LARGE DEFORMATION DIFFEOMORPHIC METRIC MAPPING PROVIDES NEW INSIGHTS INTO THE LINK BETWEEN HUMAN EAR MORPHOLOGY AND THE HEAD-RELATED TRANSFER FUNCTIONS

Reza Zolfaghari
April 2016

School of Electrical and Information Engineering
The University of Sydney

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy
The above figure shows a NURBS model of the template head and torso shape which can be used for 3D printing. The NURBS model was produced from the surface mesh of the template head, torso and ear. The procedure for computing the template shape is discussed in this thesis. Image is a courtesy of Dr Peter Jones.
This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Name: Reza Zolfaghari
To my children

To my wife

To my mother, father and sister
List of Publications

The following lists the publications that I have contributed while completing this PhD thesis:

"Guillon, Pierre; Zolfaghari, Reza; Epain, Nicolas; van Schaik, André; Jin, Craig T; Hetherington, Carl; Thorpe, Jonathan; Tew, A; "Creating the sydney york morphological and acoustic recordings of ears database, 2012 IEEE International Conference on Multimedia and Expo, 461-466, 2012, IEEE

"Jin, Craig T; Guillon, Pierre; Epain, Nicolas; Zolfaghari, Reza; van Schaik, André; Tew, Anthony I; Hetherington, Carl; Thorpe, Jonathan; "Creating the sydney york morphological and acoustic recordings of ears database, IEEE Transactions on Multimedia, 16, 1, 37-46, 2014, IEEE

"Zolfaghari, Reza; Epain, Nicolas; Jin, Craig T; Glaunes, Joan; Tew, Anthony; "Large deformation diffeomorphic metric mapping and fast-multipole boundary element method provide new insights for binaural acoustics,"2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)".,2863-2867,2014,IEEE

"Zolfaghari, Reza; Epain, Nicolas; Jin, Craig T; Tew, Anthony; Glaunés, Joan; "A multiscale LDDMM template algorithm for studying ear shape variations,"Signal Processing and Communication Systems (ICSPCS), 2014 8th International Conference on",1-6,2014,IEEE

"Zolfaghari, Reza; Epain, Nicolas; Jin, Craig T; Glaunés, Joan; Tew, Anthony; "Generating a morphable model of ears,"2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)"1771-1775,2016,IEEE
"Zolfaghari, Reza; Epain, Nicolas; Jin, Craig T; Glaunés, Joan; Tew, Anthony; "Kernel principal component analysis of the ear morphology,"2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)"(Accepted)
Acknowledgements

Many individuals helped me in the process of completing my PhD, without whom the completion of this work would not have been possible. First I would like to thank my supervisor A/Prof Craig Jin for his great support and help during the course of my PhD. I learnt a lot from Craig during the five year period of my PhD, most importantly my approach to research is now different and my research skills have greatly been improved. His perspective on various matters is particularly at times interesting and novel.

I would like to thank A/Prof Joan Glaunés for his great support and help during the course of my PhD. Without his help and knowledge specially in the field of shape deformations the completion of this PhD would not have become possible. He is an excellent person to work with and offered much support and help when it was required. I would like to thank Dr Nicolas Epain, his support and input to my PhD was very useful. I would like to thank Tony Tew from the University of York for his input and support into this project. I would also like to thank Dr Peter Jones for the support he offered in CARLab for the figure at the start of this PhD and for the many conversations that we had.

I would like to thank the many friends and people that I met and worked with in CARLab and the university of Sydney while completing my PhD, Sepideh, Mahendra, Farzad, Abdullah, Joseph, Xian, Mohammad Shahnawaz, Arun Sebastian, Nadim and Ashnil, Michael and A/Prof Jinman Kim from the school IT.

I would like to also thank the three reviewers of my PhD thesis for providing me with valuable feedback for completing this thesis.

Finally I would like to thank my children, my wife, my parents, sister and other extended family members for the excellent support they gave and the patience they had while I was working on my PhD. Without the support of my wife and parents the completion of this work might have not become possible.
Abstract

This thesis aims to provide a better understanding on the relationship between the human head and ear morphology to a set of individualized filters known as the head related impulse response functions (HRIRs). The HRIR filters contain all the information required by the auditory system to localize sound in 3D space. The analyses presented in this thesis uses the Large Deformation Diffeomorphic Metric Mapping (LDDMM) technique and the Fast-Multipole Boundary Element Method, better known as FM-BEM, to study the effect of morphological changes on the ears to the corresponding acoustic changes. This work relies heavily on the morphological and acoustic data that is available in the SYMARE database, which contains acoustic and morphological information for 62 subjects.

The research findings presented in this thesis is composed of four sections. In the first section of this thesis, the LDDMM framework and its applications to analysing ear shapes is explained. Further, important functions describing various operations for deforming shapes are introduced and elaborated. In order to examine the quality of the ear deformations, tools are developed to compare and measure the differences between the acoustic responses obtained from BEM simulations for two ear shapes. This part of the work also involved the development of techniques to compare and measure differences in 3D shapes using the framework of currents. Finally this section introduces the multi-scale approach for mapping ear shapes.

The second section of the thesis estimates a template ear, head and torso shape from the shapes available in the SYMARE database. This part of the thesis explains a new procedure for developing the template ear shape. The procedure builds upon the multi-scale approach, which is described and used in the first part of this research investigation. The template ear and head shapes were examined and verified by comparing the features in the template shapes to corresponding features in the population. In addition, the acoustic and morphological differences of the multi-scale template ear shape to a template ear that was estimated at a single-scale
are compared and highlighted.

The third section of the thesis examines the quality of the deformations from the template ear shape to target ears in SYMARE from both an acoustic and morphological standpoint. As a result of this investigation, it was identified that ear shapes can be studied more accurately by the use of two physical scales and that scales at which the ear shapes were studied were dependant on the parameters chosen when mapping ears in the LDDMM framework. It also examined the effect of selecting the parameters on the quality of the deformations. As a result of this work regions in the ear shapes were identified that can potentially contribute to differences in the observed acoustic responses. Finally, this section concludes by noting how shape distances vary with the acoustic distances using the developed tools.

In the final part of this thesis, the variations in the morphology of ears were examined using the Kernel Principle Component Analysis (KPCA) and the changes in the corresponding acoustics were studied using the standard principle component analysis (PCA). These examinations involved identifying the number of kernel principle components that were required in order to model ear shapes with an acceptable level of accuracy, both morphologically and acoustically. In addition, a further study was conducted which examined the underlying mechanisms that caused changes to the Direction Transfer Functions (DTF) spectrum when the ear shapes were changed. Specifically, as a result of this study, important features in the acoustic principle components were identified that caused the changes in the DTF spectrum. It also identified how coherently the DTFs corresponding to a few angles in the median plane changed as the ear shapes changed. Finally, a morph-able model for ear shapes based on the KPCA technique is presented. The strengths and limitations of the morph-able model are examined when trying to model external ear shapes that were not included in the dataset.
## Contents

### Glossary

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Brief Background</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Motivation</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Problem Statement and Scope of Work</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Contribution</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Challenges</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Outline of the Thesis</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>BACKGROUND</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spatial Hearing</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Spectral Cues</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Head Related Impulse Response And Head Related Transfer functions</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Methods for Acquiring HRIRs</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Numerically Calculating The HRIRs</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Boundary Element Method</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Validating The Computation of HRIRs Through BEM and FM-BEM Simulations</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Virtual Auditory Space</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Applications Of VAS</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Principal Component Analysis</td>
<td>33</td>
</tr>
</tbody>
</table>
## 2.8 Large Deformation Diffeomorphic Metric Mapping (LDDMM) framework

- **2.8.1 Induced distances between shapes in the LDDMM framework**
  - **2.8.1.1 The Role of $\sigma_V$ Parameter**
- **2.8.2 Measuring shape differences using currents**
- **2.8.3 Geodesic Shooting**

## 3 REVIEW OF LITERATURE

- **3.1 Shape Parametrization: Spherical Harmonics**
- **3.2 Shape Parametrization: Elliptical Fourier Transform**
- **3.3 Modelling The Acoustics Of The Human Morphology Using Simple Geometrical Objects**
- **3.4 Morphoacoustic Perturbation Analysis**
  - **3.4.1 Differential Pressure Synthesis (DPS)**
  - **3.4.2 Morphoacoustic Perturbation Analysis Frequency Domain (MPA-FD)**
  - **3.4.3 Acoustic sensitivity to micro-perturbations of KEMAR's pinna surface geometry**
- **3.5 HRTF personalization using anthropometric measurements of the head and ear shape**
  - **3.5.1 Frequency Scaling**
  - **3.5.2 Database Matching**
  - **3.5.3 Regression Analysis**

## 4 MORPHOACOUSTIC ANALYSIS FRAMEWORK

- **4.1 LDDMM Method For Mapping Shapes**
- **4.2 Diffeomorphically Mapping the Head and Ears**
- **4.3 Shape Differences Based On The Metric of Currents**
  - **4.3.1 Hausdorff Distance in 3D Ear Shape Analysis**
4.3.2 Shape Differences based on CSDA ........................................... 84
  4.3.2.1 The Effect of \( \sigma_R \) on measuring shape differences ........ 89
4.4 LDDMM Scale Parameters ......................................................... 92
  4.4.1 Scale Parameter \( \sigma_V \) ....................................................... 92
  4.4.2 Scale Parameter \( \sigma_W \) ....................................................... 95
  4.4.3 Optimization weight parameter \( \gamma \) ..................................... 97
4.5 Producing FM-BEM Ready Meshes ............................................ 99
4.6 Conducting FM-BEM Simulations ............................................. 101
4.7 Extracting FM-BEM Data ......................................................... 104
4.8 Globally Comparing the Acoustical Responses for Ear Shapes ....... 109
  4.8.1 Spatial Frequency Response Surface (SFRS) ......................... 110
  4.8.2 Spatial Correlation Metric (SCM) ........................................ 111
  4.8.3 Absolute Spatial Gain Error (SGE) ...................................... 115
  4.8.4 Standard Deviation of the Spectral Difference (SDS) ................ 116
4.9 SMS-LDDMM Shape Matching .................................................. 118
  4.9.1 Conclusion ........................................................................ 120

5 ESTIMATING A TEMPLATE HEAD TORSO AND EAR SHAPE .............. 121
5.1 Method ................................................................................. 124
  5.1.1 Barycentre ................................................................. 124
  5.1.2 Single Scale Template Estimation .................................... 125
  5.1.3 Sequential Multi Scale Template Estimation ......................... 128
5.2 Results ................................................................................. 130
  5.2.1 Single and multi-scale template ear shapes ......................... 130
  5.2.2 Template head, torso and ear shape .................................. 133
  5.2.3 Comparing anthropometric features in the template head and ear shape to the SYMARE population .......................... 136
5.2.3.1 2D Anthropometric Measurements .................................. 137
5.2.3.2 3D Anthropometric Measurements .................................. 141
5.2.4 Angle and scale factor measurements ................................. 146
5.2.4.1 Rigidly Aligning Ear Shapes Using Measures ...................... 147
5.2.5 Morphoacoustic Comparison of the Single and Multi-Scale Template Shapes .................................................................. 148
5.3 Conclusion ............................................................................ 153

6 EXAMINING DEFORMATIONS FROM THE TEMPLATE SHAPE .......... 155
6.1 Single-Scale and Multi-Scale LDDMM Morphoacoustic Analysis of Ear Shapes .................................................. 159
6.1.0.1 second-scale deformations .............................................. 166
6.2 Morphoacoustic Distance Analysis Along the Deformation Path ...... 174
6.2.1 Morphoacoustic Distance Analysis ....................................... 181
6.3 Conclusion ............................................................................ 183

7 KERNEL PRINCIPAL COMPONENT ANALYSIS OF EAR MORPHOLOGY ................................................ 186
7.1 Kernel Principal Component Analysis (KPCA) ......................... 190
7.1.1 KPCA On Ear Shapes ...................................................... 191
7.2 Quantifying the Reduction in Acoustic Variation Obtained With Alignment of the Ear Shapes to The Template Ear .................................. 194
7.3 Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes .................................................. 197
7.3.1 Choosing The Ear Shapes in SYMARE ................................. 198
7.3.2 Methodology ..................................................................... 200
7.3.3 Analysis Of Results .......................................................... 203
7.3.4 Distance Analysis Between Morphology and Acoustics .......... 220
7.4 Examining the Kernel Principal Components ........................... 224
7.4.1 Methodology ..................................................................... 224
7.4.2 Experimental Setup .......................................................... 228
## Contents

7.4.3 Results ................................................. 229  
7.4.4 Analysis of Morphological Results ..................... 257  
7.4.5 Morphology and Acoustic Distance Analysis .......... 262  
7.4.6 Acoustic Variations in the Median Plane ................ 264  
7.5 Mechanisms That Generate Changes In The DTF Spectrum ... 267  
7.5.1 Gain Analysis On Selected Frequencies and Angles .... 267  
7.5.2 Methodology: Modelling the DTFs in the Median Plane Using PCA ... 268  
7.5.2.1 Iteratively Reconstructing a DTF and Calculating the Reconstruction Error ....... 270  
7.5.3 Results: Modelling the DTFs In The Median Plane Using PCA ... 272  
7.5.4 Examining The Relationship Between Shape and Acoustic Models ... 275  
7.5.5 Results: Examining the Changes in the Acoustic Principal Components as the Ear Shape Changes ... 277  
7.6 How Coherently Do DTFs Change In The Median Plane .... 294  
7.7 Morphable Model of Ears ................................ 302  
7.7.1 Method ............................................. 302  
7.7.2 Experimental Setup .................................. 304  
7.7.3 Results ............................................. 304  
7.8 Conclusion and Summary Of Chapter ....................... 306  

8 CONCLUSION AND FUTURE WORK .......................... 310  
8.1 Questions and Answers .................................. 311  
8.2 Future Work ............................................. 317  

Bibliography .................................................. 334  

Appendices ..................................................... 335
List of Figures

1.1 Deforming Ears ..................................................... 6
2.1 External Ear ...................................................... 14
2.2 The HAT model ................................................... 15
2.3 HRTFs vs PRTFs ................................................ 16
2.4 Auditory planes .................................................. 17
2.5 Example of HRIR/HRTF ....................................... 18
2.6 Example of SFRS ............................................... 19
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.7</td>
<td>Apparatus for measuring HRTFs</td>
<td>20</td>
</tr>
<tr>
<td>2.8</td>
<td>Domain For BEM</td>
<td>22</td>
</tr>
<tr>
<td>2.9</td>
<td>Katz 2001, sphere, head only, head and ear</td>
<td>25</td>
</tr>
<tr>
<td>2.10</td>
<td>Kahana Reciprocity</td>
<td>26</td>
</tr>
<tr>
<td>2.11</td>
<td>Kahana 2007, DB60 Low resolution</td>
<td>26</td>
</tr>
<tr>
<td>2.12</td>
<td>Kahana 2007, DB60 High res and low res scanner acoustics</td>
<td>27</td>
</tr>
<tr>
<td>2.13</td>
<td>Kahana 2007, Baffled pinna vs Pinna on head</td>
<td>28</td>
</tr>
<tr>
<td>2.14</td>
<td>Kahana 2006, pinna modes</td>
<td>28</td>
</tr>
<tr>
<td>2.15</td>
<td>Kahana 2006, pinna modes</td>
<td>29</td>
</tr>
<tr>
<td>2.16</td>
<td>Kahana 2006, pinna directivity</td>
<td>29</td>
</tr>
<tr>
<td>2.17</td>
<td>FM-BEM mesh shapes</td>
<td>29</td>
</tr>
<tr>
<td>2.18</td>
<td>FM-BEM Validation</td>
<td>30</td>
</tr>
<tr>
<td>2.19</td>
<td>Production of VAS</td>
<td>31</td>
</tr>
<tr>
<td>2.20</td>
<td>Surround sound speaker configuration</td>
<td>32</td>
</tr>
<tr>
<td>2.21</td>
<td>Mean DTF and principal components</td>
<td>36</td>
</tr>
<tr>
<td>2.22</td>
<td>Reconstruction of DTF</td>
<td>37</td>
</tr>
<tr>
<td>2.23</td>
<td>DTF Reconstruction Using PCA</td>
<td>38</td>
</tr>
<tr>
<td>2.24</td>
<td>LDDMM Flow</td>
<td>39</td>
</tr>
<tr>
<td>2.25</td>
<td>LDDMM cost function chart</td>
<td>41</td>
</tr>
<tr>
<td>2.26</td>
<td>Discrete LDDMM Flow</td>
<td>42</td>
</tr>
<tr>
<td>2.27</td>
<td>Geodesic Path</td>
<td>42</td>
</tr>
<tr>
<td>2.28</td>
<td>Gain of the Cauchy function</td>
<td>44</td>
</tr>
<tr>
<td>2.29</td>
<td>LDDMM shooting</td>
<td>47</td>
</tr>
<tr>
<td>3.1</td>
<td>Spherical Coordinate System</td>
<td>50</td>
</tr>
<tr>
<td>3.2</td>
<td>Spherical Harmonics Simplified Head Model</td>
<td>51</td>
</tr>
<tr>
<td>3.3</td>
<td>Head Simplification Using Spherical Harmonics</td>
<td>53</td>
</tr>
</tbody>
</table>
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>Parametrizing the Head By Slicing</td>
<td>54</td>
</tr>
<tr>
<td>3.5</td>
<td>Effects of Changing Coefficients in EFT</td>
<td>55</td>
</tr>
<tr>
<td>3.6</td>
<td>Snowman Model</td>
<td>56</td>
</tr>
<tr>
<td>3.7</td>
<td>HRTF For Snowman Model</td>
<td>57</td>
</tr>
<tr>
<td>3.8</td>
<td>Increased acoustic response of rectangular flang</td>
<td>58</td>
</tr>
<tr>
<td>3.9</td>
<td>External Ear</td>
<td>58</td>
</tr>
<tr>
<td>3.10</td>
<td>Average response of real ears versus geometrically modelled ear shapes</td>
<td>59</td>
</tr>
<tr>
<td>3.11</td>
<td>Morphoacoustic Frequency domain</td>
<td>63</td>
</tr>
<tr>
<td>3.12</td>
<td>Peaks and notches in the PRTF used for perturbation analysis</td>
<td>64</td>
</tr>
<tr>
<td>3.13</td>
<td>Sensitivity patterns in the ear for peak formation</td>
<td>65</td>
</tr>
<tr>
<td>3.14</td>
<td>Sensitivity patterns in the ear for notch formation</td>
<td>65</td>
</tr>
<tr>
<td>3.15</td>
<td>Frequency scaling used for reducing the ISSD between subjects</td>
<td>68</td>
</tr>
<tr>
<td>4.1</td>
<td>LDDMM Head and Torso Mapping Chart</td>
<td>74</td>
</tr>
<tr>
<td>4.2</td>
<td>Source Shape with Momentum Vectors</td>
<td>76</td>
</tr>
<tr>
<td>4.3</td>
<td>Mapping algorithm, source and target shapes</td>
<td>79</td>
</tr>
<tr>
<td>4.4</td>
<td>Hausdorff distance</td>
<td>83</td>
</tr>
<tr>
<td>4.5</td>
<td>Ears Hausdorff distance</td>
<td>85</td>
</tr>
<tr>
<td>4.7</td>
<td>New technique for measuring shape differences</td>
<td>89</td>
</tr>
<tr>
<td>4.8</td>
<td>The effect of $\sigma_R$ value, close shapes</td>
<td>90</td>
</tr>
<tr>
<td>4.9</td>
<td>The effect of $\sigma_R$ value, different shapes</td>
<td>91</td>
</tr>
<tr>
<td>4.10</td>
<td>Template Ear Shape</td>
<td>93</td>
</tr>
<tr>
<td>4.11</td>
<td>Influence of $\sigma_V$ scale parameter on LDDMM matching</td>
<td>94</td>
</tr>
<tr>
<td>4.12</td>
<td>Influence of $\sigma_V$ scale parameter on shape distance</td>
<td>95</td>
</tr>
<tr>
<td>4.13</td>
<td>Influence of $\sigma_W$ parameter on matching</td>
<td>96</td>
</tr>
<tr>
<td>4.14</td>
<td>Influence of $\sigma_W$ on shape distance</td>
<td>97</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>4.15</td>
<td>Influence of $\gamma$ values on LDDMM matching</td>
<td>98</td>
</tr>
<tr>
<td>4.16</td>
<td>Influence of $\gamma$ value on shape distance</td>
<td>99</td>
</tr>
<tr>
<td>4.17</td>
<td>Location of sensors</td>
<td>102</td>
</tr>
<tr>
<td>4.18</td>
<td>The location of the vibrating element on a sample mesh</td>
<td>103</td>
</tr>
<tr>
<td>4.19</td>
<td>Auditory Planes and Angles</td>
<td>105</td>
</tr>
<tr>
<td>4.20</td>
<td>Median Plane Angles</td>
<td>107</td>
</tr>
<tr>
<td>4.21</td>
<td>Median Plane Image Plot</td>
<td>108</td>
</tr>
<tr>
<td>4.22</td>
<td>Median Plane Surface Plot</td>
<td>109</td>
</tr>
<tr>
<td>4.23</td>
<td>Example Plot For Globally Comparing the DTF's</td>
<td>110</td>
</tr>
<tr>
<td>4.24</td>
<td>Example SFRS</td>
<td>111</td>
</tr>
<tr>
<td>4.25</td>
<td>Example SFRS</td>
<td>113</td>
</tr>
<tr>
<td>4.27</td>
<td>Example SFRS</td>
<td>116</td>
</tr>
<tr>
<td>4.28</td>
<td>Example SFRS</td>
<td>117</td>
</tr>
<tr>
<td>4.29</td>
<td>SMS-LDDMM Analogy on 2D shape</td>
<td>119</td>
</tr>
<tr>
<td>4.30</td>
<td>SMS-LDDMM on Ears</td>
<td>120</td>
</tr>
<tr>
<td>5.1</td>
<td>Barycentre and Fréchet mean on a sphere</td>
<td>127</td>
</tr>
<tr>
<td>5.2</td>
<td>Single scale template ear shapes</td>
<td>131</td>
</tr>
<tr>
<td>5.3</td>
<td>Single-scale template distance versus iteration</td>
<td>131</td>
</tr>
<tr>
<td>5.4</td>
<td>Ear shape evolutions 1</td>
<td>132</td>
</tr>
<tr>
<td>5.5</td>
<td>Ear shape evolutions 2</td>
<td>132</td>
</tr>
<tr>
<td>5.6</td>
<td>Ear shape evolutions 2</td>
<td>132</td>
</tr>
<tr>
<td>5.7</td>
<td>Ear shape evolutions 3</td>
<td>132</td>
</tr>
<tr>
<td>5.8</td>
<td>Multi-scale template distance versus iteration</td>
<td>133</td>
</tr>
<tr>
<td>5.9</td>
<td>Multi-scale template ear shape evolution</td>
<td>134</td>
</tr>
<tr>
<td>5.10</td>
<td>3D low resolution template head and torso</td>
<td>135</td>
</tr>
<tr>
<td>5.11</td>
<td>3D template head and torso shape with multi-scale template ear shape</td>
<td>136</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>5.12</td>
<td>3D template head and torso shape with single-scale template ear shape</td>
<td>137</td>
</tr>
<tr>
<td>5.13</td>
<td>CIPIC Anthropometric Ear Measurements</td>
<td>138</td>
</tr>
<tr>
<td>5.14</td>
<td>2D anthropometric of the template ear shape</td>
<td>139</td>
</tr>
<tr>
<td>5.15</td>
<td>2D ear shape statistics</td>
<td>140</td>
</tr>
<tr>
<td>5.16</td>
<td>Landmark points for 3D head and torso measurements</td>
<td>142</td>
</tr>
<tr>
<td>5.17</td>
<td>3D ear shape statistics</td>
<td>144</td>
</tr>
<tr>
<td>5.18</td>
<td>3D head shape statistics</td>
<td>145</td>
</tr>
<tr>
<td>5.19</td>
<td>Ear coordinate system and rotation angles</td>
<td>146</td>
</tr>
<tr>
<td>5.20</td>
<td>Rotation and scale factor statistics multi-scale template</td>
<td>149</td>
</tr>
<tr>
<td>5.21</td>
<td>Rotation and scale factor statistics single-scale template</td>
<td>150</td>
</tr>
<tr>
<td>5.22</td>
<td>Morphological comparison single and multi-scale template ear</td>
<td>151</td>
</tr>
<tr>
<td>5.23</td>
<td>Acoustical comparison single and multi-scale template ear, head only</td>
<td>151</td>
</tr>
<tr>
<td>5.24</td>
<td>Acoustical comparison single and multi-scale template ear, head and torso shape</td>
<td>152</td>
</tr>
<tr>
<td>6.4</td>
<td>First Order LDDMM Transformation $S_{03}$</td>
<td>162</td>
</tr>
<tr>
<td>6.5</td>
<td>First Order LDDMM Transformation $S_{05}$</td>
<td>163</td>
</tr>
<tr>
<td>6.6</td>
<td>First Order LDDMM Transformation $S_{29}$</td>
<td>164</td>
</tr>
<tr>
<td>6.7</td>
<td>First Order LDDMM Transformation $S_{41}$</td>
<td>165</td>
</tr>
<tr>
<td>6.8</td>
<td>Second Order LDDMM Transformation $S_{3}$</td>
<td>170</td>
</tr>
<tr>
<td>6.9</td>
<td>Second Order LDDMM Transformation</td>
<td>171</td>
</tr>
<tr>
<td>6.10</td>
<td>Second Order LDDMM Transformation $S_{29}$</td>
<td>172</td>
</tr>
<tr>
<td>6.11</td>
<td>First Order LDDMM Transformation</td>
<td>173</td>
</tr>
<tr>
<td>6.12</td>
<td>Template Head With no Ears</td>
<td>177</td>
</tr>
<tr>
<td>6.13</td>
<td>HRTFs along the deformation path $E_{1}$</td>
<td>178</td>
</tr>
<tr>
<td>6.14</td>
<td>HRTFs along the deformation path $E_{41}$</td>
<td>179</td>
</tr>
<tr>
<td>6.15</td>
<td>HRTFs along the deformation path $E_{411}$</td>
<td>180</td>
</tr>
</tbody>
</table>
6.16 Geodesic Distance versus Acoustical Distance ............... 182
6.17 Mean CSDA Distance versus Acoustical Distance .......... 182

