



Binaural to your ears!

**EUROCC: MAIS COMPETÊNCIAS, MELHOR ACESSO À
COMPUTAÇÃO**

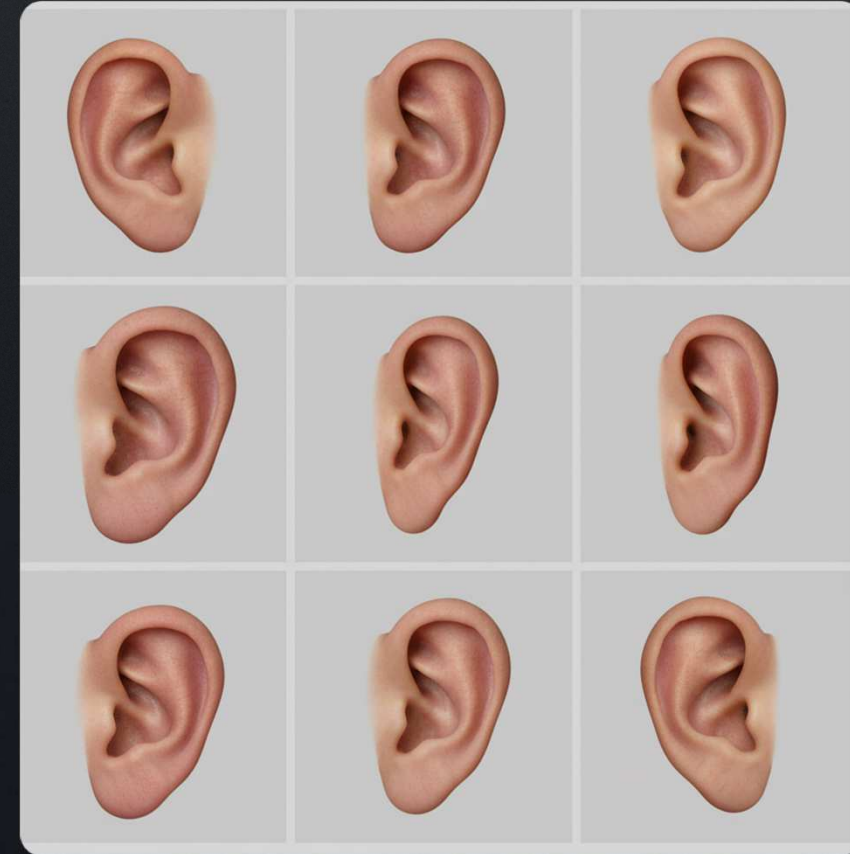
Luis Marcelino (PhD) | 2025/10/23

Overview

- The problem
- Measuring
- Computing
- Dataset creation
- Inferring
- The solution

