



*Binaural virtual auditory display
for music content
recommendation and navigation*

Rishi Shukla
Rebecca Stewart, Mark Sandler
Queen Mary University of London

Principle and use cases



Consumer (curation)

Image: [Gauthier DELECROIX](#) (CC BY 2.0)



Creative (composition)

Image: [Ars Electronica](#) (CC BY-NC-ND 2.0)

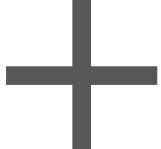
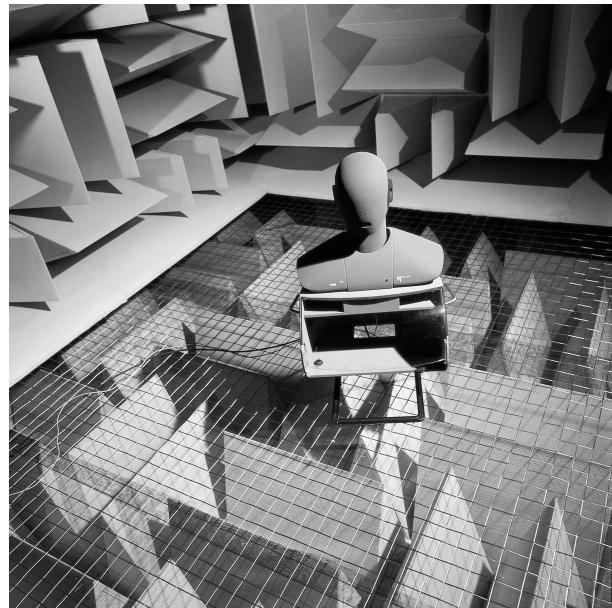


Commercial

Image: [Ryan Harvey](#) (CC BY-SA 2.0)

Virtual vector base amplitude panning

HRTFs



VBAP

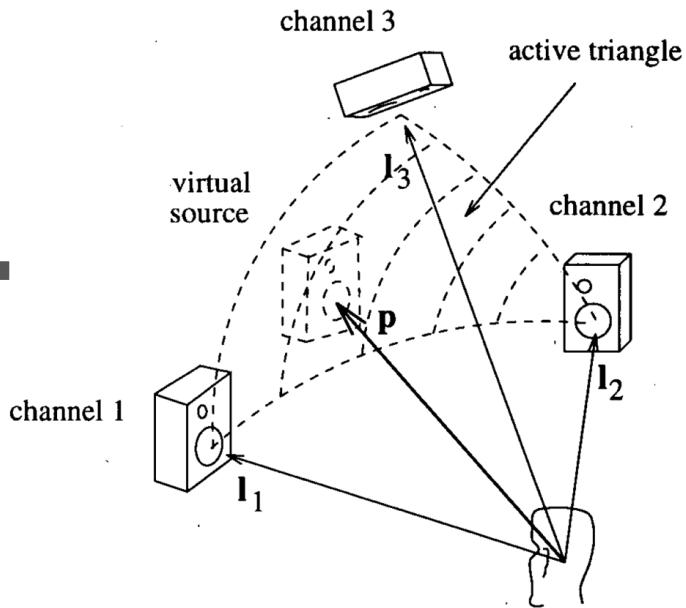


Image: [Andrew Eckel](#) (CC BY-SA 4.0)

Source: Pulkki, Ville (1997). "Virtual sound source positioning using vector base amplitude panning". In: J Audio Eng. Soc., 45.6, pp. 456-466.