7.1 Affine transformed Versus Normal Ears Shapes ............... 196
7.2 Affine Transformed Ears Versus Normal Ears Acoustical Distance .... 197
7.3 Histogram of acoustic distances SYMARE .................. 200
7.4 Coefficients Evolving Ear $S_5$ ............................ 204
7.5 Coefficients Evolving Ear $S_{13}$ ........................... 204
7.6 Evolution of source to target ear $S_5$ to $S_{55}$ .............. 205
7.7 Evolution of source to target ear $S_{13}$ to $S_{23}$ ............. 206
7.8 Evolving Acoustical Response $S_5$ to $S_{55}$ 1 .......... 208
7.9 Evolving Acoustical Response $S_5$ to $S_{55}$ 2 ............. 209
7.10 Evolving Acoustical Response $S_{13}$ to $S_{23}$ 1 ......... 210
7.11 Evolving Acoustical Response $S_{13}$ to $S_{23}$ 2 ............. 211
7.12 $S_5$ to $S_{55}$ DTFs $P=0$ .................................. 212
7.13 $S_5$ to $S_{55}$ DTFs $P=5$ .................................. 213
7.14 $S_5$ to $S_{55}$ DTFs $P=10$ ................................ 214
7.15 $S_5$ to $S_{55}$ DTFs $P=20$ ................................ 215
7.16 $S_{13}$ to $S_{23}$ DTFs $P=0$ ................................ 216
7.17 $S_{13}$ to $S_{23}$ DTFs $P=5$ ................................ 217
7.18 $S_{13}$ to $S_{23}$ DTFs $P=10$ ................................ 218
7.19 $S_{13}$ to $S_{23}$ DTFs $P=20$ ................................ 219
7.20 Moving Between Ears Acoustics ............................... 221
7.21 Moving Between Ears Morphology Distances $S_5, S_{55}$ .......... 222
7.22 Moving Between Ears Morphology Distances $S_{13}, S_{23}$ .......... 223
7.23 Distribution of scores first five kernel principal components .... 226
7.24 KPCA scores standard deviation ............................ 227
7.26 Template Shape Acoustics .............................................. 230
7.27 KPC 1 shapes ............................................................... 232
7.28 KPC 1 acoustics 1 ........................................................... 233
7.29 KPC 1 acoustics 2 ........................................................... 234
7.30 KPC 1 DTFs -2.5 to -0.5 .................................................. 235
7.31 KPC 1 DTFs 0.5 to 2.5 .................................................... 236
7.32 KPC 2 shapes ............................................................... 237
7.33 KPC 2 acoustics 1 ........................................................... 238
7.34 KPC 2 acoustics 2 ........................................................... 239
7.35 KPC 2 DTFs -2.5 to -0.5 .................................................. 240
7.36 KPC 2 DTFs 0.5 to 2.5 .................................................... 241
7.37 KPC 3 shapes ............................................................... 242
7.38 KPC 3 acoustics 1 ........................................................... 243
7.39 KPC 3 acoustics 2 ........................................................... 244
7.40 KPC 3 DTFs -2.5 to -0.5 .................................................. 245
7.41 KPC 3 DTFs 0.5 to 2.5 .................................................... 246
7.52 External Ear Shape Terminology ....................................... 258
7.53 Median Plane Angles ...................................................... 262
7.54 Shape Distances From Template Using Currents ..................... 263
7.55 Acoustic Distance From Template using SDS ......................... 264
7.56 Shape Distance Versus Acoustic Distance ............................... 264
7.57 Variation Of Gain Versus Weights of Shape principal components .... 266
7.90 Morphable model for ear shapes ....................................... 305
7.91 Morphable model of ears accuracy in reconstruction ................. 306

A.1 First intermediate template shape used in the acoustic study of ears. .... 337

The template ear is attached to a head shape from SYMARE.
List of Figures

A.2 Second intermediate template shape used for the acoustic study of ears. The template ear is attached to the template head. . . . . . . . 338

F.1 Morphable model for ear shapes . . . . . . . . . . . . . . . . . . . . 440
F.2 Morphable model for ear shapes . . . . . . . . . . . . . . . . . . . . 441

H.1 AMVE-GUI . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 445
H.2 AMVE-GUI KPCA analysis . . . . . . . . . . . . . . . . . . . . . . . . 446
H.3 AMVE-GUI translated . . . . . . . . . . . . . . . . . . . . . . . . . . . 447
H.4 AMVE-GUI rigidly aligned . . . . . . . . . . . . . . . . . . . . . . . . 447
H.5 AMVE-GUI PCA Component Number . . . . . . . . . . . . . . . . . . 448
H.6 AMVE-GUI rigidly aligned . . . . . . . . . . . . . . . . . . . . . . . . 449
H.7 AMVE-GUI Slider 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 449
H.8 AMVE-GUI Morphology only . . . . . . . . . . . . . . . . . . . . . . . 450
H.9 AMVE-GUI Slider 2 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 450
H.10 AMVE-GUI Text Boxes . . . . . . . . . . . . . . . . . . . . . . . . . . . 451
Glossary

**B**: Typically used to signify the barycentre shape when estimating a template or average shape.

**BEM**: Boundary Element Method.

**CSDA**: Current Shape Difference Analysis technique compares two sets of shapes based on the framework of currents. It can highlight regions of difference and can further provide an overall measure of dissimilarity between the shapes.

**DTF**: Directional transfer function.

**ET_{Ms}**: The Multi-scale template ear shape.

**ET_{ss}**: The Single-scale template ear shape.

**FM**: Fast Multipole Method is used widely in Physics and Engineering to speed up many numerical simulations. For acoustic simulations the computation for the interaction of the sound waves scattered by different elements is made faster. Typically it reduces the computational time from order $O(N^2)$ to $O(N)$ or $O(N\log N)$.

**FM-BEM**: Fast Multipole Method for calculating the BEM.

**FIR**: Finite impulse response filter.
**FREE FIELD:** A sound field without reflection. This can be an infinite 3D space or in the form of an anechoic chamber.

**GAAF:** The Global Acoustic Analysis figure compares the differences between two sets of HRTFs spatially and in the frequency domain. The figure contains plots that show where, at what frequencies, and by what magnitude two sets of HRTFs are different. This figure is further explained in the body of this thesis.

**HRIR:** Head Related Impulse Response. The acoustic transfer function from a point in space to the tympanic membrane of either the left or right ear.

**HRTF:** Head related-transfer function. Fourier transform of the head-related impulse response. It is the frequency response corresponding to the head-related impulse response.

**HE:** The template head shape.

**HTE:** The template head and torso shape.

**HTE_{T_{ms}}:** The average head and torso shape with the multi-scale average ear shape attached.

**HTE_{T_{ss}}:** The average head and torso shape with the single-scale average ear shape attached.

**IPSILATERAL:** Refers to the same side of the body.

**ITD:** Inter-aural time difference refers to time delay of sound arrival between the two ears.

**IID:** Inter-aural intensity difference in signal intensity for each frequency between the two ears.
**ILD:** Inter-aural level difference is the overall difference in signal intensity the two ears averaged across frequency.

**PCA:** Principal Component Analysis

**KPCA:** Kernel Principle Component Analysis.

**KPC:** Kernel Principle Component.

**LDDMM:** Large Deformation Diffeomorphic Metric Mapping is a diffeomorphic technique for deforming a source shape to a target shape such that the deformation is smooth and also invertible. It is widely used in curve, image and shape analysis.

**MEDIAN PLANE:** A vertical plane that divides the body into equal halves. Also referred to the mid-saggital plane.

**HORIZONTAL PLANE:** A plane that contains the inter-aural axis and is vertical to the median plane.

**MIDSAGITTAL PLANE:** refers to the vertical plane containing the inter-aural axis which divides the body into two approximately equal and symmetric halves.

**MINIMUM PHASE SYSTEM** : A minimum phase system in digital signal processing is a causal and stable system with the same properties for its inverse.

**PINNA:** External ear.

**SCM:** Spatial correlation metric.

**SDS:** Standard deviation on the spectral difference.
**VAS:** Binaural virtual auditory/acoustic space, refers to the synthesis of spatial audio over headphones using the head-related transfer functions.

$\bar{V}$: A value based on the SDS measure for comparing the overall acoustic similarity of two shapes.

$\bar{C}$: A value based on the SCM measure for comparing the overall acoustic similarity of two shapes.

$\bar{d}$: A measure quantifying the overall dissimilarity of two shapes based on the framework of currents.

$K_V$: The deformation kernel used in the LDDMM cost function.

$k_W$: The kernel used for comparing shapes using the framework of currents in the LDDMM cost function.

$\sigma_V$: The deformation scale parameter used in the LDDMM cost function, signifying the width of the kernel when deforming shapes.

$\sigma_W$: The shape mismatch scale parameter used in the LDDMM cost function, signifying the width of the kernel when comparing shapes.

$\sigma_R$: The width of kernel used when comparing shapes using the CSDA technique.

$a(t)$: Used for representing the time dependant momentum vectors that parameterize the deformation between the source and target ear shapes.

$\gamma$: The regularization parameter used in the LDDMM cost function.

$v(t)$: Used for representing the time dependant vector fields.
Chapter 1

INTRODUCTION

This thesis reports findings related to the simulations and analyses of 3D ear shapes and its acoustics. This research is in the field of signal processing and shape analysis and it draws from the knowledge accumulated in the area of computational anatomy. It involves the study of ear shapes statistically by means of advanced shape processing techniques. It is part of a larger ongoing project that aims to find a function to rapidly map the morphology of the head and external ear shape to the acoustics of the ear. More specifically, the research involves the study of ear shapes statistically using the LDDMM framework and investigates the corresponding acoustics by means of large scale FM-BEM simulations on high resolution head, torso and ear meshes. The following section will present a brief background for the present research and subsequently will identify the motivation and reasons for the research conducted in this thesis.

1.1 Brief Background

The research that is conducted in this thesis builds on the significant knowledge base accrued in the past few decades from many areas of science. The first relevant area relates to spatial hearing. Spatial hearing refers to human auditory perception of space, which includes spaciousness, perception of room size and perception of sound
location including direction and distance. A significant amount of psychophysical experimentations and testing has been conducted in past decades to obtain a better understanding on the mechanisms that enable an individual to localize sound in 3D space. The importance of binaural cues and spectral cues on sound localization is now well-known (refer to Sec. 2). Both binaural and spectral cues help an individual localize sound in 3D space. The binaural cues assist an individual to localize sound at low frequencies, namely, below < 5kHz, while spectral cues are broadband and contain cues at both high and low frequencies. Spectral cues are very important when localizing sound in the vertical domain, that is, in the up/down direction.

The external mechanisms that enable humans to perceive sound spatially is known to be related to the morphology of the external ear. The external ear shape is a complicated surface which contains unique features that belong specifically to an individual. The physical features in the external ear filter the incoming sound signal in a way that is personalized and unique to an individual. In other words, incoming sounds are filtered differently across individuals based on the shape of their external ear. The acoustic filtering properties of the external ear can be completely understood by a set of impulse responses, known as the head related impulse response (HRIR)\(^1\). The HRIRs are mathematical functions that describe the spatial filter properties of the external ear shape and they exist for every direction. The frequency domain representation of the HRIRs are referred to as head related transfer functions (HRTF) (for further details refer to Sec. 2.3).

Although both binaural and spectral cues help an individual localize sound in 3D space, unfortunately, the mechanisms for the generation of spectral cues is much less understood when compared to the binaural cues. The production of binaural cues over headphones is currently easily undertaken. This is because binaural cues are mainly dependent on the radius of the head, which can easily be measured. On the other hand, the mechanisms that generate the spectral cues are very complicated and depend greatly on the morphology of the ear, head and torso.

\(^1\)It should be noted that the HRIRs capture the acoustic filtering properties of not just the external ear but also other body parts such as the head and torso.
shape. Experiments have shown that an occluded external ear will cause significant errors when localizing sound (Gardner and Gardner, 1973). In summary, we can say that the morphology of the ear and its effect on the HRIR/HRTF is very important.

1.2 Motivation

The main motivating factor for this research is to obtain individualized HRIRs. Individualized HRIRs are crucial for the production of high fidelity 3D sound over head phones, which is also known as virtual auditory space or VAS (Shinn-Cunningham, 1998; Wenzel et al., 1988; Wightman and Kistler, 1989). Typically, current headphone technology does not offer quality 3D sound. In most cases the sound that is produced over head phones seems to originate from the centre of the head and does not contain any spatial cues. In a more sophisticated set-up, a generic HRIR is used instead of the individual’s HRIR (Wenzel et al., 1991). The use of non-individualized HRIRs have shown to produce larger front-back and up/down confusions when localizing sound (Wenzel et al., 1993) in comparison to individualized HRIRs. Therefore an important impetus for the present research is for the production of high quality VAS which is not possible without individualized HRIRs.

At present it is difficult for listeners to obtain individualized 3D audio filters. Currently there are two established methods for obtaining the HRIRs. In the first method if a user requires his/her HRIRs he/she would have to measure them in an anechoic chamber with the use of specialized equipment, which could be both a laborious and time consuming task. In the second method the HRIRs are computed via numerical simulations conducted on the head and ear morphology. Numerically computing the HRIRs via numerical simulations has become feasible in recent years due to the advancement of computer processing and algorithms, but, on the down side, this procedure could be equally or more time consuming. To elaborate, in order to obtain the HRIRs for an individual numerically we need to, among other tasks, obtain accurate surface or volume meshes of the individuals ear, head and
1.3 Problem Statement and Scope of Work

The relationship between ear shape and ear acoustics is complicated and is an important aspect for research investigating individualisation of the 3D audio filters. Both the morphology of the ear shape and the features in the HRIR functions are complicated. The external ear shape contains many cavities which interact in a complicated manner with the incoming sound wave. This causes the HRIR filter functions to have many complicated features in the form of peaks and sharp notches. Given this it is a very challenging task to find a mapping between the morphology of the ear and its corresponding acoustics.

Given the importance of the above, there is no comprehensive function that maps the features in the external ear shape to features seen in the HRIR/HRTFs in the prevailing literature. The current body of work that links the head and ear morphology to the HRTF spectrum has a limited use in the production of VAS (refer to Sec. 3.5). Nevertheless, despite this limitation, our understanding on the mechanisms that link the morphology to the acoustic spectrum has come a long way. For example, Lopez Poveda and Meddis (1996) modelled the physical mechanisms that generate the first and third notches in the HRTF spectrum using a simple parabolic sheet and other researchers (Mokhtari et al., 2010; TEW et al., 2012) linked specific regions of the ear shape to peaks and notches seen in the HRTF spectrum using small perturbations to a template shape. Whilst these findings are
very important, the need for a more in-depth understanding of the mechanisms that
relate the ear morphology to the acoustics is especially need. This is because the ear
shape is an intricate surface and the morphological and acoustic variations in the
 morphology of the ear shape is quite large (refer to Sec. 7).

In order to overcome this shortcoming, one of the aims of this research is to
provide further insight into how the HRIR/HRTF spectrum changes as ear shapes
change for a given population of ears. Understanding the large changes seen in the
ear shape is a crucial step towards building a rapid model that maps the morphology
 of the ear and head shape to the acoustics. Keeping this aim in mind the simulations
are of two parts. In the first part, the simulations involve morphological data in the
form of 3D surface meshes of the ear, head and torso shapes. Here, ear shapes will be
analysed using large diffeomorphic (smooth and invertible) deformations. Further,
the changes in the ear shapes with respect to a template or average ear shape will
also be measured. The production of the template shape will be described in this
thesis. As previously noted, the SYMARE database of ears is used in this research,
which includes the examination of the changes of the ear morphology.

In the second part, the current research involves examining the variations in
the acoustics based on the variations of the morphology. In order to perform this
work extensive numerical simulations are conducted. More specifically, a significant
number of FM-BEM simulations are performed on specialized computer hardware
utilizing specialized software. These simulations are performed on high resolution
ear, head and torso meshes using the Coustyx software (Ansol, Ansol). A novel and
unique feature of this research relates to the scale at which both the morphological
and acoustic simulations are conducted. The performance of simulations at this
scale would have been almost impossible to achieve a few years ago. However, with
the recent advancements in hardware such as GPUs, and also the advancement of
algorithms such as the FM-BEM, the study of the variation in shapes with respect to
the variation in the corresponding acoustics is more realistically achievable these
days.
1.4 Contribution

In this research shape and acoustic simulations using advanced techniques will be performed and this will assist in elucidating some of the probable functions and mechanism that relate the morphology of the head and ear shape to the HRIRs. More specifically, the investigation is aimed at identifying the variation in the ear shapes by utilizing the Large Deformation Diffeomorphic Metric Mapping framework (refer to Chapter 2) and the Kernel Principal Component Analysis, referred to as the KPCA (refer to chapter 7). In addition, by utilizing the Fast Multi-pole Boundary Element Method, known as FM-BEM, acoustic simulations and analysis will be performed. The work in this thesis relies on the data available in the SYMARE database which contains morphological and acoustic information for 61 subjects.

Analysing the variations of the ear shapes using the Large Diffeomorphic Deformations Metric Mapping (LDDMM) (see Sec. 2.8) is a very exciting approach to the study of ear shapes because it enables one to observe how the structure of the ear shape changes when smoothly deforming a source ear shape to a target ear shape. Fig. 1.1 shows how ear A is deformed towards ear B using the LDDMM framework. In particular, it is observed that the deformations are smooth and the changes occurring to the structure of the ear shape are visible along the deformation path.

![Figure 1.1: Changes in the ear shape can be observed as the source ear Ear A is deformed towards a target ear Ear B. In particular, the above figure shows how the structure of the ear changes. These structural changes that are seen in the deformation process can be captured and analysed, ultimately providing us with valuable information on how the source ear is different from the target ear. The colours on the plot indicate how much the vertices have moved with respect to the source Ear A. Further, the colours can be used as an indication of which regions of the ear shape is changing.](image)

Analysing the structural changes when deforming an ear shape can offer signifi-
ant insight into the relationship between the morphology of the ear shape and its acoustics. Studying the acoustic changes when deforming a template ear shape is not new and has been done in previous works (refer to Sec. 3.4.2-Sec. 3.4.3). However, the work presented here is distinctive because in the present study large deformations are applied to the template shape. Large deformations are necessary because it can capture large variations seen in the population of ears. This is in contrast to previous studies where only small perturbations were applied to a template or reference ear shape.

One of the important contributions of this thesis is the estimation of a template ear shape using LDDMM from the population of ears in the SYMARE database. The template ear shape is used as the reference ear shape when measuring and analysing the variations in the ears. The current investigation verifies the template ear shape by examining the features in the template ear using 2D and 3D images to corresponding features seen in the population of ears (Sec. 5.2.3). These features provide information on the size of the Concha, the width and length of the ear shape, and the size of the Lobe and Helix, to name a few. It is shown in this research that the template ear shape contains features that resemble the average features of the population of ears. The selection of these features are explained in detail later.

A novel feature of this study is that the quality of deformations from the template or reference ear shape are analysed both acoustically and morphologically. In order to analyse the quality of deformations from the template ear shape the necessary tools were built to compare shapes and the acoustic responses. Specifically, a new tool was built to compare shapes based on the framework of currents (refer to Sec. 4.3). By the aid of these tools it is shown that ear shapes can be best analysed at multiple physical scales (Sec. 6.1). The physical scale that was used to analyse the ear shapes depends on a set of input parameters to the LDDMM algorithm (Sec. 4.4). It will be shown that at the first physical scale the important structural information in ears is captured and, if more accuracy is needed, then deformations at a smaller scale can then be subsequently conducted and analysed.
The other major highlight and contribution of this research work is that the variations of the ear shapes is analysed using KPCA (Sec. 7). Utilizing the KPCA a systematic examination of the changes in the ear shapes and the associated acoustic variations is studied. As a result of this analysis the structural differences that exist for the ear shapes in the SYMARE population will be shown and, importantly, the effect of these structural differences on the HRTFs will be presented. In addition, through a series of innovative experiments the number of kernel principal components that are required to represent a shape acoustically and morphologically will be identified.

Another contribution that this thesis presents relates to examining the strength of a morph-able model for the ear shapes based on the LDDMM framework and the KPCA technique. The morph-able model on the ear shapes examines how accurately an external ear shape that is not included when building the model can then be constructed (i.e. modelled). It will be shown that the model constructs some ear shapes with a high level of accuracy, while for some other ear shapes the model does not construct accurate shapes.

Other aspects of the contribution of the current research is the analysis that is performed on the relationship between a set of ears and their acoustics. Important features in the HRTF spectrum are identified that change as the ear shapes change. This analysis is significant as it provides insight into the mechanisms that underlie the changes in the morphology and the mechanisms that cause changes in the acoustic spectrum. Finally, the analysis shows how coherently the HRTFs for a few angles in the median plane change with respect to each other as the ear shapes change.

1.5 Challenges

Analysing the morphology and acoustics of the ear shapes at such a large scale has had its challenges. Probably one of the biggest challenges inherent to this project
has been the time taken to run the acoustic and morphological simulations at the present scale\textsuperscript{2}. Although FM-BEM simulations can be done much quicker now than a decade ago, calculating the full acoustic spectrum, that is, for the frequency range of 20Hz to around 17kHz with a spatial resolution of approximately 3 degrees is very time consuming. Simulations currently take around 1 day to complete if the full head and torso shape is used and shorter if only the head shape is used. Analysing and processing shapes using LDDMM is also very challenging because estimating a template ear shape for the simulation can take up to a month to conduct.

Modelling of the ear shapes has also presented it own unique challenges. This is because the LDDMM framework operates at a certain physical scale which is determined by a few input parameters. The quality of the mapping is greatly effected by these parameters and finding an optimal value for the parameters can be a very time consuming task, especially if both the morphology and acoustics are taken into consideration. Given this, the multi-scale approach was proposed to overcome this problem. This is further explained in Sec. 6.1.

\textbf{1.6 Outline of the Thesis}

This thesis is composed of a total of 8 chapters. The first 3 chapters consist of the introduction, background and literature review; the next four chapters are the contributing chapters that contain the research work performed. Then eight chapter provides concluding remarks to the thesis.

The background chapter (Sec. 2) contains important information on the standard principal component analysis and also the LDDMM framework. The LDDMM framework is used extensively in this thesis document. The section on LDDMM (Sec. 2.8) aims to inform the reader on important concepts and equations used within this framework. The section on PCA (Sec. 2.7) introduces the standard principal

\textsuperscript{2}For example producing a template ear shape using LDDMM could take several weeks on a Tesla K40 GPU. Further a dedicated Intel sever with six cores is being used for running FM-BEM simulations, however every FM-BEM simulation could take one day for a high resolution mesh shape.
component analysis technique and shows an application of PCA to the modelling of HRTFs. This chapter also contains a brief introduction to VAS (Sec. 2.6) and spatial hearing (Sec. 2.1) in order to give the reader some information as to how the present study fits into the larger research work conducted in this field.

Chapter 3 contains a review of the literature. This review contains shape modelling techniques that are used within the context of HRTF personalization and modelling (Sec. 3.1, Sec. 3.2). Further, this section details a review on the literature known as morphoacoustic perturbation analysis. In morphoacoustic perturbation analysis changes in the ear features are examined together with changes in the HRTF spectrum (Sec. 3.4). This part also details some of the techniques that have been carried out to date on the personalization of HRTFs using various methods like frequency scaling and regression based techniques (Sec. 3.5).

