A method for efficiently calculating head-related transfer functions (HRTFs) directly from head scan point clouds

Presented at the 143rd AES Convention
October 21, 2017, New York, NY, USA

Rahulram Sridhar (presenter)
Edgar Y. Choueiri

3D Audio and Applied Acoustics (3D3A) Laboratory
Princeton University, Princeton, NJ, USA
www.princeton.edu/3D3A
Outline

• Introduction
• Motivation
• Method Formulation
• Application
• Validation
• Conclusions
• Future Work
Outline

• Introduction
  • Motivation
  • Method Formulation
  • Application
  • Validation
  • Conclusions
  • Future Work
Introduction

HRTFs from acoustical measurements

- Currently most accurate
- Benchmarks for comparing computed HRTFs
- Requires anechoic chamber
- Cumbersome to set up
- Can be tiresome for subject
- Not ideal for commercial implementation

HRTF measurement setup at 3D3A lab
Paper Number 9892

Outline

• Introduction - Acoustically-measured HRTFs are accurate but not ideal for commercial use.

• Motivation

• Method Formulation

• Application

• Validation

• Conclusions

• Future Work
Motivation

HRTFs from morphological data: Existing methods

I. Refs: [3 - 6]

3D meshed scan of head/torso → Numerical technique
Ex: BEM, FDTD, ARD, etc. → Computed HRTFs

- Computationally expensive.
- Inaccurate without an accurate 3D mesh.

All reference numbers correspond to those in the associated convention paper.
Motivation

HRTFs from morphological data: Existing methods

II. Ref: [2]

Anthropometric data → Data-driven model
Ex: PCA & regression-based → Computed HRTFs

Requires explicit identification and measurement of anthropometric features.

All reference numbers correspond to those in the associated convention paper.
Motivation

HRTFs from morphological data: Proposed method

- Requires only point cloud data.
- Computationally inexpensive.
- Does not require explicit identification and measurement of anthropometric features.
Introduction - Acoustically-measured HRTFs are accurate but not ideal for commercial use.

Motivation - Meshed scans and anthropometric features for computing HRTFs using existing methods are difficult to obtain with required accuracy.

Method Formulation

Application

Validation

Conclusions

Future Work
Method Formulation

Basic idea

Point clouds of head/torso of training subjects → Spherical-harmonic decomposition → Spherical-harmonic coefficients, $c_S$

Measured HRTFs of training subjects → Spherical-harmonic decomposition → Spherical-harmonic coefficients, $c_H$

Spherical-harmonic coefficients, $c_S$ for new subject → Linear, least-squares regression → Mapping from $c_S$ to $c_H$ → $c_H$ for new subject → Synthesize HRTFs

Same type of basis functions
Method Formulation

Spherical harmonic representation of point cloud data

Assumption: Scan is already aligned as shown above.

Represent $r(\theta, \phi)$ using spherical harmonics
Method Formulation

Spherical harmonic representation of HRTF data

Choose HRTF feature and represent as spatial function, $H(\theta, \phi)$

Represent $H(\theta, \phi)$ using spherical harmonics

Example features:
(1) ITDs
(2) HRTF frequency responses
(3) HRTF magnitude responses
Paper Number 9892

Outline

• **Introduction** - Acoustically-measured HRTFs are accurate but not ideal for commercial use.

• **Motivation** - Meshed scans and anthropometric features for computing HRTFs using existing methods are difficult to obtain with required accuracy.

• **Method Formulation** - Point clouds $\rightarrow$ Spherical harmonics $\rightarrow$ Matrix multiplications $\rightarrow$ HRTFs

• **Application**

• Validation

• Conclusions

• Future Work
**Application**

Data acquisition

<table>
<thead>
<tr>
<th>Measured HRTF and head scan database</th>
<th>RIEC* [12]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of “training” subjects, $U$</td>
<td>23</td>
</tr>
<tr>
<td>HRTF features</td>
<td>(1) HRTF magnitude spectra in dB (cf. [10])</td>
</tr>
<tr>
<td></td>
<td>(2) ITD computed by thresholding (cf. [13])</td>
</tr>
</tbody>
</table>

*http://www.rieck.tohoku.ac.jp/pub/hrtf/index.html*

All reference numbers correspond to those in the associated convention paper.
Application

Data pre-processing

- Make measured HRIRs minimum-phase and truncate to 5.8 ms [10].

- Align head scan such that y-axis = interaural axis and x-axis lies in both horizontal and median planes.