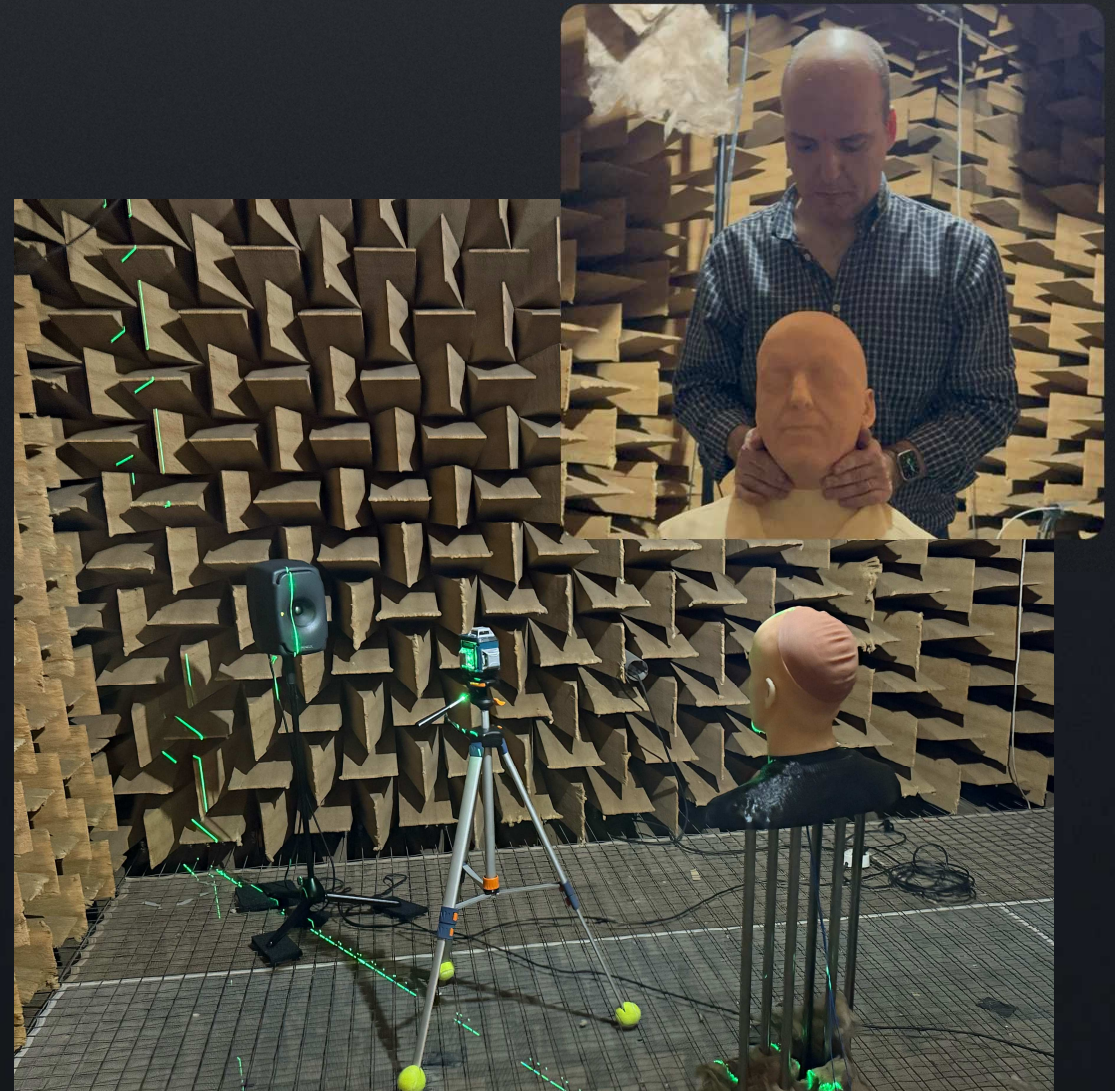
The problem

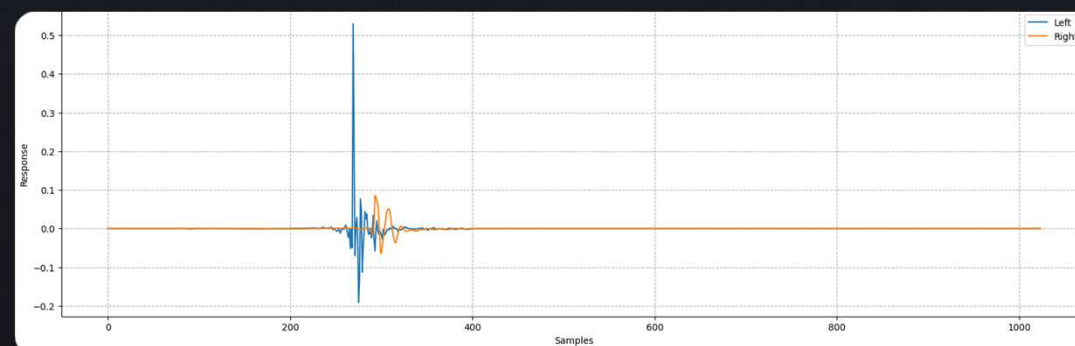
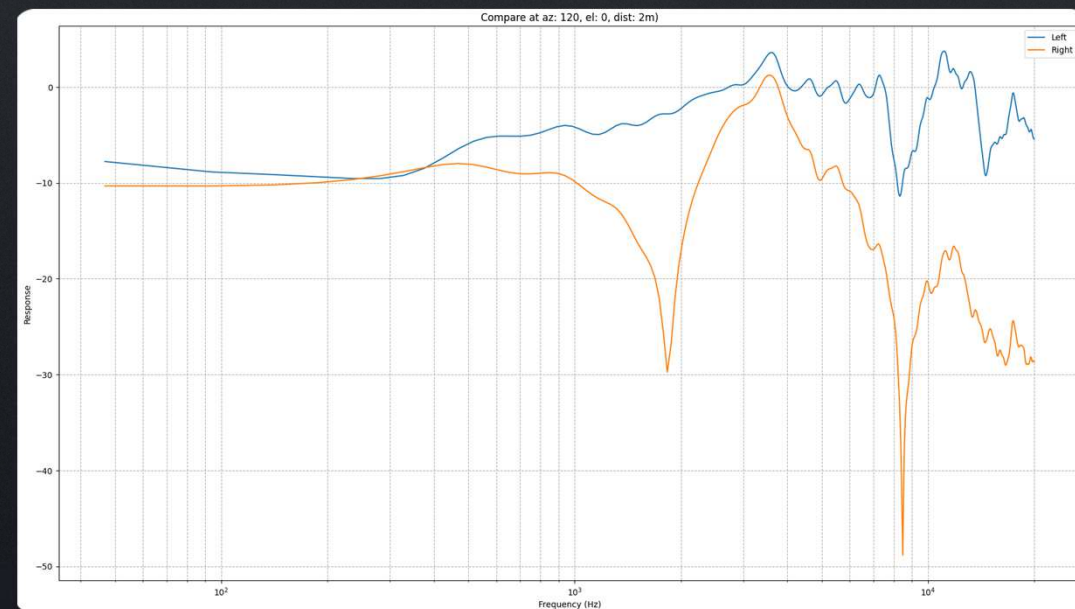
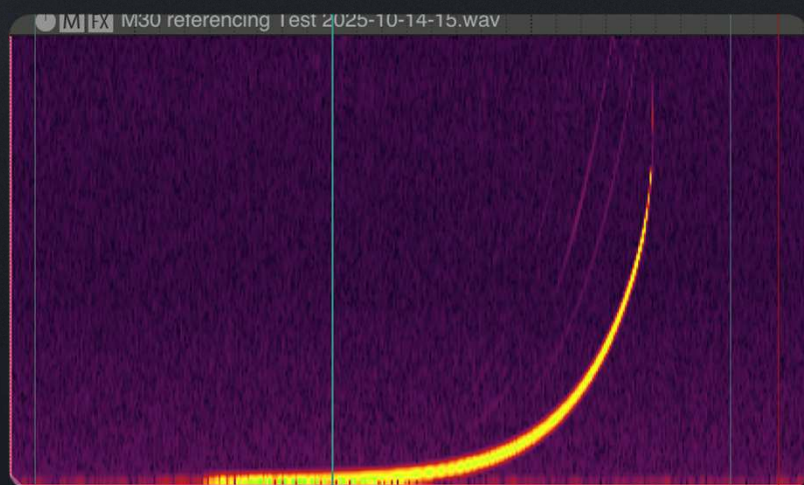
- How is the sound that reaches our timpanums?
- Each person has unique ears
- ... and head shape
- Produces unique sound signature
- The brain can determine the source localisation
- ... using headphones?