Chapter 4 is an important chapter as it details the framework by which morphological and acoustic studies of the ears are conducted. Specifically, in this chapter, tools are developed to compare both the morphology of the ear shapes (Sec. 4.3) as well as the acoustic responses of the ears (Sec. 4.8). These tools will be used extensively in this thesis. Further this chapter introduces important mathematical functions that describe the deformations between a source and target shape (Sec. 4.2). Explaining the process for deforming a source shape to a target shape is not an easy task and these mathematical functions are required in order to explain and elucidate the process. By using these functions it will be identified how deformations between a source ear, head and torso shape to a target ear, head and torso shape have been performed. Importantly, LDDMM scale and optimization parameters together with their corresponding effects on deforming ear shapes will be examined (Sec. 4.4). This chapter will also describe the procedure used for obtaining high resolution meshes (Sec. 4.5) for conducting FM-BEM simulations as these procedures are significantly used for the simulations carried out in this research work. Finally, the multi-scale approach to mapping shapes, particularly mapping ear shapes is presented (Sec. 4.9). The multi-scale approach will also be used in the production of the template ear
Chapter 5 details a new procedure for estimating a template ear shape. The new template estimating procedure (Sec. 5.1.3) that is described here is based on a multi-scale approach that was previously discussed (Sec. 4.9). This chapter also validates the template shape by comparing measurements of important features in 2D and 3D images of the template ear shape to that of the SYMARE (Jin et al. 2014) and CIPIC (Algazi et al. 2001) population of ears (Sec. 5.2.3).

Using the tools, the functions and the template shape developed in the previous chapters the next chapter (Chapter 6) examines both the acoustic and morphological quality of deformations from the template ear shape. By examining the deformations this section also identifies how the multi-scale approach Sec. 6.1 can be used to study ear shapes more effectively. Finally, this section contains a distance analysis in which morphological distances between ear shapes and the corresponding acoustic distances are compared using the tools developed in Sec. 6.2.

The last chapter of the thesis, Chapter 7, shows the variation in the shape of the ears by conducting a KPCA analysis. In particular, this section shows the variation in the ears captured by the first five kernel principal components, and also examines the corresponding acoustic responses (Sec. 7.4). It examines the number of kernel principal components that is required for modelling the morphology and acoustic responses of the ear shapes, (Sec. 7.3). In addition, a morph-able model for the ear shapes is constructed and examined (Sec. 7.7). Moreover, this chapter builds a simplified model of the HRTF spectrum in the median plane by using the PCA technique (Sec. 7.5). By employing this model it identifies important features in the principal components that cause changes to the HRTF spectrum as the shape of the ear changes. Finally, this chapter also contains an interesting research study that explores how coherently the HRTFs belonging to a few angles in the median plane change as the shape of the ear changes (Sec. 7.6). The study uses the PCA analysis to identify features in the acoustic spectrum related to a few angles that change coherently.
Chapter 2

BACKGROUND

This chapter explains key concepts that are useful for understanding the research presented in this thesis. The research presented here ultimately aims to enable an individual to hear sounds spatially over headphones. Topics covered in this chapter will assist the reader to better understand the present study. More specifically, it will familiarize the reader with relevant concepts and techniques related to spatial hearing, human acoustics and statistical shape processing.

This chapter will start by introducing the reader to spatial hearing and the cues that enable an individual to localize sound in 3D space (Sec. 2.1-Sec. 2.2). Following the section on Spatial Hearing a set of filter functions that enable an individual to listen to spatial sounds over headphones, known as HRIRs, are described in Sec. 2.3. Subsequently, details are provided on how to measure the HRIRs in a laboratory (Sec. 2.4) and to numerically compute them from an individual’s head and ear geometry (Sec. 2.5). In Sec. 2.6 the virtual auditory space (VAS) is described together with the method for the production of VAS over headphones. Further, some applications of VAS are detailed in Sec. 2.6.1. The principal component analysis technique (PCA) utilized to statistically analyse the DTFs, and the LDDMM framework for analysing and processing shapes is described in Sec. 2.7 and Sec. 2.8 respectively.

A reader familiar with topics related to spatial hearing or VAS can skim the relevant sections or refer to it later. Notations and theorem given in the PCA analysis
2.1 Spatial Hearing

Spatial hearing refers to the human ability to perceive the location of sound sources in space (Blauert, 2013). Spatial hearing is important because it enables humans to interact much more effectively with their surrounding environment. Further, it complements our other senses such as vision. Human vision is a powerful but limited organ of sense as it is not able to observe objects behind the head or at situations where there is a deficiency in light. In contrast, humans are able to perceive and localize sound originating from all directions in space. For this reason spatial hearing is a powerful sense in humans as it helps them to interact more effectively with the surrounding environment. This can be illustrated by imagining a scenario from a cocktail party (Bronkhorst, 2000) where many competing sound sources exist around an individual. Now, if humans were not able to distinguish the sound sources and respond to them appropriately, it would severely limit the ability of the individual to interact socially.

People are capable of localizing sound in 3D space quite remarkably. In the horizontal plane, humans on an average can distinguish sounds to within 1° to 3° (Perrott and Saberi, 1990). On the other hand, localization of sound in the vertical planes, which is the up/down direction, is harder. The angle at which humans are capable of distinguishing sound in the vertical domain is around 4° for white noise and 17° for speech (Perrott and Saberi, 1990; Xie, 2013). This angle is also known as the minimum audible angle (MAA). Humans are able to localize sound in 3D space using two binaural cues (Rayleigh, 1907; Wightman and Kistler, 1992) and also monaural spectral cues.

Binaural cues are cues that the human auditory system picks up from both the left and right ears. Binaural cues are one of the oldest cues discovered for localizing sound in humans. They were first introduced a century ago by Lord
Rayleigh (Rayleigh 1907). Binaural cues arise from the physical separation of the ears by the head. This separation causes the sound signals to have different time of arrival at the two ears, referred to as the ITD cue. Further, there is a difference in the sound intensity level at the two ears referred to as ILD. Both ITD and ILD are very important cues in humans for localizing sound in the horizontal domain. The ITD is used for localizing low frequency sound while ILD is used for localizing high frequency sound (Macpherson and Middlebrooks 2002).

![Figure 2.1](image.png)

**Figure 2.1:** Different sections of the human external ear. Hunter et al 2009

### 2.2 Spectral Cues

The other set of cues that contribute towards spatial hearing is spectral cues (Asano et al. 1990; Butler 1969; Butler and Belendiuk 1977; Hebrank and Wright 1974; Musicant and Butler 1984; Roffler and Butler 1968). These cues are also referred to as monaural spectral cues and arise from the acoustic filtering of the outer ear. Spectral cues are formed by the interaction of the incoming sound wave with the morphology of the listener’s ear, head and torso shape (Algazi et al. 2001, 2007). Spectral cues exist at both low frequencies and high frequencies. For low frequency
sound it is shown that reflections from the head and torso contribute significantly towards spectral cues (Algazi et al., 2001), and for high frequency sound the ear shape contributes most towards the formation of the spectral cues (Algazi et al., 2007).

The external ear shape is a very intricate surface. The interaction of the sound waves with different sections of the external ear shape in the form of reflections and diffractions cause the formation of the spectral cues. Fig. 2.1 shows a plot of the external ear shape and the terminology used for the different sections of the ear shape (Hunter et al., 2009; Hunter and Yotsuyanagi, 2005).

![Figure 2.1](image)

Figure 2.2: The above figure compares the acoustic responses for three shapes for a cone of confusion located at an azimuth of 25 degrees, the measurements were taken at 5 degree intervals (a) The acoustic response of the KEMAR head and torso shape measured without the pinna, (b) The acoustic response of the KEMAR pinna shape alone (PRTF) measured on a circular plate, (c) The sum of the responses shown in (a) and (b), (d) the measured acoustic response of pinna, head and torso shape. Image reprint from Algazi et al. 2001

Experiments have shown that the pinna shape contributes greatly to the formation of the monaural spectral cues (Butler and Belendiuk, 1977; Gardner and Gardner, 1973; Hebrank and Wright, 1974). This has inspired researchers to model the response of the pinna shape without the head (Algazi et al., 2007; Spagnol et al., 2010).
2.2. Spectral Cues

Fig. 2.2 shows the separate acoustic responses of the KEMAR head and torso shape without the pinna (Fig. 2.2(a)), the KEMAR pinna shape alone (Fig. 2.2(b)), the combination of the KEMAR pinna plus the KEMAR head and torsos shape (Fig. 2.2(c)), the full KEMAR shape (Fig. 2.2(d)). In a separate examination, Takemoto et al. (2012a) examined the acoustic responses of the head and ear shape versus only the ear shape. Fig. 2.3 shows the acoustic responses of the head and ear shape versus the ear shape only. Again, much of the acoustic features seen in the head and ear responses (HRTFs) can also be seen in the ear responses alone (i.e PRTFs).

One of the prominent spectral cues that is formed by the ear shape is the notches and peaks that can be observed in the frequency spectrum of the sound sig-
nal (Hebrank and Wright, 1974) (Butler and Belendiuk, 1977). A significant amount of research suggests that removing these features from the sound signal spectrum can degrade the ability of humans to localize sound considerably, particularly when localizing sound in the median plane (Asano et al., 1990; Langendijk and Bronkhorst, 2002). Further, Gardner and Gardner (1973) performed an interesting experiment on localization using occluded ear shapes. Results showed that localization by means of an occluded ear was almost impossible.

![Image of auditory angles using head centred spherical coordinates.](image)

Figure 2.4: The auditory angles using the head centred spherical coordinates is shown in the above figure. The $\theta$ signifies the horizontal angle, while $\phi$ signifies the elevation angle.

### 2.3 Head Related Impulse Response And Head Related Transfer functions

The head-related impulse response (HRIR) is an acoustic transfer function that describes the transformation of an acoustic signal from a particular direction in space to the left or right ear drum (Cheng and Wakefield, 1999b; Middlebrooks, 1992). The HRIR is a finite impulse response filter and mathematically explains the spatial acoustic filtering properties of the external ear shape for a given direction in space. A pair of HRIRs for the left and right ear shapes contains all the binaural and spectral cues which are used by an individual to localize sound sources. The HRTF is the frequency response associated with the HRIR and is obtained by performing a
Fourier transform on the HRIR. It is a complex valued function and has a magnitude and phase at each frequency. The HRTF shows the gain of the ear shape for a given direction in space with respect to the frequency. There is a single HRIR/HRTF for each direction in space. Directions in space are identified by the head centred spherical coordinates, which are denoted by $\theta$ and $\phi$. These angles are shown in Fig. 2.4. Each of the left and right ear shapes contain a unique set of HRIR/HRTFS for all directions in space. Fig. 2.5 shows a plot of the HRIR and HRTF for a particular direction of space for a given individual.

![HRIR and HRTF plots](image)

Figure 2.5: The head related impulse response (HRIR) and the head related transfer function (HRTF) is shown in the above plots for azimuth angle of $31.7174^\circ$ and elevation of $0^\circ$ for a given subject.

While the HRIR/HRTF shows the variation in the gain with respect to frequency for a given direction, we can also show the gain or directivity of the ear shape with respect to various directions in space. This representation is known as the Spatial Frequency Response Surface (SFRS) (Cheng and Wakefield, 1999a). The SFRS is plotted for a single frequency and shows the directivity of the external ear shape. Fig. 2.6 shows an SFRS plot at 12kHz.
2.4 Methods for Acquiring HRIRs

There are two popular methods for obtaining individualized HRIRs. The first method is to acoustically measure the HRIRs \cite{Algazi2001, Jin2014, Majdak2007, Moller1995}, and the second method is to compute the HRIRs by means of numerical computations on an individual's head geometry in the form of a surface or volume mesh \cite{Gumerov2007, Katz2001a, Mokhtari2007, Mokhtari2008}. The method to numerically compute the HRIRs will be discussed in more detail in Sec. 2.5.

Traditionally, individualized HRIRs were acoustically measured in an anechoic chamber that incorporated a robotic arm that moved the sound source in space \cite{Moller1995}. In this method microphones are placed in the ears of the listener and impulse responses are recorded for various directions in space by moving the robotic arm. Although this method provides accurate measurements of an individual’s HRIRs it is relatively expensive and time consuming.
Fig. 2.7 shows a typical set-up for measuring the HRIRs in an anechoic chamber. An anechoic chamber is a room without echo for sound signals. Different parts of the apparatus are also shown in the same figure. The cost of setting up such an apparatus is significant as it requires the set-up of the anechoic chamber, loudspeaker array located on a robotic arm, and/or other specialist equipment for recording the HRIRs for an individual.
2.5 Numerically Calculating The HRIRs

As indicated previously, an alternative technique for the computation of personalized HRIRs is through numerical simulations using the boundary element method (BEM) (Kahana and Nelson, 2007; Katz, 2001a). Computing HRIRs using numerical simulations has only recently become feasible with the advancement of algorithms such as the fast multipole method and hardware power. Numerically computing the HRIRs also requires personal measurements of the head and ear shape in the form of surface or volumetric meshes. Computation of the HRIRs using BEM is a very interesting topic in human acoustics for a number of reasons (Kreuzer et al., 2009; Mokhtari et al., 2007, 2008; Rui et al., 2013; Ziegelwanger et al., 2015). Firstly, numerical simulations can be done much faster now than just a decade ago, both because of the increased power in computers and also due to the advancement of algorithms such as the fast multipole method (FMM) (Gumerov et al., 2007; Gumerov and Duraiswami, 2009; Gumerov et al., 2010; Kreuzer et al., 2009). Secondly, obtaining head meshes is feasible these days using various forms of scanning such as MRI (Guillon et al., 2012; Jin et al., 2014), laser scanning (Katz, 2001a; Kreuzer et al., 2009) or computerized tomography (De Mey et al., 2008), among other options. The next section will briefly discuss how BEM is formulated.

2.5.1 Boundary Element Method

Numerical solutions for obtaining the HRIRs involve numerically solving and computing the PDE that governs the scattering of sound waves around objects in three dimensional space. The PDE for scattering of sound waves is known as the Helmholtz PDE (Sommerfeld, Som; Chertock, 1964; Strutt, 2011). In order to describe the Helmholtz PDE lets assume that there is a radiating sound source, \( m \), at a sufficiently large distance in a domain \( W \). This domain is shown in Fig. 2.8. Subsequently, the pressure at any point of this domain is obtained by the sum of the incident and
scattered sound pressures:

$$\Phi = \Phi^{sc} + \Phi^{in}$$  \hspace{1cm} (2.1)

where $\Phi$ is a function representing the total pressure field in the domain $W$. The superscripts "sc" is for the scatter field and the superscript "in" is for the incident field. The total pressure field $\Phi$, the scatter field $\Phi^{sc}$, and the $\Phi^{in}$ in the domain $W$ are modelled and satisfy the **Helmholtz** equation:

$$\nabla^2 \Phi + k^2 \Phi = 0,$$  \hspace{1cm} (2.2)

the operator $\nabla^2$ is the Laplace operator and $k$ is the wave number and is equal to $k = 2\pi f / c$, and $f$ is the frequency of the sound wave, and $c$ is the speed of sound in the domain $W$. Fig. 2.8 shows the domain $W$, the radiating sound source $m$, the scatter point $y$ on the boundary, and a sample point $p$ where the pressure is to be known. The point $p$ can be anywhere in the domain.

![Figure 2.8: 2D plot showing the scattering of sound waves in a domain. The boundary of an object in this domain is represented by $B$. The PDE that computes the pressure field in the domain is known as the Helmholtz PDE. In the above figure $W$ is the external domain where the scattering problem is solved, $D$ is the internal domain and $B$ is the surface of the scatter, $m$ is an external radiating source, $y$ is a reflecting source on the boundary and $p$ is the point where the pressure is being considered.](image)

The fundamental solution to the Helmholtz PDE (Eq. (2.2)) is given by the **Greens**
function $G_k$: 

$$G_k(y, p) = \frac{e^{ik|y-p|}}{4\pi|y-p|}$$ (2.3) 

where $k = \frac{2\pi f}{c}$ and $c$ is the speed of sound. Using the Greens function a Boundary integral equations (BIE) can be formulated in order to solve the PDE (Burton and Miller [1971] Copley [1967]). The Burton-Miller (Burton and Miller [1971]) formulation of the BIE for solving the Helmholtz PDE is given by the following:

$$\frac{1}{2} \Phi - M_k[\Phi] + h \frac{\partial M_k[\Phi]}{\partial n(p)} = \Phi^{in} - h \frac{\partial \Phi^{in}}{\partial n(p)} \quad (p \in B)$$ (2.4) 

where $M_k$ is the "double layer potential" defined as:

$$M_k[\Phi] = \int_B \Phi(y) \frac{\partial G_k(y, p)}{\partial n(y)}(y) dS(y) = \int_B \Phi(y) F_k(y, p)$$ (2.5) 

The coefficient $h$ is complex and must be chosen arbitrary with a non-zero imaginary part. Kirkup (2007) showed that the Burton-Miller formulation Eq. (2.4) can be implemented using the collocation method:

$$\left[\frac{1}{2} I - M_k + hN_k\right] \Phi = \Phi^{in} - h\nu^{in}$$ (2.6) 

where:

- $\Phi$ is a vector of the total acoustic pressure at each element of the boundary. The boundary can be 2D or 3D.
- $\Phi^{in}$ and $\nu^{in}$ are the vectors of the incident acoustic pressures and incident velocity respectively.
- $I$ is the identity matrix.
- $M_k$ is the matrix representing the discrete version of the double-layer potential operator.
- $N_k$ is the matrix representing the discrete version of the normal derivative of
2.5. Numerically Calculating The HRIRs

the double-layer potential operator.

Eq. (2.6) is in the form of a linear system of equations:

\[ A_k \Phi_k = b_k \] (2.7)

where \( k \) signifies the frequency under consideration. In order to compute the pressure for a range of frequency and source positions Eq. (2.7) needs to be constructed and solved for each of the given frequencies and source directions.

2.5.2 Validating The Computation of HRIRs Through BEM and FM-BEM Simulations

This section will describe a few studies from Gumerov et al. (2007); Gumerov and Duraiswami (2009); Kahana and Nelson (2007); Katz (2001a,b) on the use and verification of BEM and FM-BEM for the computation of the HRIR functions numerically.

The first exploration of HRTFs using BEM simulations conducted by Katz (2001a) compared the acoustic responses of four shapes (1) a full pinna, head and neck mesh, (2) a pinna less head and neck shape (3) a rigid sphere with a diameter equal to the spacing of the ears, "IAS sphere" (4) and rigid sphere with the same volume as the head shape "Eqvol sphere" using BEM simulation on a computer. The mesh resolution for the head shapes was valid up to a critical frequency of 5.4kHz given four elements (mesh elements) per sound wavelength. The Geometrical data was obtained using an optical 3D scanner and then the meshes were taken through many stages of refinement.

Using BEM simulations Katz identified (a) the head contributes primarily to the low frequency components of the HRTF’s (b) the sphere had a similar acoustic response to the head without the pinna (c) the pinna contributes significantly at frequencies greater than 3kHz. Figure 2.9 shows the responses of the four meshes for a source positioned directly in front.

Following the work by Katz (2001a,b), Kahana and Nelson (2007) carried
2.5. Numerically Calculating The HIRs

out various BEM simulations on the KEMAR manequin and the KEMAR pinna shapes DB60, DB65, and DB90 to investigate the computations of HIRs using BEM simulations. Some of these findings are summarized below:

- Resonance modes of the pinna obtained by BEM simulations agreed well with resonance modes of the pinna shown by Shaw [Shaw and Teranishi, 1968].

- The principle of reciprocity was validated. Using a baffled DB60 pinna shape in the first case, the source was positioned in the far field and the sensor was positioned at a distance of 1mm from the entrance to the blocked ear canals; in the second case, the position of the source and sensor was interchanged. Fig. 2.10 shows results which confirm that these two scenarios produce almost identical results.

- It was shown that the low resolution pinna mesh shape degrades the acoustic response for a wide range of angles and for frequencies above 8kHz in comparison to a high resolution pinna mesh. The pinna used in this investigation was the DB60 KEMAR pinna. Fig. 2.11 shows the low resolution DB60 pinna shape, and Fig. 2.12 shows the acoustic frequency responses for the low and high resolution baffled pinna shapes for the horizontal plane.
2.5. Numerically Calculating The HIRs

Figure 2.10: Indirect BEM simulations are performed in two cases to test the reciprocity principle. In the first case the source is positioned in the far field and sensors were positioned in the blocked ear canal; in the second case the source is positioned in the blocked ear canal and the sensor is in the same position of the source as in the first case. -o- source in the far field, -x- reciprocity, pressure 1 mm away from the blocked ear canal. The -*- plot with jitter in the high frequency is the simulation with reciprocity with no over determination points. Image taken from Kahana et al.(2007)

- The baffled pinna shape (i.e. the pinna attached to a rectangular plane) and the pinna on a head had different acoustic responses for sources in the horizontal plane and similar responses for sources in the vertical plane. Fig. 2.13 illustrates acoustic responses for the baffled and pinna on the head for a source originating in the vertical plane versus a source originating in the horizontal plane.

Figure 2.11: The image of the DB60 pinna shape scanned by a low resolution scanner. The shape consists of approximately 6000 nodes and 12 000 elements. Image taken from Kahana et al.(2007)
2.5. Numerically Calculating The HRIRs

Figure 2.12: The acoustical responses for a pinna shape obtained by a low resolution scanner "low-res" versus a pinna shape obtained by a high resolution scanner "high-res". The "low-res" pinna had approximately 12000 elements, and the "high-res" pinna had 13488 elements, and a critical frequency of 15560 given 6 elements per wavelength. Image taken from Kahana et al. (2007)

In a different study, Kahana and Nelson (2006) examined the acoustic transfer function of the baffled DB-65 pinna (pinna attached to a rectangular plane) shape shown in Fig. 2.14. BEM was used to obtain the acoustic responses for the baffled pinna shapes. The BEM’s were calculated using the reciprocity principle, which was validated in a separate paper (Kahana and Nelson, 2007). The maximal edge length for the meshes was at \( e_{\text{max}} = 3.7 \text{ mm} \) corresponding to a critical frequency of approximately 15 kHz. Resonance frequencies were identified for the pinna shapes; the resonance frequencies reported in this study are at frequencies of 4.0kHz, 7.2kHz, 9.5kHz, 11.6kHz, 14.8kHz and 18kHz. Fig. 2.15 demonstrates the plot of the normalized acoustic responses at grazing angles for the baffled pinna. Further, Fig. 2.16 shows the directivity plot of the pressure at the ear canal and the excitation of the pressure field within the pinna for the direction where this pressure is maximum.

The fast multipole method (FM) for computing the BEM has made the computation of the HRIR functions using BEM faster by many orders of magnitude (Gumerov et al., 2007, 2010; Kreuzer et al., 2009). Gumerov et al. (2010) performed a study on the quality of the HRIR produced by FM-BEM on a head and pinna shape belonging to the KEMAR mannequin. Fig. 2.17 shows the mesh shapes used in the simulations. The criteria for suitable mesh shapes for FM-BEM reported by Gumerov et al. (2010)
Figure 2.13: The acoustic responses for a pinna attached to a rectangular plane and the pinna shape on the head are shown. The two sources are (a) originating in the vertical plane and (b) originating in the horizontal plane. The correlation in the acoustic responses between the baffled pinna and the pinna attached to the head is greater for the vertical plane than the horizontal plane. Image taken from Kahana et al. (2007)

Figure 2.14: The DB-65 baffled pinna shape and the coordinate system used for the simulations for computing the acoustic transfer functions. Image taken from Kahana et al. (2006)

were: (1) six elements per wavelength, (2) triangles should be as close to possible to equilateral, (3) the number of triangles associated with each vertex should be as close to six as possible and, (4) the rate of change for triangle areas should be gradual. It was shown that the HRIR/HRTFs produced by the FM-BEM simulations matched the experimental results well. Fig. 2.18 presents the results for the computed and experimental HRTFs in the median plane.
2.5. Numerically Calculating The HRIRs

Figure 2.15: Acoustic transfer function for the DB-65 pinna attached to a rectangular plane, at grazing angles. The frequency is in steps of 200Hz and the angular resolution is at 1°. The modes corresponding to resonance frequencies of the pinna are seen in warm intense colours. The modes are at frequencies of 4.2, 7.2, 9.5, 11.6, 14.8 and 18.0kHz. Image taken from Kahana et al. (2006)

Figure 2.16: Directivity plot (left) for the DB-65 pinna mode at 9.5kHz. The pressure field (right plot) on the surface of the ear is also shown. Image taken from Kahana et al. (2006)

Figure 2.17: The KEMAR mannequin and the ear shape used in the FM-BEM experiments. Image taken from Gumerov et al. (2009)
2.6 Virtual Auditory Space

Virtual auditory space is the process of rendering spatial sounds over a loudspeaker array (Gardner, 1998) or headphones (Begault and Trejo, 2000; Blauert, 2013; Carlile and Hyams, 1997; Zotkin et al., 2004). Binaural VAS is the production of 3D spatial audio over headphones and is the focus of the discussion presented here. The production of VAS over loudspeakers has common steps to the production of VAS over headphones with the use of an additional set of filters. VAS can be used to emulate either a single direction or a number of directions in space. For example, VAS can be used to emulate 5.1 and 7.1 surround sound over headphones. Fig. 2.20 shows speaker configuration for the 5.1 and 7.1 sound. The production of high quality VAS over headphones is currently very limited because personalized HRIRs (Sec. 2.3), which are essential for the production of realistic 3D sounds, is not easily obtained. Simple schemes such as using the HRIRs of a different individual or using a generic HRIR have been proposed to compensate for the lack of personalized HRIRs. However, the usefulness of such schemes is very limited as the sound signal produced over headphones or loudspeakers using non-personalized or generic HRIRs will appear smeared and non-natural (Gardner and Gardner, 1973; Wenzel et al., 1993). Simple techniques for personalizing the HRIRs are described in more
2.6. Virtual Auditory Space

detail in Sec. 3.3.

