CREATING THE SYDNEY YORK MORPHOLOGICAL AND ACOUSTIC RECORDINGS OF EARS DATABASE

P. Guillon, R. Zolfaghari, N. Epain, A. van Schaik and C.T. Jin

THE UNIVERSITY OF SYDNEY

C. Hetherington, J. Thorpe and A. Tew

THE UNIVERSITY of York

Presented by Craig Jin

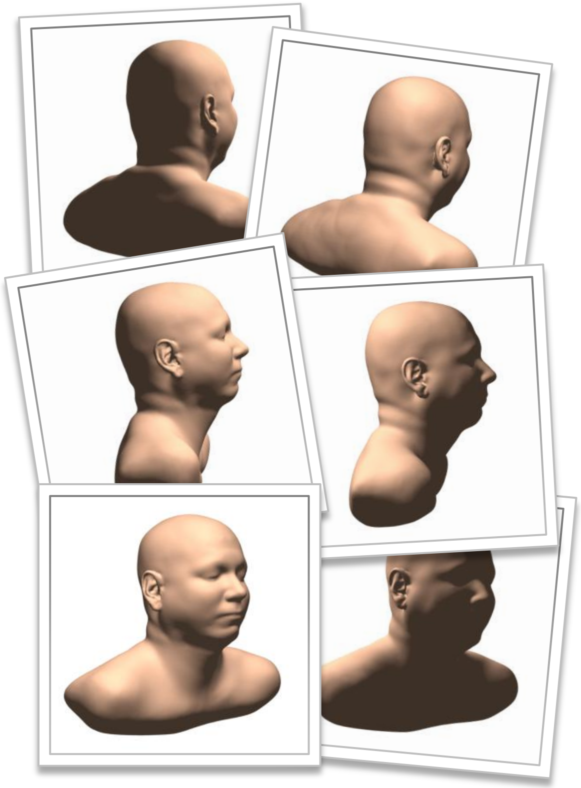
THE UNIVERSITY OF SYDNEY

OUTLINE

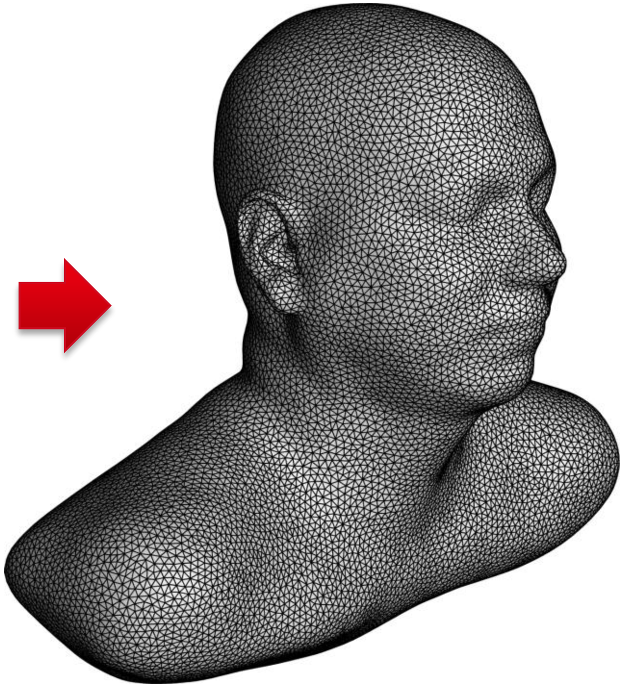
- Introduce a “Holy Grail” in 3D Audio
- Describe the SYMARE database
- Genesis of the database
- Preliminary validation of the data.

INTRODUCTION

› The Problem: Customize 3D audio for listeners based on 2D photos



2D Photos



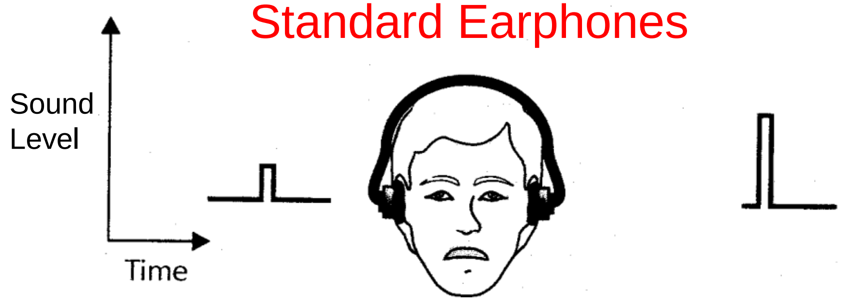
Parameterized
3D Head model

Personalised filters
for 3D audio

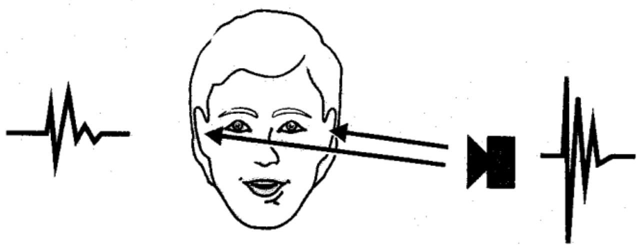
HEAD-RELATED IMPULSE RESPONSES

› Head-Related Impulse Responses (HRIRs): are acoustic filters for earphones that produce 3D Audio

