

Adding 3D sound to 3D cinema: new challenges and perspectives

Cédric André, Jean-Jacques Embrechts, Jacques G. Verly
Laboratory for Signal and Image Exploitation (INTELSIG), University of Liège
Liège, Belgium



Context

Virtually no research has been performed to date on the production of a 3D soundscape consistent with the visual content of a 3D-stereoscopic movie. Yet, there are 3D audio “spatialization” techniques that would allow one to make sound originate from any location in the 3D space.

Objective

We want to isolate the most important characteristics of 3D cinema related to sound reproduction.

An average cinema layout

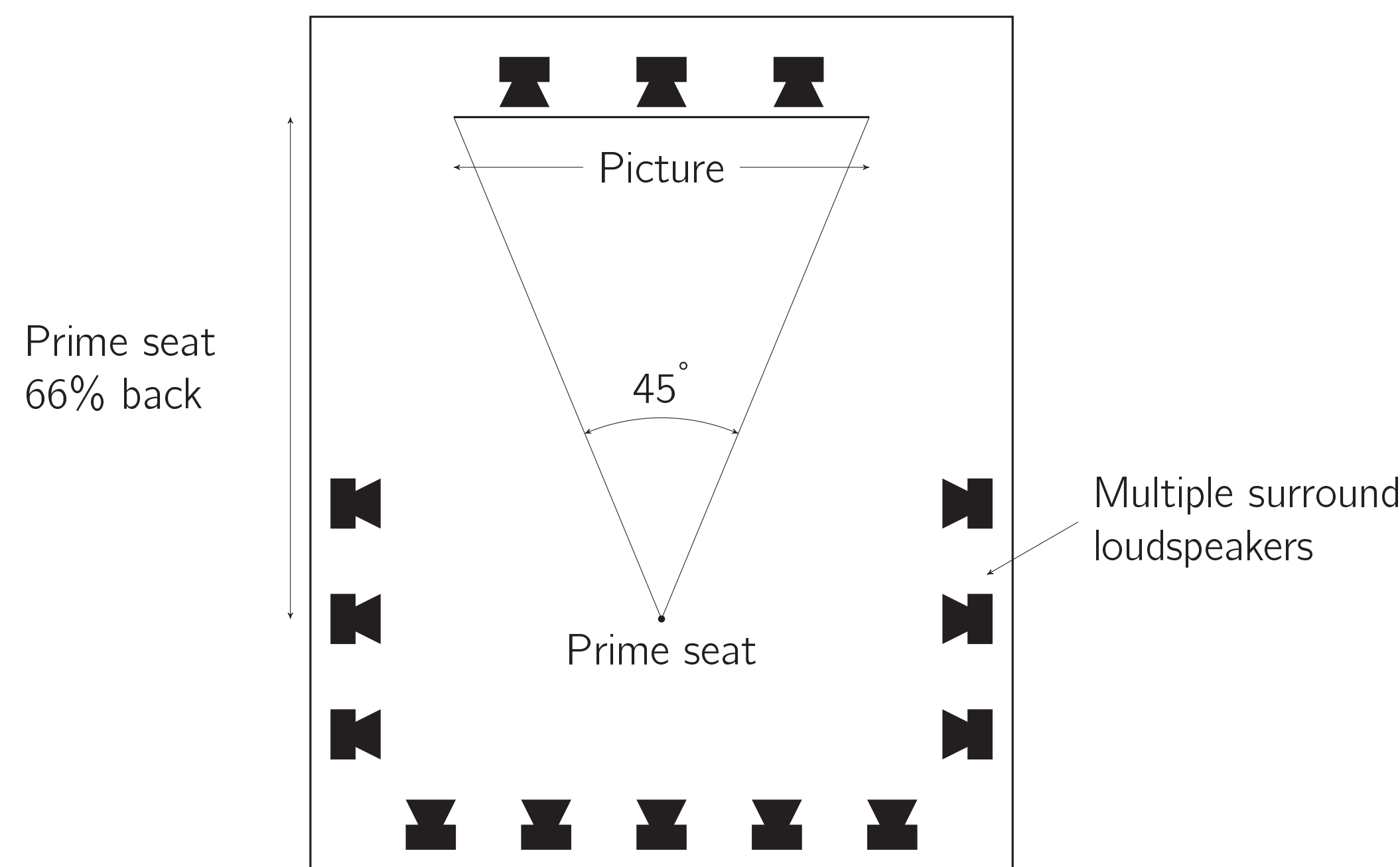


Figure 1: An average cinema layout (after [1]).