![Diagram of VAS production]

Figure 2.19: The above figure shows a typical set-up for the production of VAS over headphones. The input sound signals $e_1, e_2$ are convolved with the appropriate left and right HRIR filter functions and then added separately for the left and right ear shapes to produce the spatial sounds.

The basic mechanism for the production of VAS over headphones involves emulating a sound signal, $e_1$, from a particular direction in space given by the spherical coordinate angles, $\theta_1$ and $\phi_1$, the radius, $r$, for the spherical coordinate system is ignored for this discussion. The sound signal, $e_1$, needs to be convolved with the corresponding left and right ear personalized HRIR filters, i.e., $h_L(\theta, \phi)$ and $h_R(\theta, \phi)$:

$$e_1 \ast h_L(\theta, \phi) = L \quad (2.8)$$
$$e_1 \ast h_R(\theta, \phi) = R \quad (2.9)$$

Where in the above $\ast$ denotes the convolution function. Fig. 2.19 shows the production of VAS over headphones for two input sources, $e_1$ and $e_2$, originating from directions $(\theta_1, \phi_1)$ and $(\theta_2, \phi_2)$. In Eq. (2.8) and Fig. 2.19 the production of "free-field VAS" using the HRIR functions is presented. This does not take into consideration
effects such as room reverberation. A more accurate VAS can be generated using the binaural-room impulse responses (BRIR) (Zotkin et al., 2004), which is the transfer function from a point in space to the ear drum for a non-anechoic room.

![Figure 2.20: A typical 5.1 and 7.1 surround sound speaker configuration.](image)

Humans often move their heads in different directions when trying to localize sound sources and when trying to focus their senses to a particular direction in space. The procedure for the production of VAS given in Eq. (2.8) did not incorporate the effects of head movement. When an individual moves his head the relative direction of the sound source with respect to the ears also moves with the movement of the head. In dynamic VAS systems, head movement is incorporated using a head tracker system (Begault et al., 2000, 2001; Zotkin et al., 2004). The head tracker system will send information for the position of the head to the computer device that controls the production of VAS and the filter functions used in the production of VAS is adjusted accordingly.

### 2.6.1 Applications Of VAS

VAS has a significant history of known applications in health, research, defence and entertainment. For example, in the research sector, VAS has been used in clinical experiments to study the spatial hearing abilities of subjects that have hearing impair-
ments (Häusler et al., 1983). VAS has been used in localization experiments in order to obtain a better understanding on the mechanisms of spatial hearing (Wightman and Kistler, 1989). Socially as well, at the emotional health level, VAS can assist many people who have hearing impairments to interact more effectively with their environment. For instance, patients with implanted cochlea devices can greatly utilize VAS to interact more effectively with other people.

From an entertainment perspective broadcasting companies such as, for example, BBC and Microsoft are investing a significant amount of resources into the production of VAS. Games are also becoming much more involved and 3D scenes are currently being rendered more realistically with the use of hardware acceleration and specialist glasses. The same can also be said about movies, in that the quality of movies is also rapidly increasing. Currently, movies can be watched in 3D using specialist glasses, and 4K movies are becoming increasingly more popular. Nevertheless, due to this sophistication, there is a significance short coming related to the lack of an engaging sound system, such as VAS, to complement the visual improvements seen in the entertainment industry. This issue also extends to mobile phones and communication software on computers as well. VAS can enhance consumers ability to communicate more effectively with each other using these devices (Algazi and Duda, 2011).

2.7 Principal Component Analysis

This section describes the standard principal component analysis (PCA) technique. The PCA is used for the analysis of the acoustic responses of the ear shapes in Sec. 7.5 - Sec. 7.6 of this thesis. In this section a set of DTF data will be used to explain the PCA procedure. The notation and procedures developed here will be referred to and also used in Sec. 7.

By performing PCA on the data set, the data is transformed into a new coordinate system. PCA is an orthogonal transformation technique in which a set of orthogonal
bases vectors are estimated from the data set.

To begin the discussion, we define an observation vector, \( H(f, \theta_1, \phi_1) \), which represents a single DTF that belongs to a particular direction in space identified by the angles \( \theta_1 \) and \( \phi_1 \) (Please refer to Sec. 4.7 for details on extracting and computing the DTF data). The angle \( \theta_1 \) is the azimuth angle, and angle \( \phi_1 \) is the elevation angle. The DTF has \( K \) frequency components or features which are the complex gains for the \( K \) frequency bins of the DTF spectrum. The discrete frequencies are denoted as \( f_i \) in the DTF spectrum. For the PCA we use only the magnitude of the DTF spectrum, i.e., we take the absolute value of the complex gain. The PCA is performed on a set of DTF data that consists of \( L \) DTF magnitude spectrums belonging to \( L \) directions in space which can be arranged into a data matrix \( X \):

\[
X = \begin{pmatrix}
H(f_1, \theta_1, \phi_1) & H(f_2, \theta_1, \phi_1) & \ldots & H(f_K, \theta_1, \phi_1) \\
\vdots & \vdots & \ddots & \vdots \\
H(f_1, \theta_L, \phi_1) & H(f_2, \theta_L, \phi_1) & \ldots & H(f_K, \theta_L, \phi_1)
\end{pmatrix} \in \mathbb{R}^{L \times K}
\]

There are four steps involved in computing the PCA on the data matrix \( X \) and to obtain the coordinates using the PCA bases. These steps are listed below:

**One** The mean for every feature for our observations is computed. In the case of the DTFs we compute the mean for every frequency bin for the \( L \) directions in space:

\[
\bar{H}_k = \frac{1}{L} \sum_{i=1}^{L} H(f_k, \theta_i, \phi_1)
\]  

(2.10)

**Two** The zero mean data matrix \( \bar{X} \) is computed by removing the mean from each observation:

\[
\bar{X} = \begin{pmatrix}
H(f, \theta_1, \phi_1) \\
\vdots \\
H(f, \theta_L, \phi_1)
\end{pmatrix} - \begin{pmatrix}
\bar{H} \\
\vdots \\
\bar{H}
\end{pmatrix} \in \mathbb{R}^{L \times K}
\]

(2.11)
A singular value decomposition on the zero mean data matrix $\hat{X}$ can be performed:

$$G_\hat{X}O_\hat{X}F_\hat{X}^T = \hat{X} \tag{2.12}$$

The SVD shown in Eq. (2.12) will be used later in this section.

**Three** The covariance matrix is computed where every entry in the covariance matrix is:

$$C_{kp} = \sum_{j=1}^{L} H(f_k, \theta_j, \phi_j) H(f_p, \theta_j, \phi_j) \tag{2.13}$$

In matrix notation this can more simply be written as:

$$C = \hat{X}^T \hat{X} \tag{2.14}$$

**Four** The singular value decomposition (SVD) on the covariance matrix $C$ is computed:

$$FOF^T = C \tag{2.15}$$

The covariance matrix can also be written using Eq. (2.12):

$$F_\hat{X}O_\hat{X}^T G_\hat{X}^T G_\hat{X} O_\hat{X} F_\hat{X}^T = C \tag{2.16}$$

$$F_\hat{X}O_\hat{X}^T OF_\hat{X}^T = C \tag{2.17}$$

From Eq. (2.16) and Eq. (2.14) it is deduced that $F_\hat{X} = F$ and $O = O_\hat{X}^T O_\hat{X}$. The dimension for the matrices $F, O \in \mathbb{R}^{K \times K}$. The matrix $F$ constitutes the **new orthogonal bases** for our data, and each column $f_i$ in $F$ is a **principal component**. Each pair of principal components $f_i$ and $f_j$ of the matrix $F$ are orthogonal with respect to each other:

$$\langle f_i, f_j \rangle = 0 \tag{2.18}$$

Further, the matrix $O$ is a diagonal matrix that contain all the **eigenvalues** from the SVD decomposition.
The new coordinates or weights, \( w \), for the DTF \( H(f, \theta_x, \phi_x) \) is given by:

\[
\mathbf{w} = (H(f, \theta_x, \phi_x) - \bar{H})^T \mathbf{F}
\]  

(2.19)

The vector \( \mathbf{w} \in \mathbb{R}^k \). Further, the complete weights in the form of a matrix \( \mathbf{Q} \) for our data matrix \( \mathbf{X} \) is obtained by:

\[
\mathbf{W} = \hat{\mathbf{X}} \mathbf{F}
\]  

(2.20)

To clarify the PCA method, we now show an example of applying PCA to DTF data. The data matrix \( \mathbf{X} \in \mathbb{R}^{1342 \times 342} \) was generated by taking the DTF’s belonging to the median plane for 11 ear shapes. There were 122 DTFs for each ear shape (the DTFs were approximately 2.95° apart in the median plane), hence the data consisted of a total of 1342 DTFs. The DTFs each had 342 frequency bins, ranging from 20Hz
2.7. Principal Component Analysis

...to 16kHz for a 1024 point FFT computation.

![DTF and Reconstructed DTF plots](image)

Figure 2.22: The DTF for a given angle in the median plane and the same DTF reconstructed using all the acoustical principal components.

Fig. 2.21 shows the mean ($\bar{H}$) and the first three principal components $f_1$, $f_2$ and $f_3$ for the example DTF data. A full and accurate reconstruction for any of the DTFs in our data set can be obtained by linearly combining all of the principal components $f_i$ with the appropriate $w_i$. Fig. 2.22 shows a specific DTF and the reconstructed DTF when all principle components are used. When using PCA we typically do not require the full set of $L$ weights to represent the data with reasonable accuracy. In this sense, PCA provides a lossy compression of the data. Fig. 2.23 shows the reconstruction of the PCA’s using a subset of the weights which are indicated above the plots.

\[\text{More details for extracting and manipulating the acoustic data are given in Sec. 4.7}\]
2.8 Large Deformation Diffeomorphic Metric Mapping (LDDMM) framework

The LDDMM framework is presented in Grenander and Miller (1998); Miller and Younes (2001) and specifically developed and applied to shapes in Vaillant and Glaunès (2005); Vaillant et al. (2007). It is based on theories from functional analysis, variational analysis, and reproducible kernel Hilbert spaces. In order to describe surface matching using LDDMM, we define the vertices and connectivity information between vertices for two surface meshes, as $S_1(X)$ and $S_2(Y)$, where $X \subset \mathbb{R}^{3 \times N}$ and $Y \subset \mathbb{R}^{3 \times M}$ specify the vertices of the two surfaces, respectively. LDDMM models
the mapping or morphing of \( S_1(X) \) to \( S_2(Y) \) as a dynamic flow of diffeomorphisms of the ambient space, \( \mathbb{R}^3 \), in which the surfaces are embedded. This flow of diffeomorphisms, \( \phi^v(t, \cdot) \), is defined via the partial differential equation:

\[
\frac{\partial \phi^v(t, X)}{\partial t} = v(t) \circ \phi^v(t, X),
\]

where \( v(t) \) is a time-dependent vector field with a vector defined for each point in space, for \( t \in [0, 1] \), which models the infinitesimal efforts of the flow, and \( \circ \) denotes function composition. At this stage the vector field represented by \( v(t) \) can be any vector field belonging to a Hilbert space of regular vector fields () denoted by \( V \) which is equipped with a kernel, \( k_V \), and a norm \( \| \cdot \|_V \) that models the infinitesimal cost of the flow.

Note that the superscript \( v \) on \( \phi^v(t, X) \) simply denotes the flow of diffeomorphisms defined for a particular time-dependent vector field \( v(t) \). In the LDDMM framework, we determine \( v(t) \) by minimizing the cost function, \( J_{S_1, S_2} \):

\[
J_{S_1, S_2} (v(t)) = \gamma \int_0^1 \| v(t) \|_V^2 dt + E (S_1(\phi^v(1, X)), S_2(Y)),
\]

The cost function is composed of an energy term \( \int_0^1 \| v(t) \|_V^2 dt \) and a shape comparison term \( E (S_1(\phi^v(1, X)), S_2(Y)) \). The energy term provides a measure of the energy used when deforming surface \( S_1 \) to surface \( S_2 \). The shape comparison term \( E \) provides a measure for the discrepancy between the mapped and target surface. More details for each term will be provided in Sec. 2.8.1 and Sec. 2.8.2.
the cost function shown in Eq. (2.22) ensures that the deformation matches the target shape and that the energy of the deformation as measured in the Hilbert space $V$ is also minimal. The parameter $\gamma$ appearing in the cost function is a weight parameter to the energy term:

$$
\int_0^1 ||v(t)||_V^2 dt
$$

(2.23)

Higher values for $\gamma$ will cause the minimization algorithm on the cost function $J$ to penalize the energy term Eq. (2.23) more heavily. Further, explanation on this term will be provided later in this section. The energy term in the cost function is also known as the regularization term as it insures that the deformations are diffeomorphic. Fig. 2.25 highlights various components of the LDDMM cost function.

**LDDMM Flow For Discrete Meshes:** It can be shown (Camion and Younes, 2001; Younes, 2010) that the solution of Eq. (2.22) in the form of the time-dependent vector field, $v(t)$, can be expressed as the convolution of the momentum vectors, $\alpha_n(t)$ with the kernel $k_V$, with one momentum vector defined for each of the $N$ vertices in $X$:

$$
v(t) = \frac{dx(t)}{dt} = \sum_{n=1}^{N} k_V(x_n(t), x(t)) \alpha_n(t),
$$

(2.24)

Eq. (2.24) is a continuous-time differential equation; however, when solving differential equations numerically time is discretized. Also, an Euler scheme is assumed for solving the differential equation. An intuitive interpretation is provided now for the flow operation on discrete meshes with discrete time steps. In the Euler method, the diffeomorphic flow is characterized by a sequence of deformations, which are uniformly ordered in time with a single deformation occurring at each time step. Fig. 2.24 shows the evolution of Ear A to Ear B using the LDDMM flow operation with a vector field that has been computed specifically for mapping the two ears with one another. Further, Fig. 2.26 shows an example particle and its displacement in space over seven time steps. The displacement vectors, $v(t)$, are obtained by applying the convolution given in Eq. (2.24).
In practice, we use the centred Euler scheme for solving the Ordinary Differential Equation (ODE) in Eq. (2.24), which is more accurate than the basic Euler method (Atkinson et al., 2011). Empirically 11 time steps seem appropriate when mapping one ear shape to another ear shape. In order to provide an analogy for the geodesic path between shapes we can consider points on the surface of a sphere.
The surface of the sphere is a non-linear Riemannian space. Fig. 2.27 shows two points on a sphere denoted by green stars. The shortest path, known as the geodesic path, connecting these two points is coloured in black and lies along the great circle containing the two points. Other paths on the sphere can also be constructed that join these two points, but these paths are longer.

Figure 2.27: The above figure illustrates the concept of the geodesic path using the surface of a sphere. The surface of a sphere is non-linear Riemannian space the two points on a sphere are shown using green stars, the optimal geodesic path between the points is shown in black.
2.8.1 Induced distances between shapes in the LDDMM framework

In this section, the cost-function used in the LDDMM analysis is described in detail. The first term in the cost function shown in Eq. (2.22) is a measure of the energy to deform the source shape $S_1$ to the target shape $S_2$:

$$d_{De f}(S_1, S_2) = \int_0^1 \|v(t)\|_V^2 dt$$ (2.25)

when the optimal momentum vectors are obtained by minimizing the cost function the energy term shown in Eq. (2.25) is also a measure of the geodesic distance between the source and target shapes. The term $\|v(t)\|_V^2$ is a normed squared value that can be expanded using the discrete flow equation Eq. (2.24):

$$\|v(t)\|_V^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \alpha_i(0), k_V(x_i(t), x_j(t)) \alpha_j(t) \rangle$$ (2.26)

There is a wide range of kernels $k_V$ that can be used for the purposes of LDDMM mappings. The kernels that are selected in LDDMM mappings have to adhere to a set of mathematical conditions. Details on the selection of the deformation kernel can be seen in Younes (2010). Radially decaying kernels such as the Cauchy or Gauss kernels are widely used for LDDMM surface mappings and are also used for mapping shapes in this thesis. The Cauchy kernel is defined as:

$$k_V(x, y) = \frac{1}{1 + \frac{\|x-y\|^2}{\sigma_V^2}},$$ (2.27)

where the $\sigma_V$ parameter is the deformation scale parameter that determines through the kernel, $k_V$, the range of influence of the momentum vectors. The deformation scale parameter $\sigma_V$ is an important parameter when mapping shapes using LDDMM.
and in Sec. 4.4 we provide more details as to how the value of $\sigma_V$ is set when mapping ear shapes. Fig. 2.28 shows the gain of the Cauchy kernel as a function of the distances between the points $x$ and $y$ in millimetres.

![Figure 2.28: The gain of the Cauchy kernel is plotted versus the absolute distances between points $x$ and $y$. Each colour in the plot shows the gain for a different $\sigma_V$ value.](image)

### 2.8.1.1 The Role of $\sigma_V$ Parameter

The $\sigma_V$ parameter defines the coupling between the vertices of the source shape as they move along the deformation path from the source shape to the target shape. This greatly impacts the quality and characteristic of the mapping between the two shapes. Consider the discrete mesh-flow equation shown in Eq. (2.24) and a point on the source mesh denoted by $x(t)$. As the $\sigma_V$ parameter is made larger, the momentum vectors surrounding this point will have a greater impact on the movement at $x(t)$, and consequently the vertices will move more coherently along the deformation path. The energy for deforming a shape in the LDDMM framework is greatly influenced by the $\sigma_V$ parameter.

### 2.8.2 Measuring shape differences using currents

The second term of the cost function, $E(S_1(\phi^X(1, X)), S_2(Y))$, measures the difference in the surface geometry of $S_1(\phi(1, X))$ and $S_2(Y)$ and is calculated based on the theory of currents (Vaillant and Glaunès, 2005). Currents can represent surfaces...
and are linear functionals on the space of differential 2-forms. The intuition behind using currents to represent surfaces is that differential 2-forms can be integrated over a surface to give a real value, and when two surfaces have similar geometry the difference in the value of the surface integrals should be close to zero. We use the notation $[S]$ to denote the current representing the surface $S$. $E(S_1, S_2)$ is defined as $E(S_1, S_2) = ||[S_1] - [S_2]||^2_W$, and is the squared norm of the difference of the two currents for a dual norm (i.e. $||.||_W^*$) corresponding to a Hilbert space $W$ of differential forms. The dual norm is a norm that is defined over the space of currents (Vaillant and Glaunès, 2005).

In the discrete setting, a surface, $S$, is approximated by a triangular mesh in $\mathbb{R}^3$. Given a face $f$ of $S$, let $c_S(f)$ denote the centre of the face and $n_S(f)$ denote the normal vector to the face with a length equal to the area of the face. We can then express $E(S_1, S_2)$ using the mesh elements as:

$$E(S_1, S_2) = \sum_{f,g} \langle n_{S_1}(f), k_W(c_{S_1}(g), c_{S_1}(f)) n_{S_1}(g) \rangle - 2 \sum_{f,q} \langle n_{S_2}(f), k_W(c_{S_2}q, c_{S_1}(f)) n_{S_1}(q) \rangle + \sum_{p,q} \langle n_{S_2}(p), k_W(c_{S_2}(p), c_{S_2}(q)) n_{S_2}(q) \rangle,$$

where in the above $\langle ., . \rangle$ represents a vector dot product and the kernel $k_W$ is typically chosen as the Gauss or Cauchy kernel. In this work, we use the Cauchy kernel for measuring shape mismatches:

$$k_W(x, y) = \frac{1}{1 + \frac{||x - y||^2}{\sigma_W^2}},$$

The shape comparison scale parameter $\sigma_W$ determines the physical scale at which the shapes are compared. Larger values for $\sigma_W$ result in a comparison of shapes at a coarse level of detail and small values of $\sigma_W$ result in a comparison of shapes.
at a fine detail. The effect of the $\sigma_W$ value on the quality and characteristic of the matching is explained and shown in Sec. 4.4.2. Please note that because the shape difference measure $E$ is a metric it can also be defined using the notion of the inner or scalar products:

$$E(S_1, S_2) = \| [S_1] - [S_2] \|^2_{W^*} = \langle [S_1] - [S_2], [S_1] - [S_2] \rangle_{W^*}$$

$$= \| [S_1] \|^2_{W^*} + \| [S_2] \|^2_{W^*} - 2 \langle [S_1], [S_2] \rangle_{W^*}$$  \hspace{1cm} (2.30)

2.8.3 Geodesic Shooting

Once the initial momentum vectors, $\alpha_n(0)$ are computed by minimizing the cost function $J(S_1, S_2)$ the diffeomorphic mapping between $S_1$ to $S_2$ is entirely determined (Vaillant et al., 2004). This follows from the fundamental principle in the LDDMM framework known as the conservation of momentums, which was proved in a seminal work by (Miller et al., 2006). In other words, $S_2$ can be represented (modelled) as a deformation of $S_1$ through the diffeomorphic flow defined by the initial momentum vectors $\{\alpha_n(0)\}_{1 \leq n \leq N}$.

Geodesic shooting consists in using a set of initial momentum vectors, $\{\alpha_n(0)\}_{1 \leq n \leq N}$, to morph a shape $S_1$ into another shape, $S_3$. This is done by solving the shooting equations, which couple the momentum vectors to the vertex positions across time and are given by:

$$\frac{d\alpha_r(t)}{dt} = - \sum_{n=1}^{N} \langle \alpha_r(t), \alpha_n(t) \rangle \nabla_{x_r(t)}(k_V(x_r(t), x_n(t)))$$

$$\frac{dx_r(t)}{dt} = \sum_{n=1}^{N} k_V(x_n(t), x_r(t)) \alpha_n(t)$$  \hspace{1cm} (2.31)

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^3$ and $\nabla_{x_r(t)}(\cdot)$ denotes the gradient operator and $1 \leq r \leq N$. Note that the initial conditions for Eq. (2.31) are given by the initial positions of the vertices and the corresponding momentum vectors. Fig. 2.29 shows three ear shapes that have been generated using different $\sigma_V$ values. When generating the three ear shapes the same initial momentum vector located on the
source shape $S_1$ was used.

Figure 2.29: The above figure shows three ear shapes that are produced by the use of the shooting and flow equations using a single non-zero initial momentum vector located on the source shape $S_1$. The $\sigma_V$ parameters for generating the shapes ranged from 2.5 to 25. When a small $\sigma_V$ is used the deformations are very local resulting in abnormal ear shapes having sharp features.
Chapter 3

REVIEW OF LITERATURE

The previous chapter provided information on how to obtain the acoustic properties of the head and ear shape using the Boundary Element Method (BEM) and the Fast Multipole BEM (FM-BEM). Both BEM and FM-BEM have shown to provide accurate acoustic responses for the head and ear shapes (refer to previous chapter). However, despite the availability of FM-BEM the time taken to perform simulations can be very long when the triangulated surface contains a large number of elements.

The literature review that is presented in this section is related to the morphoacoustic study of the head and ear shapes. The ultimate goal for the morphoacoustic study of the head and ear shapes is to obtain a rapid mapping between the morphology of the ear and head to the acoustic responses.

In this chapter a review of the literature related to parametrizing the shapes using two techniques will be presented. In Sec. 3.1 the use of the spherical harmonics for the modelling and parametrizing of shapes will be described and in Sec. 3.2 the elliptical Fourier transform technique for the study of the head and ear shapes will be reviewed. Both these techniques are used in the acoustic analysis of the head and ear morphology.

Also, in this chapter, the modelling of the ear shape using simple geometrical objects (Sec. 3.3), such as a cylindrical cavity or a parabolic metallic surface, will be presented. The work conducted on these simple models of the human pinna
shape are informative despite their simplicity because they can still explain some of the important features seen in the acoustic spectrum of the pinna shape such as the formation of the peaks and the primary spectral notches that are seen in the acoustic spectrum.

The other technique that is described in this chapter is known as morphoacoustic perturbation analysis (Sec. 3.4). When conducting perturbation analysis a small section of the template/reference ear shape is perturbed and changes in the acoustics of the perturbed shape are observed with respect to the reference or template shape. By the use of such studies regions of the ear shape have been identified that contribute most towards the formations of peaks and notches in the spectrum.

Another popular approach that has been investigated for individualizing the HRTFs is known as the frequency scaling approach (Sec. 3.5.1). Frequency scaling uses simple measurements of the head and pinna shape to scale the HRTF spectrum. It is shown that this technique can improve the similarity seen in the HRTF spectrum’s. Another method, based on simple anthropometric measurements, is known as database matching Sec. 3.5.2. In this method a HRTF is chosen for an individual based on a matching criterion between his anthropometric measurements and the anthropometric measurements for individuals in the database.

