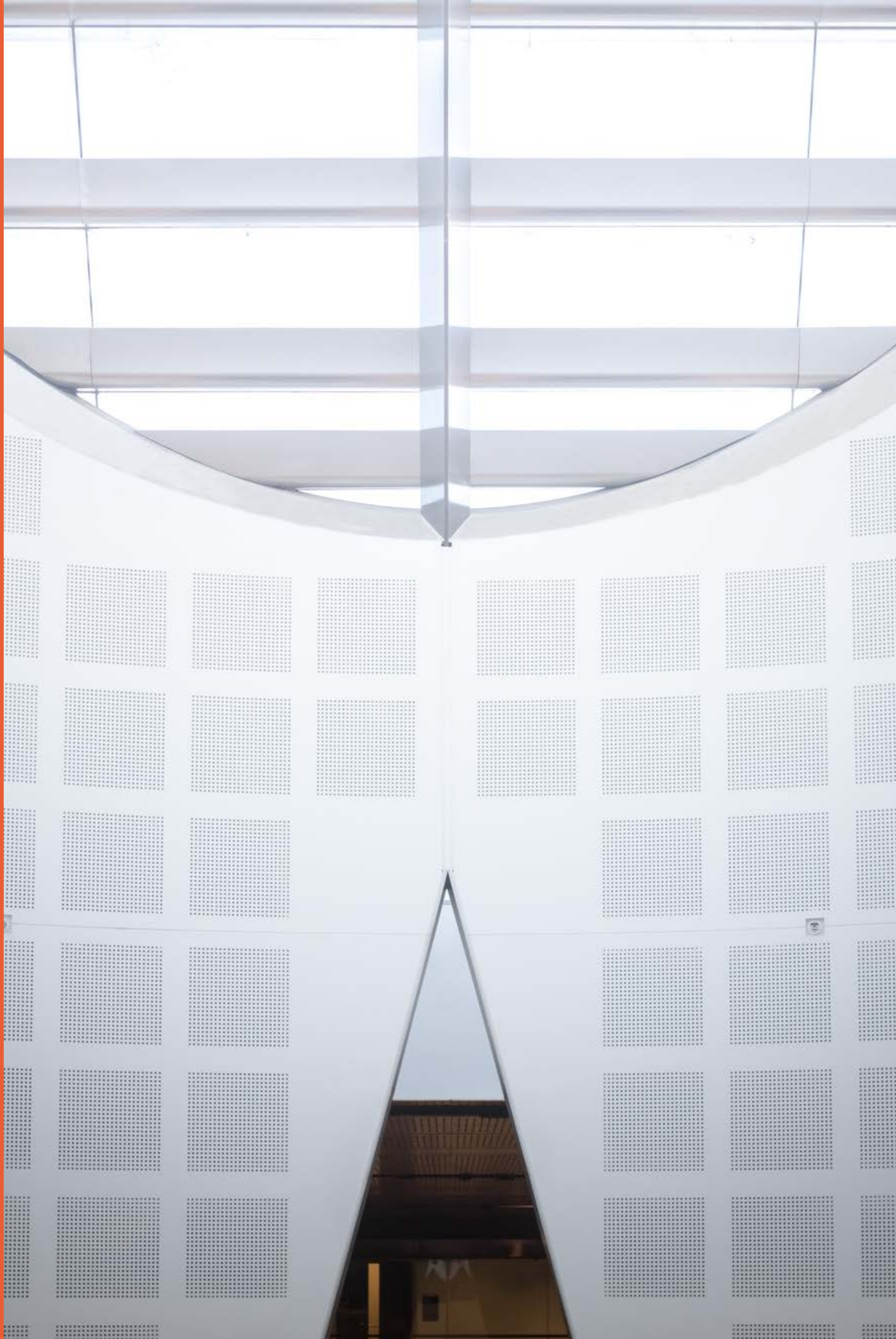
CONSIDERATIONS REGARDING INDIVIDUALIZATION OF HEAD-RELATED TRANSFER FUNCTIONS

(ICASSP 2018)

Sydney: Craig Jin, Reza Zolfaghari, Xian Long, Arun Sebastian, Shayikh Hossain
York: Tony Tew
Paris: Alexis Glaunes
Milan: Muhammad Shahnawaz, Augusto Sarti

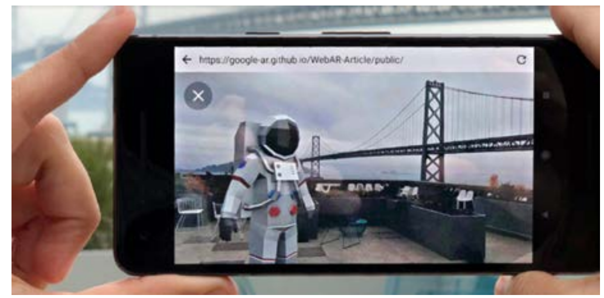
University of York
POLITECNICO MILANO 1863
THE UNIVERSITY OF SYDNEY
UNIVERSITÉ PARIS DESCARTES



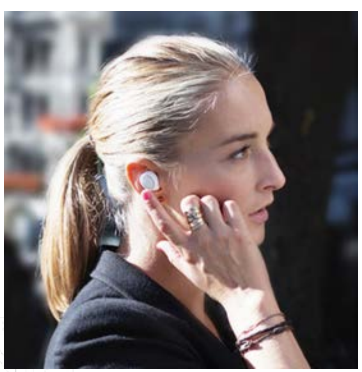
THE DRIVE TOWARD ENHANCED PERCEPTION VIA MIXED-REALITY SYSTEMS – RENEWS INTEREST IN BINAURAL VIRTUAL AUDITORY PERCEPTION

Industry matures and understands head-tracking is required.

Microsoft 3D Soundscape



Google Resonance Audio



Bragi Dash (Hearables)

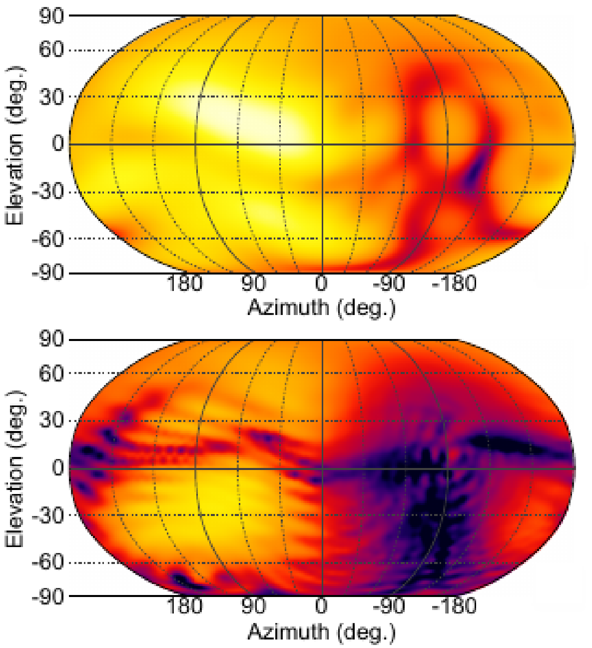
Oculus and 3D Sound Spatialization

DEMAND FOR BETTER CONTROL OF SPATIAL HEARING AND TIMBRAL PERCEPTION -- CUSTOMIZATION OF OUTER EAR ACOUSTIC FILTERS

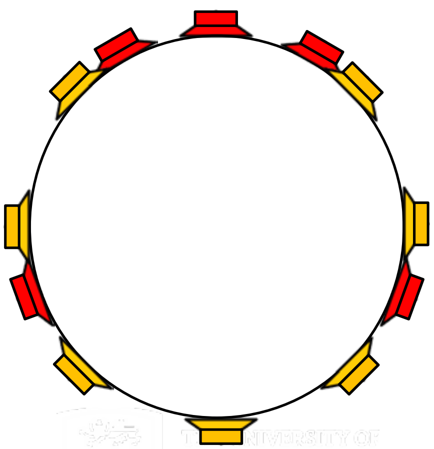
Two-sides: spatial perception and timbral perception.