Sound volumes and image volumes do not match

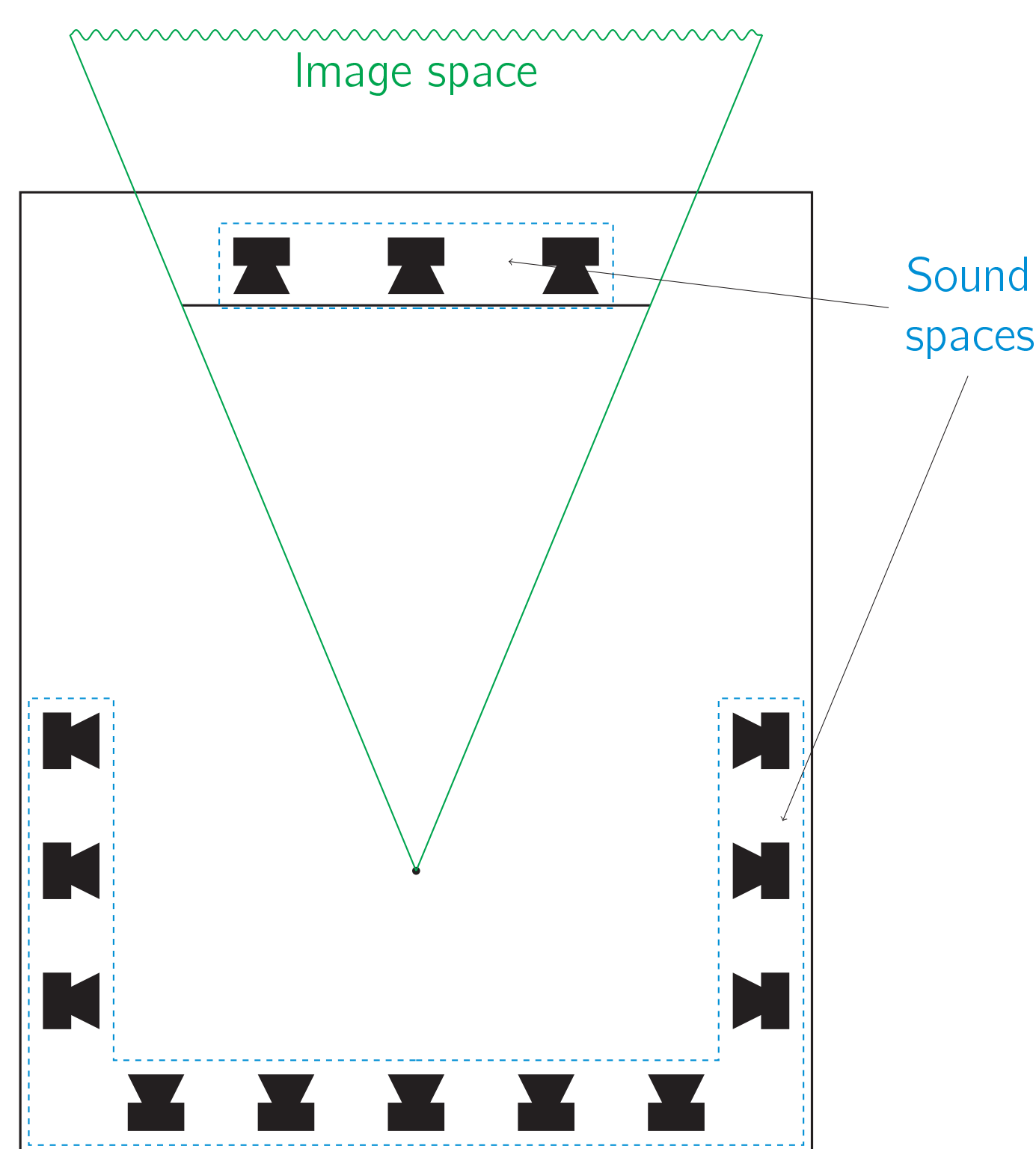


Figure 2: The sound and image spaces (after [2]).

Visual localization is personalized

- Geometry
In 3D stereo cinema, all viewers see the exact same 3D world, but from their own seat. Different spectators therefore localize the same object at different places in some common coordinate frame [3].
- Interocular distance
The distance between the two eyes plays a significant role in the positioning. Persons of different ethnicity, gender, and/or age will see things at different locations [3].