### 3.1 Shape Parametrization: Spherical Harmonics

The first technique for modelling three dimensional surfaces that will be discussed here is the application of spherical harmonics to parametrizing shapes (Kazhdan et al., 2003; Muller, 1966; Saupe and Vranic, 2001). Spherical harmonics are used to parametrize as well as perform orthogonal transformations to a template head shape in morphoacoustic studies (Lao et al., 2003a,b).

In order to use spherical harmonics to parametrize shapes, the shape needs to be defined using the spherical coordinate system as shown in Fig. 3.1. The function \( r(\theta, \phi) \) that maps the angles \( \theta \) and \( \phi \) to the surface of the shape needs to be one-to-one.
3.1. Shape Parametrization: Spherical Harmonics

For surfaces where the above conditions are satisfied we can parametrize the surface using the angles $\theta$ and $\phi$ and the function $r$ as follows (Tao et al., 2003b):

$$ r(\theta, \phi) = \sum_{n=0}^{\infty} a_n P_n(\cos \theta) + \sum_{n=0}^{\infty} \sum_{m=1}^{n} P^m_n(\cos \theta) \times [a_{nm}\cos(m\phi) + b_{nm}\sin(m\phi)] $$

(3.1)

Where in the above $P^m_n$ are Legendre polynomials of degree $n$ and order $m$ and further $0 \leq m \leq n$. In Eq. (3.1) we are modelling the shape using Legendre polynomials of all degrees which is neither efficient or necessary. A low pass filtered shape can be obtained by truncating the higher degree Legendre polynomials. Eq. (3.1) can be written using the truncation order $N$ as follows:

$$ r(\theta, \phi) = a_{00} + \sum_{n=0}^{N} \sum_{m=1}^{n} P^m_n(\cos \theta) \times [a_{nm}\cos(m\phi) + b_{nm}\sin(m\phi)] $$

(3.2)

Where in Eq. (3.2) Legendre polynomials up to the $N^{th}$ degree are being used. The coefficients $a_{nm}$ and $b_{nm}$ are spherical harmonic coefficients which are calculated using the following equations:

$$ a_{nm} = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} r(\theta, \phi) Y^1_n(\theta, \phi) \sin(\theta) d\theta d\phi $$

$$ b_{nm} = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} r(\theta, \phi) Y^0_n(\theta, \phi) \sin(\theta) d\theta d\phi $$

(3.3)
the function \( \tilde{Y}^1 \) and \( \tilde{Y}^0 \) are the real spherical harmonics representations and can be written as follows:

\[
\tilde{Y}^1 = \frac{1}{N_{nm}} \cos(m\phi)P_n^m(\cos\theta) \\
\tilde{Y}^0 = \frac{1}{N_{nm}} \sin(m\phi)P_n^m(\cos\theta)
\] (3.4)

in the above \( N_{nm} \) is a normalization factor for the real spherical harmonics and is given by [Blanco et al., 1997]:

\[
N_{nm} = \sqrt{\frac{4\pi}{\epsilon_m 2n + 1}} \frac{(n + m)!}{(n - m)!} \\
\epsilon_m = \begin{cases} 
1 & \text{if } m = 0, \\
2 & \text{if } m \neq 0,
\end{cases}
\] (3.5)

The value of \( N \), which is also called the truncation order, can be adjusted depending on the level of accuracy with which we require to model our surface. Fig. 3.2 shows a reconstructed KEMAR head shape using a truncation order of \( N=17 \). Once the coefficients \( a_{nm} \) and \( b_{nm} \) are calculated for a given shape we can then obtain deformations for the shape by perturbing these coefficients. In the context of acoustic study and morphoacoustic perturbation analysis (Sec. 3.4) there are restrictions for
3.1. Shape Parametrization: Spherical Harmonics

the magnitude of the perturbation that can be applied to these coefficients. In acoustic studies perturbations to the shape should result in linear change to the acoustic response.

Tao et al. (2003b) also performed an analysis in which the KEMAR head shapes were simplified (low pass filtered) using \( n = 0 \) to \( n = N = 34 \) degree Legendre polynomials in the spherical harmonic expansions. Both shape errors as well as errors for the corresponding pressure field around the head were examined. The error for the generated pressure field will be shown in Sec. 3.4.1. The Root Mean Square shape error (RMS) between the reference shape and the shape reconstructed with degree \( n \) Legendre polynomials is given by:

\[
\epsilon_{r_{\text{rms}}} (n) = \sqrt{\frac{\sum_{i=1}^{2N^2} [r_n(i) - r_N(i)]^2}{2N^2}}
\]  

(3.6)

In Eq. (3.6) the domain of the function \( r_n(\theta, \phi) \) for representing shapes is composed of a grid of \( N \) distinct points for the vertical angles \( \phi \) and \( 2N \) points for horizontal angles \( \theta \) with a total of \( 2N^2 \) mesh grid of points for representing the full shape. Fig. 3.3 shows the percentage error for the reconstruction of the head shape using various values of \( n \). It can be observed that except for the nose regions, which require higher degree Legendre polynomials for accurate reconstruction the rest of the head shape can be reconstructed accurately using \( N = 15 \) degree Legendre polynomials.

A big disadvantage in using spherical harmonics for parametrizing the head shapes is that the pinna cannot be modelled using Eq. (3.2) as the surface of the head and ear shape cannot be defined using a one to one function.
3.2 Shape Parametrization: Elliptical Fourier Transform

The second technique which has been used for parametrizing and also deforming 3D surfaces is the Elliptical Fourier Transform (EFT) (Hetherington et al., 2003; Hetherington and Tew, 2003; Park and Lee, 1987).

EFT was initially used for parametrizing 2D closed contours. However, with the inclusion of additional steps, it can also be used for the study of the human pinna and head shapes too. One of the advantages in using the EFT technique over spherical harmonics for parametrizing the head and pinna shapes is that it does not suffer from the problems faced by the spherical harmonics in parametrizing the pinna region of the head shapes.

The EFT method described here is an adaptation of the EFT method proposed in Park and Lee (1987). In order to apply EFT to the parametrization of the head and pinna shapes, first the head and pinna shapes have to be aligned with each other such that the y-axis passes through the ear canal. Then the pinna (and possibly the head shapes) are rotated for multiples of the angles $2\pi / S$ and then intersections of the pinna shape are taken with respect to the $xy$ plane. In other words, the pinna shape is rotated by angles $\theta$ where $\theta$ takes the values $0, \alpha, 2\alpha, \ldots, 2\pi - \alpha$. Fig. 3.4(a)
shows an intersecting plane with the KEMAR head and pinna shape, and Fig. 3.4(b) displays the cross section produced. Once the slices are obtained they are made regular by means of a linear interpolation function and are then parametrized using a parameter $t$. The parameter $t$ takes values $t = 0, 1, 2, 3, 4, \ldots T$. Using the parameter $t$ two functions $f^x_s[t]$ and $f^y_s[t]$ are constructed for each of the $x$ and $y$ components. The EFT is then computed on the $f^x_s[t]$ and $f^y_s[t]$ functions separately using two sequential Fourier transforms shown in Eq. (3.7)–Eq. (3.8)

$$A_x[s, n] = \sum_{t=0}^{T-1} f^x_s[t] e^{-jnt}$$

(3.7)

$$B_x[m, n] = \sum_{s=0}^{S-1} A_x[s, n] e^{-jms}$$

(3.8)

The result of the 2D Fourier analysis is a set of complex coefficients that model the pinna and head shape. By perturbing any of the coefficients $A_x[s, n]$ or $B_x[m, n]$ obtained by using EFT we can apply deformations to the head and pinna shapes [Hetherington and Tew, 2003]. However, this technique is shown to be problematic because smooth and evenly distributed spatial deformations cannot be generated as in the case of spherical harmonics. In particular, making deformations to the head and pinna shape using the EFT result in non-linear changes in the acoustics which is
not desirable for the purposes of morphoacoustic analysis.

A further adaptation was proposed to the EFT technique in [TEW et al. 2012] in which deformations were applied perpendicular to the contour of a slice (Fig. 3.4(b)). In the analysis performed by [TEW et al. 2012] coefficients $A_{u,v}$ and $B_{u,v}$ are generated which are termed surface harmonic amplitudes. Perturbing these set of coefficients results in deformations to the shape. Fig. 3.5 shows the deformations obtained by perturbing coefficients $A_{u,v}$ and $B_{u,v}$ for two values, $u$ and $v$.

![Figure 3.5: The above figure shows the effect of perturbing the surface harmonic amplitudes which is detailed in Tew et al. 2012 for a range of $u$ and $v$. $u$ is cross harmonic and $v$ is the slice harmonic. Image taken from Thorpe(2009).](image)

### 3.3 Modelling The Acoustics Of The Human Morphology Using Simple Geometrical Objects

One of the earliest models for modelling an individuals acoustic responses was provided by Lord Rayleigh [Kuhn 1977; Struβ 2011]. In the early model the head
and pinna effects were modelled using a sphere. This model provided a simple yet adequate explanation for important binaural cues such as the ITD and ILD.

Algazi et al. (2002) extended the spherical head model to a snowman model. The snowman model consisted of a spherical head on top of a spherical torso. This is shown in Fig. 3.6. Using the snowman model, Algazi et al. (2002) showed that important elevation cues contributed by the torso can be modelled using a spherical torso shape. He showed that the spherical head and spherical torso models provide good approximation to the HRTF’s of a pinna less KEMAR mannequin. Fig. 3.7 shows the computed HRTFs on the snowman model.

Teranishi and Shaw (1968) tried modelling the acoustic responses of an ear shape by first using a cylindrical cavity of radius R and depth L, which was representative of the Concha shape, and then adding a simple rectangular flang to the inclined cylindrical Concha. This enhancement improved the directivity of the acoustic spectrum compared to just a cylindrical Concha. Fig. 3.8 shows the increased response of a cylindrical Concha with the added rectangular flang compared to just a Cylindrical Concha. Fig. 3.10 shows a comparison of modes and resonances between the average responses of real ears to that of replicated ear shapes using
3.3. Modelling The Acoustics Of The Human Morphology Using Simple Geometrical Objects

Figure 3.7: HRTF spectrum computed for frontal elevation angles $\delta$ and for frequencies in the range of 0-5kHz. The arch shaped notches are symmetric about the elevation angle $\delta = 90^\circ$ and are due to the reflections from the torso region. Image taken from Algazi, et al.(2002)

various geometrical objects, including a cylindrical Concha and with the addition of a rectangular flang. Teranishi and Shaw (1968) found that the responses and directivity of the modelled and real ear shapes looked very similar for frequencies up to 7kHz. More recently, Lopez Poveda and Meddis (1996) modelled the acoustic properties of the Concha cavity using a parabolic metallic surface and showed that important features seen in the spectrum of real ears, such as the first and third notches, can be modelled adequately using a formulation that included reflections, diffractions, and interferences applied to a parabolic metallic sheet. Fig. 3.9 shows the HRTFs obtained from the KEMAR mannequin using the DB61 pinna shape and the HRTFs calculated using the diffraction and reflection model in Lopez Poveda and Meddis (1996).
3.3. Modelling The Acoustics Of The Human Morphology Using Simple Geometrical Objects

Figure 3.8: Increase in response when a rectangular flang is added to cylindrical Concha for various angles of incidence $\theta$. Plot (A) is for sources originating in front and plot (B) is for sources originating at the back. Image taken from Teranishi, et al.(1967)

Figure 3.9: The above is a reprint from the paper Lopez-Poveda et al.1996. It shows the frequency response of the KEMAR ear with the DB61 pinna attached (top plot) and the modelled pinna using the diffraction and reflection model. It can be seen that the first notch N1 and the third notch N3 is well modelled using the diffraction and reflection model.
3.3. Modelling The Acoustics Of The Human Morphology Using Simple Geometrical Objects

Figure 3.10: The above shows (A) the average frequency response of the real ear shapes, averaged over six subjects. (B) cylindrical Concha, (C) tilted cylindrical Concha (D) cylindrical Concha with tilted segmented pinna (E) tilted cylindrical Concha with rectangular flang. It can be observed that adding the flang adds to the directivity of the cylindrical Concha. Reprint from Teranishi et al.1968.
3.4 Morphoacoustic Perturbation Analysis

Morphoacoustic perturbation analysis (MPA) was a term that was first used in a paper by TEW et al. (2012) and it is a method by which changes in the morphology and acoustic changes with respect to a base template shape are examined. The perturbations applied to the template shape are usually small such that the effect of changing the morphology to the change in the acoustic response is linear. Many researchers have used MPA based techniques to link features in the HRTF spectrum to regions in the ear shape (Mokhtari et al., 2010; Takemoto et al., 2012b; Tao et al., 2003a,b). These features can be in the form of peaks and notches. The base template shape used in these studies is typically the KEMAR mannequin (Burkhard and Sachs, 1975).

3.4.1 Differential Pressure Synthesis (DPS)

The discussion here follows from the discussion on spherical harmonics for the modelling of head shapes (Sec. 3.1). Differential Pressure Synthesis (DPS) (Tao et al., 2003a,b) was designed with the goal of rapidly computing the acoustic pressure field around a deformed template shape. By using DPS costly and time consuming numerical computations of the pressure field by means of BEM would not be required. In DPS the pressure field around the deformed template shape is calculated by means of scaling and summing the pressures stored in a precomputed database of orthogonal transformations of the template shape. The orthogonal deformations from the template shape were obtained using spherical harmonics.

Constraints exist on the size of deformations that can be applied to the template ear shape which as this is dependant on the wavelength of the incident sound wave. Given this constraint a first order relationship between shape changes and a change in the pressure field is formulated. Let's denote the pressure field around the template shape as $\Phi_0$, and when the template shape is deformed the new pressure field $\Phi$ for
3.4. Morphoacoustic Perturbation Analysis

<table>
<thead>
<tr>
<th>Frequency</th>
<th>250 Hz</th>
<th>500 Hz</th>
<th>1 kHz</th>
<th>2 kHz</th>
<th>3 kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Error</td>
<td>0.92</td>
<td>0.22</td>
<td>0.69</td>
<td>18.9</td>
<td>34.6</td>
</tr>
</tbody>
</table>

Table 3.1: Percentage error for the pressure field of a deformed spherical template shape to the pinna less KEMAR mannequin, when calculated using DPS versus direct BEM. (Taken from Tao et. al. 2003)

The deformed template shape can be written in simple mathematics as:

$$\Phi = \Phi_0 + \Delta \Phi$$  \hspace{1cm} (3.9)

In brief the method for calculating the $\Delta \Phi$ is obtained by a first order Taylor expansion that relates spherical harmonic coefficients representing the shape changes to the pressure changes (Tao et al., 2003a). Tao et al. (2003a) conducted an experiment in which a spherical template shape was deformed to a pinna-less KEMAR head. Table 3.1 tabulates the percentage error for the pressure field calculated for the deformed spherical template shape when using DPS versus direct BEM simulations. Tao et al. (2003a) reported that the large errors seen at higher frequencies, that is, 2kHz and 3kHz were due to the comparable sizes of the sound wavelength with respect to the deformation size from the spherical template shape.

In summary, the DPS can be used to compute the pressure field around deformed objects that can be modelled using orthogonal transformations such as spherical harmonics. Importantly, DPS requires the deformations from the template ear shape to be moderate with respect to the sound wavelength under consideration. This is because large errors in the computed pressure field are observed if this criterion is not met.

3.4.2 Morphoacoustic Perturbation Analysis Frequency Domain (MPA-FD)

As previously discussed, the DPS technique is aimed at rapidly computing the BEM pressure around a deformed template shape (refer Sec. 3.4.1). However because deformations from the template shape need to be linear with respect to the acoustical
changes this method cannot be used to study large variations in the ear shape.

The morphoacoustic perturbation analysis frequency domain (MPA-FD) (TEW et al., 2012) is a method that can be used to analyse the changes in the acoustics when a small perturbation is applied to a template shape. On the other hand if a change is seen in a particular feature of the HRTF, MPA-FD can identify morphological locations that contribute most towards the change (i.e. in the HRTF spectrum). The MPA-FD avoids one of the limitations inherent in the DPS technique proposed by Tao et al. (2003a) as it can model the pinna regions of the head and ear shapes.

The MPA-FD can rapidly identify interesting relationships between the acoustic features and morphological features by constructing a database of orthogonal shape changes and the corresponding acoustic changes. The orthogonal shape changes in MPA-FD are conducted using elliptical Fourier transforms. This was described in more detail in Sec. 3.2. The acoustic responses for the deformed template shapes are computed using BEM simulations. Fig. 3.11 shows how MPA-FD was used to identify regions in the template ear shape that contribute most towards the notch seen in the HRTF spectrum.

In summary, MPA-FD is very useful for identifying and relating regions in the ear shape that influence particular HRTF features and vise versa. However, similar to the DPS technique, one of the limitations of the MPA-FD is that the formulation requires a linear relationship between the morphological changes and the acoustic changes. For this reason the database of shape changes is constructed by applying only small perturbations. Consequently, this will limit any morphoacoustic study on ear shapes because the morphology of the ear shapes change substantially across the population of humans.

### 3.4.3 Acoustic sensitivity to micro-perturbations of KEMAR’s pinna surface geometry

The work in Mokhtari et al. (2010) examined the acoustic sensitivity of different regions of the ear shape to small perturbations of a voxelated ear and small head
patch region belonging to the KEMAR head. FDTD was used in the simulations to produce the HRTF data for the base and perturbed shapes. The voxel grid had a uniform resolution of 2 mm. Consequently a total of 1784 unique voxels was used for the perturbation analysis thereby resulting in 1785 FDTD simulations. Peaks and notches in the resulting PRTF’s were then examined with respect to the base PRTF belonging to the KEMAR head. Fig. 3.12 shows the identification of a series of peaks and notch patterns for a couple of PRTFs. Some of the key points relating to Mokhtari et al. (2010) research findings include:

1. The first peak, P1, with a centre frequency of around 5kHz was mainly effected by perturbations in the Concha and Cymba cavities. Other peaks labelled P1-b and P1-c were also identified for particular direction in space. The P1-b sensitivity shows a shift from the Concha base towards its anterior.

2. The third peak, P3, has a centre frequency of around 9.8kHz. This peak ap-
3.4. Morphoacoustic Perturbation Analysis

Figure 3.12: The above figures shows the labelling of the peak and notch patterns for a series of PRTFs. Image taken from Mokhtari et al. 2010

peared consistently across all direction in space. They reported that different regions in the Cymba, Concha and triangular Fossa have positive and negative sensitivity towards P3 (i.e they shift the peak P3 to a lower or higher frequency value).

3. The notch labelled, N1, was divided into three notches: N1-a, N1-b and N1-c. Each of these notches appeared for different directions in space. The N1-a notch appeared for sources from the lower front regions and had a centre frequency of 7.1kHz. It had positive sensitivity to a region that covered the Cymba and triangular Fossa. The N1-b notch appeared for sources in the front hemisphere near the horizontal plane and it had sensitivity to regions corresponding to the Cymba and the upper back wall of the Concha. The N1-c occurred for twelve locations at high elevations and had a mean frequency of 8.8kHz. The centre frequency for the N1-c notch appeared to be affected by a large area of the pinna.

Fig. 3.13 - Fig. 3.14 show the regions of the ear shape that had positive and negative sensitivity to the peaks and notches in the PRTF spectrum.
Figure 3.13: The above figure shows regions of positive sensitivity (warm colours) and negative sensitivity (cold colours) in the formations of peaks P1-P4. Image taken from Mokhtari et al. 2010.

Figure 3.14: The above figure shows regions of positive sensitivity (warm colours) and negative sensitivity (cold colours) in the formation of the notches N1-N3. Image reprint from Mokhtari et al. 2010.
3.5 HRTF personalization using anthropometric measurements of the head and ear shape

3.5.1 Frequency Scaling

One of the intuitive yet simple schemes used for personalizing the HRTF from anthropometric measurements was proposed in Middlebrooks (1999a,b); Middlebrooks et al. (2000). In the study conducted by Middlebrooks the HRTF spectrum was scaled proportional to the ratio of the anthropometric measurements between the head and pinna shapes from the two individuals. The frequency scaling procedure proposed by Middlebrooks stemmed from two observations:

- Using the analytical solution for the wave scattering problem on a spherical head model, it was observed that the HRTF for a smaller or bigger spherical head can be obtained by appropriately scaling the HRTF of a spherical head with a given radius for which the HRTF is already known (Xie, 2013).

- Given that many features, such as the appearance of notches and peaks, are common across the HRTF spectrum for humans, it was observed that these notches and peaks occur in varying locations in the frequency spectrum of the HRTF, which in turn could be related to the diameter of the head.

Middlebrooks (1999b) studied the relation between head and pinna anthropometric measurements and the optimal scale factor for use in scaling the HRTFs between the two individuals so that the identified features in the HRTF spectrum can match each other more closely.

The dissimilarity between pairs of HRTFs was quantified using the inter-subject spectral differences (ISSD). The ISSD in Middlebrooks (1999b) is calculated in the frequency range of 3.7kHz to 12.9kHz using 64 ERB filterbanks. The procedure for computing the ISSD between two DTFs corresponding to two subjects, \( S_1 \) and \( S_2 \), involves three steps. In the first step DTF differences are computed at each of the 64
bands generated by the ERB filterbanks:

\[ \Delta DTF(\theta, \phi, f_i) = 20\log_{10}|DTF_{S_1}(\theta, \phi, f_i)| - 20\log_{10}|DTF_{S_2}(\theta, \phi, f_i)| \] (3.10)

Using the \( \Delta DTF(\theta, \phi, f_i) \) the mean for all the 64 frequency bands \( f_i \) are computed:

\[ \bar{\Delta DTF}(\theta, \phi, f_i) = \frac{1}{64} \sum_{i=1}^{64} \Delta DTF(\theta, \phi, f_i) \] (3.11)

In the second step the variance in \( \Delta DTF(\theta, \phi, f_i) \) for every direction in space \( \theta \) and \( \phi \) are computed:

\[ \sigma^2(\theta, \phi) = \frac{1}{64} \sum_{i=1}^{64} \| \Delta DTF(\theta, \phi, f_i) - \bar{\Delta DTF}(\theta, \phi, f_i) \|^2 \] (3.12)

Finally, the ISSD between subjects \( S_1 \) and \( S_2 \) is obtained in the third step:

\[ ISSD_{S_1, S_2} = \frac{1}{M} \sum_{\theta, \phi} \sigma^2(\theta, \phi) \] (3.13)

Middlebrooks (1999b) made the following observations from his study:

1. The range for the optimal scale factors was in the range of 1.061 and 1.38.

2. The optimal scale factors for the left and right ear shapes were highly correlated with a correlation value of 0.95.

3. Frequency scaling applied to the HRTFs halved the ISSD for 9.5% of individuals and was not effective for 20% of the population. It had a varying degree of effectiveness for the rest of the population. A histogram plot of the effect of the ISSD is shown in Fig. 3.15.

4. Given that the head widths \( hw_{S_1}, hw_{S_2} \) and pinna cavity heights \( ph_{S_1}, ph_{S_2} \) for two subjects \( S_1 \) and \( S_2 \) are known then the optimal scale factor can be
approximated using the following equation:

\[
\log_2 \gamma_{opt} = 0.34 \log_2 \left( \frac{ph_{S1}}{ph_{S2}} \right) + 0.527 \log_2 \left( \frac{hw_{S1}}{hw_{S2}} \right)
\]  

(3.14)

5. The scale factor for the left and right ear shapes are very highly correlated.

6. After frequency scaling an average of 6.18\text{dB}$ of error between the HRTFs was observed, signifying the need for a more sophisticated model for personalizing the HRTFs.

7. The ratio in the ITD between a pair of subjects was significantly correlated to the optimal frequency scale factor between the subjects.

The optimal scale factors obtained by analysing 990 pair of subjects and the reduction in the ISSD after applying the frequency scaling is displayed in Fig. 3.15.

![Figure 3.15](image.png)

Figure 3.15: Image taken from Middlebrooks(1999), showing the optimal scale factor and the percentage reduction in the ISSD after applying the scale factor to the HRTF.

### 3.5.2 Database Matching

Another technique that is used for personalizing the HRTFs is known as Database Matching [Zotkin et al., 2004, 2003]. Database matching involves personalizing the HRTF’s for a given individual by finding the HRTF for a subject from a database of
HRTF and anthropometric measurements whose anthropometric measurements are the closest to the given subject.

In order to provide a personalized HRTF for a given subject anthropometric measurements are taken from the pinna shape of the individual. Subsequently, by using a similarity function, the anthropometric measurement of the individual is compared to all pinna shapes in a database. The personalized HRTF for the individual is then chosen to be the HRTF of the individual with the largest similarity in the anthropometric measurements given the function used.