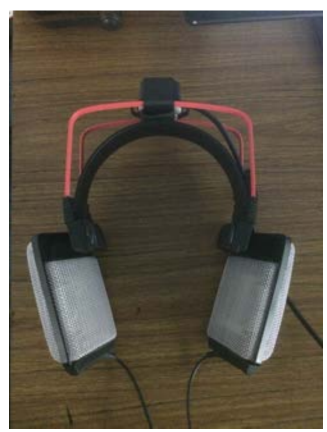
Outer Ears

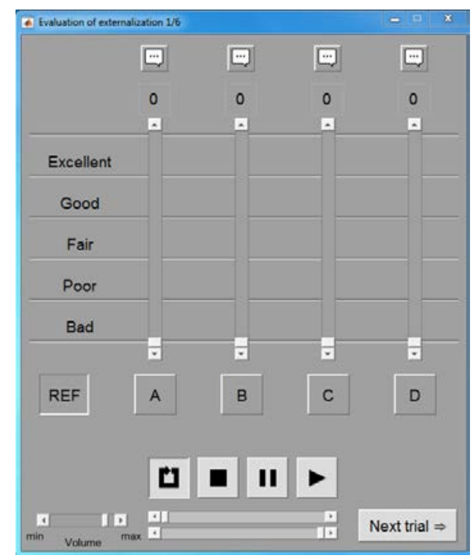
HRTFs and Acoustic Directivity



MUSIC LISTENING and
HRTF INDIVIDUALIZATION

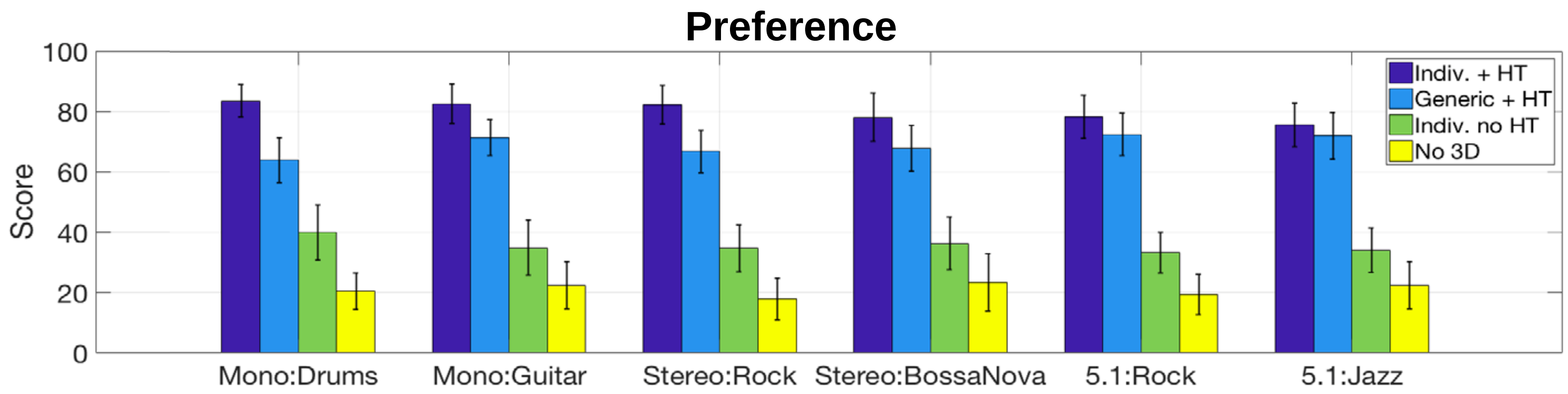


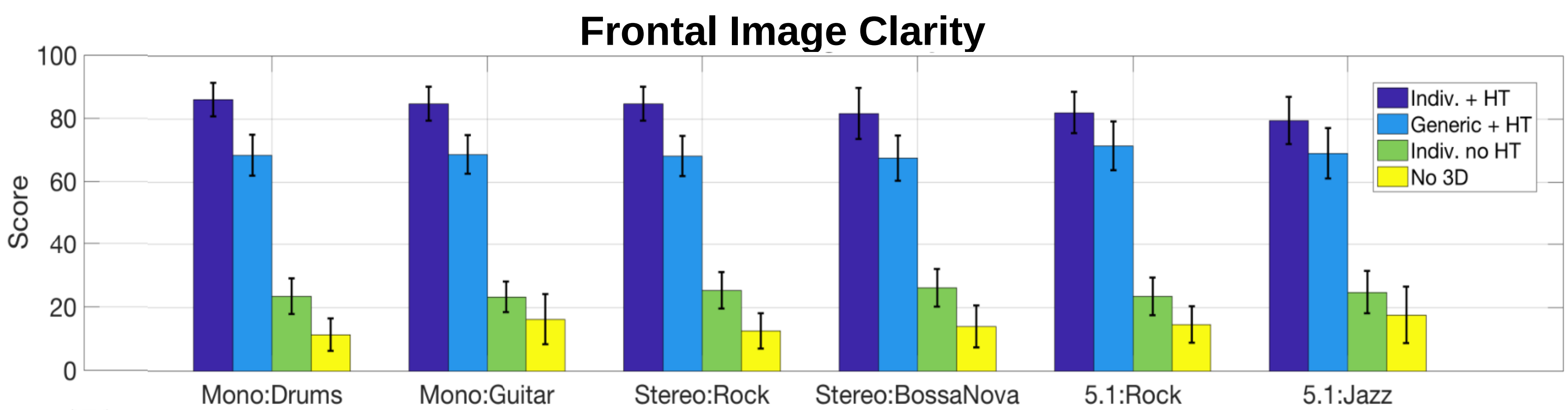




MUSIC LISTENING and HRTF INDIVIDUALIZATION

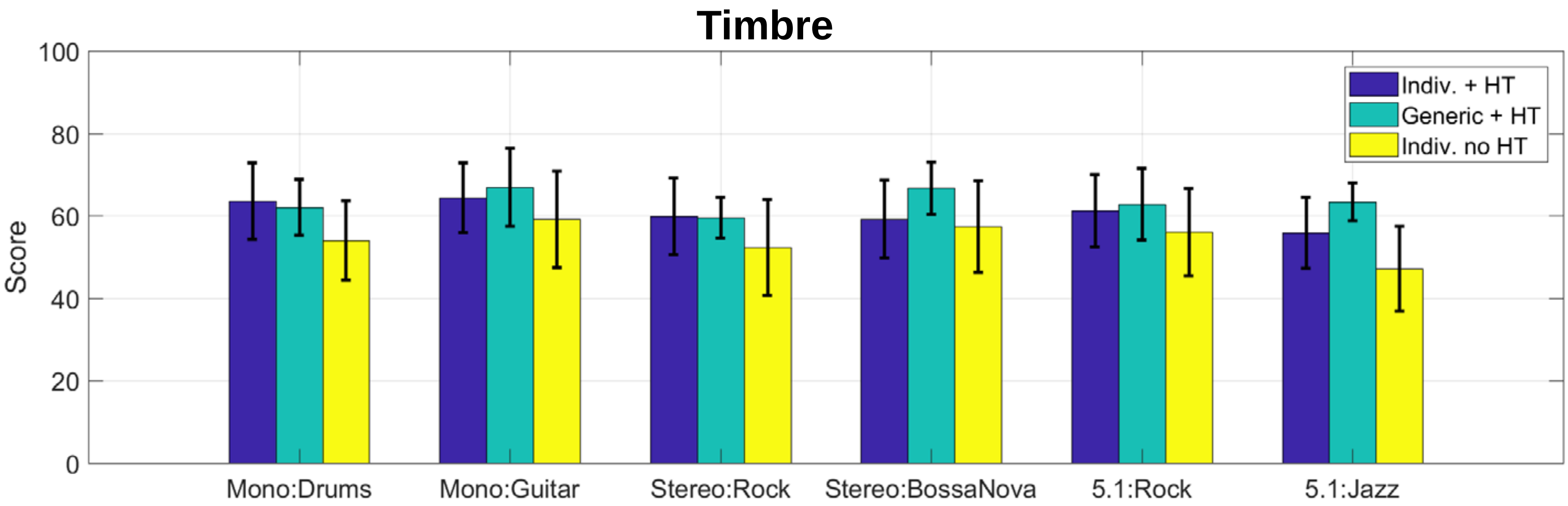
Mean data for
23 subjects





MUSIC LISTENING and HRTF INDIVIDUALIZATION

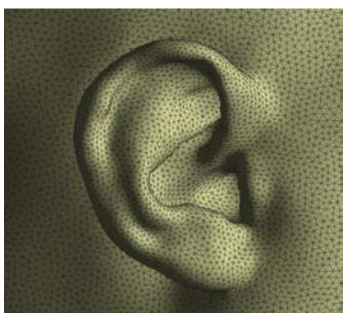
Mean data for
23 subjects

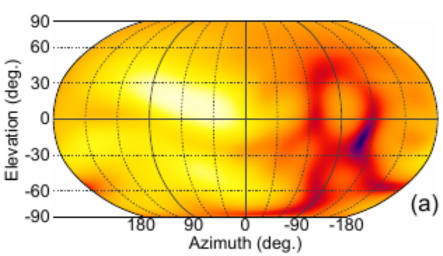


MORPHOACOUSTICS

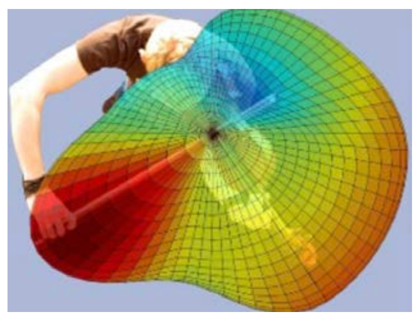
Morphoacoustics – the study and exploration of the inter-relationship between physical structure, acoustic properties, and perception.

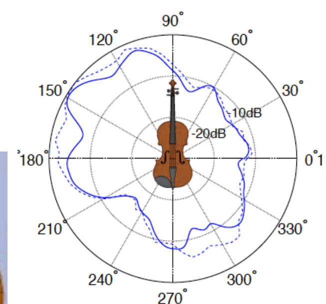
Outer Ears



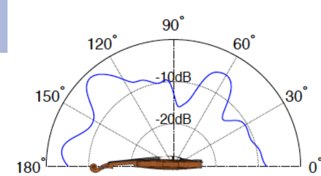


Violins



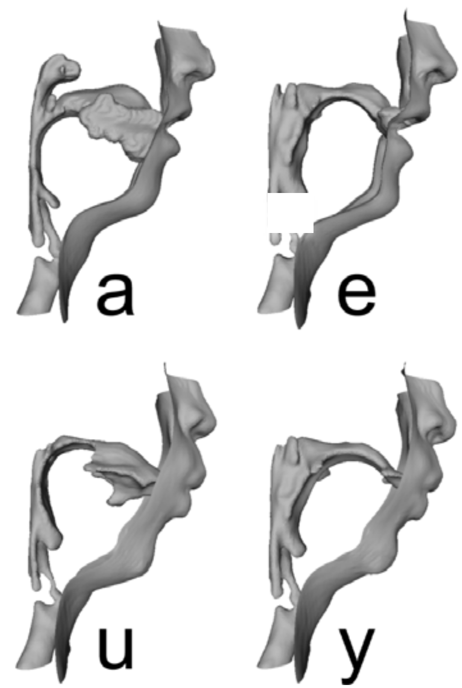


Horiz. plane
(882 Hz)



Lateral plane
(882 Hz)

Vocal Tract



MORPHOACOUSTICS

Key Concept: ***Deformations*** in one space relate to ***deformations*** in another space.