Measuring

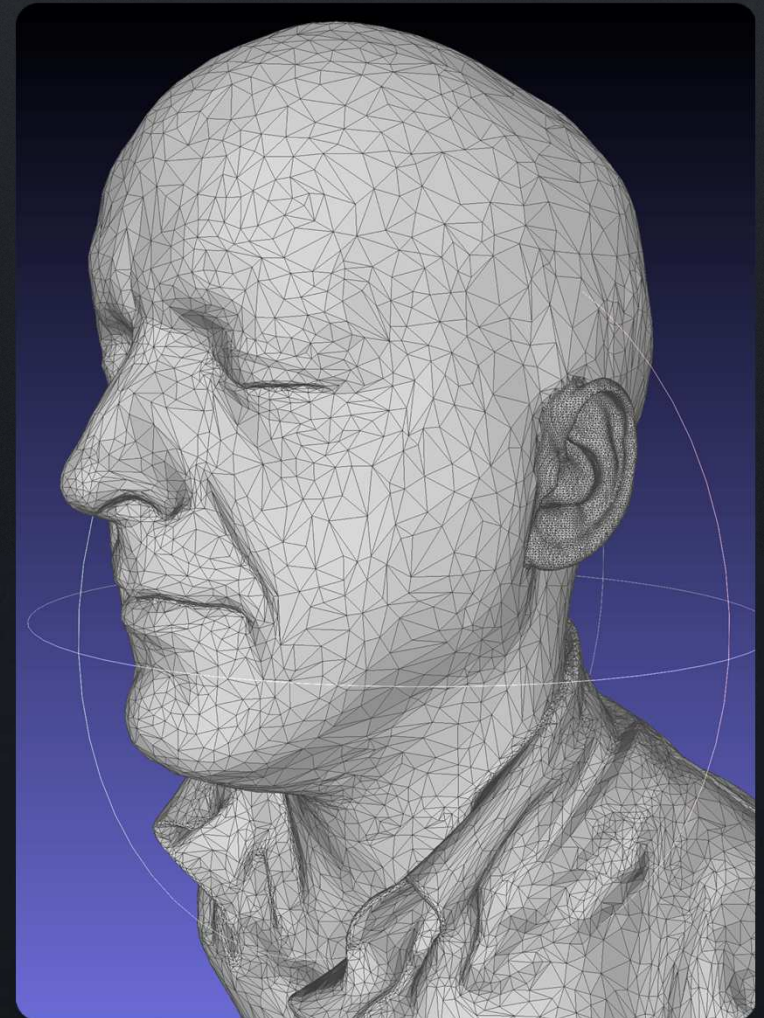
- Anechoic chamber
- Reproduce “sweeps”
- Convert to response
- Our “ground truth”
- Error metric (dBO)