In Zotkin et al. (2004, 2003) seven anthropometric measurements were used $d_1, \ldots, d_7$. They are respectively the Cavum Concha height, Cymba Concha height, Cavium Concha width, Triangular-Fossa height, pinna height, pinna width, Intertragal incisure width (Fig. 2.1). The function used for finding the similarity between the anthropometric measurements of the given individual with individual $k$ in the database is shown below:

$$E^k = \sum_{i} d_i - d_i^k \sigma_i^2$$  \hspace{1cm} (3.15)

in the above, $\sigma_i$ is the standard deviation for anthropometric measurement $i$ across all subjects in the database. Because the anthropometric measurements $d_i$ are for the pinna region, a blending regime is used such that at low frequencies the HAT model proposed in Duda et al. (2002) is utilized at high frequencies the best matched HRTF is used.

The HAT model is used for modelling the HRTFs at low frequencies where the shadowing and or reflections from the head and torso play an important role in saggittal plane localization. The HAT model uses the "snowman model" for the head and torso shapes. Denoting the HAT model HRTF as $H_h$ and the HRTF selected from the database as $H_c$, then the output HRTF $H_o$ has the same phase $H_h$; however, the
magnitude of the output HRTF $H_o$ is given by the following blending operation:

$$
A_o(w) = \begin{cases} 
A_h(w) & w < w_l \\
A_h(w) + \frac{A_c(w) - A_h(w)}{w_h - w_l} & w_l < w < w_h \\
A_c, & w > w_h
\end{cases}
$$

(3.16)

$$
A_o = \log |H_o(w)|, \quad A_h(w) = \log |H_h(w)|, \\
A_c(w) = \log |H_c(w)|, \quad w_l = 500 \text{Hz}, w_h = 3000 \text{Hz}.
$$

### 3.5.3 Regression Analysis

Another popular method that has been suggested for personalizing the HRTFs is performing regression analysis between the HRTF parameters and anthropometric parameters for the head, ear and torso shapes (Inoue et al., 2005; Jin et al., 2000; Nishino et al., 2007). (Inoue et al., 2005).

Jin et al. (2000) conducted a regression analysis between the PCA coefficients of HRTF data and anthropometric data for 36 subjects and showed that regression models can be used for the purposes of predicting HRTF data from anthropometric data. Jin et al. (2000) also performed localization tests to examine how many HRTF PCA coefficients were required for localizing sound to a given level of accuracy. It was shown that with seven PCA coefficients the localization of sound was accurate for the examined subjects.
Chapter 4

MORPHOACOUSTIC ANALYSIS FRAMEWORK

This chapter details the framework for performing morphoacoustic analysis on ear shapes using LDDMM and FM-BEM. Concepts and procedures developed here are used extensively in this chapter and also in later chapters. LDDMM and FM-BEM are intricate tools and a detailed explanation for the terminology and methodology to use them is provided. More specifically, the framework provided in this chapter describes: (1) the procedure to match one set of head and ear surface meshes to another set using the LDDMM framework (2) how to measure and compare difference between two ear surfaces (3) the procedure for conducting FM-BEM simulations on high-resolution surface meshes of head and ear shape to obtain head-related impulse response functions and (4) a procedure to measure the differences between the acoustic responses of two head and ear surfaces and (5) how to robustly deform and match ear shapes using the LDDMM framework.

Contributions: Five main contributions are made in this chapter as follows: (1) A procedure is described which maps one head and ear mesh to another head and ear mesh using the LDDMM framework. The mapping is performed in such a way that the head shapes are matched at relatively low spatial resolution and the ear shapes are matched at much higher spatial resolution. The benefit of such a multi-
resolution approach to the mapping is that it significantly reduces the computation time required for matching the head shapes. There is, of course, an underlying assumption made in this approach that the relatively low resolution matching of the head shapes does not significantly influence the acoustics of the ears, i.e., the corresponding HRIR filters. That this assumption is valid will be demonstrated later in Sec. 6.1. To facilitate a precise description and discussion of the LDDMM matching process, a concise set of mathematical functions are defined that completely characterize the LDDMM matching process. (2) A new technique is described by which to compare two head and ear surfaces with one another. This technique is termed Current Shape Difference Analysis (CSDA) as it uses the framework of currents for comparing shapes. It may be recalled that the framework of currents was used for mapping two surfaces using LDDMM in Sec. 2.8.2. The CSDA technique for comparing the head and ear surfaces is shown to be more effective when comparing the ears and associated acoustics when compared to euclidean based measures. (3) The optimal ranges for several scale and regularization parameters required for the LDDMM matching procedure are studied and described. These parameters have to be set appropriately in order to obtain the best results. (4) Given a head and ear surface mesh that has been obtained from the LDDMM matching procedure, a semi-automated procedure is described for removing any irregularities in the surface mesh and preparing it for FM-BEM simulation. Importantly, the resolution of the surface mesh, i.e. the maximum edge length of any triangular element in the surface mesh, should be less than a specific value in order to achieve FM-BEM simulations that are valid up to a specific acoustic frequency. (5) The method for performing the FM-BEM simulations using the commercial software Coustyx (Ansol, Ansol) is described. (6) A comprehensive/global method for comparing the acoustics related to two surface meshes in the frequency domain as well as spatially is described. The global acoustic analysis compares two sets of DTF (i.e. Directional Transfer Functions) and a result for this comparison is a figure which is termed GAAF. This will be explained in detail within this chapter. (6) A new procedure is described to robustly
4.1. LDDMM Method For Mapping Shapes

This section describes the methods for mapping surfaces using LDDMM. More specifically, three set of mathematical functions are defined in order to express the complicated procedures involved when analysing shapes using LDDMM. The three functions are (1) mapping, (2) flow, and (3) geodesic shooting. These functions are used
4.1. LDDMM Method For Mapping Shapes

Figure 4.1: The above figure lists steps to diffeomorphically map a source head and ear shape to a target head and ear shape at LDDMM scales of $\sigma^H_V$, $\sigma^H_W$, $\sigma^V_H$, and $\sigma^W_H$.

extensively within this and the following chapters to describe various algorithms for diffeomorphically mapping ear shapes. Fig. 4.1 is a chart that illustrates the inputs and outputs to some of the functions explained in this section. The figure also lists steps in mapping two head and ear shapes with each other. We will refer to Fig. 4.1 throughout this section.

(1) The function $\mathcal{M}$ is the LDDMM mapping function and is used to obtain diffeomorphic maps between a source and target shape. The function $\mathcal{M}$ is defined as:

$$\{\alpha(t)\}_0^1: S_1 \to S_2 \leftarrow \mathcal{M}(S_1, S_2, \sigma^H_V, \sigma^H_W)$$ (4.1)

The use of the function $\mathcal{M}$ in mapping two shapes is shown in step (1) of the chart shown in Fig. 4.1. The function $\mathcal{M}$ has four inputs and a single output. It takes
in the source shape \( S_1 \) and the target shape \( S_2 \). The source and target shapes used here are as defined in Sec. 2.8. Further the third and fourth inputs to the function \( \mathcal{M} \) are the deformation scale parameter \( \sigma_V \) and the shape comparison parameter \( \sigma_W \).

Recall that the parameters \( \sigma_V \) and \( \sigma_W \) were also defined in Sec. 2.8.

The function \( \mathcal{M} \) signifies the process of minimizing the cost functional \( J \):

\[
J_{S_1,S_2}(v(t)) = \gamma \int_0^1 \| v(t) \|^2 V dt + E(S_1(\phi^V(1,X)), S_2(Y)), \tag{4.2}
\]

By minimizing Eq. (4.2) we obtain a set of time dependant initial momentum vectors \( \{ \alpha(t) \}_{S_1 \rightarrow S_2} \) which parametrize the deformation path between the source and target shape. The selection and assignment of the deformation parameter \( \sigma_V \) and shape comparison parameter \( \sigma_W \) are detailed Sec. 4.4.1 and Sec. 4.4.2 respectively.

There is a single momentum vector \( \alpha_i \in \mathbb{R}^3 \) (see Fig. 4.2) for every vertex in the source shape \( S_1 \) at every time step \( t \) between \( 0 \leq t \leq 1 \). Fig. 4.2 shows a shape with the momentum vectors plotted at each vertex. For the purposes of clarity and simplicity the time interval \( 0 \leq t \leq 1 \) label on the set of momentum vectors will not be shown in future text unless it plays a significant role.

(2) \( \mathcal{F} \) is the flow function which can diffeomorphically deform one shape into another:

\[
\hat{S}_2 \leftarrow \mathcal{F}(\{ \alpha(t) \}_{S_1 \rightarrow S_2}, S_1, \sigma_V, t_s, t_e) \tag{4.3}
\]

It has five input parameters consisting of the time dependant momentum vectors \( \{ \alpha(t) \}_{S_1 \rightarrow S_2} \), the source shape \( S_1 \), the deformation scale parameter \( \sigma_V \) and two time parameters \( t_s \) and \( t_e \) signifying the start set of momentum vectors \( \{ \alpha(t_s) \}_{S_1 \rightarrow S_2} \) and end set of momentum vectors \( \{ \alpha(t_e) \}_{S_1 \rightarrow S_2} \). The default values used for the start and end time for the deformations are \( t_s = 0 \) and \( t_e = 1 \) for this reason the flow function \( \mathcal{F} \) will be written more compactly as:

\[
\hat{S}_2 \leftarrow \mathcal{F}(\{ \alpha(t) \}_{S_1 \rightarrow S_2}, S_1, \sigma_V) \tag{4.4}
\]
Figure 4.2: The above figure shows a source shape involved in a LDDMM mapping. The initial momentum vectors are plotted on surface mesh. The momentum vectors have been scaled by the same constant factor for clarity, however it can be observed that the momentum vectors have large and small magnitudes and further point at different directions in space indicating the direction and magnitude of displacement of the vertices for the source shape at the first time instance.
The function $F$ is shown in steps 2, 4 and 5 in the chart in Fig. 4.1. The flow function performs a deformation by solving the following differential equation:

\[
v(t) = \frac{dx(t)}{dt} = \sum_{n=1}^{N} k_V(x_n(t), x(t)) \alpha_n(t), \tag{4.5}\]

Eq. (4.5) was described in Sec. 2.8 and performs a diffeomorphic flow by integrating between the start time $t_s$ and end time $t_e$ using the time dependant momentum vectors (i.e. $\{\alpha(t)\}_{S_1 \rightarrow S_2}$) and the deformation scale parameters $\sigma_V$. Fig. 2.24 was used in Sec. 2.8 to show the flow of the ear shape and its evolution with the time parameter. We also define the reverse flow function denoted by $F^{-1}$:

\[
\hat{S}_1 \leftarrow F^{-1} = F(\{\alpha(t)\}_{S_1 \rightarrow S_2}, S_2, \sigma_V, t_s = 1, t_e = 0) \tag{4.6}\]

The function $F^{-1}$ applies the reverse diffeomorphic flow, deforming $S_2$ backwards in time starting with the momentum vectors at time $t_e$ and ending with the momentum vectors defined at time $t_s$.

(3) $S$ is the geodesic shooting function which calculates the set of momentum vectors for a given geodesic path. The function $S$ is shown below:

\[
\{\alpha(t)\}_{S_1 \rightarrow S_2} = S(\{\alpha(0)\}_{S_1 \rightarrow S_2}, S_1, \sigma_V) \tag{4.7}\]

It takes as input a set of initial momentum vectors denoted by $\{\alpha(0)\}$, the source shape $S_1$ and the deformation scale parameter $\sigma_V$ and outputs the time dependant momentum vectors $\{\alpha(t)\}_{S_1 \rightarrow S_2}$. The function $S$ signifies the solution to the coupled differential equations, known famously as the shooting equations shown:

\[
\frac{d\alpha_r(t)}{dt} = - \sum_{n=1}^{N} \langle \alpha_r(t), \alpha_n(t) \rangle \nabla_{x_r(t)} (k_V(x_r(t), x_n(t)))
\]

\[
\frac{dx_r(t)}{dt} = \sum_{n=1}^{N} k_V(x_n(t), x_r(t)) \alpha_n(t) \tag{4.8}\]

This function is not shown in the chart Fig. 4.1 however is used extensively in
4.2 Diffeomorphically Mapping the Head and Ears

This section details the methodology for diffeomorphically mapping a source head and ear shape to a target head and ear shape. The source and target head and ear shapes are shown in Fig. 4.3. A similar algorithm was detailed in the paper (Zolfaghari et al., 2014) in which head, ear and torso shapes (HTE) were mapped with each other. The algorithm described in this section is designed to overcome the following two constraints:

1. Mapping complete head and ear shapes (or head, torso and ear shapes) using fine meshes with more than hundred thousand elements is not practical when conducting morphoacoustic studies. Given the significant number of elements involved in the meshes, mappings can take many hours to complete.

2. Further it is desirable to map ear shapes at a high resolution to preserve important shape features. This is because the ear shape contributes the most to the features found in the HRTF spectrum. Please refer to Sec. 2.2 for further details.

Given the above constraints, the mapping method described here uses high resolution meshes for mapping the ear shapes and low resolution meshes for mapping the head shapes in order to further speed the mapping process while simultaneously maintaining a high level of accuracy in the mapping of the shapes. Sec. 6.1 Sec. 6.2 in the next chapter will show the accuracy of the head and ear mapping discussed in this section, both from a morphological and acoustical perspective.

In order to explain the method for mapping a source head and ear shape to target head and ear shape five mesh surfaces will be used these are:

1. The high resolution source head and ear mesh: $HE_T$

2. The low resolution source head and ear mesh: $HE_{T,LR}$
4.2. Diffeomorphically Mapping the Head and Ears

Figure 4.3: The above figure shows the source and target head and ear shape used for describing the diffeomorphically mapping the head and ears.

3. The low resolution head and ear target mesh: $HE_{x,LR}$

4. The high resolution source ear mesh: $E_T$

5. The high resolution target ear mesh: $E_x$

In the above list the source shape is denoted by a subscript $T$ and the target shape is denoted by the subscript $x$, further we have abbreviated low resolution as (LR) and high resolution as (HR) to simplify the notations used.

Diffeomorphically mapping the head and ear shapes is shown in the chart depicted in Fig. 4.1 and this is further articulated in Algorithm 1. The discussion to follow will describe in detail each of the five steps shown in the chart and listed in Algorithm 1.

**Step One:** A LDDMM mapping $\mathcal{M}(HE_{T,LR}, HE_{x,LR}, \sigma^H_V, \sigma^H_W)$ is performed between the low resolution source head and ear shape $HE_{T,LR}$ and low resolution target head and ear shape $HE_{x,LR}$ at a deformation scale of $\sigma^H_V$ and shape comparison scale of
Algorithm 1 Diffeomorphically Mapping the Head and Ears

**Input:** \( H_{ET}, H_{ET,LR}, H_{E_s,LR}, E_T, E_s, \sigma^E_V, \sigma^H_V, \sigma^H_W \).

**Output:** \( H_{ETm} \)

1. \( \{a(t)\}^{H_{ET,LR} \rightarrow H_{E_s,LR}} \leftarrow \mathcal{M}(H_{ET,LR}, H_{E_s,LR}, \sigma^H_V, \sigma^H_W) \)
2. \( E_{T1} \leftarrow \mathcal{F}(\{a(t)\}^{H_{ET,LR} \rightarrow H_{E_s,LR}}, E_T, \sigma^H_V) \)
3. \( \{a(t)\}^{E_{T1} \rightarrow E_s} \leftarrow \mathcal{M}(E_{T1}, E_s, \sigma^E_V, \sigma^E_W) \)
4. \( H_{ET1} \leftarrow \mathcal{F}(\{a(t)\}^{H_{ET,LR} \rightarrow H_{E_s,LR}}, H_{ET}, \sigma^H_V) \)
5. \( H_{ETm} \leftarrow \mathcal{F}(\{a(t)\}^{E_{T1} \rightarrow E_s}, H_{ET1}, \sigma^E_V) \)

\( \sigma^H_W \)

To obtain a set of time dependant momentum vectors:

\[
\{a(t)\}^{H_{ET,LR} \rightarrow H_{E_s,LR}} \leftarrow \mathcal{M}(H_{ET,LR}, H_{E_s,LR}, \sigma^H_V, \sigma^H_W) \tag{4.9}
\]

The superscript \( H \) on the scale parameters \( \sigma^H_V \) and \( \sigma^H_W \) signify the scales used for the low resolution head and ear mapping.

**Step Two:** Using the set of time dependant momentum vectors (i.e. \( \{a(t)\}^{H_{ET,LR} \rightarrow H_{E_s,LR}} \)) obtained in the first step an LDDMM flow operation \( \mathcal{F} \) is then performed on the high resolution source ear shape \( E_T \) at a deformation scale of \( \sigma^H_V \) to obtain an intermediate ear shape \( E_{T1} \):

\[
E_{T1} \leftarrow \mathcal{F}(\{a(t)\}^{H_{ET,LR} \rightarrow H_{E_s,LR}}, E_T, \sigma^H_V) \tag{4.10}
\]

**Step Three:** A mapping \( \mathcal{M} \) at a scale \( \sigma^E_V \) and \( \sigma^E_W \) between the intermediate high resolution ear shape \( E_{T1} \) and the target ear shape \( E_s \) is performed and a set of time dependant momentum vectors is obtained \( \{a(t)\}^{E_{T1} \rightarrow E_s} \) for the ear shapes.

\[
\{a(t)\}^{E_{T1} \rightarrow E_s} \leftarrow \mathcal{M}(E_{T1}, E_s, \sigma^E_V, \sigma^E_W) \tag{4.11}
\]

**Step Four:** A LDDMM flow \( \mathcal{F} \) is performed using the momentum vectors \( \{a(t)\}^{H_{ET,LR} \rightarrow H_{E_s,LR}} \), on the high resolution source shape \( H_{ET} \) at the deformation scale of \( \sigma^H_V \) in order to

---

In practice the \( \sigma_W \) scale parameter is gradually lowered at multiple scales starting from a large value when conducting an LDDMM matching i.e. \( \sigma^W_N > \sigma^W_{N-1} > \ldots > \sigma^W_0 > \sigma_W \). The \( \sigma_W \) value shown in the algorithm is the final \( \sigma_W \) used when performing LDDMM matchings. The \( \sigma_V \) stays constant throughout the matching. This multi-scale approach to LDDMM matching by changing the \( \sigma_W \) parameter assists the optimization algorithm to converge quicker when mapping two surfaces using the fine \( \sigma_V \) value in the LDDMM cost function. This is different to the multi-scale approach discussed in later sections of this thesis document where both the \( \sigma_W \) and \( \sigma_V \) value change together.
obtain an intermediate high resolution shape $H_{ET_1}$

$$H_{ET_1} \leftarrow F\left(\{a(t)\}^{HE_T,ls \rightarrow HE_x,ls}, HE_T, \sigma^H_V\right) \quad (4.12)$$

**Step Five:** A subsequent flow is performed using the momentum vectors $\{a(t)\}^{ET_1 \rightarrow E_x}$ on the intermediate high resolution shape $HE_T$ to obtain the high resolution mapped head and ear shape.

$$H_{ETm} \leftarrow F\left(\{a(t)\}^{ET_1 \rightarrow E_x}, HE_{Tm}, \sigma^E_V\right) \quad (4.13)$$

The process for diffeomorphically mapping the high resolution source shape to target shape entails a composition of two separate flows. To elaborate further let's denote the flow for the head shape as $\phi^{HE}$ and the flow for the ear shape as $\phi^E$ then the process of deforming the head and ear shape given above can be expressed as:

$$\psi^{HE} = \phi^E \circ \phi^{HE} \quad (4.14)$$

The following is a summary of the main features of the mapping algorithm discussed in this section:

1. All mapping $\mathcal{M}$ and flow functions $\mathcal{F}$ are implemented and executed on a Tesla K40 GPU [Nvidia Corporation Tesla] running Cuda 6 [Nvidia Corporation Cud]. Further the computer is 8 core Intel 64-bit CPU running Linux Ubuntu. It is to be noted that the execution of the program on the GPU unit is many orders of magnitude faster than the execution on the CPU.

2. Given the importance of ear shape features in the formation of acoustical cues found in the HRTFs, the ears are mapped at a high resolution while the head shapes are mapped at a low resolution.

3. The diffeomorphic mapping of head and ears can be seen as a composition of flows, with the first flow corresponding to the low resolution head mapping...
and the second flow corresponding to the high resolution ear mapping.

4. In the work presented in this thesis the $\sigma^E_V = \sigma^H_V$ and $\sigma^E_W$ and $\sigma^H_W$ values vary based on the size of the elements for the mesh under consideration. Choosing these values will further be elaborated in Sec. 4.3

### 4.3 Shape Differences Based On The Metric of Currents

The previous section described how a source head and ear shape can be deformed and mapped to a target head and ear shape. In this section a new tool is developed to highlight and measure the differences between two surfaces. The main application for the technique developed here is to highlight and measure differences between the morphed and target shapes (please refer to Sec. 4.4 for examples). Using this tool we can obtain better insight into the causes of acoustical differences between the deformed and target ear shapes (please refer to Sec. 6.1 for further discussions). The new technique developed in this section for analysing shape differences is based on the framework of currents, which was previously explained in Sec. 2.8.2.

Currently a popular technique that is used for measuring shape differences is the Hausdorff distance (Ziegelwanger et al., 2015). When analysing differences between 3D surfaces the Hausdorff distance is calculated using the Euclidean metric between the vertices corresponding to the two shapes. The Euclidean distances between the vertices for the two shapes can be used to highlight differences in ear shapes, however, it will be shown in this section the new technique which is based on the framework of currents is able to provide a better insight into the important differences seen when deforming ear shapes.

What follows is a background on the Hausdorff distance and more specifically the Euclidean distance maps which are used for highlighting shape differences. Subsequently some of the problems with the Euclidean distance maps is highlighted and finally, the new technique based on currents is introduced. Particularly it will be
shown how the new measure can provide an improvement on the Euclidean distance maps for highlighting shape differences which in turn shows the superiority of the new measure to the Hausdorff distance when analysing and comparing ear shapes.

### 4.3.1 Hausdorff Distance in 3D Ear Shape Analysis

The Hausdorff distance is used extensively in the study of ear biometrics (Yan and Bowyer 2005) and more specifically for the morphoacoustic analysis of ear shapes (Ziegelwanger et al. 2015). Given two surfaces $S_1$ and $S_2$ (as defined in Sec. 2.8) with corresponding vertices $X, Y \in \mathbb{R}^3$, the Hausdorff distance between the two shapes is denoted as $d^H(S_1, S_2)$. Using the Euclidean distance the Hausdorff distance is defined as:

$$d^H(S_1, S_2) = \max\{d^Y_H = \sup_{y \in Y} \inf_{x \in X} \sqrt{||x - y||^2}, d^X_H = \sup_{x \in X} \inf_{y \in Y} \sqrt{||x - y||^2}\} \quad (4.15)$$

The Hausdorff distance is a scalar value that quantifies the distances between two shapes. When comparing surfaces $S_1$ and $S_2$ with one another it is desirable to identify regions on the surface $S_1$ or $S_2$ that are different between the two surfaces.
In order to do this a distance map based on the distances of the vertices on the two shapes is formed. In this thesis the distance map will be denoted as $Q$. Given a single vertex $x \in X$ in the source shape $S_1$, the value for $Q$ is calculated as:

$$Q(S_1, S_2, x) = \inf_{y \in Y} \sqrt{||(x - y)||^2}, \quad (4.16)$$

where in the above $y \in Y$ is a vertex in the target surface $S_2$. Eq. $(4.16)$ is computed for each vertex $x$ in the surface $S_1$. Also note that in general if the surfaces $S_1$ and $S_2$ are different the function $Q(S_2, S_1, y) \neq Q(S_1, S_2, x)$. We will denote the distance map for the vertices in the source shape $S_1$ as simply $Q(S_1, S_2)$.

Fig. 4.5 displays two pairs of deformations from the template ear $E_T$ to target ears $S_{48}$ and $S_1$. For the mapping between the template ear to target ear $S_{48}$ the acoustical properties of the ears are very different due to an anomaly seen in the target ear shape $S_{48}$. This region is not mapped in the morphed template ear shape and, consequently, the acoustical properties of the ears are significantly different. On the other hand, for the deformation between the template ear shape and ear $S_1$, the mapping is satisfactory from an acoustical view point. However $Q(E_T, S_1)$ still highlights differences between the shapes.

The acoustical responses for the deformation of the template to $S_1$ and $S_{48}$ can be seen in Appendix A.

### 4.3.2 Shape Differences based on CSDA

Ear shapes are intricate surfaces and the distance maps discussed previously accounts only for the proximity between vertices in the shapes through the Euclidean distance. The new shape difference analysis technique presented here is based on the representation of surfaces as currents (Vaillant and Glaunès, 2005). For convenience it will be called currents shape difference analysis (CSDA). Fig. 4.7 shows a typical use of the CSDA measure for highlighting differences in ear shapes. Regions highlighted by the red colour indicate significant differences between the ear shapes. In
4.3. Shape Differences Based On The Metric of Currents

Figure 4.5: Two pairs of matchings are shown between the template ear shape and target ears $S_1$ and $S_{48}$. For subject $S_{48}$, due to an anomaly in the target ear the mapped shape is different and the acoustical properties are very different. For subject $S_1$ on the other hand the mapping is successful from an acoustical view point, however the Euclidean distance map indicates regions of shape differences which are not important acoustically. The colorbar has been threshold to 2mm for both shapes.