Requirements: Mathematics and tools to model large deformations and to explore the inter-relationship between deformations in different spaces.

Research Questions:

1. How to establish or identify corresponding features or landmarks.

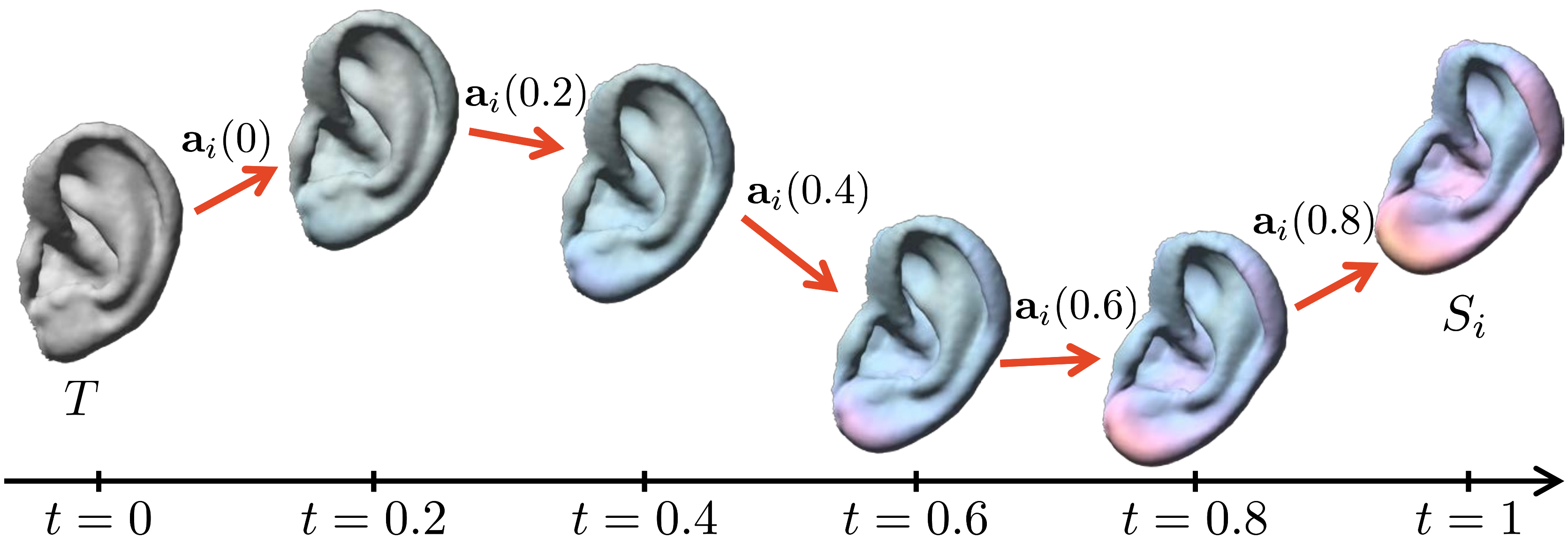
2. How to measure and quantify deformations.

3. How to characterise and define the inter-relationship between deformations in different spaces.

METRIC SPACE OF DEFORMATIONS

Large Deformation Diffeomorphic Metric Mapping

LDDMM provides a metric to measure shape deformations in a Riemannian space. Linearization of the Riemannian space provides a tangent space for statistical analyses.



LDDMM MINIMIZATION PROBLEM

Find $\mathbf{v}(t)_{t\in[0,1]}$ that minimises $J_{T,S}\left(\mathbf{v}(t)\right)$:

$$J_{T,S}\left(\mathbf{v}(t)\right) = \underbrace{\gamma \int_0^1 \|\mathbf{v}(t)\|_V^2 dt}_{\text{Smoothness of Deformation}} + \underbrace{E\left(\phi^{\mathbf{v}}(t,T)|_{t:[0\to 1]}, S\right)}_{\text{Data matching using a geometric measure referred to as currents.}} \ .$$

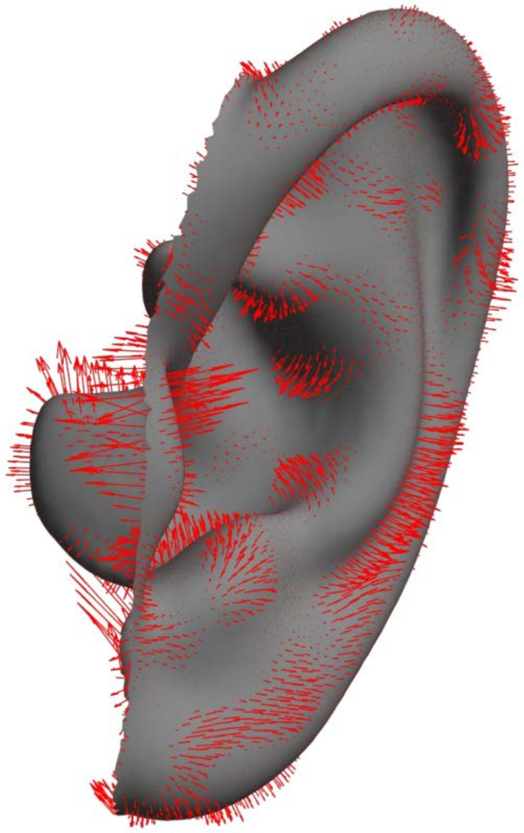
Smoothness of Deformation

Data matching using a geometric measure referred to as currents.

