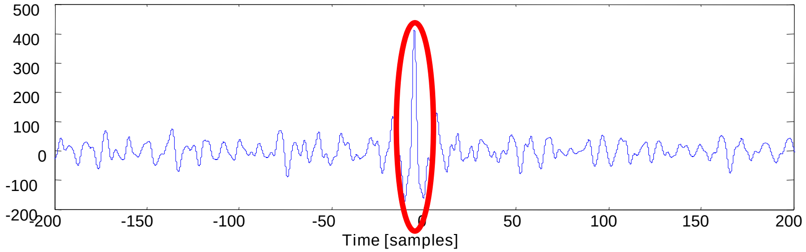
$$TDOA(i, j) = \Delta_t(\vec{r}_s - \vec{r}_j) - \Delta_t(\vec{r}_s - \vec{r}_i)$$

TDOA Measurement





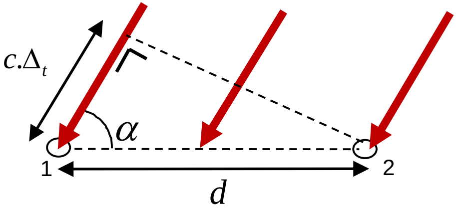
(Generalised) Cross Correlation:



Time-domain maxima in the GCC indicate possible TDOA values

From TDOA to Location

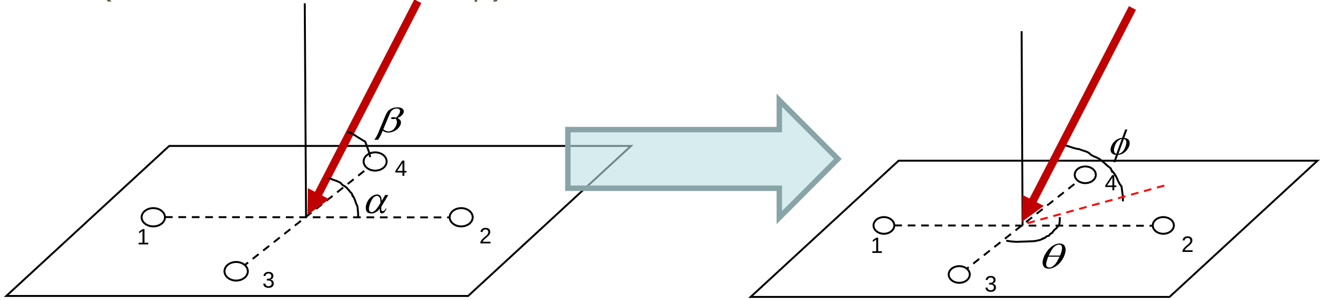
- Based on TDOA, a microphone pair (1,2) allows determination of the incidence angle relative to the pair's axis:



$$\alpha = \arccos\left(\frac{c.\Delta_t}{d}\right) \qquad \Delta_t = TDOA$$

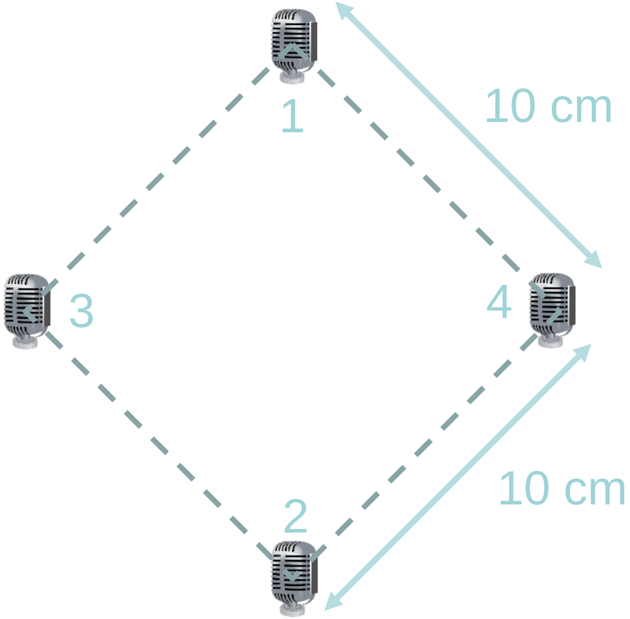
$$c = 340 m/s \quad \text{speed of sound}$$

