

3D Audio Acquisition and Reproduction Systems

Cédric André, Marc Evrard,
Jean-Jacques Embrechts, Jacques Verly
Laboratory for Signal and Image Exploitation (INTELSIG),
Department of Electrical Engineering and Computer Science,
University of Liège, Liège, Belgium

23 February 2011

Journée ABAV : « Qui fait quoi dans la recherche acoustique en Belgique ? »
(Neder-over-Heembeek)

Outline

- Introduction
 - Context
 - Aim of our projects
 - Framework
- 3D audio techniques
- New challenges of 3D cinema
- Joint 3D audio and 3D video for cinema
- Conclusions

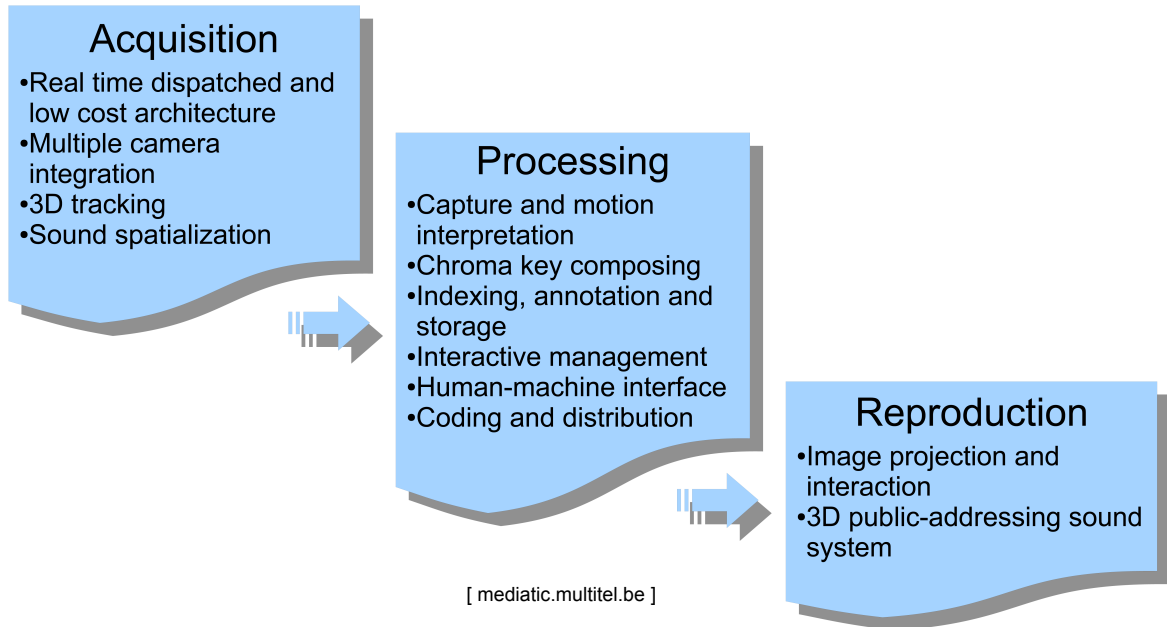
Context

- Many current film productions include spatial audio content, but...
- No current cinema standard is able to capture and reproduce in real-time true 3D audio scene
- While temporal aspect of audio reaches the limit of human hearing capabilities (DVD-Audio)
- Spatial aspect of current technologies remains limited

Aim of our projects

- The aim is to develop a true 3D audio perception using a high spatial resolution
- Our projects are respectively to:
 - ◊ Study and develop a system capable of acquiring a full 3D soundscape and processing it in real-time
 - ◊ Study the reproduction of a 3D soundscape consistent with the visual content of a 3D-stereoscopic movie

Framework: 3D Media Project



23/02/2011

5

Outline

- Introduction
- **3D audio techniques**
 - Definition
 - Binaural techniques – Transaural
 - Virtual Base Amplitude Panning (VBAP)
 - Wave Field Synthesis (WFS)
 - Ambisonics
- New challenges of 3D cinema
- Joint 3D audio and 3D video for cinema
- Conclusions