Ambisonics

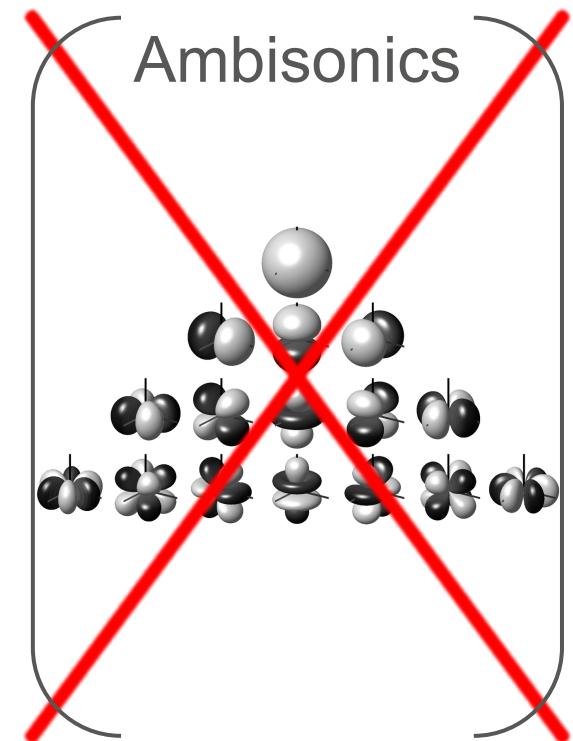
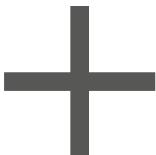
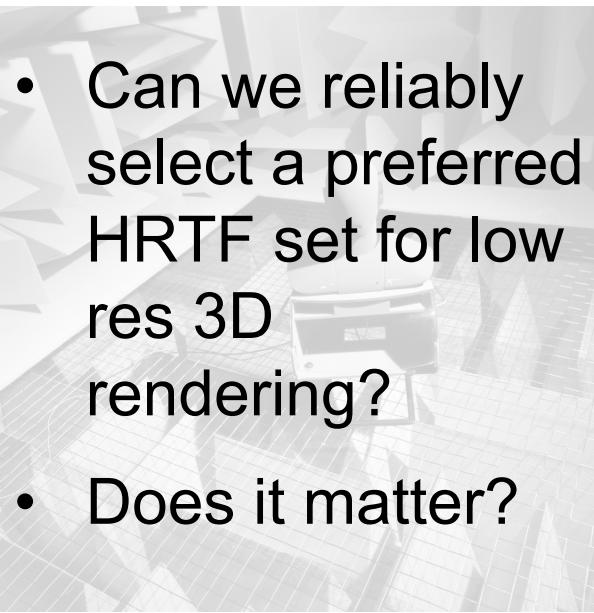


Image: [Franz Zotter](#) (CC BY-SA 3.0)

Questions

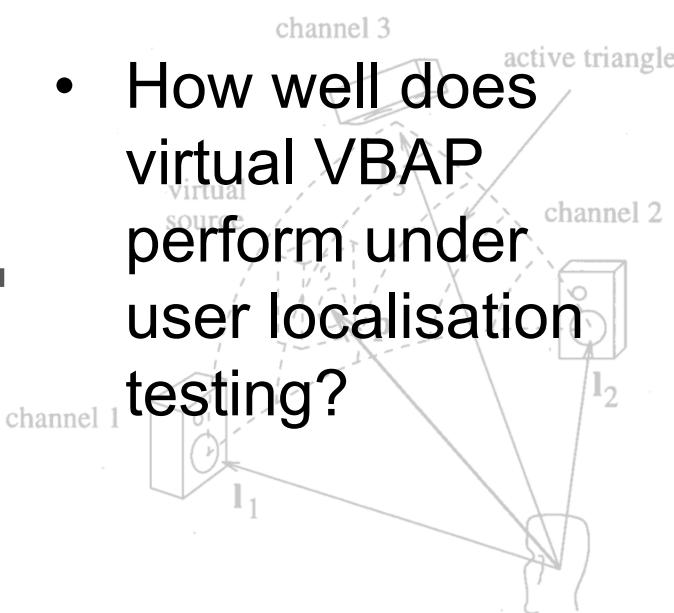
HRTFs

- Can we reliably select a preferred HRTF set for low res 3D rendering?
- Does it matter?



VBAP

- How well does virtual VBAP perform under user localisation testing?



Source: Pulkki, Ville (1997). "Virtual sound source positioning using vector base amplitude panning". In: J Audio Eng. Soc., 45.6, pp. 456-466.

Ambisonics

- Is there a technical case for using virtual VBAP over Ambisonics?