LDDMM INTUITION AND IMAGES

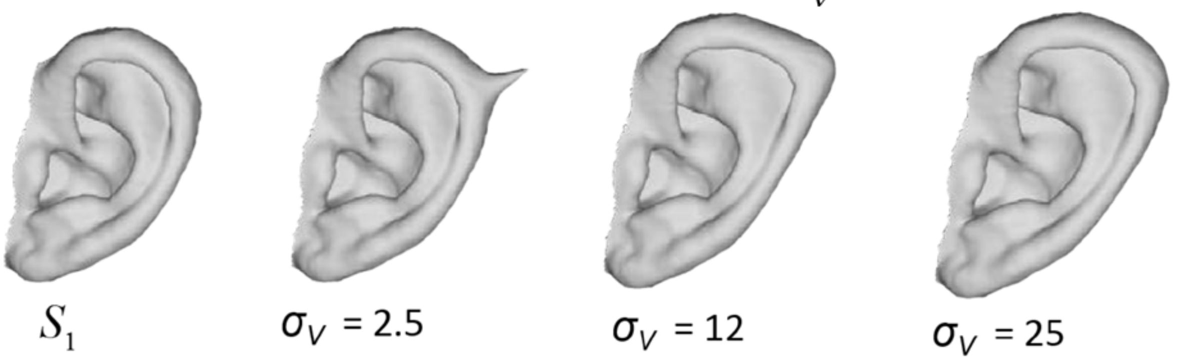
Momentum vector field at each vertex of the surface mesh model.

$$\mathbf{v}(t) = \frac{d\mathbf{x}(t)}{dt} = \sum_{n=1}^{N} k_V(\mathbf{x}_n(t), \mathbf{x}(t))\, \boldsymbol{a}_n(t)$$



Single momentum vector, with varying σ_V in the kernel.

$$k_V(\mathbf{x}, \mathbf{y}) = \frac{1}{1 + \frac{\|\mathbf{x}-\mathbf{y}\|^2}{\sigma_V^2}}$$



THREE FUNDAMENTAL LDDMM OPERATIONS

- **Mapping** is the operation of calculating the deformation from shape T to shape, S_i:

$$\mathbf{a}_i(t)^{0\leq t\leq 1} = \mathscr{M}(T, S_i)$$

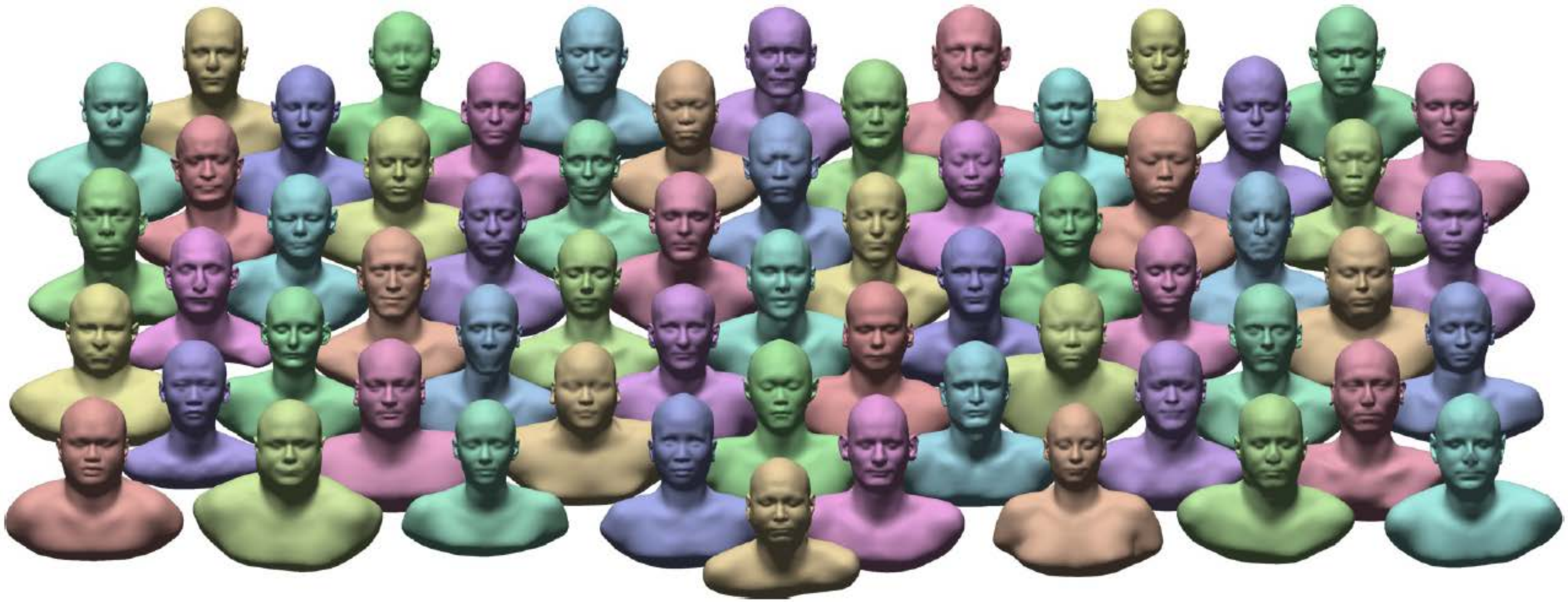
- **Shooting** is the operation of morphing T into an approximation of, S_i, given the initial momentum vectors $\mathbf{a}_i(0)$:

$$\{S_i^{'}, \mathbf{a}_i(t)^{0\leq t\leq 1}\} = \mathscr{S}(T, \{\mathbf{a}_i(0)\}) \ .$$