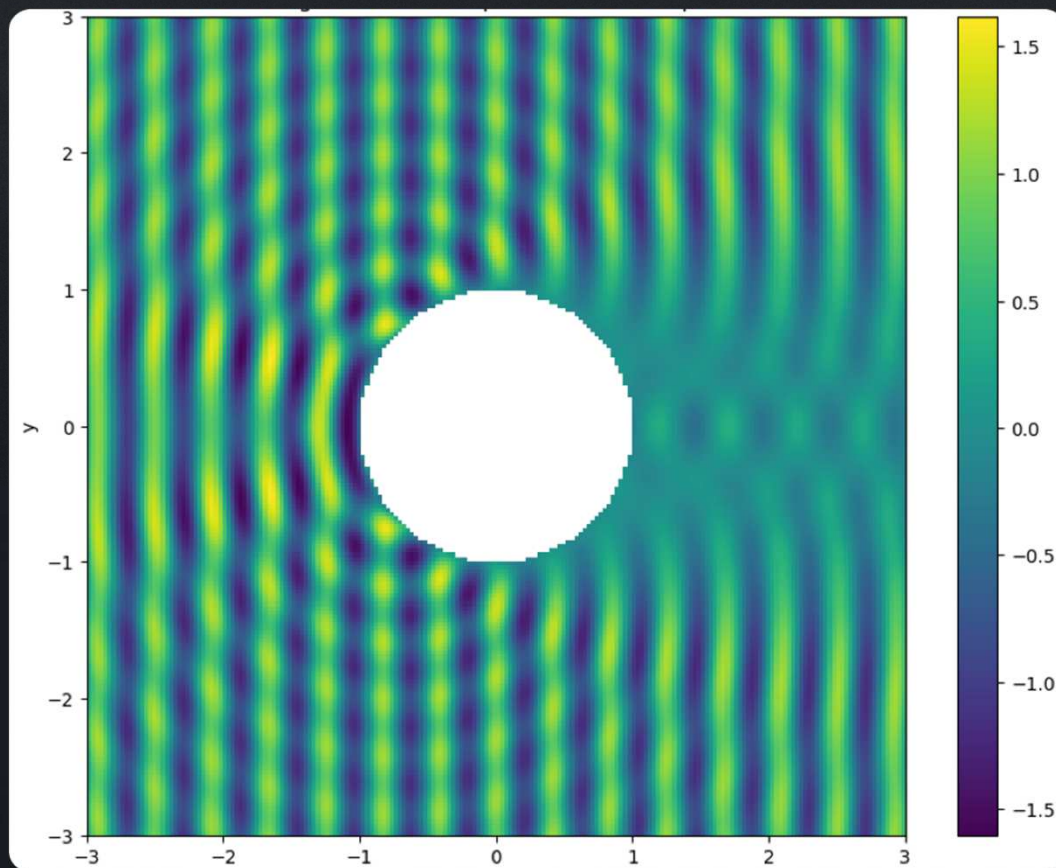




Computing

- BEM simulation
- Left and right ear simulation
- 512 frequencies
- Multiple distances
- Multiple absorptions





$$u^{\text{inc}}(\mathbf{x}) = e^{ikx}.$$

where $\mathbf{x} = (x, y, z)$.

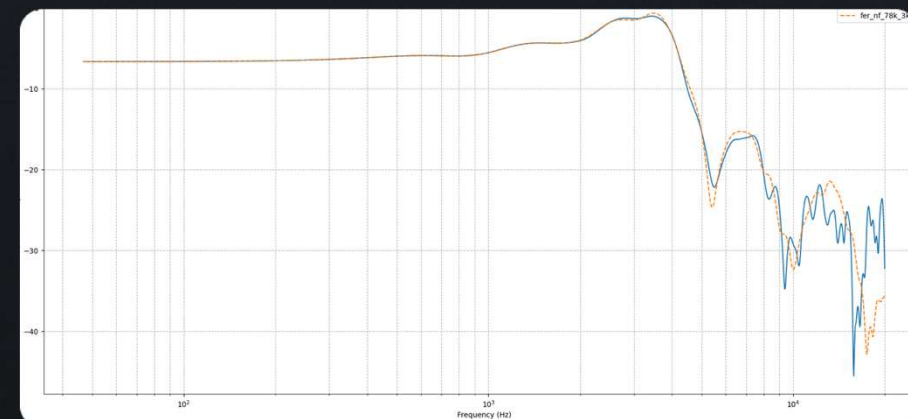
The PDE is given by the Helmholtz equation:

$$\Delta u + k^2 u = 0, \quad \text{in } \mathbb{R}^3 \setminus \Omega,$$

where $u = u^s + u^{\text{inc}}$ is the total acoustic field and u^s satisfies the Sommerfeld radiation condition

$$\frac{\partial u^s}{\partial r} - iku^s = o(r^{-1})$$

for $r := |\mathbf{x}| \rightarrow \infty$.



