Research projects in reproduction and synthesis of spatial sound for virtual reality

Ville Pulkki

Department of Signal Processing and Acoustics
School of Electrical Engineering
Aalto University, Helsinki, Finland

August 16, 2018
This talk

- Background
This talk

Background

- Early project with amplitude panning
- Ambisonics recording methods
- Parametric time-frequency-domain spatial audio tools
This talk

- Background
  - Early project with amplitude panning
  - Ambisonics recording methods
  - Parametric time-frequency-domain spatial audio tools
- Spatial audio effects
This talk

- Background
  - Early project with amplitude panning
  - Ambisonics recording methods
  - Parametric time-frequency-domain spatial audio tools

- Spatial audio effects
  - Synthesis of spatial width
  - Spatial modulation of sound
A music student with MSc (Eng) needs extra income (1995)

- Sibelius Academy chamber music hall had lots of loudspeakers on walls and ceiling
- SibA wanted to have a "panning tool" for their loudspeaker system (one month salary for student)

Fig. 9. Possible use of three-dimensional VBAP panning tool. Number of sound sources can vary up to eight; loudspeaker placement is arbitrary; virtual sources may be moving or stationary.

©Audio Engineering Society, Inc. 1997
A music student with MSc (Eng) needs extra income (1995)

- Helsinki Univ Tech had a self-made 8-channel AD/DA for music instrument synthesis
- Paid student project with 1-month salary
Vector base amplitude panning

N loudspeakers in arbitrary positioning. Applications: virtual reality, computer music, theatres. Key question: How well and why at all does the listener perceive the virtual source?

PhD degree in 2001.
Products with "VBAP inside"

- ITU MPEG-H audio standard (broadcast)
- DTS:X audio format (cinema + blueray) (88 movies already)
- Sony Playstation VR (gaming)
- Dedicated audio programming softwares
Time after PhD (2001–)

Spatial sound recording captured my mind
How could a sound field be reproduced

Problems with existing techniques

Research projects in reproduction and synthesis of spatial sound for virtual reality
Pulkki
Dept Signal Processing and Acoustics

August 16, 2018
First-order B-format recording

Captures signals with zeroth-order and first-order spherical harmonics
First-order B-format recording

- Captures signals with zeroth-order and first-order spherical harmonics
- Pressure signal W. 3D velocity signals XYZ.

www.soundfield.com
First-order Ambisonics

- Weighted sum of WXYZ signals (mixing, matrixing)
First-order Ambisonics

- Weighted sum of WXYZ signals (mixing, matrixing)
- High coherence between loudspeaker signals
First-order Ambisonics

- Weighted sum of WXYZ signals (mixing, matrixing)
- High coherence between loudspeaker signals
- Spectral and spatial issues, very small listening area
First-order Ambisonics

- Weighted sum of WXYZ signals (mixing, matrixing)
- High coherence between loudspeaker signals
- Spectral and spatial issues, very small listening area
- Blurred images in headphone listening
Higher-order B-format recording

- Signals with directional patterns following to spherical harmonics
Higher-order B-format recording

- Signals with directional patterns following to spherical harmonics
- Reproduce plane-wave expansion over loudspeakers
More spherical harmonics captured
Higher-order Ambisonics

- More spherical harmonics captured
- Better resolution, more expensive devices
Higher-order Ambisonics

- More spherical harmonics captured
- Better resolution, more expensive devices
- Good quality in limited frequency window
Higher-order Ambisonics

- More spherical harmonics captured
- Better resolution, more expensive devices
- Good quality in limited frequency window
- Emphasized problems with low-frequency noise and high-frequency aliasing
Parametric time-frequency-domain techniques

Directional audio coding / COMPASS / Other similar techniques
Parametric time-frequency-domain techniques

Directional audio coding / COMPASS / Other similar techniques

- Analyze/synthesize the directional parameters of sound field
- Non-linear signal-dependent signal processing method
- In 90% of recordings, the audio quality is improved prominently
Commercial application

Fraunhofer upHear Spatial Audio Microphone Processing

Overview

About
The Fraunhofer spatial audio capturing algorithm has been designed to significantly improve the sound capture capabilities of professional and consumer 360° cameras and mobile devices using built-in microphones.

It is the first audio technology to be delivered under Fraunhofer's upHear brand of immersive audio innovations.

The algorithm automatically transforms the captured sound in real-time to any popular surround or immersive audio reproduction format, such as FOA, HOA, 5.1 channels, and 7.1+4 height channels, while preserving the authenticity of the audio scene.
Head-mounted audiovisual displays

Reproduction

- Head-mounted visual display + headphones
- Both video and spatial audio are updated with head tracking information
- Generic representation of audio in DirAC is well-suited for this
Insert a B-format microphone on the position of the avatar!
Audio engine based on B-format stream

First-order / Higher-order B-format bus
Synthesis of B-format signal

- Level
- Propagation delay
- Panning direction
- Spatial width of source
- Direct-to-reverberant ratio
- Distribution of reverberant energy (?)
Synthesis of B-format signal

- higher-order synthesis also possible
- multiply each signal with corresponding spherical harmonic

\[
\begin{align*}
W_{\text{syn}} &= W_{\text{in}} \\
X_{\text{syn}} &= 1 - \sqrt{1 - (1 - \psi)^2} \\
Y_{\text{syn}} &= \frac{1 - \sqrt{1 - (1 - \psi)^2}}{1 - \psi} \\
Z_{\text{syn}} &= \frac{1}{2} \psi
\end{align*}
\]

Research projects in reproduction and synthesis of spatial sound for virtual reality

Aalto University

Pulkki
Dept Signal Processing and Acoustics

August 16, 2018
Audio engine based on B-format stream

- Can perform all tasks needed in typical virtual world rendering
- Demo
Synthesis of spatial extent of virtual sources

- Different frequencies of mono input to different directions
- Demo

Research projects in reproduction and synthesis of spatial sound for virtual reality
Pulkki
Dept Signal Processing and Acoustics
Projection of real B-format recordings into virtual reality objects

B-format signal → projection → Virtual object

Listener
Spatial audio effects

- Spatial modulation
Spatial audio effects

- Spatial modulation
- Modification of diffuse component of sound
Spatial audio effects

- Spatial modulation
- Modification of diffuse component of sound
- Spatial zooming, rotation,
Spatial modulation

- Spatial information comes from a real situation, and audio from another recording
- Demo
Ambience extraction

- Possible to effect only reverberant parts of sound
- Demo
A reference

- 15 chapters, 416 pages
- Matlab code
- Published in Dec 2017