SYMARE DATABASE

Sydney York Morphological and Recording of Ears Database

High-Resolution Meshes, HRIRs, FM-BEM HRIR Simulations: 60 Listeners

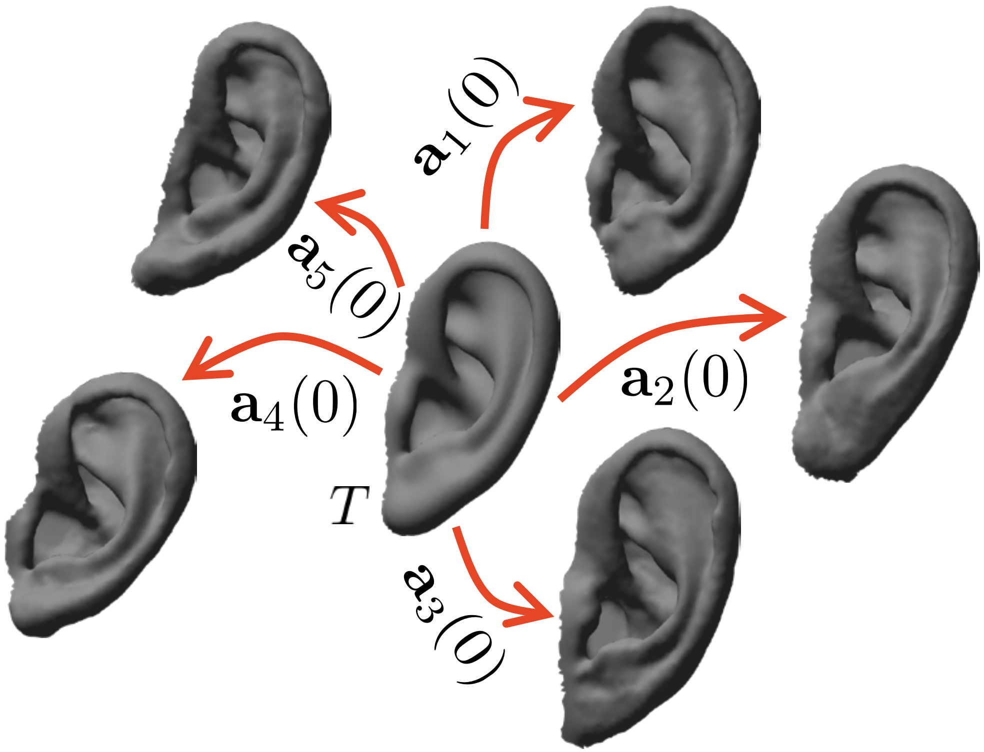


TEMPLATE ESTIMATION – POPULATION AVERAGE

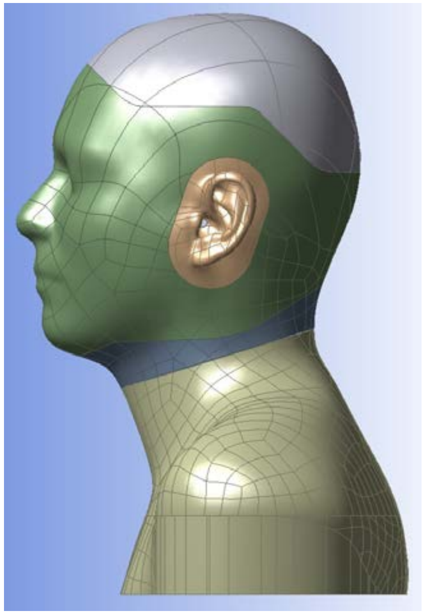
Template Estimation Principle:

$$\sum_{i=1}^{L} \boldsymbol{\alpha}(0, T, S_i) = 0 \;\; ,$$

$$\text{where } \{\boldsymbol{\alpha}(t, T, S_i)\} = \mathscr{M}(T, S_i, \sigma_V, \sigma_W)\,.$$



TEMPLATE ESTIMATION – POPULATION AVERAGE

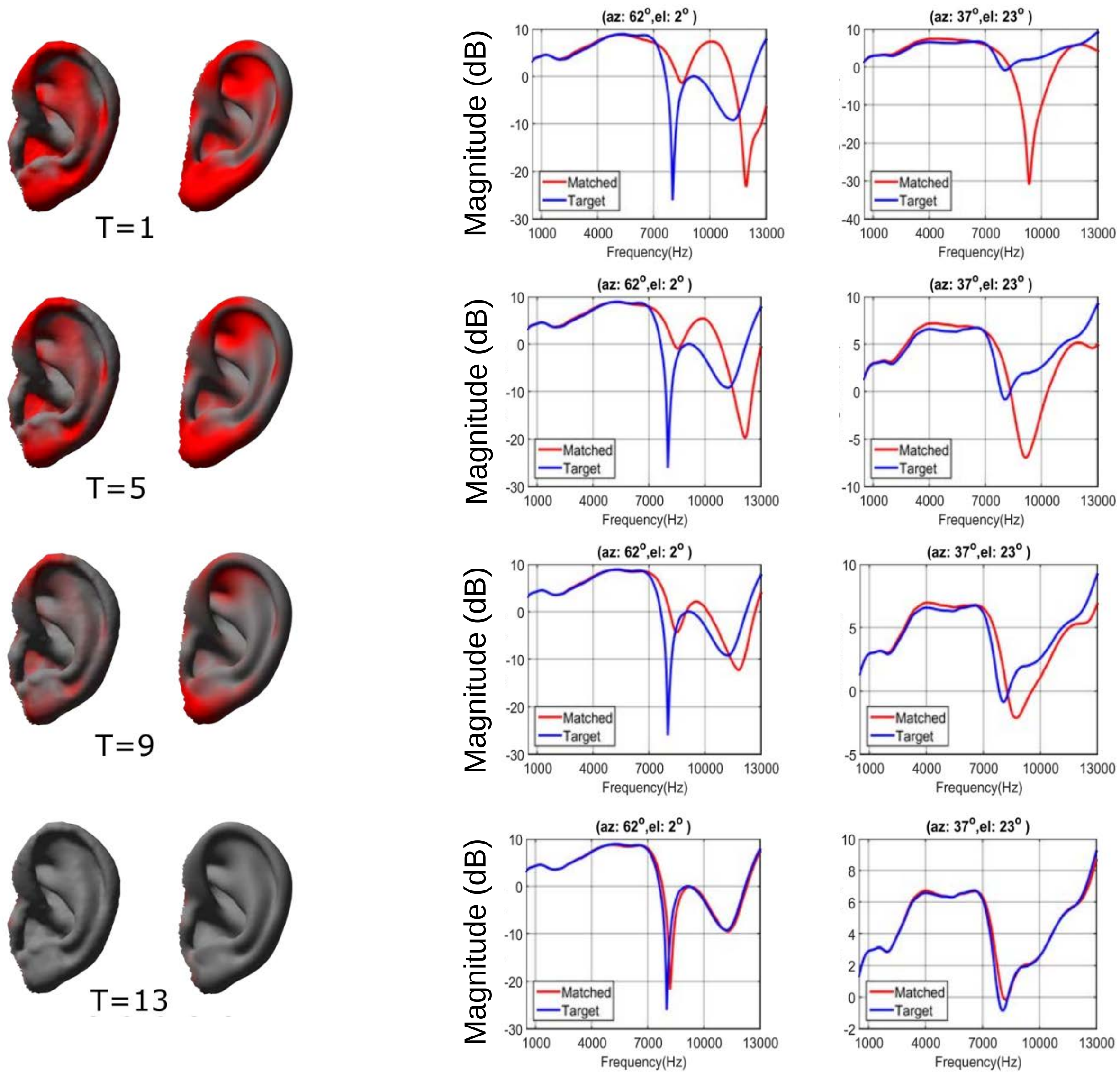


Variable Name	x=Mean(mm)	std(mm)	$y = \langle H(20kHz)\rangle_{8kHz}\langle E(20kHz)\rangle_{20kHz}$	Deviation$(100(1-\frac{y}{x})))$
de_1	37.4684	3.5251	37.2704	-0.5314
de_2	13.5886	1.4749	13.6559	0.4924
de_3	14.1253	1.7303	14.1459	0.1454
de_4	10.2448	1.5407	10.2631	0.1785
de_5	48.3337	4.1452	47.0130	-2.8093
de_6	17.4733	2.8634	17.2845	1.0109
de_7	18.8608	2.7212	18.3512	-2.7770
de_8	44.8029	4.2071	44.1009	-1.5919
de_9	26.4235	2.6073	26.7447	1.2011
de_{10}	18.9115	1.5685	18.6310	-1.5057
de_{11}	24.6659	2.6271	24.9915	1.3031
de_{12}	10.9502	1.7323	10.9830	0.2987
dh_1	145.2545	9.7463	144.9769	-0.1915
dh_2	47.8141	4.2460	46.4775	-2.8760
dh_3	123.1606	7.4160	123.7949	0.5123
dh_4	196.1966	9.5911	196.9262	0.3705
dt_1	393.954	29.298	406.070	-3.0755

Variable Name	Description	x =Mean (cm)	std (cm)	$y = \langle E(20kHz)\rangle_{20kHz}$	Deviation$(100(1-\frac{y}{x}))$	B&K (DZ 9764)	KEMAR (Left)
$d_1 + d_2$	Concha height	2.4465	0.2418	2.40	-2.18	2.20	2.60
d_3	Concha width	1.5705	0.2618	1.50	1.30	2.00	2.00
d_5	Pinna height	6.3853	0.4563	6.50	-2.57	7.00	7.10
d_6	Pinna width	2.9467	0.2927	2.90	1.58	3.60	3.10