- Combine two orthogonal pairs (1,2) and (3,4) to obtain spherical coordinates (azimuth θ / elevation ϕ):

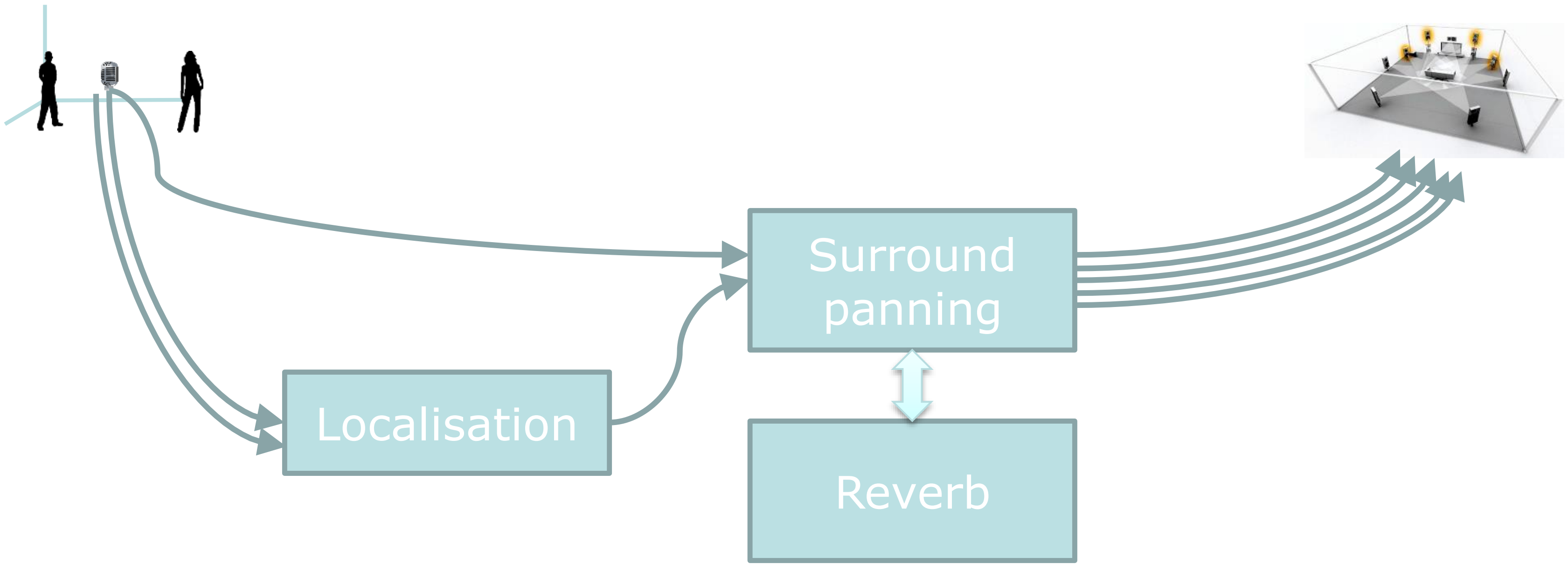


Sound source localisation

The array:

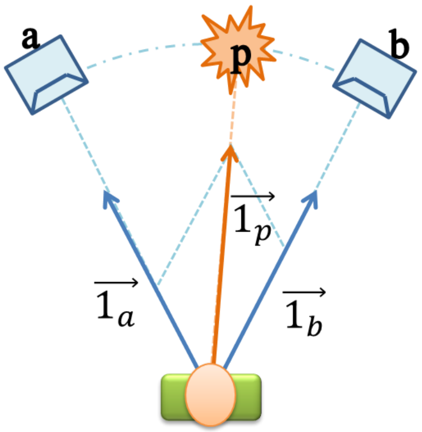


Surround sound scene reconstruction



Sound scene reconstruction

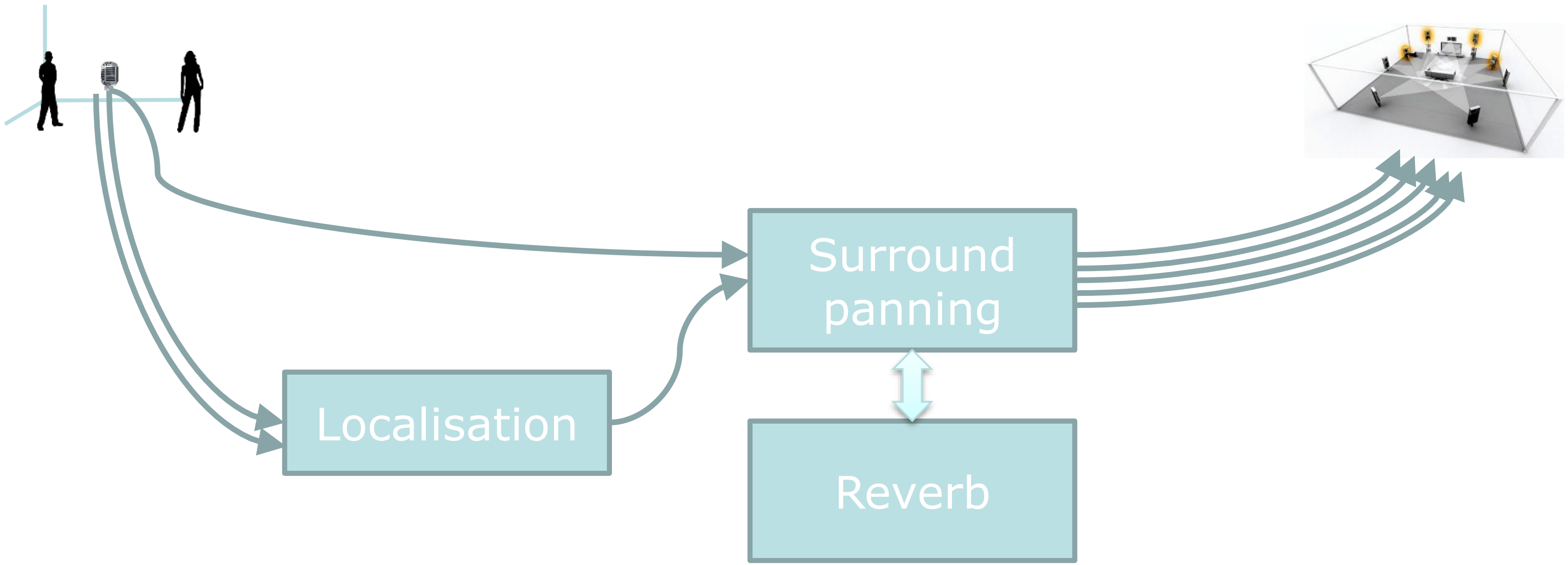
1. Surround panning: Vector-Base Amplitude Panning



- Virtual sound source p, reproduced in a position between two real loudspeakers a and b
- Gains g_a and g_b, applied to the signal for loudspeakers a and b (normalise g_a, g_b):

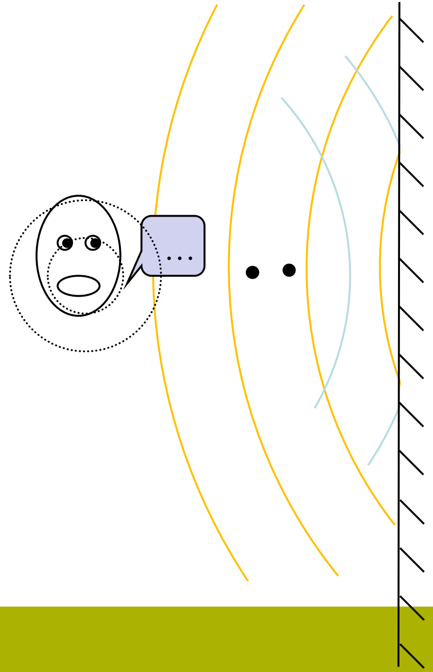
$$\begin{cases} \vec{1}_a = (x_a, y_a) \\ \vec{1}_b = (x_b, y_b) \\ \vec{1}_p = (x_p, y_p) \end{cases} \Rightarrow \quad (g_a \quad g_b) = (x_p \quad y_p) \begin{pmatrix} x_a & y_a \\ x_b & y_b \end{pmatrix}^{-1}$$