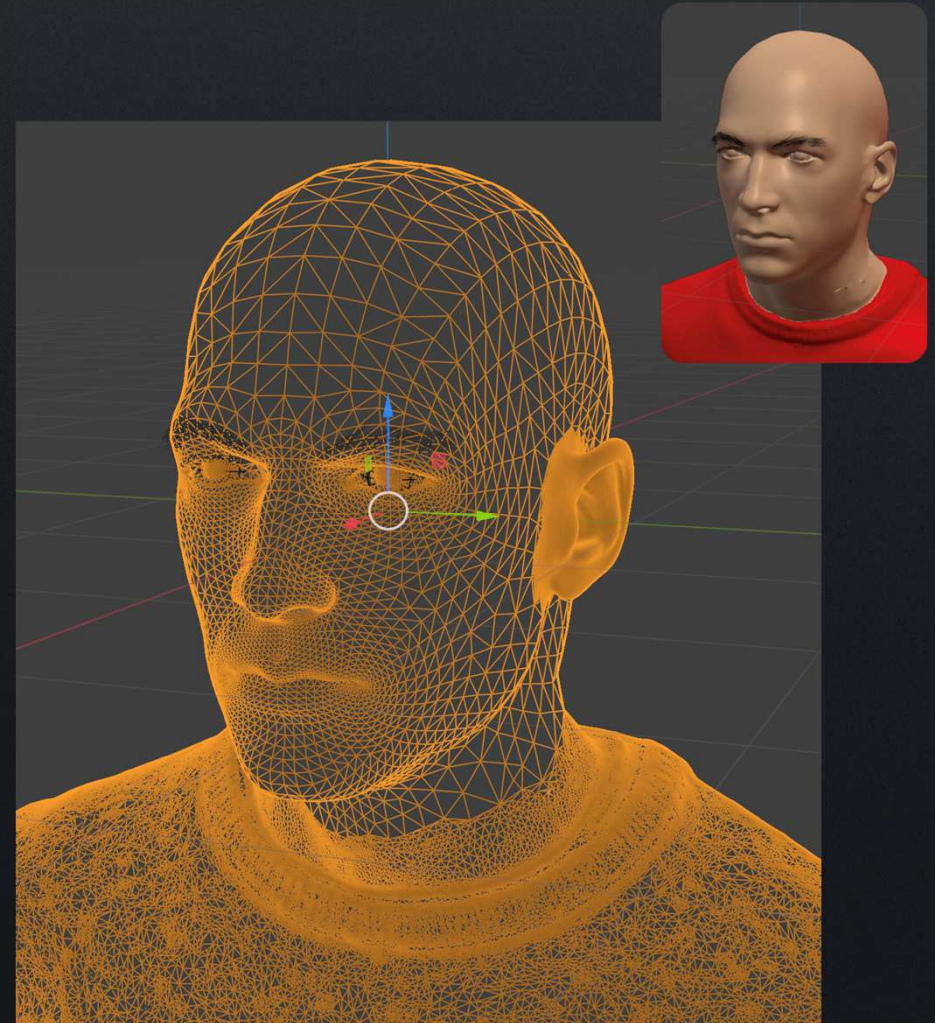
https://nbviewer.org/github/bempp/bempp-cl/blob/notebooks/helmholtz/helmholtz_combined_exterior.ipynb

Computation Limits

- < 100k polygons => 250GB RAM
 - $O(n^2)$
- Simulation time: approx. 300s on 48 core CPU
 - Single frequency
 - Single density
 - One ear
- Computing infrastructure:
 - HPC (Oblivion, Deucalion, MareNostrum?)
 - Custom

Simulations Dataset

- Synthesise heads and torsos
- Define simulation configuration
- Acoustic simulations (1000s)