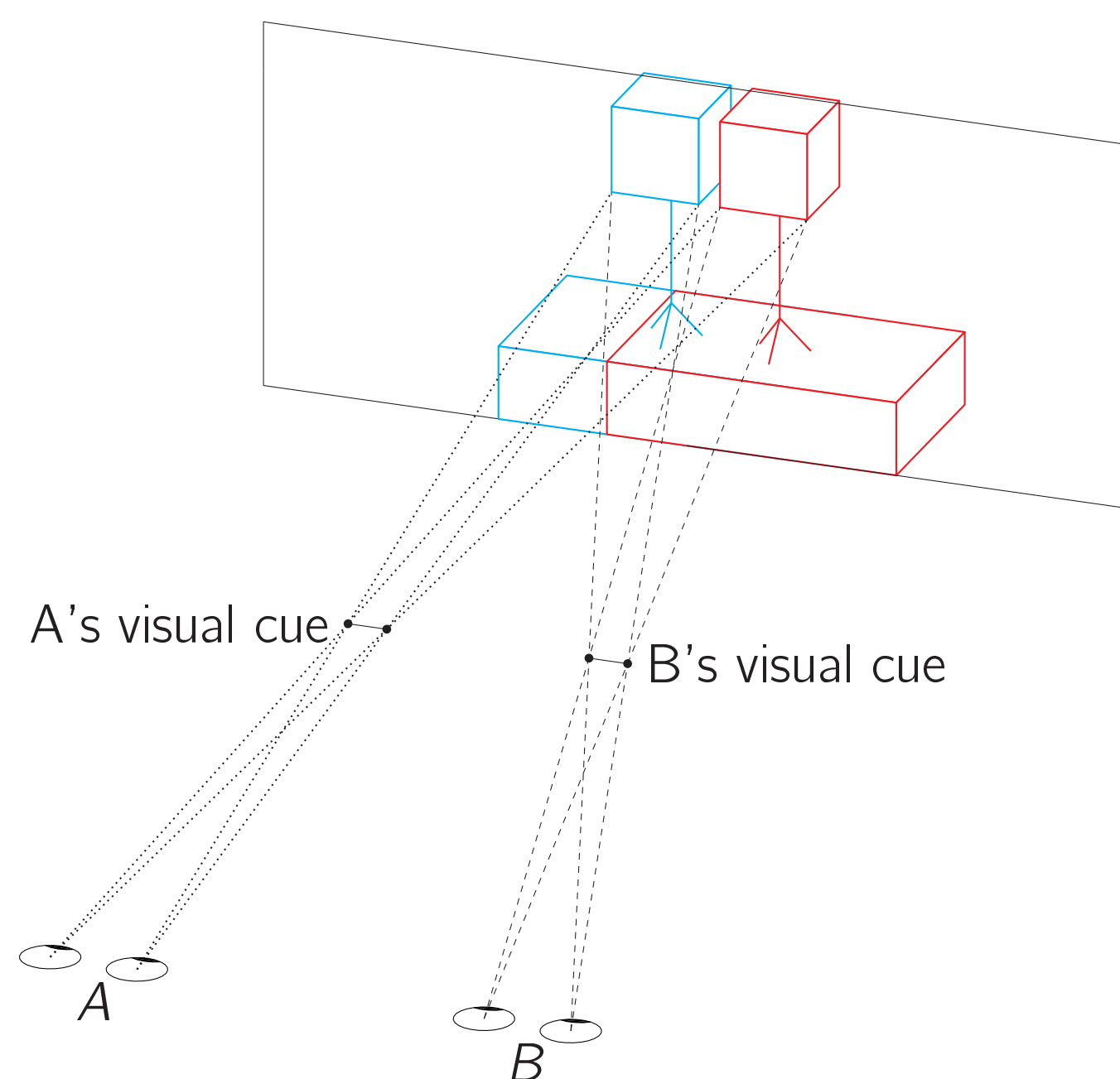


Figure 3: Visual localization for two viewers *A* and *B* seated at two different locations.

Summary

Matching the volumes in which sound objects and visual objects can appear, together with personalizing the sound field in a highly multi-user environment like a cinema, are the most important aspects of the problem considered here.

A review of 3D audio technology

We have reviewed the following 3D audio reproduction techniques:

- Vector Base Amplitude Panning (VBAP),
- Binaural techniques,
- Transaural techniques,
- Wave Field Synthesis (WFS),
- Ambisonics.

Among these, only binaural techniques through headphones offer the technical possibility to tackle all the challenges of 3D-stereo movies. This conclusion is not surprising, considering that binaural reproduction is the audio equivalent of 3D stereo for video.

