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
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

2 kHz

16 kHz
MUSIC LISTENING and HRTF INDIVIDUALIZATION
MUSIC LISTENING and HRTF INDIVIDUALIZATION

Mean data for 23 subjects

Preference

Frontal Image Clarity
MUSIC LISTENING and HRTF INDIVIDUALIZATION

Mean data for 23 subjects

![Graph showing timbre score for different audio configurations and HRTF individualization types](image)
MORPHOACOUSTICS

Morphoacoustics – the study and exploration of the inter-relationship between physical structure, acoustic properties, and perception.
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 $v(t)_{t \in [0,1]}$ that minimises $J_{T,S}(v(t))$:

$$J_{T,S}(v(t)) = \gamma \int_0^1 \|v(t)\|_V^2 dt + E(\phi^v(t, T)|_{t:[0\to1], S}).$$

- **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.

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

Single momentum vector, with varying \( \sigma_V \) 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, $S_i$:
  \[ a_i(t)_{0 \leq t \leq 1} = M(T, S_i) \]

• **Shooting** is the operation of morphing $T$ into an approximation of, $S_i$, given the initial momentum vectors $a_i(0)$:
  \[ \{S_i', a_i(t)_{0 \leq t \leq 1}\} = \mathcal{S}(T, \{a_i(0)\}) . \]
SYMARE DATABASE
Sydney York Morphological and Recording of Ears Database
High-Resolution Meshes, HRIRs, FM-BEM HRIR Simulations: 60 Listeners
Template Estimation Principle:

\[ \sum_{i=1}^{L} \alpha(0, T, S_i) = 0 , \]

where \( \{ \alpha(t, T, S_i) \} = \mathcal{M}(T, S_i, \sigma_V, \sigma_W) \).
### TEMPLATE ESTIMATION – POPULATION AVERAGE

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Description</th>
<th>$x = \text{Mean (cm)}$</th>
<th>$\text{std (cm)}$</th>
<th>$y = (E(20kH\pi z))_{20Hz}$</th>
<th>Deviation $[100(1-\frac{y}{x})]$</th>
<th>BkK (DZ 9764)</th>
<th>KEMAR (Left)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_1 + d_2$</td>
<td>Concha height</td>
<td>2.4465</td>
<td>0.2418</td>
<td>2.40</td>
<td>-2.18</td>
<td>2.20</td>
<td>2.60</td>
</tr>
<tr>
<td>$d_3$</td>
<td>Concha width</td>
<td>1.5705</td>
<td>0.2618</td>
<td>1.50</td>
<td>1.30</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>$d_5$</td>
<td>Pinna height</td>
<td>6.3853</td>
<td>0.4563</td>
<td>6.50</td>
<td>-2.57</td>
<td>7.00</td>
<td>7.10</td>
</tr>
<tr>
<td>$d_6$</td>
<td>Pinna width</td>
<td>2.9467</td>
<td>0.2927</td>
<td>2.90</td>
<td>1.58</td>
<td>3.60</td>
<td>3.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>$x = \text{Mean (mm)}$</th>
<th>$\text{std (mm)}$</th>
<th>$\text{Deviation (100(1 - \frac{y}{x}))}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_1$</td>
<td>37.8684</td>
<td>13.5254</td>
<td>-0.5314</td>
</tr>
<tr>
<td>$d_2$</td>
<td>13.5886</td>
<td>1.4749</td>
<td>0.4024</td>
</tr>
<tr>
<td>$d_3$</td>
<td>14.1253</td>
<td>1.7388</td>
<td>0.1454</td>
</tr>
<tr>
<td>$d_4$</td>
<td>10.2148</td>
<td>1.5807</td>
<td>0.1785</td>
</tr>
<tr>
<td>$d_5$</td>
<td>8.9337</td>
<td>4.1492</td>
<td>2.8093</td>
</tr>
<tr>
<td>$d_6$</td>
<td>17.1733</td>
<td>2.8614</td>
<td>1.0109</td>
</tr>
<tr>
<td>$d_7$</td>
<td>18.8683</td>
<td>2.7412</td>
<td>1.2770</td>
</tr>
<tr>
<td>$d_8$</td>
<td>41.8929</td>
<td>4.2071</td>
<td>1.5910</td>
</tr>
<tr>
<td>$d_9$</td>
<td>26.2535</td>
<td>2.6073</td>
<td>1.2011</td>
</tr>
<tr>
<td>$d_{10}$</td>
<td>18.9115</td>
<td>1.5685</td>
<td>1.3007</td>
</tr>
<tr>
<td>$d_{11}$</td>
<td>24.6659</td>
<td>2.6271</td>
<td>1.3031</td>
</tr>
<tr>
<td>$d_{12}$</td>
<td>10.9502</td>
<td>1.7323</td>
<td>0.2987</td>
</tr>
<tr>
<td>$d_{13}$</td>
<td>145.2545</td>
<td>9.7463</td>
<td>-0.1915</td>
</tr>
<tr>
<td>$d_{14}$</td>
<td>47.8441</td>
<td>4.2000</td>
<td>-2.8760</td>
</tr>
<tr>
<td>$d_{15}$</td>
<td>123.1606</td>
<td>7.4160</td>
<td>0.5123</td>
</tr>
<tr>
<td>$d_{16}$</td>
<td>196.9066</td>
<td>9.9011</td>
<td>0.3705</td>
</tr>
<tr>
<td>$d_{17}$</td>
<td>389.354</td>
<td>29.2098</td>
<td>-3.0755</td>
</tr>
</tbody>
</table>
METRIC SPACE OF DEFORMATIONS

Large Deformation Diffeomorphic Metric Mapping

T=1

T=5

T=9

T=13
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:

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

\[ \mathbf{C} = \frac{1}{L-1} \mathbf{A}^\top \mathbf{K} \mathbf{A} \]

3. Calculate singular value decomposition:

\[ \mathbf{C} = \mathbf{VDV}^\top \]

4. Calculate principal components

\[ \mathbf{U} = \mathbf{AVD}^{-\frac{1}{2}} \]
HOW SIMILAR ARE RIGHT AND LEFT EARS?

The graph shows the correlation values for each ear number, comparing the mean correlation with other ears and the left-right ear correlation. The data points are scattered across the graph, indicating variability in the correlation values across different ear numbers.
RANGE OF MODELED EAR SHAPE VARIATION

Original Ear

Ear Reconstructed with 8 KPCA Components
RELATING EAR SHAPE TO ACOUSTIC DIRECTIVITY

6000 Hz

EAR 2  EAR 4  EAR 6  EAR 8

True:

PCA:

Predicted:
RELATING EAR SHAPE TO ACOUSTIC DIRECTIVITY
8063 Hz

True:

PCA:

Predicted:
RELATING EAR SHAPE TO ACOUSTIC DIRECTIVITY

9938 Hz

True:

PCA:

Predicted
THE END

Thanks for Listening