23/02/2011

6

Definition

- 3D audio is a technique that pursue the acquisition and reproduction with the best **spatial perception** of a soundscape
- 3D audio is characterized by:
 - the **direction** of the source
 - the **distance** to the source
 - the associated **room effect**
- According to the chosen technique, recording and reproduction configurations are either dual or decoupled

23/02/2011

7

Binaural techniques (1) Recording - Reproduction

Binaural recording

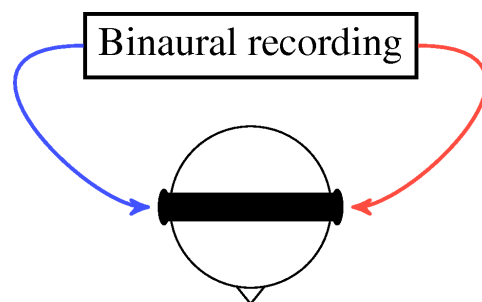



[Blauert, Springer 2005]



[www.neumann.com]

Binaural reproduction



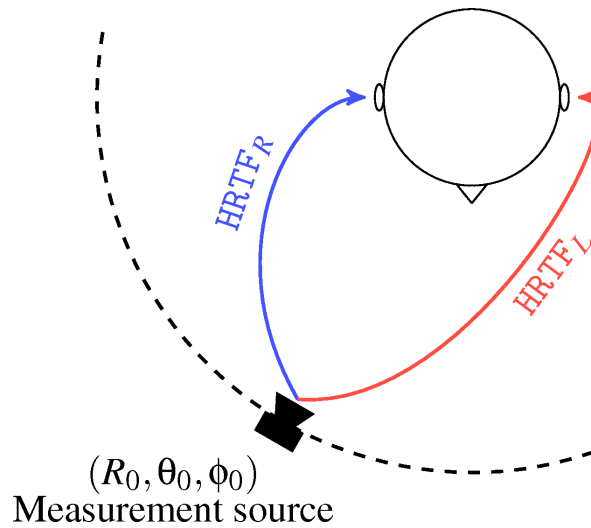

 (R_S, θ_S, ϕ_S)
 Virtual source

23/02/2011

8

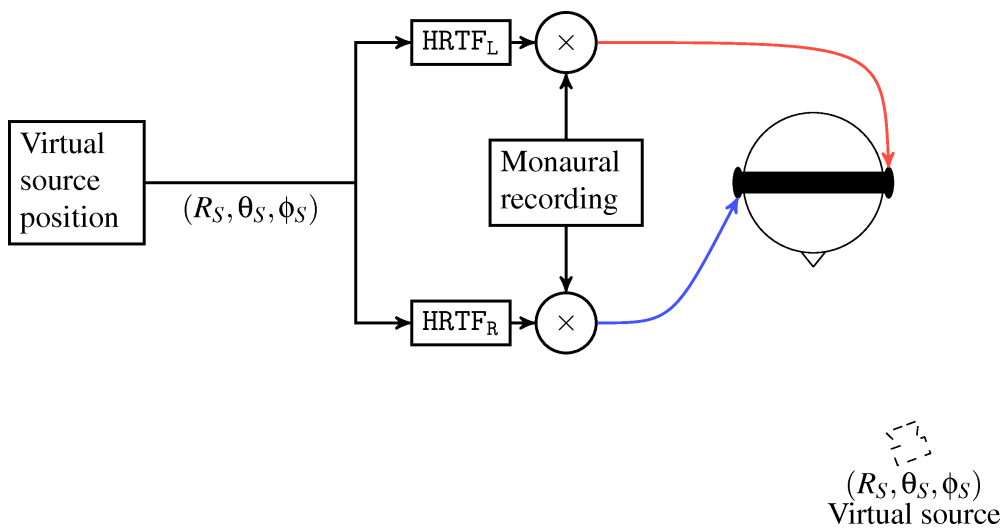
Binaural techniques (2)