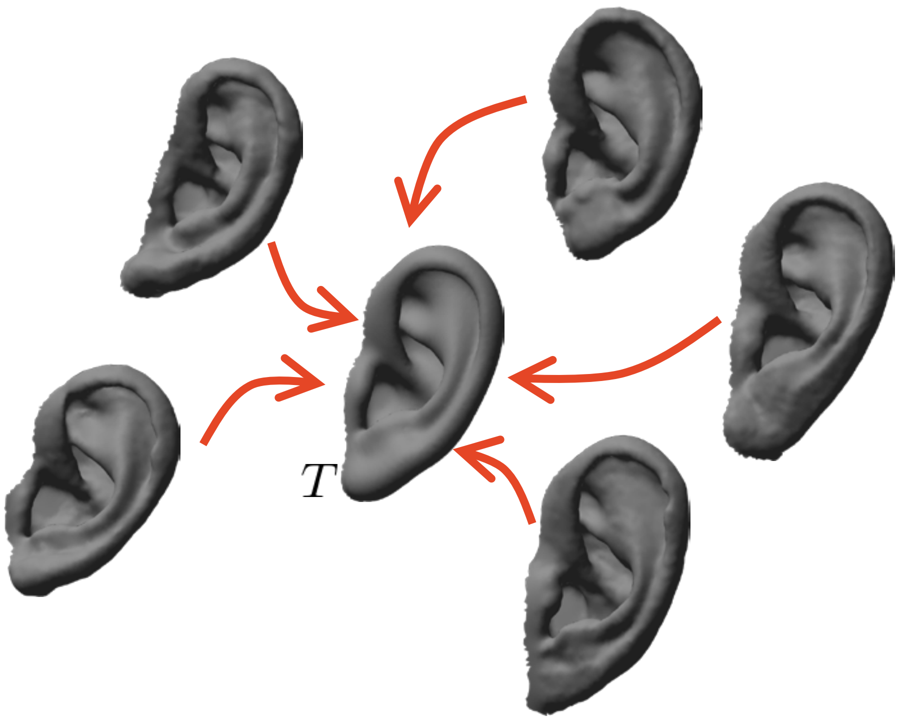
METRIC SPACE OF DEFORMATIONS

Large Deformation Diffeomorphic Metric Mapping



AFFINE-TRANSFORMATION OF EARS

All ears in the database are matched to the template ear via an affine transformation. We recalculate the HRTFs using FM-BEM.



KERNEL PCA EAR MODEL

1. Create zero-mean data

$$\bar{\mathbf{a}} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{a}_i(0)$$

$$\hat{\mathbf{a}}_i = \mathbf{a}_i(0) - \bar{\mathbf{a}}$$

2. Compute the correlation matrix:

$$\mathbf{A} = [\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \ldots, \hat{\mathbf{a}}_L]$$

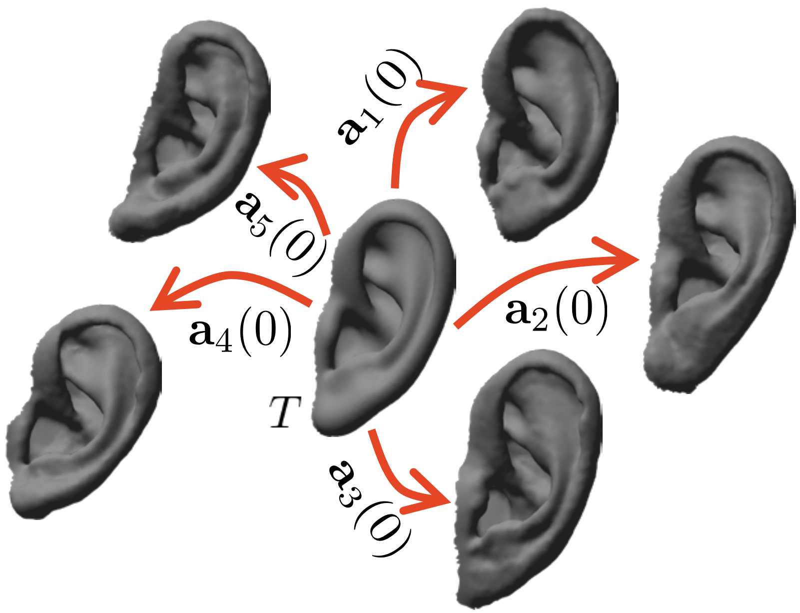
$$\mathbf{C} = \frac{1}{L-1}\hat{\mathbf{A}}^\mathsf{T}\mathbf{K}\hat{\mathbf{A}}$$

3. Calculate singular value decomposition:

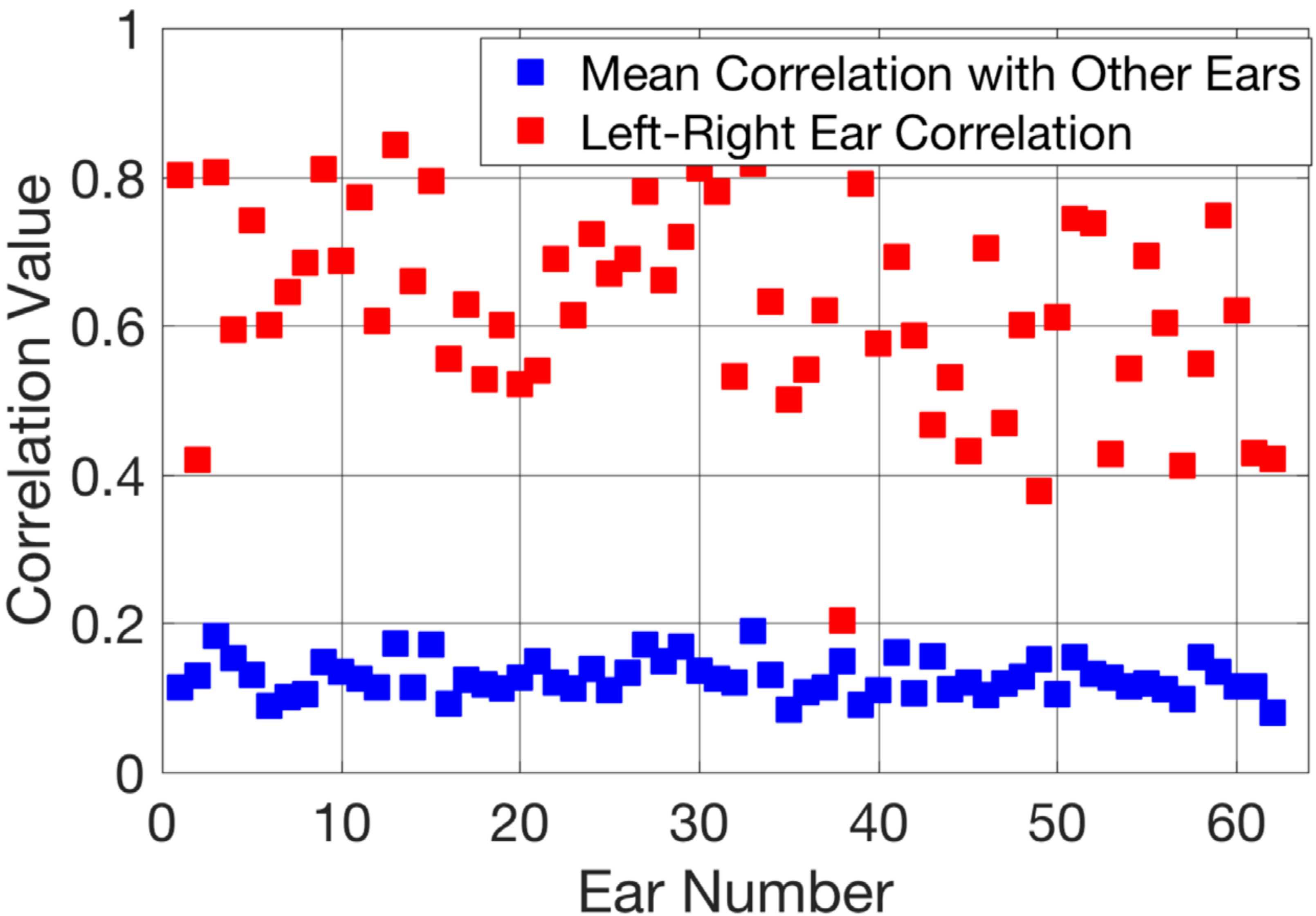
$$\mathbf{C} = \mathbf{V}\mathbf{D}\mathbf{V}^\mathsf{T}$$