In the following section the development of the CSDA technique will be discussed.

There are two motivating factors for the development of this technique. Firstly, the CSDA technique takes into consideration two measures: (1) the proximity of the regions between the two shapes through a kernel function denoted here as $k_R$, and (2) incorporates the orientation of the currents, i.e. the tangent space structure for the two surfaces for a given region. The combination of the above two measures is achieved by the use of the currents and the norm defined via a kernel $k_R$ on the space of currents. The kernel used can be a radially decaying kernel, such as the Gauss or Cauchy Kernel, and in the analysis used here the Cauchy kernel is used:

$$k_R(x, y) = \frac{1}{1 + \frac{||x-y||^p}{\sigma_R}}$$  \hspace{1cm} (4.17)

The scale parameters $\sigma_R$ defines the radius of the kernel. It can be noted again that currents encode the location of the elements, representing the surface in 3D space and the tangent space structure to the shapes. Differences in the tangent space structure for the surfaces can indicate important topological shape differences in ear
shapes.

The **second motivating factor** for the development of this technique stems from the success of the metric $E(S_1, S_2)$ shown in (Eq. (2.28)) in capturing shape differences when mapping surfaces using LDDMM. The functional given in Eq. (2.28), which is a metric in the space of currents, is a crucial element of the LDDMM cost function $J_{S_1, S_2}$ when mapping ear shapes with one another. The merits of the CSDA technique for comparing deformations is highlighted as below:

1. Incorporates the tangent space structure and proximity of corresponding regions in the shapes.

2. Fast implementation for the computation of CSDA is available through the use of GPUs.

Fig. 4.6 shows two ear shapes that have a topological difference in a small region, which is further evident via the orientation of the currents that are shown as red arrows.

To begin formulating the CSDA technique the notion of currents introduced earlier in Sec. 2.8 is used. The current, $[S_1]$, corresponding to shape $S_1$ is characterized by a set of vectors, $n_1(f)$, defined as follows: (1) there is one vector originating from the centre, $c_1(f)$, of each face $f$ of $S_1$; (2) the vectors point outwards and are normal to the faces; (3) the norm of each vector is proportional to the area of the face. In order to examine the similarity between two shapes $S_1$ and $S_2$ for the region surrounding face $f$ of $S_1$ we propose to use the measure $d(S_1, S_2, f)$ given by:

$$
\beta_1(f) = \sum_k k_R (c_1(f), c_1(h)) \langle n_1(f), n_1(h) \rangle,
$$

$$
\beta_2(f) = \sum_h k_R (c_1(f), c_2(h)) \langle n_1(f), n_2(h) \rangle,
$$

$$
d(S_1, S_2, f) = |\beta_2(f) - \beta_1(f)|,
$$

$$
\hat{d}(S_1, S_2, f) = \min \left( \frac{d(S_1, S_2, f)}{|\beta_1(f)|}, 1 \right),
$$

(4.18)
4.3. Shape Differences Based On
The Metric of Currents

Figure 4.6: Topological difference in a small region of the ear shapes are shown. The differences in the shapes are further highlighted by the orientation of the normals (red arrows) pointing out of the centres of the mesh faces. The difference in the region shown can cause significant differences in the acoustical spectrum of the shapes.

where $k_R$ denotes the Cauchy kernel and is defined as in Eq. (4.17). In Eq. (4.18), $\beta_1(f)$ corresponds to the convolution of vector $\mathbf{n}_1(f)$ with every other vector $\mathbf{n}_1(g)$.
in $S_1$, while $\beta_2(f)$ corresponds to the convolution of $n_1(f)$ with every vector $n_2(h)$ in $S_2$. When $S_1$ is very similar to $S_2$ in the vicinity of face $f$, $\beta_1(f)$ and $\beta_2(f)$ are almost equal and $d(S_1, S_2, f)$ is very small. On the other hand, when the two shapes are very dissimilar (orthogonal, or far away from each other), $\beta_2(f)$ is very small and $d(S_1, S_2, f)$ is relatively large. In order to enable meaningful comparisons across different triangular faces and different shapes, we normalise $d(S_1, S_2, f)$ by the absolute value of $\beta_1(f)$. We also limit the maximum value of $\hat{d}(S_1, S_2, f)$ to unity to ensure the measure does not blow up when $\beta_1(f)$ is very small.

The overall similarity between two ear shapes is calculated as the average similarity measure, $\bar{d}(S_1, S_2)$, given by:

$$\bar{d}(S_1, S_2) = \frac{1}{F} \sum_{f=1}^{F} \hat{d}(S_1, S_2, f),$$  

(4.19)

where $F$ denotes the total number of faces in shape $S_1$. Note that in general $\bar{d}(S_1, S_2) \neq \bar{d}(S_2, S_1)$.

Fig. 4.7 shows ear shapes where the CSDA is used to highlight regions of important shape difference. The same pair of subjects $S_1$ and $S_{48}$, which were used in Fig. 4.5 when examining the Hausdorff distance, are used in this analysis as well. The scale parameter for the kernel $k_R$ is at $\sigma_R = 1$. It can be observed that the CSDA technique indicates a good matching for target $S_1$ and further highlights a problem for subject $S_{48}$. The highlighted regions for $S_{48}$ can cause considerable differences in the acoustical responses between the deformed template shape and the target shape.
4.3. Shape Differences Based On The Metric of Currents

Figure 4.7: Two deformations from the template ear is examined using the CSDA technique, the same deformations were compared in the previous figure using the Euclidean distance maps. The CSDA measure indicates a good quality of mapping for subject $S_1$ which is verified by the close acoustical responses for the ear shapes. On the other hand the regions of mismatch which cause significant acoustical difference is highlighted for the mapping of subject $S_{48}$.

4.3.2.1 The Effect of $\sigma_R$ on measuring shape differences

This sections shows a series of experiments that examine the effect of the scale parameter $\sigma_R$ when comparing shapes using the CSDA technique. The $\sigma_R$ value plays a similar role when comparing shapes to the shape comparison parameter $\sigma_W$ used in the metric $E$ (please refer to Sec. 2.8.2).

The experiment consisted in comparing shapes that were either very similar or significantly different using various values of $\sigma_R$. The average edge length for the triangulated mesh shapes used in the experiments was just below 1mm. The values of $\sigma_R$ used for comparing the shapes was chosen from the set $\{40, 10, 1, 0.1\}$. Fig. 4.8 and Fig. 4.9 show the effect of different $\sigma_R$ values when comparing shapes using CSDA. Fig. 4.8 is an example for which the shapes are very similar. On the other hand Fig. 4.9 shows shapes that are significantly different with respect to each other. It can be observed from the figures that when $\sigma_R < 1$ and $\sigma_R \geq 10$ the CSDA shape comparison tool is ineffective in highlighting shape difference.
4.3. Shape Differences Based On
The Metric of Currents

![Diagram showing the effect of changing the $\sigma_R$ value on the quality of shape comparison using the CSDA technique when the shapes are very similar.]

Figure 4.8: The effect of changing the $\sigma_R$ value on the quality of shape comparison using the CSDA technique when the shapes are very similar.
4.3. Shape Differences Based On
The Metric of Currents

$\sigma_R = 0.1$

$\sigma_R = 1$

$\sigma_R = 2.5$

$\sigma_R = 10$

$\sigma_R = 40$

Figure 4.9: The effect of changing the $\sigma_R$ value on the quality of shape comparison using the CSDA technique when the shapes are significantly different.
4.4 LDDMM Scale Parameters

This section explains how the value for three LDDMM parameters is set when mapping ear shapes. Two of these parameters are LDDMM scale parameters $\sigma_V$ and $\sigma_W$, which are used in the kernels $k_v$ Eq. (2.27) and $k_w$ Eq. (2.29). The third parameter $\gamma$ is primarily used for optimizing shape matchings. A series of experiments is conducted in this section in which the template ear shape is deformed to the target ear shapes in the SYMARE database. The generation of the template ear shape is described in a later chapter (Sec. 5) and its shape is displayed here in Fig. 4.10. This section utilizes a technique named "Current Shape Difference Analysis" (CSDA) for comparing matched to target ear shapes. The CSDA technique is used to visually highlight differences between ear shapes and, also, to obtain a quantitative value for differences between the ears. This technique is introduced and explained in more detail in Sec. 4.3.

4.4.1 Scale Parameter $\sigma_V$

This section describes how the value for the deformation scale parameter $\sigma_V$ is set when mapping ear shapes. The $\sigma_V$ parameter is used in the deformation kernel $k_V$:

$$k_V(x, y) = \frac{1}{1 + \frac{\|x-y\|^2}{\sigma_V^2}} \quad (4.20)$$

Recall from the background section Sec. 2.8 that the $\sigma_V$ parameter determines how one vertex is influenced by the movements of neighbouring vertices. This effect was further shown to be due to the convolution of the momentum vectors using the kernel function $k_V$. Larger values of $\sigma_V$ will make the region where vertices are coupled to become greater, which causes deformations from the source shape to be crude and inaccurate. On the other hand when the $\sigma_V$ parameter is set to a very small value, the vertices in the source shape can move randomly and hence the deformations can be meaningless.
An experiment was conducted where the template ear shape $E_T$ was deformed to several ears within the SYMARE database. The value of $\sigma_W = 1.25$ and $\gamma = 5e - 4$ were kept constant and the value of $\sigma_V$ was varied. The value for the $\sigma_V$ parameter was selected from the set $[1.25, 10, 20, 80]$. Fig. 4.11 shows an example of a deformed template shape and the target ear shape in which the CSDA technique is used on the target ear shape to highlight shape differences. The reader can refer to Sec. 4.3 to see how the CSDA measure can be used to find, measure, and quantize shape differences between a source and target ear shape. The CSDA measure between the target ear $S_1$ and deformed template shape $E_T$ is denoted by $\hat{d}(S_1, E_T, f)$, where $f$ indicates the element in the target shape. The average value of $\hat{d}(S_1, E_T, f)$ over all the faces of the target shape can be used in order to obtain a single number that quantizes the difference between the target and morphed template ear shape:

$$\bar{d}(S_1, E_T) = \frac{1}{F} \sum_{f=1}^{F} \hat{d}(S_1, E_T, f), \quad (4.21)$$

in the above, $F$ is the total number of elements in the target shape. Note that in general, $\bar{d}(E_T, S_1) \neq \bar{d}(S_1, E_T)$ due to different number of elements on each of the mesh surfaces. Fig. 4.12 shows a plot of $\langle \hat{d} \rangle$ versus $\sigma_V$ which is the value of $\hat{d}(S_n, E_T)$ averaged over multiple deformations of the template ear shape to target ear shapes.
4.4. LDDMM Scale Parameters

Figure 4.11: The quality of a mapped shape is shown in comparison to the target shape. The mapping is conducted at a value of $\sigma_W = 1.25$ and $\gamma = 5e - 4$. The value for the $\sigma_V$ value is changing in the range of $[1.25, 10, 20, 80]$ . The target shape is highlighted using the CSDA technique which is explained in the next chapter. The width of the kernel in the CSDA technique for this experiment is set to $\sigma_R = 1$.

By visually inspecting the deformations a cutoff value of $\langle \bar{d} \rangle = 0.05$ was chosen over multiple mappings. For ear mappings a $\sigma_V = 10$ is selected as it is the largest value in the curve adhering to this cut off value.

in the SYMARE database. That is:

$$\langle \bar{d} \rangle = \frac{1}{N} \sum_{n=1}^{N} \bar{d}(S_n, E_T),$$  

(4.22)
Figure 4.12: The above plot shows the value of $\langle \tilde{d} \rangle$ versus $\sigma_V$ which is the value of $\tilde{d}(S_n, E_T)$ averaged over multiple deformations of the template ear shapes to ears in SYMARE. The value for $\sigma_W = 1.25$ and $\gamma = 5e - 4$ were kept constant and $\sigma_V$ varied in the range of 1.25 to 80.

4.4.2 Scale Parameter $\sigma_W$

This section explains how the value of the shape comparison parameter $\sigma_W$ is chosen for mapping ear shapes. The $\sigma_W$ parameter is used in the kernel $k_W$ for comparing shapes:

$$k_W(x, y) = \frac{1}{1 + \frac{||x - y||^2}{\sigma_W^2}}$$

(4.23)

The $\sigma_W$ parameter specifies the granularity in comparing 3D surfaces with one another. This can ultimately effect the accuracy of the mapping between the source and target ear shapes. Small values of $\sigma_W$ compare the tangent space structure of the shapes to a very fine level, while large values coarsely compare the shapes with each other. Similar to the methodology employed in the previous section an experiment was conducted in which the template ear shape was deformed to many target ear
shapes in the SYMARE database. The deformation scale parameters \( \sigma_V = 10 \) and optimization parameter \( \gamma = 5e - 4 \) were kept constant and the value for the \( \sigma_W \) varied. The values chosen for \( \sigma_W \) were 1, 5, 10 and 40. Fig. 4.13 shows the deformed (mapped) template ear shape and a target ear for mappings at varying values of \( \sigma_W \). Similar to the plots in the previous section the CSDA measure is used to highlighting shape differences.

![Target Matched Front Back](image)

\( \sigma_W = 1 \)

\( \sigma_W = 5 \)

\( \sigma_W = 10 \)

\( \sigma_W = 40 \)

Figure 4.13: The quality of a mapped shape is shown in comparison to the target shape. The mapping is conducted at a value of \( \sigma_V = 10 \) and \( \gamma = 5e - 4 \). The value for the \( \sigma_W \) value is changing in the range of 1 to 40. The target shape is highlighted using the CSDA technique which is explained in the next chapter.

Fig. 4.14 is a plot of \( \langle \bar{d} \rangle \) versus \( \sigma_W \) which is the value of \( \bar{d}(S_n, E_T) \) averaged over multiple mappings of the template ear shape to subjects in the SYMARE database.
(please look at the previous section for more detail on how the plot was obtained). Having a nominal cut off value of $\langle \bar{d} \rangle = 0.05$ the plot in Fig. 4.14 indicates that any value of $\sigma_W \leq 5$ can be used for performing ear shape mappings. In this research to increase the accuracy in the ear shape mappings a $\sigma_W < 5$ was chosen for performing LDDMM minimization.

![Graph showing $\langle \bar{d} \rangle$ versus $\sigma_W$](image)

Figure 4.14: The above graph shows $\langle \bar{d} \rangle$ versus $\sigma_W$ which is the value of $\bar{d}(S_n, E_T)$ averaged over multiple deformations of the template ear shape to ears in SYMARE. The value for $\sigma_Y = 10$ and $\gamma = 5e^{-4}$ were kept constant and $\sigma_W$ varied in the range of 1 to 40.

### 4.4.3 Optimization weight parameter $\gamma$

This section details how the value of the optimization parameter $\gamma$ was chosen when mapping ear shapes. The parameter $\gamma$ controls the relative weight between the terms on the LDDMM cost function:

$$ J_{S_1, S_2}(v(t)) = \gamma \int_0^1 \|v(t)\|^2 dt + E(S_1(\phi^v(1, X)), S_2(Y)), \quad (4.24) $$
when the value of $\gamma$ is increased any minimization on $J$ shown above will penalize the first term (i.e. $\int_0^1 \|v(t)\|_V^2 dt$) increasingly. The template ear shape was deformed to target ear shapes in SYMARE using a constant values for $\sigma_V = 10$ and $\sigma_W = 1.25$ and by changing the value for $\gamma$ in the range of $[5e^{-4}, 5e^{-3}, 1, 5]$. Fig. 4.15 shows the deformed template and target ears when mapped at the various values of $\gamma$. The CSDA measure is used to highlight mapping differences.

![Figure 4.15: Shape differences based on the currents metric for LDDMM matchings at various $\gamma$ values with the $\sigma_V = 10$ and the $\sigma_W = 1.25$](image)

Fig. 4.16 shows a plot of $\bar{d}(S_n, E_T)$ versus $\gamma$ averaged over multiple mappings from the template ear shape to target ears in SYMARE. It can be observed that the quality of the mappings do not change greatly when compared to the scale
parameters \( \sigma_W \) and \( \sigma_V \) when the value of \( \gamma \) is changed. In particular the value of \( \langle \bar{d} \rangle \) is below the threshold of 0.05 when \( \gamma < 1 \).

Figure 4.16: The above graph shows \( \langle \bar{d} \rangle \) versus \( \gamma \) which is the value of \( \bar{d}(S_n,E_T) \) averaged over multiple deformations of the template ear shapes to ears in SYMARE. The value for \( \sigma_W = 1.25 \) and \( \sigma_V = 10 \) were kept constant and \( \gamma \) varied in the range of \( 5e^{-4} \) to 5.

4.5 Producing FM-BEM Ready Meshes

This section details an automated procedure for producing high quality meshes for FM-BEM simulations. FM-BEM simulations are generally more accurate with increasing number of mesh elements, however, the disadvantage is that the cost of running simulations with a large number of elements is very expensive. In order to obtain accurate FM-BEM simulations using triangular meshes certain restrictions need to be imposed on the triangular elements within the mesh (Gumerov et al., 2010). These restrictions ensure that triangular elements are close to equilateral and
further element sizes do not vary greatly across the mesh. The remeshing software, and the procedure detailed in this section, tries to minimize the number of elements used through a mesh coarsening procedure while at the same time producing a high quality uniform mesh. More specifically, this procedure produces meshes that satisfy the following requirements on mesh elements:

1. Maximum to minimum edge length $< 5$

2. Minimum angle of $> 15^\circ$

3. Maximum angle of less than $< 150^\circ$

4. Maximum edge length $< \frac{1}{6}$

where $\lambda$ is the minimum wavelength considered for sound waves. An open source software ACVD (CREATIS Medical Imaging Research Center, ACV; Valette and Chassery, 2004; Valette et al., 2008) is used to produce isotropic (i.e. uniform) mesh elements that satisfy the requirements 1-4 detailed above. Once a uniform mesh is created using ACVD, Meshlab is used to remove any unreferenced vertices and invert and flip face normals if required. ACVD is a software used mainly for coarsening large meshes, but it can also be used for remeshing and or subdividing triangular mesh surface. The main parameter used as input to the ACVD software is the number of vertices $N_v$. Using $N_v$, ACVD will try to create a uniform mesh using only $N_v$ vertices such that the output mesh is an approximate to the input mesh to a given threshold. If the input number of vertices is larger than the current number of vertices then surface subdivision is first applied.

The final FM-BEM ready shapes are obtained using a program in Matlab that is written to iteratively check the quality of the output mesh against the requirements 1-6 detailed above and adjust the number of vertices. Given the number of mesh vertices $N_v(I)$ in the current iteration $I$, and the current critical frequency $f_c$ of the mesh calculated using Eq. (4.25), the Matlab program adjusts the input number of vertices to the ACVD program at the next iteration $N_v(I + 1)$ to meet the target
critical frequency $f_{ctarg}$ by the use of Eq. (4.26).

$$f_c = \frac{c}{6e_{max}}$$  \hspace{1cm} (4.25)

$$N_v (I + 1) = N_v (I) (\frac{f_{ctarg}}{f_c})^2$$  \hspace{1cm} (4.26)

In Eq. (4.25) $e_{max}$ is the maximum edge length for the mesh. In this research the critical frequency for the meshes is set to be $f_c = 26kHz$. The average edge length (AEL) for output meshes at the above critical frequency were below 1.5mm. HRTFs obtained by numerical FM-BEM simulations using this AEL are shown to provide good agreement with the acoustically measured HRTFs (Ziegelwanger et al., 2015). It is to be noted that due to the iterative nature of the Matlab program and also because the output mesh shapes have to satisfy criterion’s 1-4 specified above the resultant meshes can have a critical frequency value that is higher and not lower than 26kHz.

### 4.6 Conducting FM-BEM Simulations

FM-BEM simulations are conducted using the Coustyx software by Ansol (Ansol). The simulations are conducted using the principle of reciprocity (Jin et al., 2014; Kreuzer et al., 2009). In the direct setting a microphone is placed inside the ear canal and the sound is emitted from a certain direction in space. To calculate the HRTF for different directions in space the positions of the loud speaker is adjusted. In the reciprocity approach the microphone and loudspeaker positions are interchanged. When calculating the HRTF numerically the virtual loud speaker is placed in the ear canal and virtual sensors are placed at the desired directions in space. Using this technique the HRTF can be computed simultaneously for several directions in space. In this thesis the exact position of the vibrating element (virtual sound source) is calculated by finding the intersection of the inter-aural axis with the left ear mesh. Fig. 4.18 shows the position of the vibrating element with respect to the ear shape.
4.6. Conducting FM-BEM Simulations

In Fig. 4.18 the vibrating element is shown in red and the inter-aural axis is the green line, and the intersection of the inter-aural axis with the vibrating element is shown in blue. For the experiments conducted in this thesis, virtual sensors were placed at 2562 evenly spaced directions in space at a distance of 1m from the centre of the head shape. The evenly spaced directions were obtained using Icosahedral (Wenninger [1974]) subdivision. Fig. 4.17 shows a plot of the sensor positions used when calculating the HRIRs using FM-BEM simulations.

The HRTFs were obtained using the Burton-Miller BIE formulation and further using the Galerkin implementation. This method is known to be superior to other formulations when numerically solving the Helmholtz PDE using the BEM system of equations. The BEM simulations are conducted using the multi-level FM. To compute the number of FM levels $N_{FM}$ required for the simulations, first, the dimension $r_{HE}$ and average edge length $e_{avg}$ of the object (head shape) is calculated and then Alg. 2 is used to calculate the FM levels. The FM-BEM implementation in Coustyx uses the GMRES iterative solver for the FM computation on the BEM system of equations. The GMRES is further preconditioned in order to accelerate the convergence to a solution. Appendix Sec. G shows parts of the configuration script used by Coustyx software when conducting FM-BEM simulations.
4.6. Conducting FM-BEM Simulations

Figure 4.18: A section of a head mesh is shown after being processed by the ACVD software. The vibrating element is in red color and the interaural axis is the green line. The blue dot is the point of intersection of the inter-aural axis with the mesh.

Algorithm 2 Compute FM Levels

| inputs: | $r_{HE}, e_{avg}$. |
| outputs: | $N_{FM}$. |
| 1: | $L \leftarrow 20$ |
| 2: | for $1 \leq l \leq L$ do |
| 3: | if $2(r_{HE})^l - e_{avg} \leq 0$ then |
| 4: | $N_{FM} = l$ |
| 5: | end if |
| 6: | end for |
4.7 Extracting FM-BEM Data

This section details how the DTF data is extracted from the FM-BEM simulations and further processed for the various analysis performed in this thesis. This section further builds the notations used in this thesis for representing the DTFs.

The result of FM-BEM simulations on the head and torso surfaces are the complex pressure values $\Phi$ at the sensors positions $(\theta, \phi)$ and for the desired frequency $f$. This will be denoted as $\Phi(f, \theta, \phi)$. Fig. 4.19 shows the convention used for angles $\theta$ and $\phi$ in 3D space. The angle $\theta$ is in the range $-\pi \leq \theta \leq \pi$ and the angle $\phi$ is in the range of $-\frac{\pi}{2} \leq \phi \leq \frac{\pi}{2}$.

In order to obtain the raw HRTFs the obtained complex pressures values are normalized by the free field Greens function:

$$G(f) = \frac{-i1.21f^2}{2\pi}e^{\frac{2\pi}{343}f}$$

(4.27)

Where in the above $f$ is the frequency of the sound wave in Hertz and 343 is the speed of sound in air, further it is assumed that the sensors are positioned at a distance of 1 meter from the centre of the head. As detailed in Sec. 2.3 the HRTFs are a function of frequency, azimuth and elevation angle and are obtained from the pressure values using the following equation:

$$\text{HRTF}(f, \theta, \phi) = \frac{\Phi(f, \theta, \phi)}{G(f)}$$

(4.28)

Assuming there are a total of $P$ HRTFs available then the portion of the HRTF which is common to all directions in space is known as the common transfer function CTF (Middlebrooks [1999b]; Xie [2013]) and can be removed from the HRTF spectrum. The magnitude of the CTF is computed by:

$$|\text{CTF}(f)| = \sqrt{\frac{1}{P} \sum_{i=1}^{P} |\text{HRTF}(f, \theta_i, \phi_i)|^2}$$

(4.29)
Figure 4.19: The above plot shows the two important auditory planes, namely horizontal plane and the median plane. Further the convention used for the angles $\theta$ and $\phi$ are shown. These angles are used extensively to specify a DTF belonging to a direction.
The phase information for the CTF can be obtained from its magnitude using the minimum phase properties of the CTF. The phase information is not used in this thesis and for this reason further discussion will not be given on the phase information of the CTF or HRTF. By removing the CTF from the HRTF spectrum the Directional Transfer Function (Cheng and Wakefield, 1999b; Middlebrooks, 1999b) known as DTF is obtained. In mathematical terms the log magnitude for the DTF is computed by:

\[ H(f, \theta, \phi) = 20 \log_{10} \left| \frac{\text{HRTF}(f, \theta, \phi)}{|\text{CTF}(f)|} \right| \quad (4.30) \]

Once the DTF data is extracted, spherical spline interpolation (Wahba, 1981) is applied to the data to increase the spatial resolution of the DTFs. In this thesis the DTF data has a spatial resolution of approximately 2.95° degrees. The full frequency range for the DTF spectrum is 20Hz to 24kHz given a sampling frequency of \( f_s = 48kHz \), however typically the frequency range used for analysis in this thesis is between 20Hz and 16kHz and this is represented by 342 frequency bins. The following table summarizes the properties for the DTF data.