Measurement of HRTFs

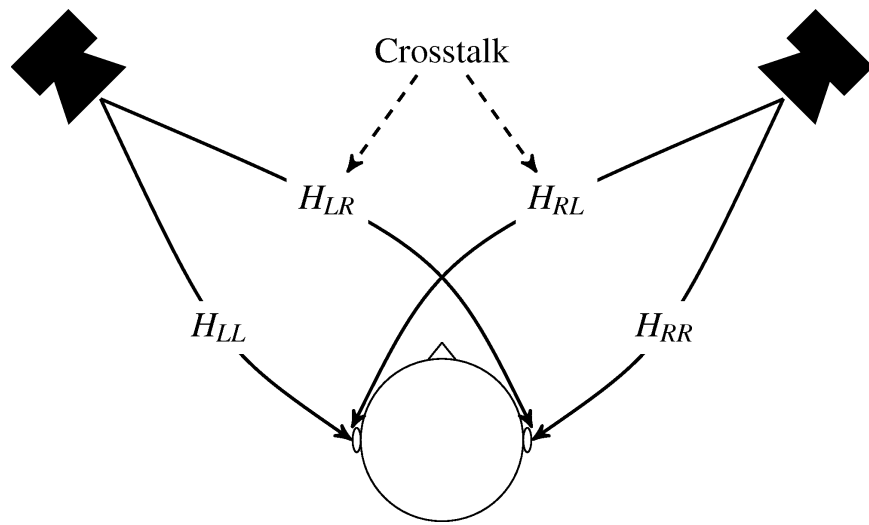


Binaural techniques (3)

Binaural synthesis

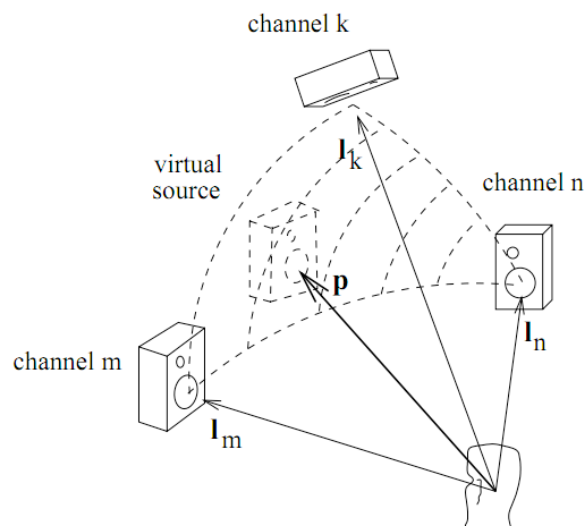


Transaural techniques



[T. Lentz, JAES vol. 54 (2006)]

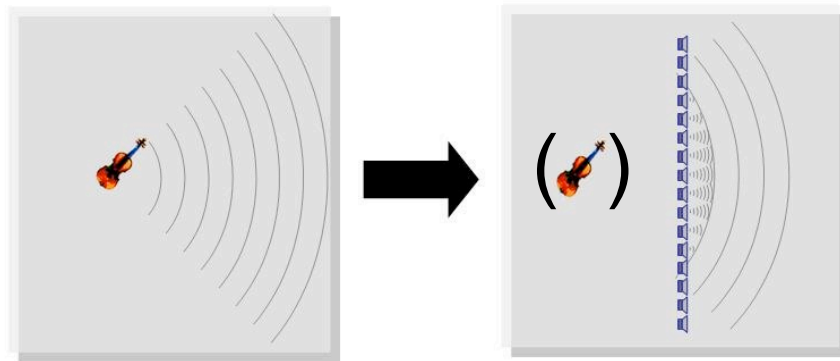
VBAP Virtual Base Amplitude Panning



[V. Pulkki, *Uniform spreading of amplitude panned virtual sources*, IEEE 1999]

WFS (1) Wave Field Synthesis

$$P(\vec{r}, \vec{r}_0) = \int_{\Lambda} \left[G(\vec{r}|\vec{r}_q) \frac{\partial P(\vec{r}_q)}{\partial \vec{n}_q} - P(\vec{r}_q) \frac{\partial G(\vec{r}|\vec{r}_q)}{\partial \vec{n}_q} \right] d\Lambda$$



[www.limsi.fr]