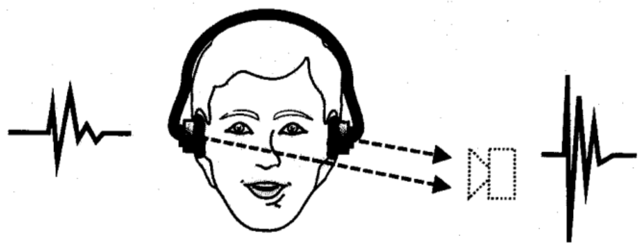
Standard Earphones



Normal Spatial Hearing



Earphones 3D Audio using HRIRs

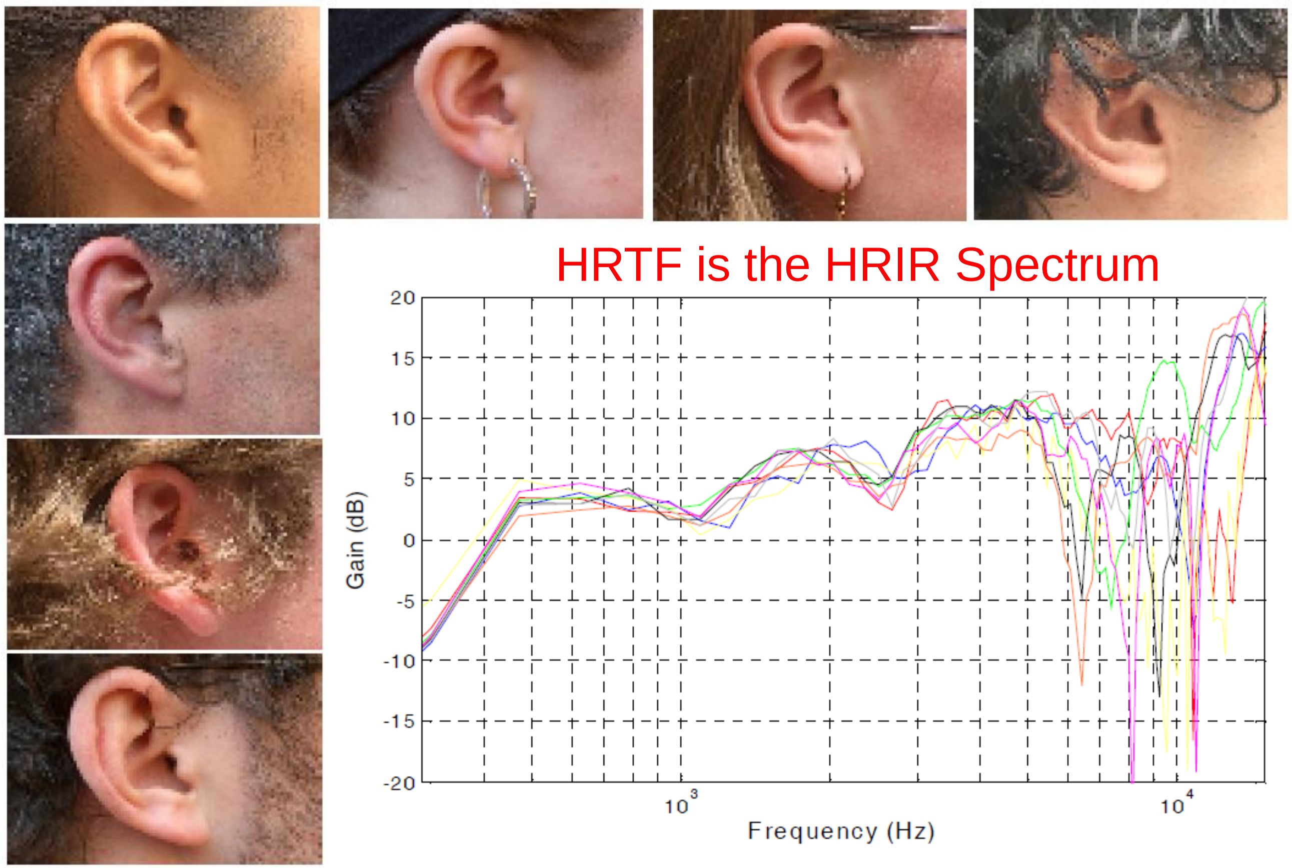


HEAD-RELATED IMPULSE RESPONSES

› Acoustically measuring HRIRs can be difficult and time-consuming. It would be better if we could predict HRIRs from morphology.

HEAD-RELATED IMPULSE RESPONSES

- Listener's ears differ and so do their HRIRs



THE SYMARE DATABASE

Sydney York Morphological and Recording of Ears Database

› The SYMARE database is a joint effort between the universities of Sydney and York.