4. Calculate principal components

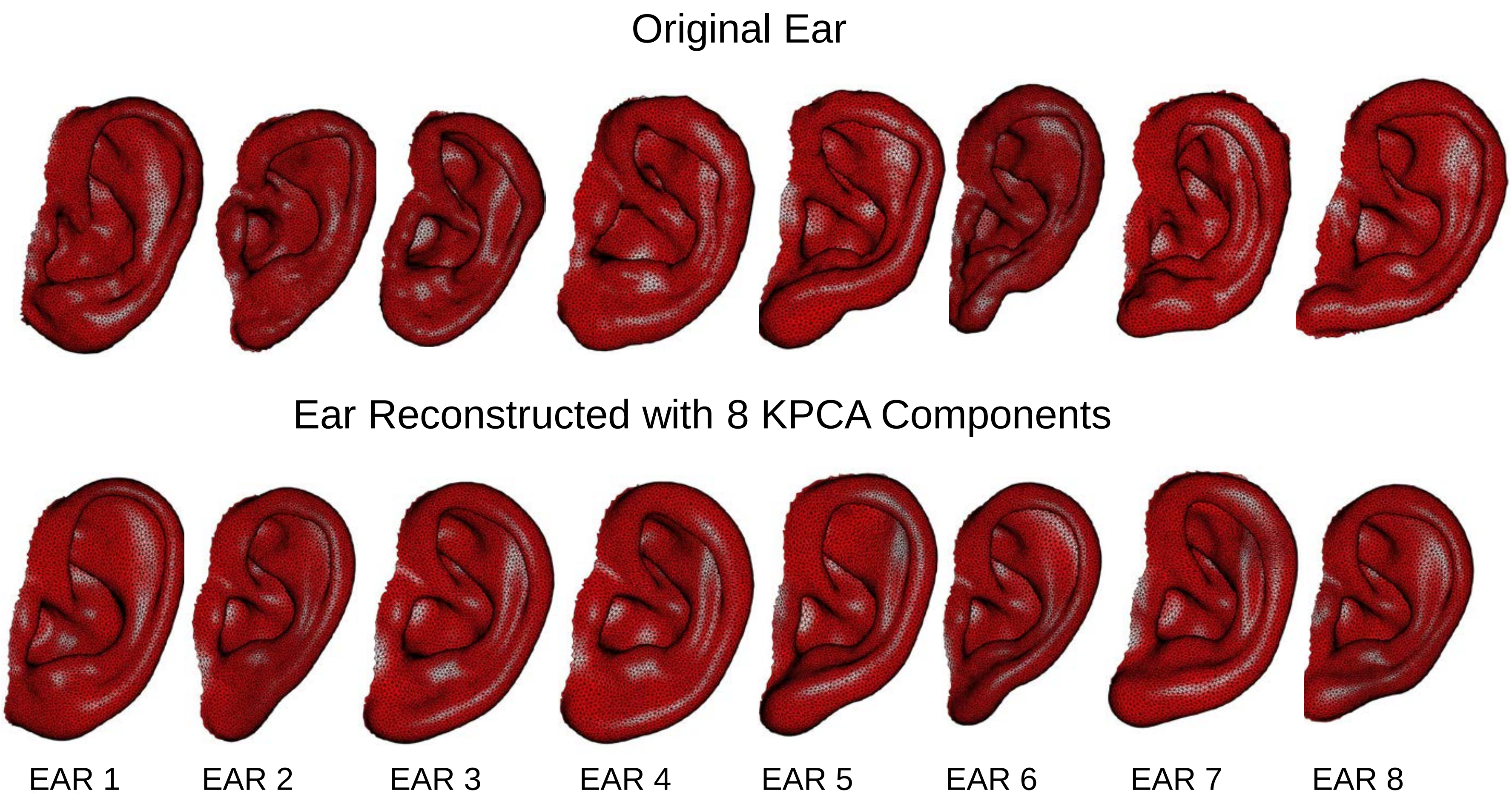
$$\mathbf{U} = \mathbf{A}\mathbf{V}\mathbf{D}^{-\frac{1}{2}}$$



HOW SIMILAR ARE RIGHT AND LEFT EARS ?