Principles of binaural techniques

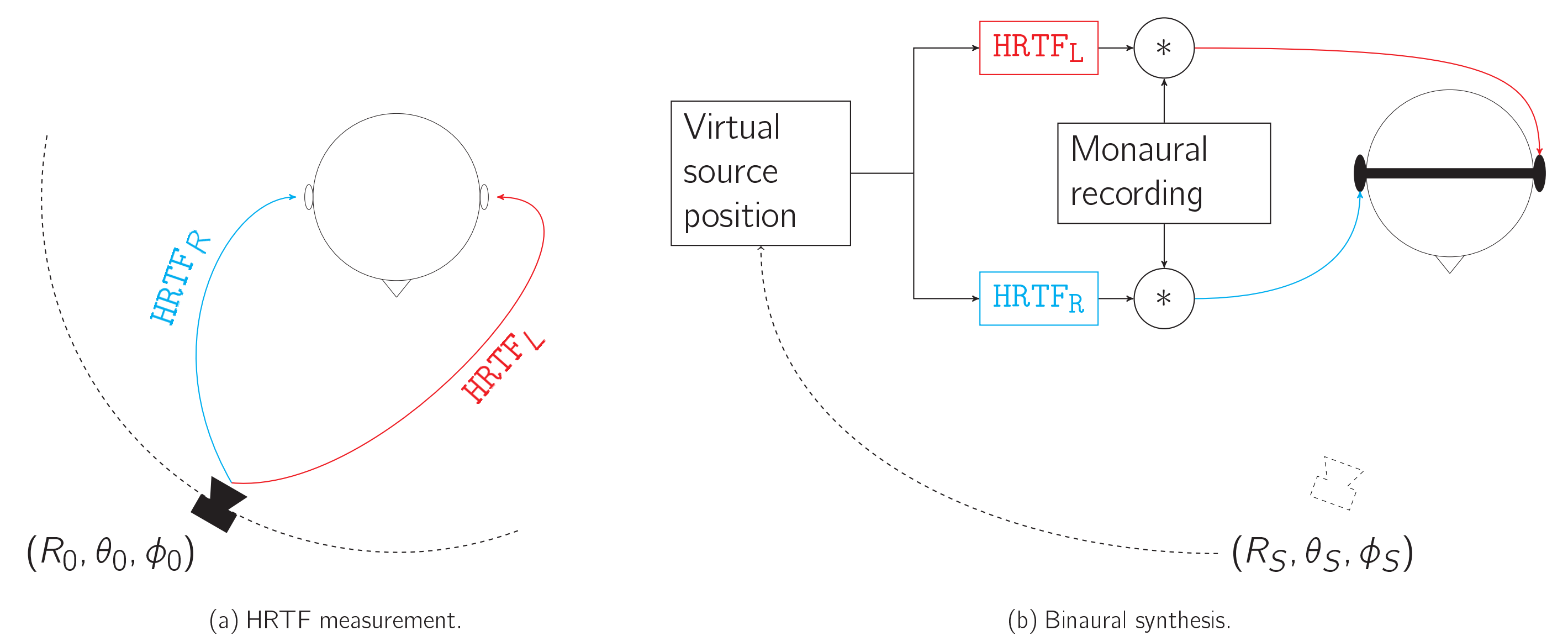


Figure 4: A first method for binaural reproduction.



Figure 5: A second method for binaural reproduction.

Head-tracking and the Fourier-Bessel formalism

Head tracking should be provided to ensure that the sound field remains fixed with respect to the room and not with respect to the listener. The associated real-time processing highly increases the required computing power. One way to reduce this power is to use the Fourier-Bessel formalism [5]. The Fourier-Bessel series expansion is the expression of a 3D sound field in spherical coordinates. In this formalism, a rotation of the soundscape is achieved by a simple linear combination of the original signal components.

Free listening-point audio

Although free-viewpoint TV (FTV) [6] is often viewed on 2D devices, it can also be used with 3D perception. The question will then arise of devising techniques for best combining 3D FTV with its audio equivalent, which is called “free listening-point audio” [6]. The perceived sound must change according to the change in viewing angle selected by the user.

Summary

For the 3D cinema application, we advocate a 3D audio reproduction based on binaural techniques and the Fourier-Bessel formalism.

Perspectives