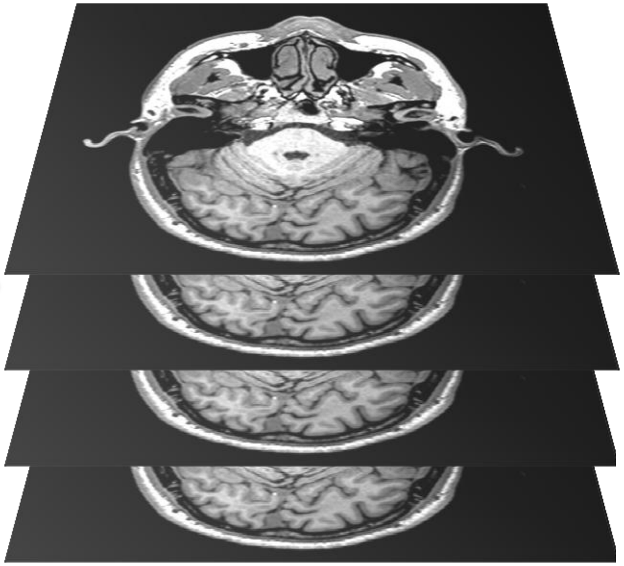
› The database contains morphological and acoustic data for 60 subjects. For each subject, the database provides:

- The subject's measured HRIRs for approximately 400 directions.
- High-resolution surface meshes of the head and combined head and torso at varying resolution *suitable for acoustic simulation*. In the meshes with the highest resolution, the size of the elements is on the order of 1 mm.
- Two sets of HRIRs simulated by applying the Fast-Multipole Boundary Element Method (FM-BEM) to the head-only and head-and-torso meshes.

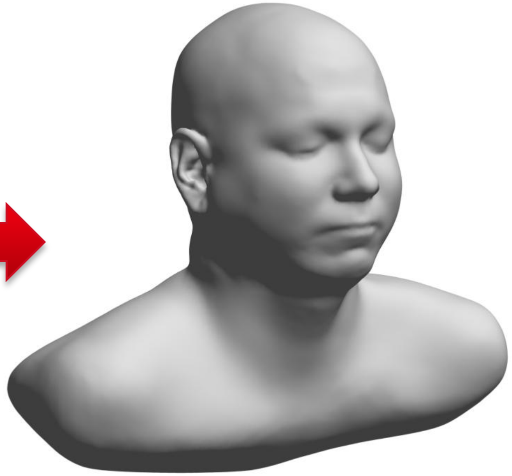
ACQUISITION OF HEAD AND TORSO MESHES

Acquiring Morphological Data

MRI scanning



DICOM images

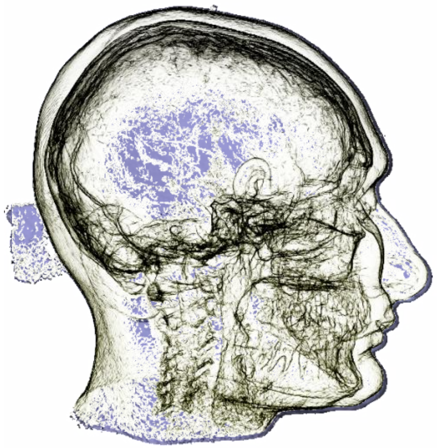


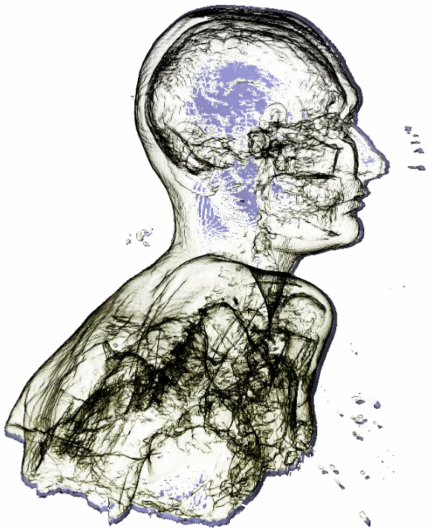
Surface mesh

FROM MRI SCANS TO 3D SURFACE MESHES

Generation of Surface Meshes

- For each subject, 2 MRI scans were performed: a high resolution scan of the head and a lower resolution scan of the head and torso.
- The surface meshes were first obtained by extracting iso-surfaces from both sets of MRI images.



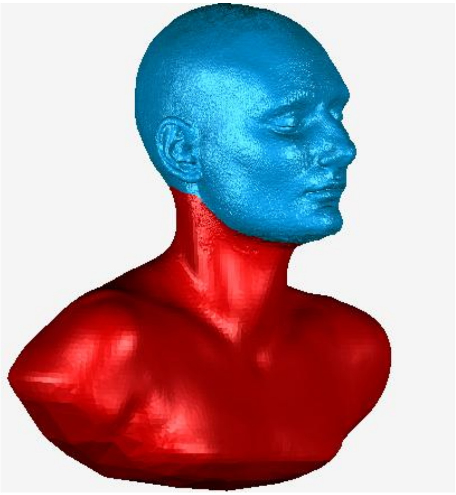


Some artifacts are present inside and outside the mesh. These artifacts were removed manually.