23/02/2011

13

WFS (2) Corresponding recording technique

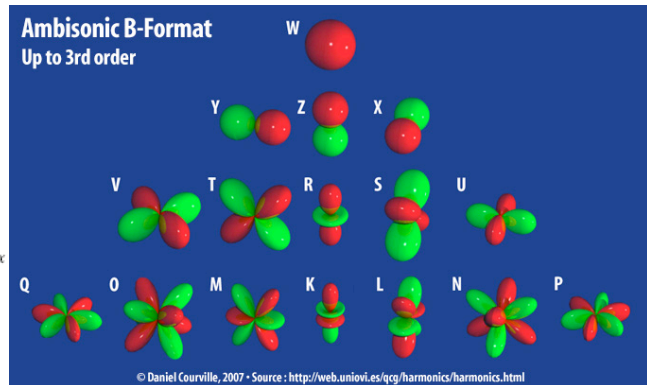
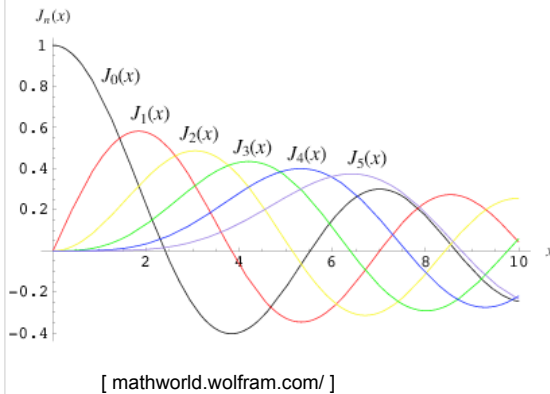
- The straightforward recording configuration is a microphone array of the same dimension
- Could become extremely big (cinema size) and thus not very convenient
- WFS is usually an object based processing technique
- Another solution is to use encoding in order to decouple the recording and the reproduction process

23/02/2011

14

Ambisonics (1)

$$P(\vec{r}) = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l P[l, m] j^l J_l(kr) Y_l^m(\theta, \phi)$$



Ambisonics (2)

- **Fourier-Bessel Transform** is one way of describing the acoustic fields of a soundscape
- The FBT decomposes an acoustic field in a superposition of elementary acoustic fields
- The coefficients $P[l, m]$ must be identified for a given sound field
- They are then used as input to the reproduction system

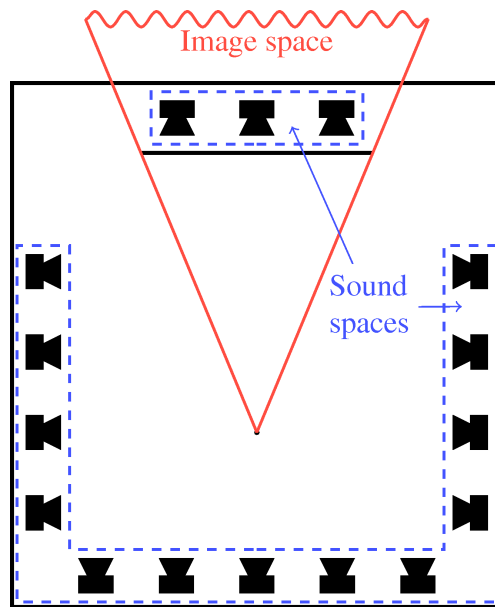
$$P(\vec{r}) = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l P[l, m] j^l J_l(kr) Y_l^m(\theta, \phi)$$

- In practice, one must **limit the order l**
- The theory states that the **minimum number of channels** required for playback (in 3D) is given by

$$L = (l + 1)^2$$

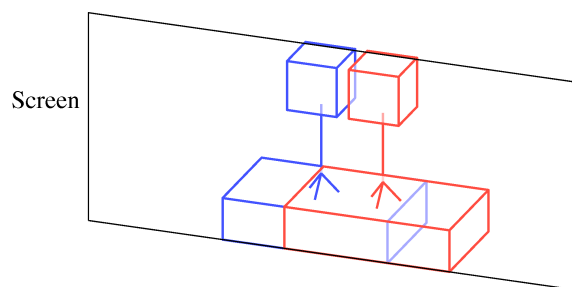
- Introduction
- 3D audio techniques
- **New challenges of 3D cinema**
 - **Sound space versus image space**
 - **Differences in visual source localization between spectators**
- Joint 3D audio and 3D video for cinema
- Conclusions