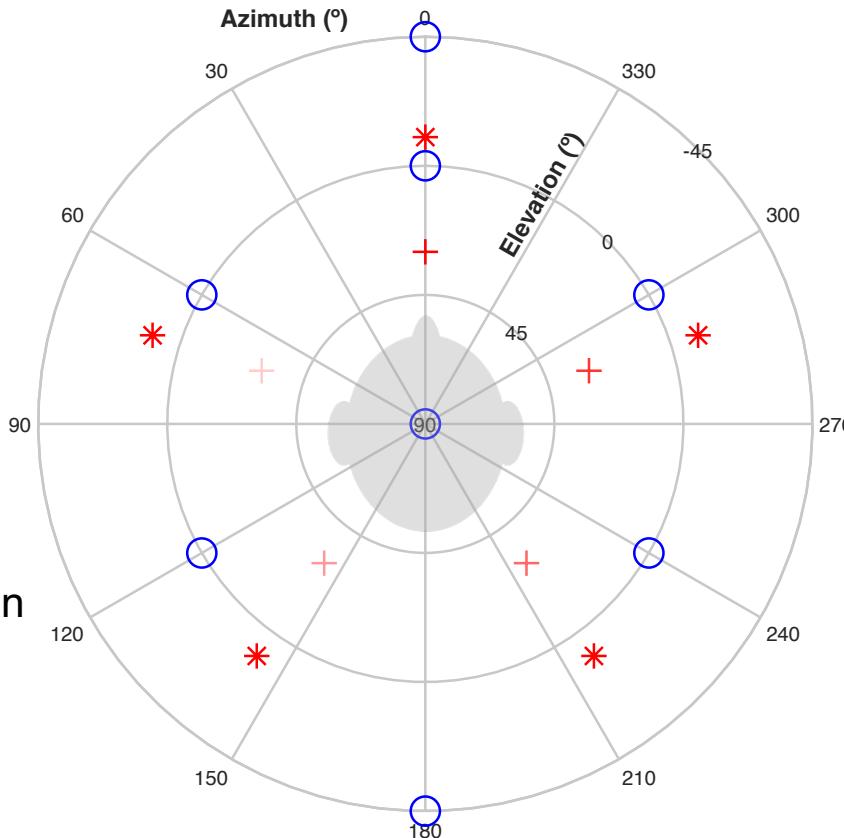


Image: [Franz Zotter](#) (CC BY-SA 3.0)

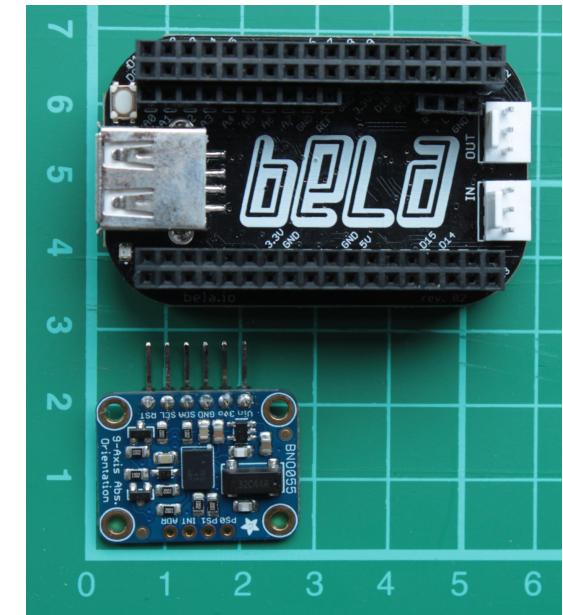
Head-tracked auditory environment prototype

Real-time 3D binaural rendering

- = virtual speaker location
- * = music source positon
- + = cycling speech source position



Bela Mini and BNO055 sensor





Concurrent playback research

Some data we have suggests:

- Presenting multiple auditory streams can be advantageous to task performance.
- First learning how to use the interface with audio only tends to result in quicker task performance, even when visual guidance is added later.
- Inadequate rendering quality may impede task performance before cognitive overload becomes a determining factor.





Example consumer use case

www.bose.com/en_us/products/wearables/frames.html