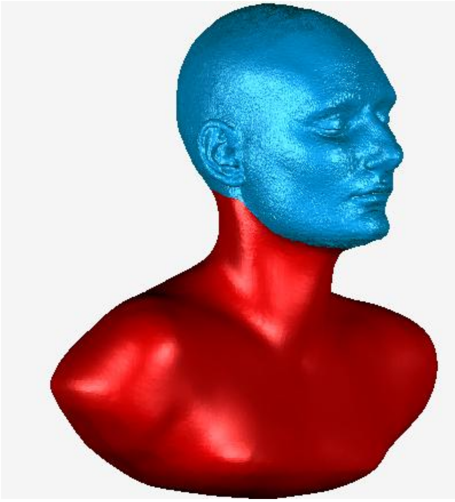
PREPARING SURFACE MESHES

Generation of Surface Meshes

- The high resolution head-only meshes were aligned and merged with the low-resolution torso mesh.
- The torsos' surfaces were smoothed.



Before smoothing the torso

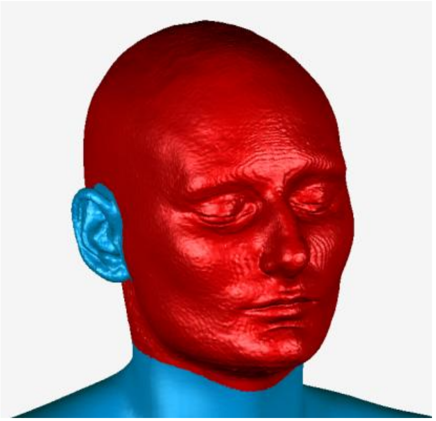


After smoothing the torso

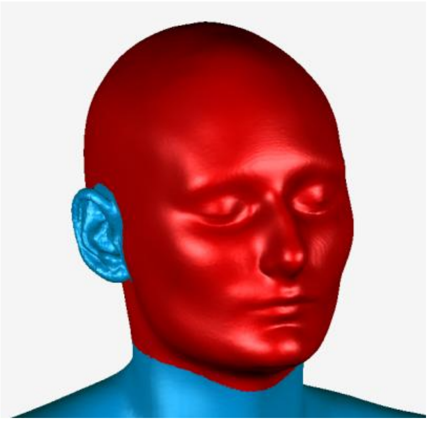
PREPARING SURFACE MESHES

Generation of Surface Meshes

- The meshes were re-meshed so that the triangular faces are as equilateral as possible.
- The heads (except the ear pinnae) were then smoothed.



Before smoothing the head



After smoothing the head

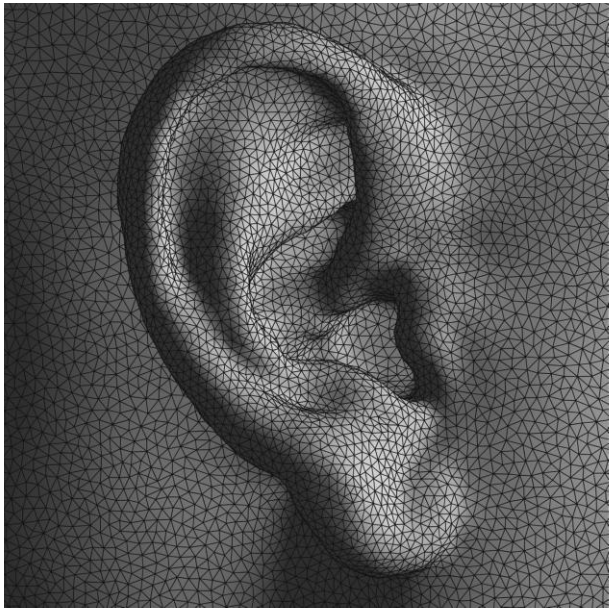
Triangle surface elements satisfy:

1. maximal to minimal edge length ratio < 5
2. Smallest angle > 15º
3. Largest angle < 160º