Sound space vs. image space



[B. Mendiburu, *3D Movie Making*, 2009]

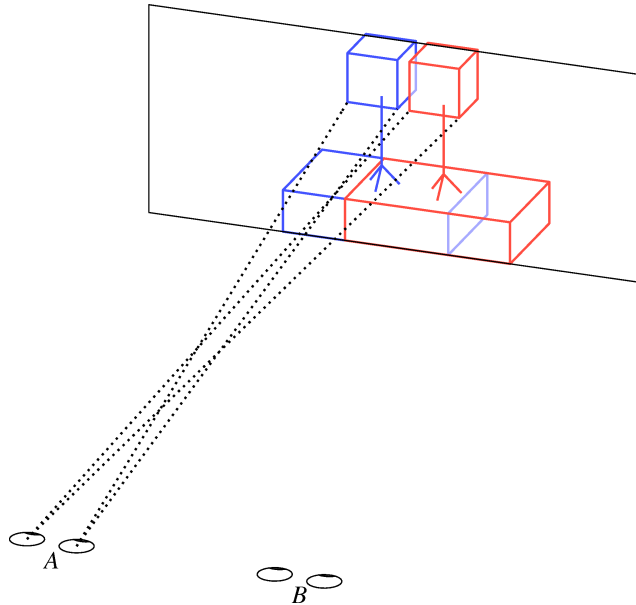
Differences in visual source localization between spectators (1)



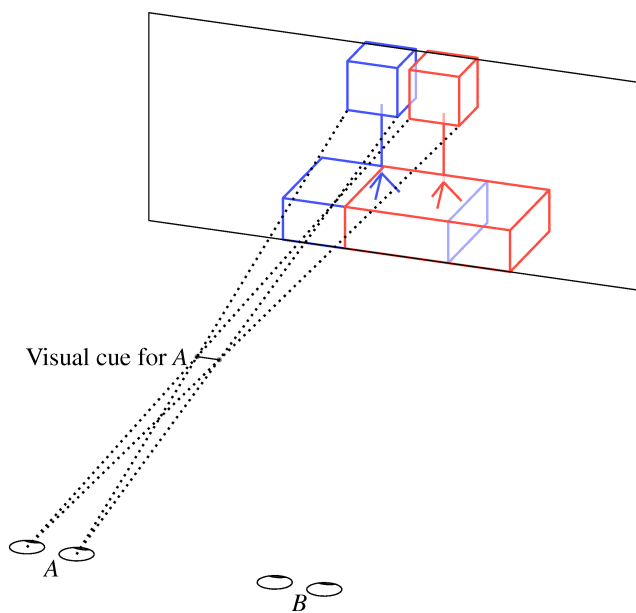
Spectators   A

  B

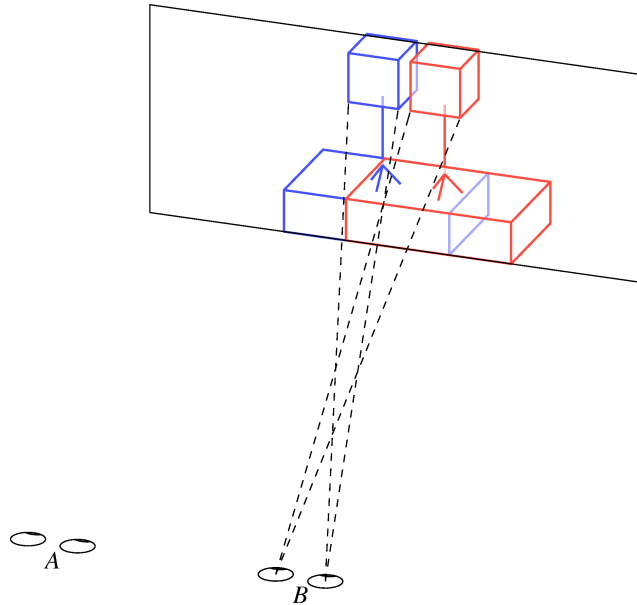
Differences in visual source localization between spectators (2)