All reference numbers correspond to those in the associated convention paper.
Application

Spherical harmonic representation of point cloud data

Maximum possible degree = 3

<table>
<thead>
<tr>
<th>Mapping to</th>
<th>Degree used</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRTF magnitudes</td>
<td>2</td>
</tr>
<tr>
<td>ITDs</td>
<td>1</td>
</tr>
</tbody>
</table>

“Degree” corresponds to degree of spherical harmonic expansion of point cloud data
Application

Spherical harmonic representation of HRTF feature data

Maximum possible degree = 14

<table>
<thead>
<tr>
<th>HRTF feature</th>
<th>Degree used</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRTF magnitudes in dB</td>
<td>6 (cf. [10])</td>
</tr>
<tr>
<td>ITDs</td>
<td>3</td>
</tr>
</tbody>
</table>
Paper Number 9892

Outline

• **Introduction** - Acoustically-measured HRTFs are accurate but not ideal for commercial use.

• **Motivation** - Meshed scans and anthropometric features for computing HRTFs using existing methods are difficult to obtain with required accuracy.

• **Method Formulation** - Derive mappings from spherical harmonic representations of point clouds to those of (1) HRTF magnitudes in dB and (2) ITDs.

• **Application** - Derive mappings from spherical harmonic representations of point clouds to those of (1) HRTF magnitudes in dB and (2) ITDs.

• **Validation**

• **Conclusions**

• **Future Work**
Validation

Approach and Metrics

• Number of “test” subjects = 2

• Metric to validate computed HRTF magnitudes: log-weighted average spectral distortion in dB.

• Metric to validate computed ITDs: absolute ITD error.
Validation

Providing Perceptual Context

Approx. perceptibility threshold for rms average spectral distortion

<table>
<thead>
<tr>
<th>Frequency range (kHz)</th>
<th>Perceptibility threshold (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 2</td>
<td>1</td>
</tr>
<tr>
<td>2 to 8</td>
<td>2</td>
</tr>
<tr>
<td>8 to 16</td>
<td>3</td>
</tr>
</tbody>
</table>

Approx. perceptibility threshold for absolute ITD error = 30 µs
Validation

Computed HRTF magnitudes

Log-weighted average spectral distortion in dB

Approx. rms perceptibility threshold = 1 dB
Validation

Computed HRTF magnitudes

Log-weighted average spectral distortion in dB

Approx. rms perceptibility threshold = 2 dB
Validation

Computed HRTF magnitudes

Log-weighted average spectral distortion in dB

Approx. rms perceptibility threshold = 3 dB
Validation

Computed ITDs

Absolute ITD error in $\mu s$
Outline

- **Introduction** - Acoustically-measured HRTFs are accurate but not ideal for commercial use.

- **Motivation** - Meshed scans and anthropometric features for computing HRTFs using existing methods are difficult to obtain with required accuracy.

- **Method Formulation** - Derived mappings from spherical harmonic representations of point clouds to those of (1) HRTF magnitudes in dB and (2) ITDs.

- **Application** - Used two test subjects to objectively show (with perceptual context) that computed HRTFs are accurate up to approx. 6 kHz.

- **Conclusions**

- **Future Work**
Conclusions

- Presented the following method to compute HRTFs from point cloud data of an individual’s morphology:

  - HRTFs directly from point clouds and no need to identify anthropometric features. This makes it suitable for commercial implementation.

  - The current implementation of our method may be used to compute HRTFs that are likely indistinguishable from measured HRTFs for $f < 6$ kHz.

  - More data is required to determine how the method performs at higher frequencies. So the 6 kHz limit above may not be a limitation of the method.
• **Introduction** - Acoustically-measured HRTFs are accurate but not ideal for commercial use.

• **Motivation** - Meshed scans and anthropometric features for computing HRTFs using existing methods are difficult to obtain with required accuracy.

• **Method Formulation** - Derived mappings from spherical harmonic representations of point clouds to those of (1) HRTF magnitudes in dB and (2) ITDs.

• **Application** - Used two test subjects to objectively show (with perceptual context) that computed HRTFs are accurate up to approx. 6 kHz.

• **Validation** - Our method is suitable to implement commercially, and shows promise for computing HRTFs accurately, but more data is needed.

• **Conclusions** - Our method is suitable to implement commercially, and shows promise for computing HRTFs accurately, but more data is needed.

• **Future Work**
Future Work

- Applying method to larger dataset.
- Trying different types of HRTF features to represent using spherical harmonics.
- Trying different mapping techniques.
- Trying to account for the fact that the head with pinnae is a non-star-shaped object.
- Validation using subjective listening tests.
A method for efficiently calculating head-related transfer functions (HRTFs) directly from head scan point clouds

Presented at the 143rd AES Convention
October 21, 2017, New York, NY, USA

Rahulram Sridhar (presenter)
Edgar Y. Choueiri

3D Audio and Applied Acoustics (3D3A) Laboratory
Princeton University, Princeton, NJ, USA
www.princeton.edu/3D3A