FINAL MESHES

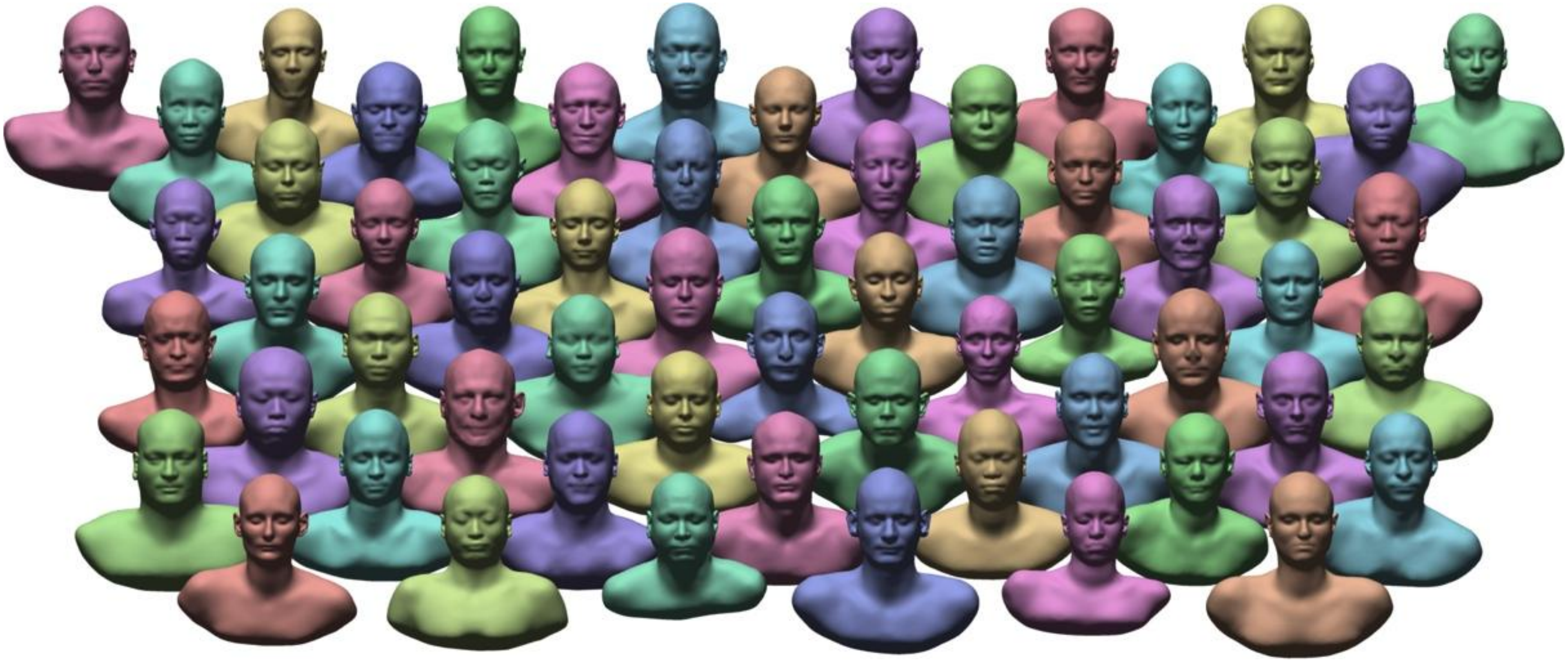
- The heads were separated from the torso to create head-only meshes.
- The Head-only and Head-and-Torso meshes were then re-meshed at various resolutions.

High-resolution Head-and-Torso mesh (~150k elts., f_{max}=16kHz)



Detail of a High-resolution Head-only mesh (~130k elts., f_{max}=20kHz)

FINAL MESHES



Terracotta Army in Xi’an China

FINAL MESHES

FM-BEM SIMULATIONS

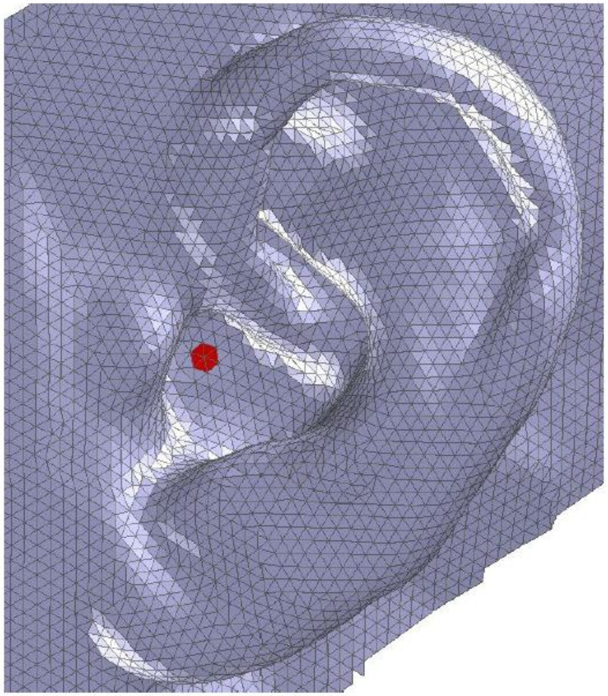
- The subjects HRTFs were simulated using Coustyx, a Fast-Multipole Boundary Element Method (FM-BEM) software.
- The size of the mesh elements determine the maximum frequency at which the BEM simulation can be done:

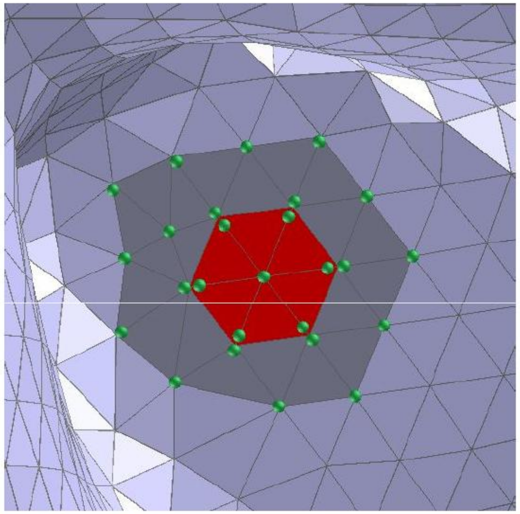
$$f_{\max} = \frac{c}{6\ l_{\max}}$$

where
- c is the speed of sound (~ 340 ms^{-1})
- l_{max} is the maximum edge length

- For each subject, two BEM simulations were performed:
  - Using a low-resolution Head-and-Torso mesh, up to 5.6 kHz.
  - Using a high-resolution Head-only mesh, up to 17 kHz.