We are currently planning experiments to evaluate the audience localization performance in the case of joint 3D-stereo video and binaural audio reproduction. We will also examine the possibility of using a hybrid solution consisting of a combination of an augmented-reality audio system, such as the one in [7], and of loudspeakers rendering low frequencies.

References

- [1] I. Allen, “Matching the sound to the picture,” in *AES 9th International Conference: Television Sound Today and Tomorrow*, Feb. 1991.
- [2] B. Mendiburu, *3D Movie Making: Stereoscopic Digital Cinema from Script to Screen*. Focal Press, Apr. 2009.
- [3] L. Verduci, “3D audio and 3D images for 3D movie theater,” in *3D Stereo MEDIA conference*, Liège, Belgium, Dec. 2009.
- [4] G. F. Maxwell. [Online]. Available: <http://en.wikipedia.org/wiki/File:Dummyhead.jpg>
- [5] A. Laborie, R. Bruno, and S. Montoya, “A new comprehensive approach of surround sound recording,” in *AES 114th Convention*, Amsterdam, The Netherlands, Mar. 2003.
- [6] M. Tanimoto, “FTV (Free-Viewpoint TV),” in *Intelligent Multimedia Communication: Techniques and Applications*, ser. Studies in Computational Intelligence, 2010, vol. 280, pp. 341–365.
- [7] A. Martin, C. Jin, and A. van Schaik, “Psychoacoustic evaluation of systems for delivering spatialized augmented-reality audio,” *Journal of the Audio Engineering Society*, vol. 57, no. 12, pp. 1016–1027, Dec. 2009.