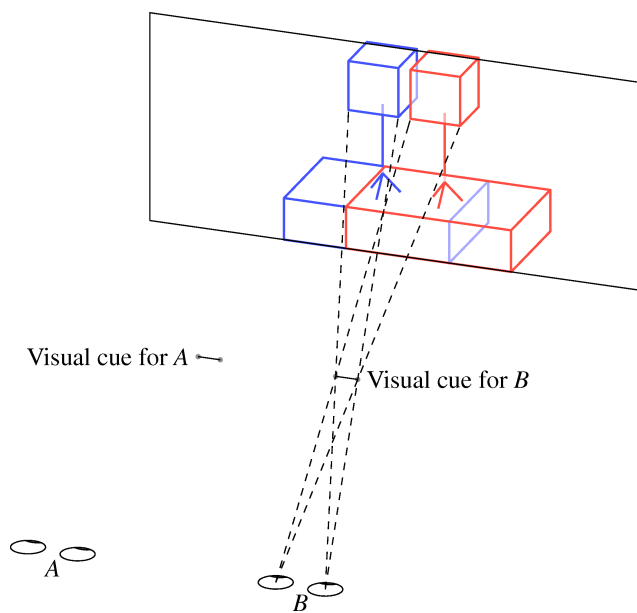
Differences in visual source localization between spectators (3)



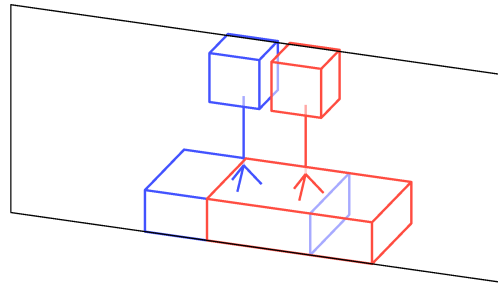
Differences in visual source localization between spectators (4)



Differences in visual source localization between spectators (5)



Differences in visual source localization between spectators (6)



Visual cue for A

Visual cue for B

A

B

Outline

- Introduction
- 3D audio techniques
- New challenges of 3D cinema
- **Joint 3D audio and 3D video for cinema**
 - Comparison of the different techniques
 - Ambisonics and Binaural
- Conclusions

Comparison of the different reproduction techniques

	Binaural	Transaural	VBAP	WFS	Ambisonics
Material - configuration	-Headphones for all + LFE -Headtracking	-2 / 4 speakers per spectator -Headtracking	Sphere of loudspeakers	Large number of very small loudspeakers	Sphere of loudspeakers
Processing requirement	Simple or Complex (HRTF)	Lower requirement on HRTF	Simple	Complex	Complex
Audience size	All spectators	All spectators	Large	All spectators	Large
Sweet spot size	All spectators	All spectators	Very small	All spectators	Sweet spot reduces with frequency
Personalized Soundscape	Possible	Possible	Not possible	Not possible	Not possible
Particularities	Best percept. if personalized HRTF provided	Cross talk between spectators	Only object based	Aliasing reduces bandwidth	Possibility of decoupling rec. & reproduction

23/02/2011

27

Ambisonics and Binaural

- Ambisonics is however interesting for its properties as a recording and coding system
- The reproduction using Ambisonic coding would be based on at least $L = (I + 1)^2$ virtual loudspeakers at order I
- In combination with binaural techniques, using Ambisonic coding allows to
 - decouple recording and reproduction configuration
 - use fixed virtual loudspeakers (only the corresponding HRTFs are needed)
 - further reduce the processing associated to head tracking

23/02/2011

28

Conclusions

- This presentation gives our first observations after bibliographical research on the state of the art of 3D sound for film
- So far, binaural techniques coupled with Ambisonic recording and processing seems the best choice for matching 3D sound to 3D movies
- Spectators soundscape can be personalized
- Offers decoupling of recording and reproduction configurations
- The virtual loudspeakers remain fixed at all times which reduce the processing requirement and could allow real-time application development

23/02/2011

29

The end

- Thank you for your attention
- Our contact addresses:
C.Andre@ulg.ac.be
Marc.Evrard@ulg.ac.be

This work was sponsored in part by
MEDIATIC - 3D MEDIA project
of the European Regional Development Fund (ERDF)

23/02/2011

30