Helmholtz Integral Equation Collocation Method with Chief Points



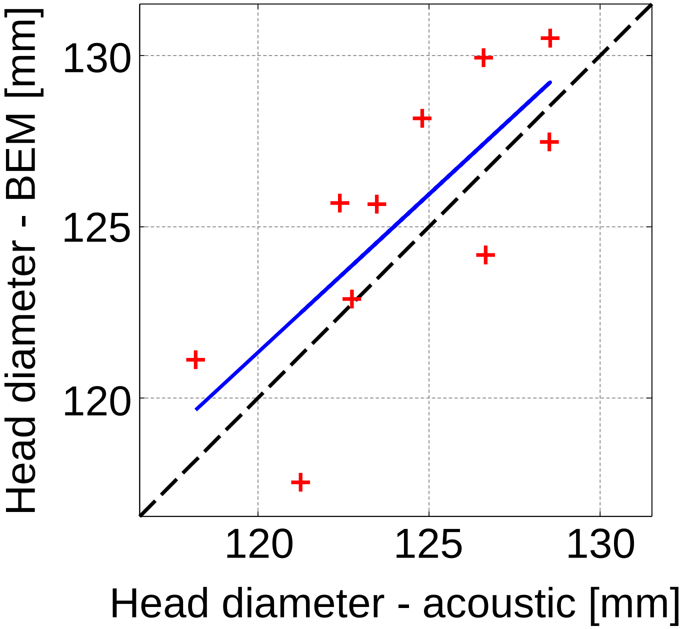


Use Reciprocity: $P(E,S_i) = P(S_i,E)$ to derive all HRTFs in one go.

Source placed in the ear; apply equivalent velocity boundary condition

ITD COMPARISON

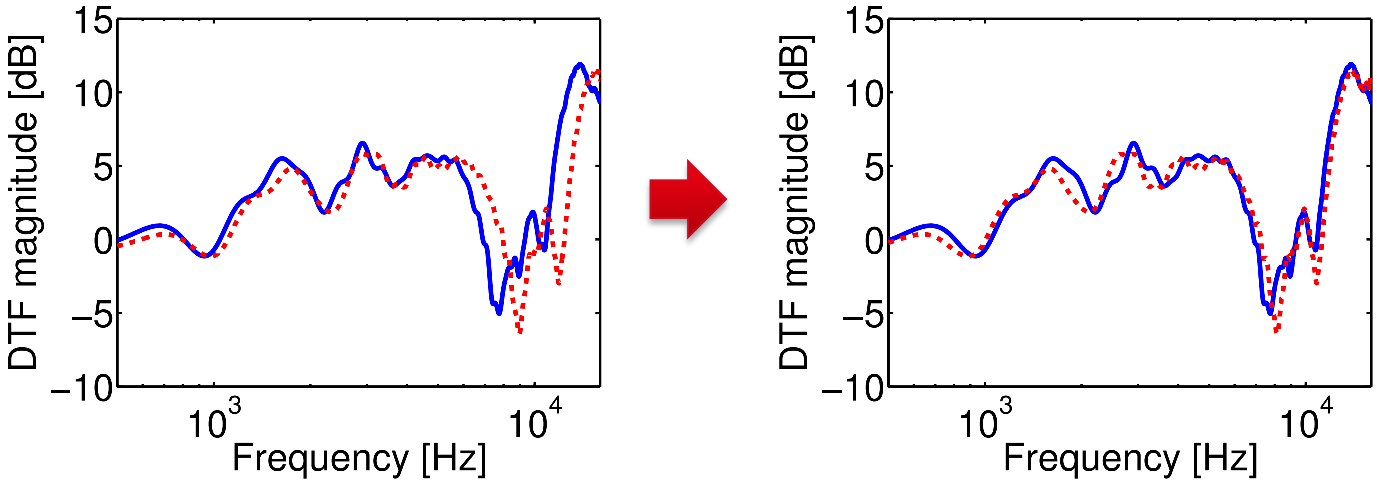
› We calculated the Interaural Time Delays (ITDs) for both the measured and simulated sets of data (Head-and-Torso meshes). We then applied Kuhn’s model to estimate the corresponding effective head diameters.



The difference in the effective diameters in on the order of a few percent.