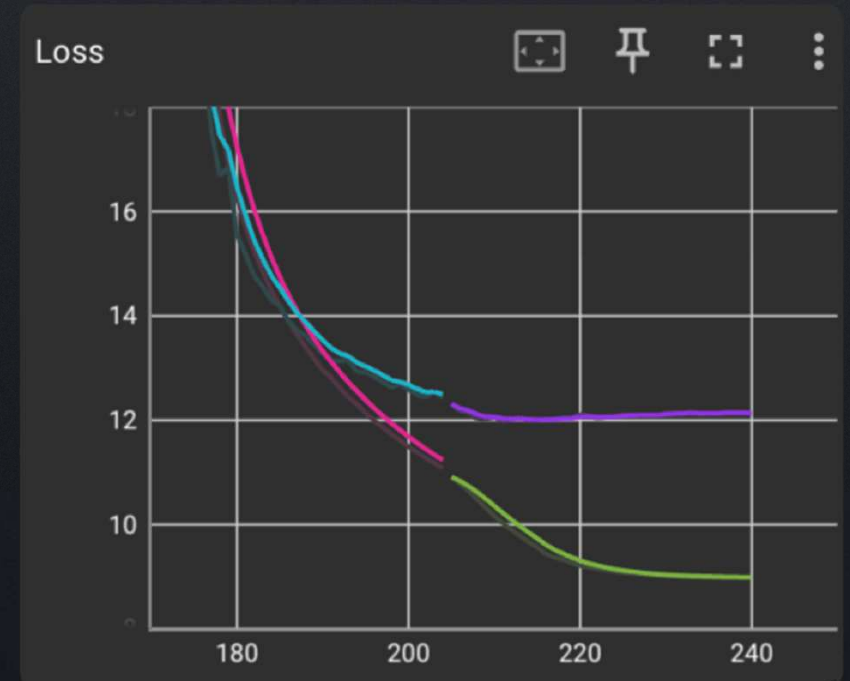
Images Dataset

- Simulated head variations
 - Skin
 - Hair
 - Environment
- “Selfie camera” FOV



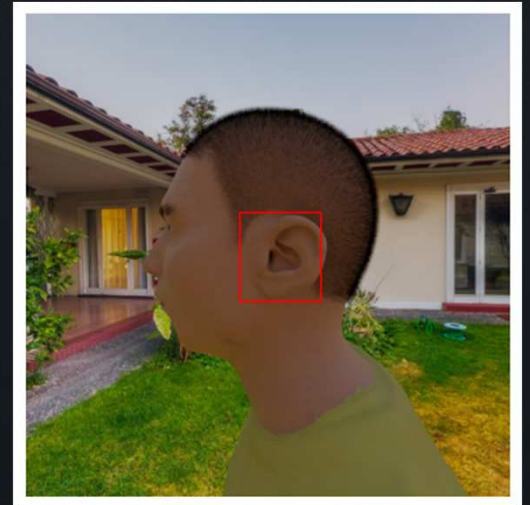
Training

- Configure the AI model:
 - Input images
 - Outputs frequency responses
- Computing infrastructure:
 - NVidia RTX4090, ...
 - HPC (Vision, Deucalion, MareNostrum?)



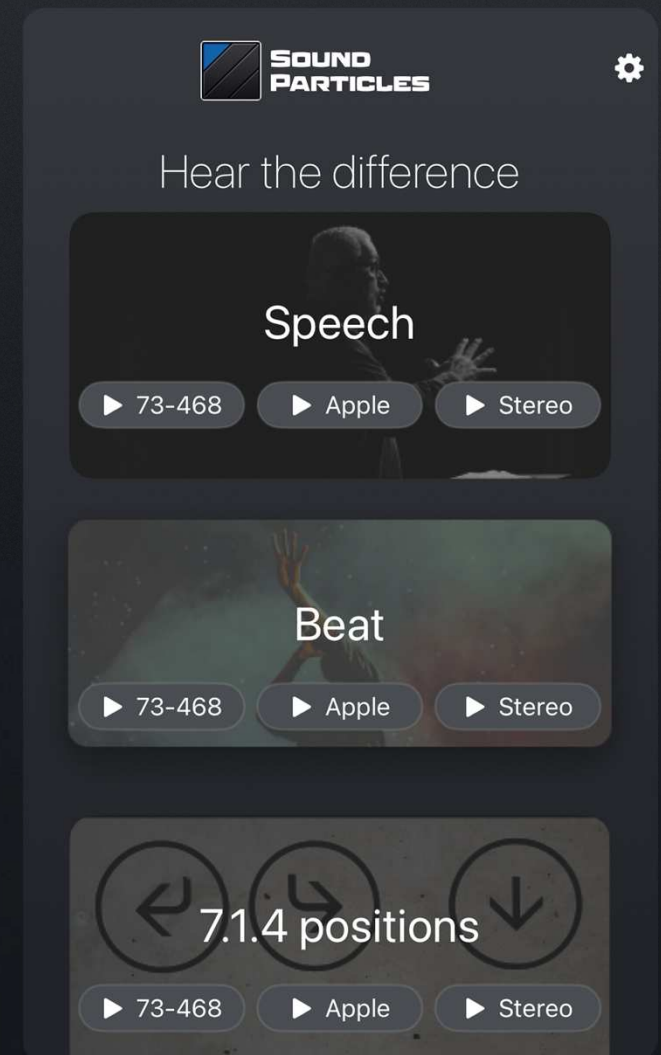
Additional Requirements

- Ear bounding box identification
- Headphone responses
- Room response (reverb)
- Cloud to upload photos and responses



The solution

- User takes pictures
 - An ID is generated
- Pictures are fed to AI model
- Frequency response is computed for relevant points
- SP BinauralSDK downloads response
- Audio channels are processed using HRTF





Thank You!

BINAURAL TO YOUR EARS!

EuroCC: mais Competências, melhor acesso à
Computação

Luis Marcelino (PhD) | 2025/10/23