Surround sound scene reconstruction



Sound scene reconstruction

2. Virtual room ambience: reverberation

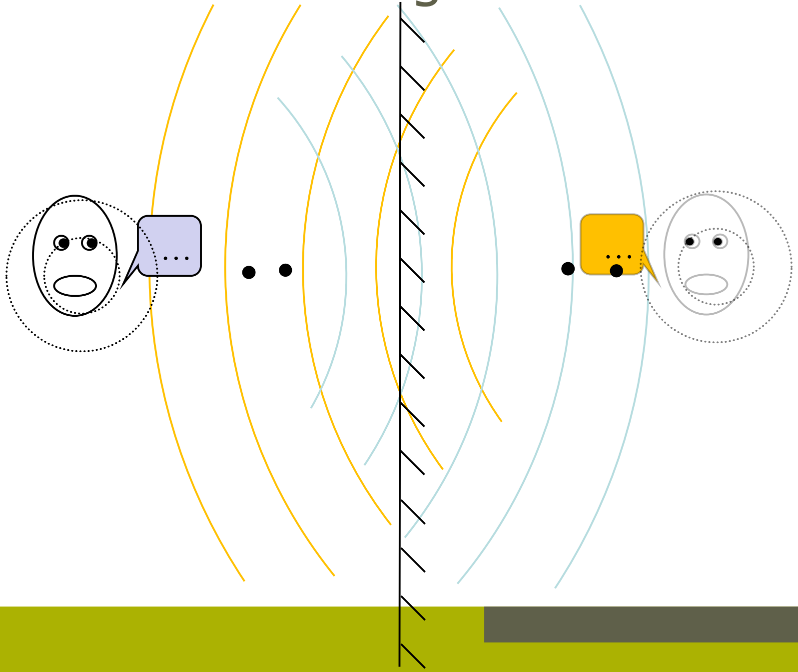


Sound waves reflect on the walls and cause reverberation

delayed and attennuated repetitions of the original sound source

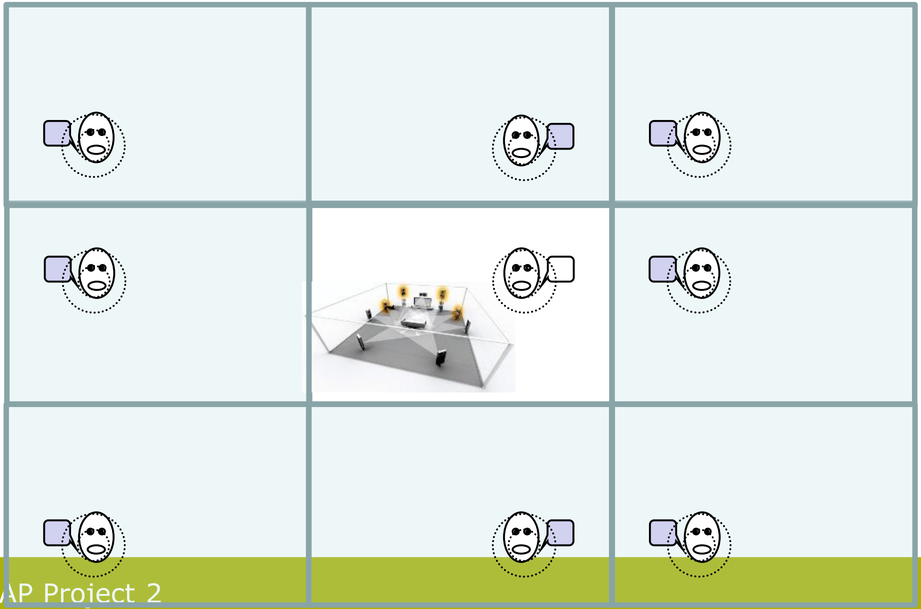
Sound scene reconstruction

2. Virtual room ambience: “image” model



Sound scene reconstruction

2. Virtual room ambience: “image” model



Reverberation adds virtual sources, each with its own:

- Location
- Delay
- Amplitude

Surround sound scene reconstruction

- Implement sound source localisation
  - Paper: "The Generalized Correlation Method for Estimation of Time Delay", C.H.Knapp et al.
  - Paper: "A closed-form location estimator for use with room environment microphone arrays", M.S.Brandstein et al.
- Implement VBAP for sound scene reconstruction
  - Paper: "Virtual Sound Source Positioning Using Vector Base Amplitude Panning", V.Pulkki link
- Combine VBAP with virtual room reverberation
  - Paper: "Image Method for Efficiently Simulating Small-room Acoustics", J.B.Allen and D.A.Berkley link
- Apply to real recordings
  - to be scheduled based on group lists