FREQUENCY SCALING

› We applied Middlebrooks' frequency scaling technique to estimate the scaling factor that minimises the mismatches between the measured and simulated HRTF data.



Subject	1	2	3	4	5	6	7	8	9	10
Head only	1.053	1.031	1.099	1.042	1.064	1.020	1.020	1.075	1.020	0.990
Head & T.	1.021	1.010	1.031	1.010	1.010	1.010	1.020	1.031	1.020	1.053

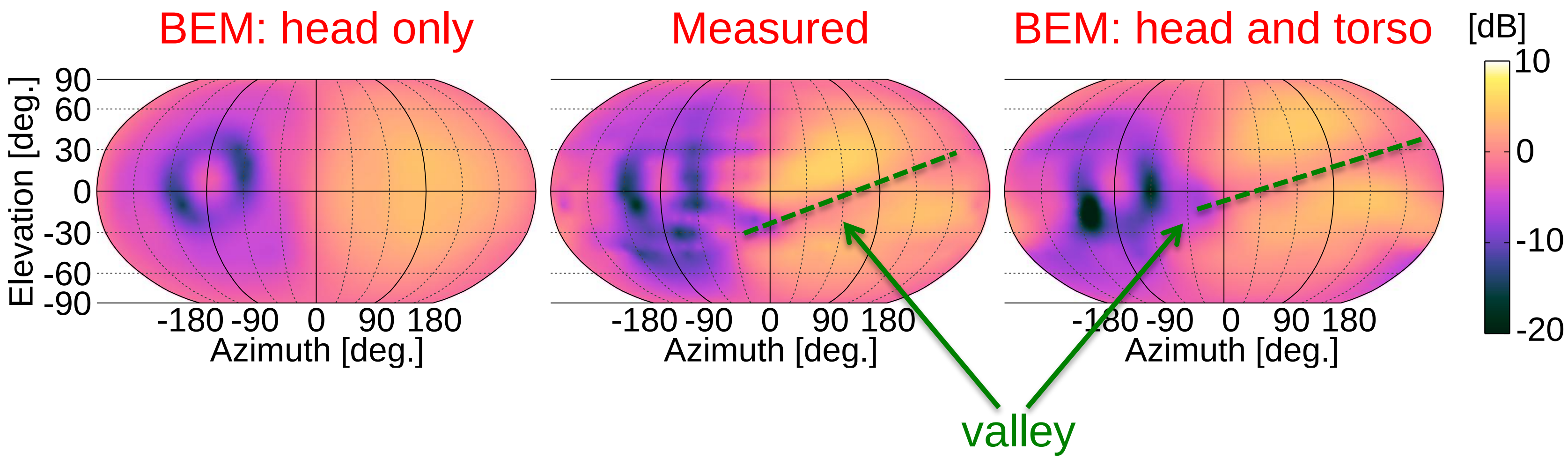
Frequency scaling factors for Subjects 1 to 10

SPATIAL FREQUENCY RESPONSE SURFACES

- A Spatial Frequency Response Surface (SFRS) is defined as the magnitude of the HRTFs for every direction at a given frequency. It characterises the directivity of the ear at this frequency.

INFLUENCE OF THE TORSO

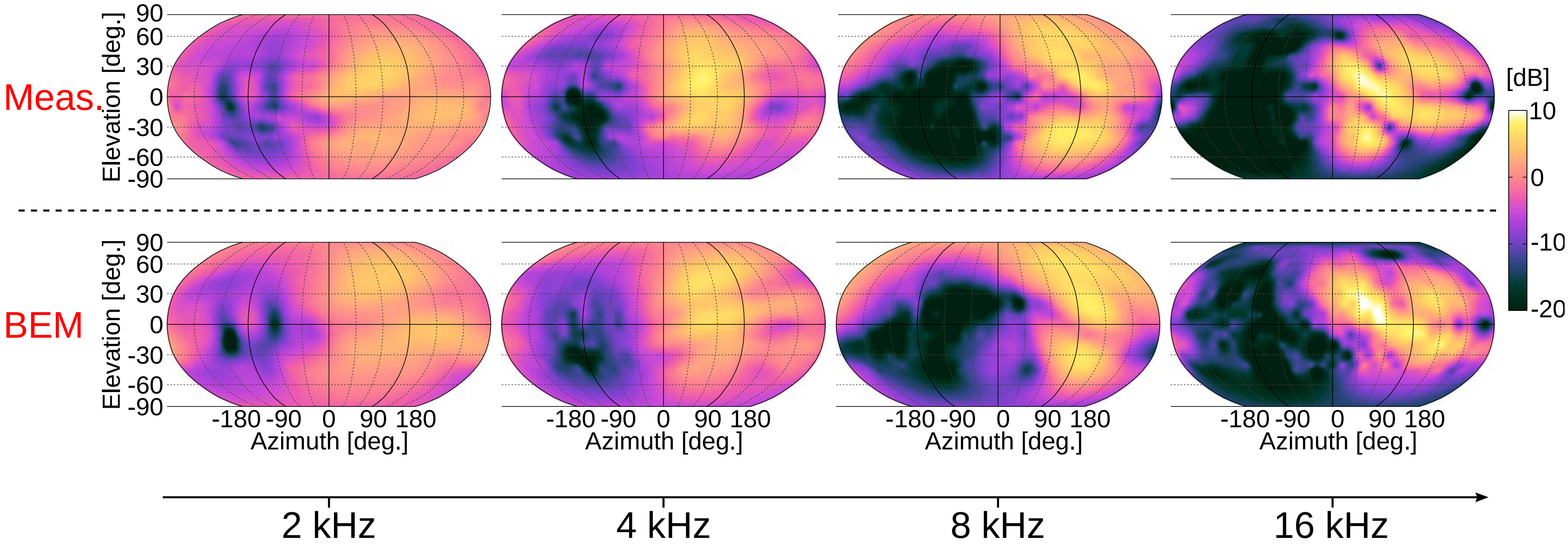
› We compared of the data obtained via BEM simulation using the Head-only meshes with that obtained using the Head-and-Torso meshes. The figure below compares Subject 1’s SFRS’s at 2kHz.