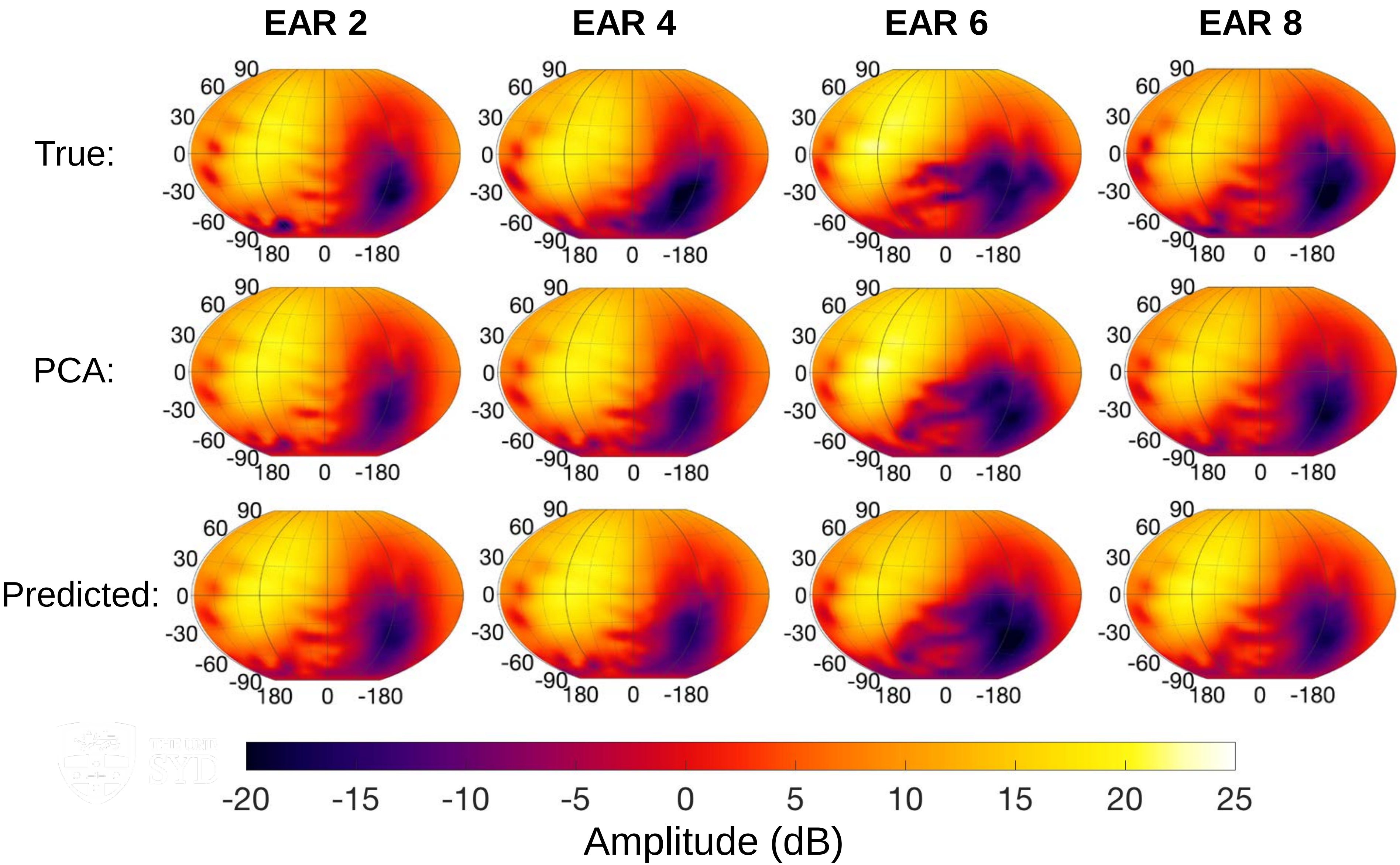


RANGE OF MODELED EAR SHAPE VARIATION



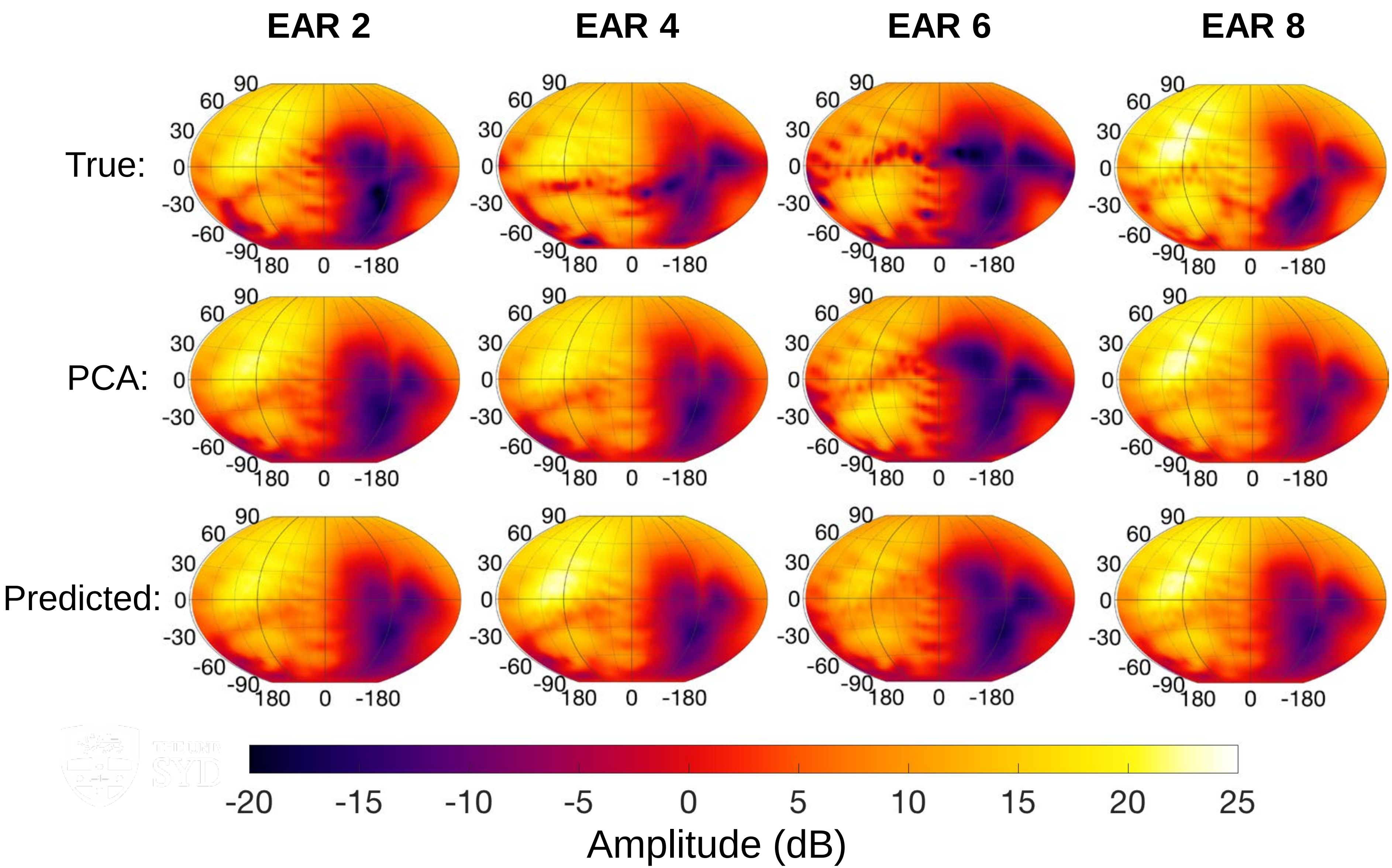
RELATING EAR SHAPE TO ACOUSTIC DIRECTIVITY

6000 Hz



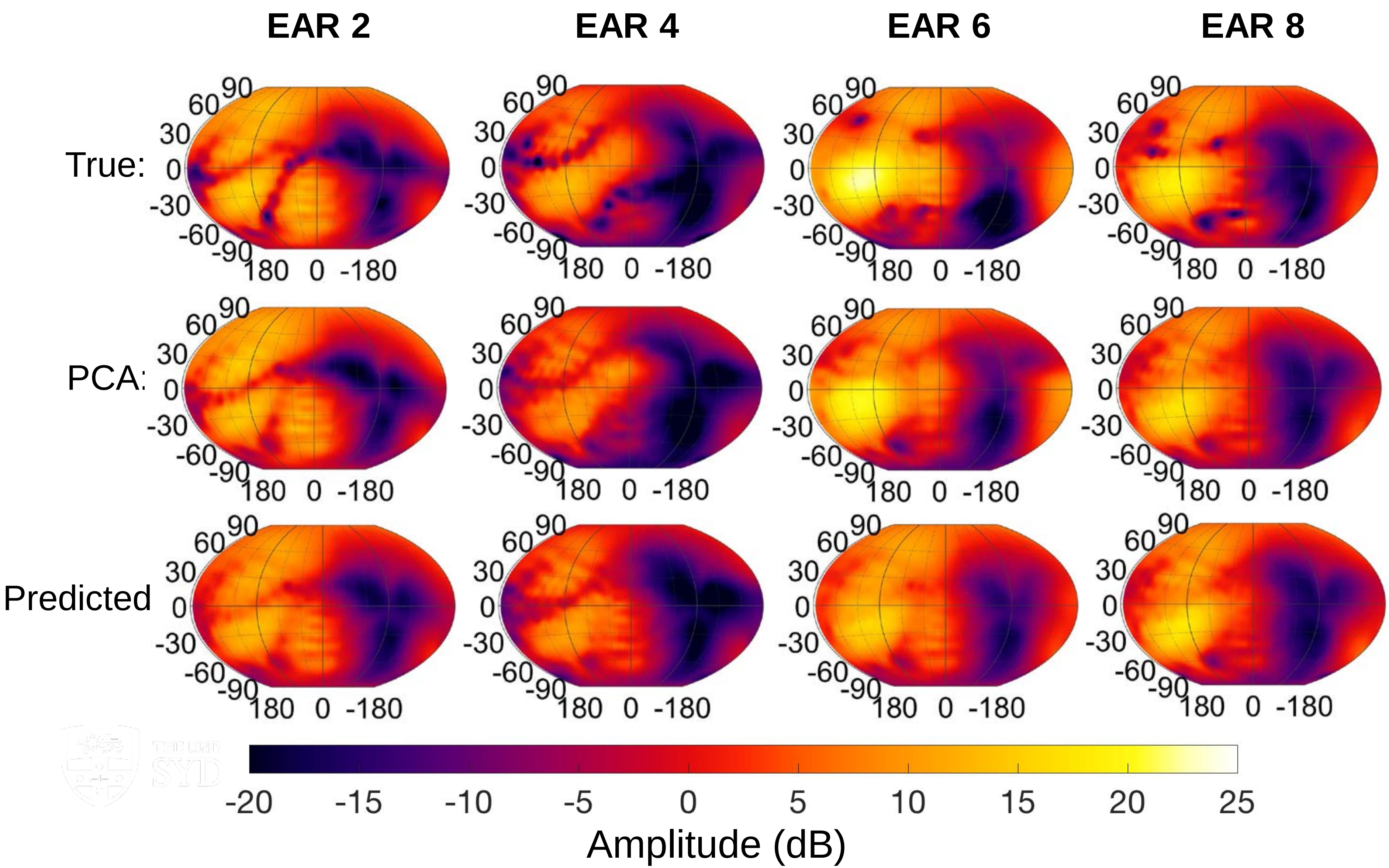
RELATING EAR SHAPE TO ACOUSTIC DIRECTIVITY

8063 Hz



RELATING EAR SHAPE TO ACOUSTIC DIRECTIVITY

9938 Hz



THE END

Thanks for Listening