Important HRIR features are present only in the data corresponding to the head-and-torso mesh.

COMPARING SFRS'S: VISUAL INSPECTION

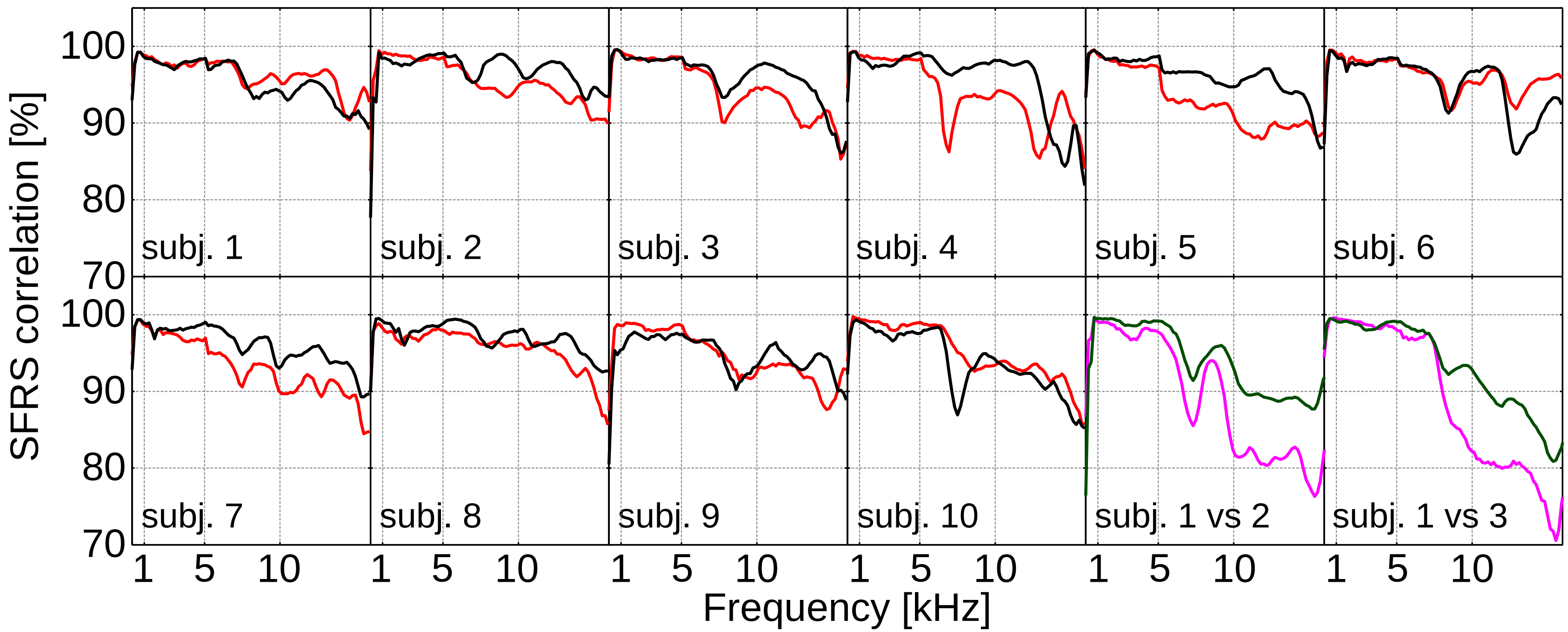
› Comparison of SFRS's corresponding to the measured and simulated data sets (the Head-and-Torso data is used below 5 kHz):



Many of the features of the SFRS corresponding to the measured data set are present in the BEM-simulated data set.

COMPARING SFRS'S: SPATIAL CORRELATION

- Spatial correlation between the SFRS's corresponding to the measured data and that corresponding to the simulated data, as a function of the frequency:



The simulated HRIR data captures some of the individual characteristics of the outer ear acoustics.

THE END

Thanks for your attention.

http://www.ee.usyd.edu.au/carlab/index.htm