<table>
<thead>
<tr>
<th>Sampling Frequency ( f_s )</th>
<th>48kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Number of FFT bins</td>
<td>1024</td>
</tr>
<tr>
<td>Typical frequency range used in DTF data</td>
<td>20Hz - 16kHz</td>
</tr>
<tr>
<td>Number of FFT bins used in the Fourier computation</td>
<td>1024</td>
</tr>
<tr>
<td>Angular resolution for the DTF’s</td>
<td>2.95°</td>
</tr>
</tbody>
</table>

In many sections of this thesis the acoustical data for the median plane is examined. The log magnitude for the DTFs for the median plane are denoted by \( W(f, \phi) \). Where now \( \phi \) varies between \(-\pi \leq \phi \leq +\pi\). When \( \phi = \pm \pi \) the elevation angle is directly below the head positive angles correspond to the front and negative angles correspond to the back of the subject. Fig. 4.20 shows the convention used for the angle \( \phi \) when working on the median plane. When analysing data for the median plane this convention will be used for \( \phi \) unless otherwise stated.

The full set of DTFs belonging to a given ear shape \( E \) in the median plane is represented as \( W_E \in \mathbb{R}^{N \times M} \), for some \( N \) and \( M \). Given that the angular resolution in
4.7. Extracting FM-BEM Data

Figure 4.20: The above figure shows the convention used for the angle $\phi$ when working on the median plane.

The DTF data is $2.95^\circ$ there will be a total of 122 angles in the median plane making the value $N = 122$ and $M = 342$ for the median plane DTFs. If we require a DTF for a particular direction $\phi_i$, we will denote it as $W_E(\phi_i) \in R^M$ and if we require the value of a DTF for a particular direction in space and frequency we will refer it to by $W_E(f_j, \phi_i)$.

There are a few ways the DTF data is shown in this thesis, these are further detailed below:

**Spatial Frequency Response Surface:** The DTF data can be represented as a map in which the gain of the ear for different directions in space is plotted for a given frequency. In order to represent the DTF in this manner the frequency value is set to $f = f_1$ and the DTF is plotted for all available directions $\theta$ and $\phi$ in space. This
representation is also known as the SFRS at frequency $f_1$.

$$SFRS_{f_1} = H(f = f_1, \theta, \phi) \quad (4.31)$$

This representation of the DTF was previously shown in Fig. 2.6.

**Frequency Response:** The DTF can also be plotted for a single direction in space $\theta_1, \phi_1$ and for all frequency values in a given range, the frequency range typically used in this thesis is from $f_1 = 20Hz$ to $f_u = 16kHz$. The frequency response plots the function $H(f, \theta_1, \phi_1)$ for some given angles $\theta_1$ and $\phi_1$. This representation of the DTF was previously shown in Sec. 2.5.

**Median Plane Acoustical Image:** A popular method for showing the DTFs corresponding to the median plane is by showing an intensity plot via an image. Fig. 4.21 shows the gain in the DTF (i.e. $W(f, \phi)$) using color intensity for a given ear shape in the median plane.

**Median Plane Surface Plot:** Another method that is used in this thesis for showing the acoustical data in the median plane is surface plots of the DTF function. Fig. 4.22 shows the gain in the DTF (i.e. $W(f, \phi)$) as a surface plot for a given ear.
shape for the median plane. This kind of surface plot enables the notches and peaks in the DTF spectrum to be seen more clearly.

4.8 Globally Comparing the Acoustical Responses for Ear Shapes

This section details a global technique for comparing the acoustical responses for ear shapes. The result of the global acoustical analysis technique is a figure which will be called GAAF. Fig. 4.23 shows GAAF when comparing the acoustical responses for two ear shapes.

The global approach for analysing the acoustical responses presented in this section stems from the need to compare and analyse the acoustical responses for deformed ear shapes in the frequency and spatial domains. More specifically, it is desirable to know what directions in space contribute to the biggest parts of the error in the frequency domain and, vice versa, how the error in the frequency domain is spread spatially. The motivating factor to the development of this technique is to identify acoustical differences between the morphed template shape and the target
Figure 4.23: The Global Acoustic Analysis Figure (GAAF), which compares the acoustical responses for two ear shapes is shown in the above figure. More specifically, GAAF provides a global insight into the differences between two sets of DTFs. The plot consists of (a) Spatial Correlation Measure (SCM) (2) SFRS map at the frequency $f^*$ where the SCM is a minimum for the first ear shape (3) SFRS map at the same frequency $f^*$ for the second ear shape, (4) A plot of the error represented by the function $G$ for the two ear shapes for the frequency $f^*$ (5)-(6) two DTF plots where $G$ indicates a large difference at the locations marked on the plot for function $G$. (7) A plot of the function $V$ and (8)-(9) two DTF plots at locations where the $V$ is large. The locations for these DTFs are also marked on the plot of the function $V$ shape.

### 4.8.1 Spatial Frequency Response Surface (SFRS)

Before detailing each of the measures used for producing GAAF, some background information is provided in this subsection on the Spatial Frequency Response Surface (SFRS).

The HRTF and DTF are complex valued functions of location and frequency. An informative view of the spatial gain for the DTFs at a particular frequency is known as the Spatial Frequency Response Surface or better know as the SFRS.
Globally Comparing the Acoustical Responses for Ear Shapes

and Wakefield [1999a]. The DTF for a given subject $S_1$ is a function of frequency $f$, azimuth $\theta$ and elevation $\phi$. In order to obtain the SFRS, the frequency value for the DTF is set to a specified value of $f_1$ (i.e $f = f_1$) and the magnitude of the complex DTF function is plotted for all directions in space. The SFRS for the subject $S_1$ is written mathematically for a given frequency $f_1$ as:

$$H_{S_1}(f = f_1, \theta, \phi)$$

(4.32)

Fig. 4.24 shows the SFRS plot for two frequencies using the Robinson projection map. The Robinson projection enables the approximate plotting of a 3D spherical map of the SFRS onto a 2D surface:

![SFRS plot](image)

Figure 4.24: An example of an SFRS plots for a given subject at frequencies corresponding to 3kHz and 12kHz. A Robinson projection is used to map the gain of the DTF for the complete sphere around the head. Brighter areas in the plot show a higher gain compared to the darker regions.

### 4.8.2 Spatial Correlation Metric (SCM)

The first metric used to compare the DTFs is the spatial correlation metric (SCM). The SCM is denoted mathematically as $C_{S_1,S_2}(f)$, signifying that it is a function of frequency $f$, and parametrised by the two input shapes $S_1$ and $S_2$. The function $C_{S_1,S_2}(f)$ is the first plot of GAAF and is shown in plot (a) of Fig. 4.23. The function $C$ provides a measure of how coherently two DTFs vary across 3D space. There is a single correlation value between 0% and 100% for every frequency under considera-
tion. A similar metric based on the correlation of HRTFs was used in (Jin et al., 2014; Ziegelwanger et al., 2015). A correlation of 0% indicates that the SFRS for the two shapes at the given frequency are statistically independent, and a 100% correlation indicates that the SFRS at the given frequency are related. Having noted this it can be said that the SCM is a function of frequency $f$ and measures the dissimilarity between two HRTFs across all directions in space for a particular frequency $f$ by using the standardized correlation metric.

Given that the log magnitude for the complex DTF function is denoted as $H_{S1}(f, \theta, \phi)$ (refer to Sec. 4.7) to compute the SCM between two shapes $S_1$ and $S_2$, the mean $\overline{H_{S1}(f)}$ and $\overline{H_{S2}(f)}$ and the empirical standard deviations $\xi(H_{S1})$ and $\xi(H_{S2})$ need to be computed using the following equations:

$$H_{S1}(f) = \frac{1}{M} \sum_{i=1}^{M} H_{S1}(f, \theta_i, \phi_i)$$  
(4.33)

$$H_{S2}(f) = \frac{1}{M} \sum_{i=1}^{M} H_{S2}(f, \theta_i, \phi_i)$$  
(4.34)

$$\xi(H_{S1}, f) = \sqrt{\frac{1}{M} \sum_{i=1}^{M} H_{S1}(f, \theta_i, \phi_i) - \overline{H_{S1}(f)}}$$  
(4.35)

$$\xi(H_{S2}, f) = \sqrt{\frac{1}{M} \sum_{i=1}^{M} H_{S2}(f, \theta_i, \phi_i) - \overline{H_{S2}(f)}}$$  
(4.36)

Using $\xi(H_{S1}, f)$, $\xi(H_{S2}, f)$ the correlation metric $C_{S1,S2}(f)$ is obtained between the two shapes $S_1$ and $S_2$ at every frequency $f$ from the following equation:

$$C_{S1,S2}(f) = \frac{\sum_{i=1}^{M} [H_{S1}(f, \theta_i, \phi_i) - \overline{H_{S1}(f)}] [H_{S2}(f, \theta_i, \phi_i) - \overline{H_{S2}(f)}]}{\xi(H_{S1}, f) \xi(H_{S2}, f)}$$  
(4.37)

A plot of the function $C_{S1,S2}$ for two subjects $S_1$ and $S_2$ for the frequency range 5kHz to 16kHz is shown in Fig. 4.25. We can also define an overall measure of
dissimilarity between the acoustics for the two subjects $S_1$, $S_2$ by averaging the measure $C_{S_1,S_2}$ over the given frequency range:

$$C_{S_1,S_2} = \frac{1}{N} \sum_{i=1}^{N} C_{S_1,S_2}(f_i)$$

(4.38)

As a further analysis to the SCM measure it is desirable to see how the SFRS for the two subjects and for a given frequency $f_1$ compare to each other. The frequency $f_1$ is chosen such that the value for $C_{S_1,S_2}(f)$ is minimum. The frequency range where $f_1$ is computed can be adjusted, however it is typically computed over the range of 5kHz and 16kHz. Fig. 4.26(a) shows the SFRS plot for subject $S_1$ at the frequency $f_1 = 8484$ and Fig. 4.26(b) shows the SFRS map at the same frequency for subject $S_2$. Fig. 4.26(a) corresponds to plot (b) in Fig. 4.23 and Fig. 4.26(b) corresponds to plot (c) in Fig. 4.23.
Figure 4.26: SFRS maps for two shapes $S_1$ (Matched shape) and $S_2$ (Target shape) at a given frequency $f_l = 8484$ where the SCM measure $C_{S_1,S_2}(f)$ is a minimum within the range of 5000Hz to 16000Hz. This minimal can be observed from the plot $C_{S_1,S_2}(f)$ in the previous figure.
4.8.3 Absolute Spatial Gain Error (SGE)

The absolute Spatial Gain Error (SGE), which will be denoted by the function $G$ is given by:

$$G_{S_1, S_2}(f = f_l, \theta, \phi) = |H_{S_1}(f = f_l, \theta, \phi) - H_{S_2}(f = f_l, \theta, \phi)|$$  \hspace{1cm} (4.39)

The function $G$ is the fourth plot in GAAF and is shown in plot(d) of Fig. 4.23. The function $G$ is used to get an understanding on the quality of mapping between two sets of SFRS maps for two subjects, $S_1$ and $S_2$, for a given frequency. In this thesis the SGE measure is used in conjunction with the information given by the SCM measure to show how the error in the corresponding DTFs for a particular frequency of interest is spread spatially. More specifically, the SGE is plotted at the frequency $f_l$, which is obtained from the function $C_{S_1, S_2}$. The information given by the function $G$ can be used to answer questions such as, whether the error in the SCM is contributed by the SFRS on the ipsilateral or contra-lateral side of the ears under investigation, or if the error is mostly contributed at very low or high elevation angles. Fig. 4.25 shows a plot of the SGE for the frequency $f_l = 8484\text{Hz}$.

In order to assist a better understanding of the acoustical differences corresponding to the error function $G$, GAAF also contains two DTF plots for locations where the magnitude of the function $G$ is high. These locations are marked by green markers (crosses) in the plot for $G$. The locations for high error were obtained automatically using a Matlab routine. Three constraints are imposed when selecting the location for the green markers:

1. $G$ should be a local maxima at the location of the marker.

2. It should be at least $20^\circ$ (this value can be adjusted in the code) apart from another marker if it has already been selected.

3. The azimuth and elevation angles for the marker should satisfy $\theta_m > 0^\circ$ and $\phi_m > -40^\circ$. In other words we only consider positions in the ipsilateral side of
4.8. Globally Comparing the Acoustical Responses for Ear Shapes

Figure 4.27: A plot of the function $G$ between two subjects $S_1$ and $S_2$ is shown above. The frequency $f_1$ at which this function is plotted is the frequency for which the function $C_{S_1,S_2}$ is minimum. The green crosses appearing in the plot are locations for which the error $G$ is large.

the ear and for elevation angles greater than $-40^\circ$.

The DTF plots for these locations correspond to plots (e) and (f) in Fig. 4.23.

4.8.4 Standard Deviation of the Spectral Difference (SDS)

The standard deviation of the spectral difference (SDS), measures the difference in the acoustical responses for two ears for all frequencies, and for a given azimuth and elevation directions in space $\theta_i, \phi_i$. The SDS measure will be denoted by the function $V$ and is given by:

$$x_n = H_{S_1}(f_n, \theta_i, \phi_i) - H_{S_2}(f_n, \theta_i, \phi_i)$$

$$u_n = \frac{1}{N} \sum_{n=1}^{N} (x_n)$$

$$V_{S_1,S_2}(H_{S_1}(f_n, \theta_i, \phi_i), H_{S_2}(f_n, \theta_i, \phi_i)) = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (x_n - u_n)^2}$$
Eq. (4.40) is the seventh plot used in GAAF and is shown in plot (g) of Fig. 4.23. The function $V$ is equal to the standard deviation of the difference between the DTFs for the two subjects and is calculated over a range of discrete frequencies denoted by $f_n$ for the given directions $\theta_i$ and $\phi_i$ in space. This frequency range can be adjusted as needed. We can also define the mean SDS measure which is useful in finding the overall acoustical quality of matching between the two shapes:

$$\bar{V}_{S_1,S_2} = \frac{1}{M} \sum_{i=1}^{M} V_{S_1,S_2}(H_{S_1}(f_n, \theta_i, \phi_i), H_{S_2}(f, \theta_i, \phi_i))$$  (4.43)

It can be noted that SDS and ISSD metric proposed by Middlebrooks [1999b] have two main differences. Firstly, SDS does not apply any preprocessing in the form of applying an Equivalent Rectangular Bandwidth (ERB) filter on the DTF data. Secondly, the SDS measure has units in dB unlike the SSD measure which has units in dB$^2$. Fig. 4.28 shows a plot of the SDS for a given pair of subjects.

**Standard deviation in Spectral Difference**

![Figure 4.28: A plot of the function $V$ between two subjects $S_1$ and $S_2$ is shown above. The green crosses appearing in the plot are locations for which the error $V$ is large.](image)

Similar to the SGE measure, it is useful to see plots of the DTFs for the two
4.9 SMS-LDDMM Shape Matching

This section details the sequential multi-scale LDDMM (SMS-LDDMM) between two shapes \( S_1 \) and \( S_2 \). SMS-LDDMM algorithm maps two shapes using intuitive and natural transformations. Fig. 4.29 illustrates the mapping between two shapes using the SMS-LDDMM approach. In the first transformation Fig. 4.29(1), the two shapes are matched in terms of their orientation, that is, the source shape is rotated clock wise (or anti-clock wise) by an angle of \( 90^\circ \). In the second transformation shown in Fig. 4.29(2), large scale features, such as the corners of the source shape, are transformed to the target shape, also in this transformation part of the fluctuations on the sides have been captured. In the third transformation the finer details for the two shapes are matched, which consists of rapid fluctuations on the side of the shapes. The third transformation is seen in Fig. 4.29(4).

The SMS-LDDMM method makes use of the LDDMM scale parameters \( \sigma_V \) and \( \sigma_W \) (Sec. 4.4) in order to vary the physical scale at which the shape matching algorithm operates. In the SMS approach we consider shape matching that is performed sequentially at successively smaller scales.

The SMS-LDDMM algorithm is described in Algorithm 3. The algorithm starts by mapping the source shape to the target shape at the scale defined by \( \sigma_V(1) \) and \( \sigma_W(1) \). The momentum vectors obtained are then used to flow the source shape diffeomorphically to the target shape resulting in a shape \( T_1 \) that matches the target shape at the first scale. \( T_1 \) then becomes the source shape for the mapping applied at the second scale. These operations are repeated for each successive scale specified by \( \sigma_V(l) \) and \( \sigma_W(l) \), where \( \sigma_V(l + 1) < \sigma_V(l) \) and \( \sigma_W(l + 1) < \sigma_W(l) \).
4.9. SMS-LDDMM Shape Matching

Figure 4.29: The above figure shows the intuition behind the SMS-LDDMM approach to shape mapping. The SMS-LDDMM approach ensures that the source and target shapes are mapped using the most natural and smooth transformations. In the first transformation (1) the orientation of the source and target shapes are mapped, while in the second transformation the large scale features of the source shape is mapped to the target shape (i.e. the corners). In the final transformation (3) the finer structure of the shapes such as the fluctuations on the side of the shapes are mapped.

Algorithm 3 SMS-LDDMM

inputs: $S_1, S_2, [\sigma_V(1), \ldots, \sigma_V(L)], [\sigma_W(1), \ldots, \sigma_W(L)]$.
outputs: $\{\alpha_l(t)\}, T_l$, for $l = (1 \ldots L)$.
1: $T_0 \leftarrow S_1$
2: for $1 \leq l \leq L$ do
3: $\{\alpha_l(t)\} \leftarrow \mathcal{M}(T_{l-1}, S_2, \sigma_V(l), \sigma_W(l))$
4: $T_l \leftarrow \mathcal{F}(T_{l-1}, \{\alpha_l(t)\}, \sigma_V(l))$
5: end for

Fig. 4.30 illustrates the SMS-LDDMM algorithm applied to match two distinct ears in the SYMARE database using four successively smaller scales. The scale parameters for the first scale, $\sigma_V(1)$ and $\sigma_W(1)$, are relatively large and have been chosen to capture differences in scale, rotation and translation. The scale parameters for the second scale, $\sigma_V(2)$ and $\sigma_W(2)$, are a factor of four times smaller than the first scale and have been chosen to capture important structural differences between the shapes. Further, for the following two scales, we continue to divide the scale parameters by a factor of four to capture increasingly finer details in the target shape.

It is to be noted that the final shape deformation, $\phi_{\text{SMS}}(t, X)$, can be described by the composition of the diffeomorphic flows at the different scales, i.e., $\phi_{\text{SMS}}(t, X) =$
\[ \phi^{V_4} \circ \phi^{V_3} \circ \phi^{V_2} \circ \phi^{V_1}(t, X). \]

Figure 4.30: SMS-LDDMM is applied at four successive scales: (a) source shape \( S_1 \) (b) shape \( T_1 \) \((\sigma_V(1) = 160; \sigma_W(1) = 20)\), (c) shape \( T_2 \) \((\sigma_V(2) = 40; \sigma_W(2) = 5)\), (d) shape \( T_3 \) \((\sigma_V(3) = 10; \sigma_W(3) = 2.5)\), (e) shape \( T_4 \) \((\sigma_V(4) = 2.5; \sigma_W(4) = 0.625)\), (g) target shape \( S_2 \). The colours indicate the distance from the target shape.

### 4.9.1 Conclusion

This chapter showed the procedure on how a source head and ear shape can be deformed using high resolution meshes to a target head and ear shape in which only the ear shapes are mapped using high resolution meshes. Further, this chapter discussed the LDDMM scale parameters \( \sigma_V \) and \( \sigma_W \) and performed morphological experiments on the ear shapes to examine a suitable range for the selection of these parameters. A threshold value was set based on the average error in the ear shapes as computed by the CSDA technique and visual inspection. The \( \sigma_V \) and \( \sigma_W \) values that satisfied the threshold and further showed a good quality of mapping for the ear shapes were \( \sigma_V = 10(\text{mm}) \) and \( 1(\text{mm}) \leq \sigma_W < 5(\text{mm}) \). Further, for the purposes of the experiments conducted in this thesis it was explained that the head mesh and the ear meshes are mapped using the same value of \( \sigma_V^E = \sigma_V^H = 10(\text{mm}) \). For the shape comparison parameter the value of \( \sigma_E^V \) and \( \sigma_E^H \) can be chosen proportional to the size of the elements in the mesh. Typically proportional to the mean edge length or maximum edge length. This said for the ear shapes examined in this thesis the value for \( \sigma_W^E = 1.25(\text{mm}) \) unless stated otherwise.
Chapter 5

ESTIMATING A TEMPLATE HEAD
TORSO AND EAR SHAPE

This chapter presents the methodology for the construction of the template, or average, ear head and torso shapes. The method for constructing the template shapes using both a single and multi-scale approach is presented here. It may be recalled that the multi-scale approach for mapping two shapes was previously discussed in Sec. 4.9 and in this chapter a similar approach is utilized to construct a multi-scale template ear shape. Further, this chapter describes the validation of the template shapes (i.e. single and multi-scale) by performing a statistical analysis on important features of the template shapes and comparing them to the population of ears.

Determining a population average ear shape is an important step in studying the statistics of ear shapes because it allows the variations in ears to be studied relative to a reference shape. There are several methods for estimating an average template shape using the LDDMM framework (Cury et al. 2013, 2014; Durrleman et al. 2008; Glaunès et al. 2006; Vaillant et al. 2004). In brief, it can be stated that these methods work at a single, fixed LDDMM scale, which is usually chosen to be small relative to the shape features of interest. In other words, the metric distance used for constructing the template ear shape is defined for shape transformations
that operate on a single, fine physical scale.

The template ear shapes described in this chapter were computed using two procedures. In the first procedure the template ear shape was computed using existing single-scale LDDMM template estimation procedures (Cury et al., 2013, 2014; Vaillant et al., 2004). The single scale template estimation utilizes only a single metric induced by the LDDMM framework. In the second procedure, a new multi-scale template estimation procedure based on the LDDMM framework is proposed (Zolfaghari et al., 2014).

The aim of the multi-scale template estimation procedure is to estimate a template ear shape that could be obtained by smooth and natural transformations by utilizing multiple metrics in the LDDMM framework. In this estimation procedure the estimated template shape is sequentially refined at increasingly finer physical scales. The construction of the multi-scale template, also referred to by the average ear shape, is closely related to the multi-scale mapping (SMS-LDDMM) (Zolfaghari et al., 2014) of two ear shapes (Sec. 4.9). It might be recalled that the SMS-LDDMM method enables one to robustly deform a source shape into a target shape with high precision. In order to speed the calculation for the single and multi-scale template estimation procedures, a barycentre shape is used as an initializer or starting shape for the template estimation procedure. It may be noted that the barycentre shape can be computed relatively quickly in comparison to the main template estimation procedure for a given population of shapes, which can assist in speeding up the generation of the final template.

An important analysis presented in this chapter is to validate the template shapes by statistically comparing some of the important features of the template ear shapes to the shapes in the SYMARE database. Anthropometric measurements on the population of ear shapes, and, also, for the template shape were conducted using 3D points similar to the points defined by Jin et al. (2000). Also, 2D measurements by means of 2D images similar to the measurements defined in the CIPIC database (Algazi et al., 2001) were obtained. By performing a statistical analysis on the anthropometric
measurements for the template ears and the SYMARE ear population, it is identified that the template ear shapes have features that are close to the population mean and consistently within [-1,+1] standard deviation from the mean. Moreover, it is shown that the template ear shapes have features that fall within [+,-1] standard deviation from the mean of the CIPIC database measurements. This chapter also describes the generation of the template head shape using the single-scale LDDMM template estimation procedure. It is shown that the template head shape has anthropometric sizes that are close to the mean size of the head shapes in SYMARE.

In addition to the previous analysis, and to further validate the template ear shapes, an analysis on the orientation (i.e angle and scale factor) of the ear shapes with respect to the template ear shape is performed. The analysis on the orientation was conducted by performing a rigid matching between the template shape vertices and vertices for each subject within the population. Specifically, Tait-Bryan angles, which will be presently described, were calculated with respect to the template shape. By examining the Tait-Bryan angles and the scale factors from the rigid matching, it was identified that the orientation and scale of the template ear shapes are also within close proximity to the population mean.

Finally, a morphological and acoustical comparison is performed between the single and multi-scale template ear shapes by using the tools developed in Sec. 6. By comparing the two template ear shapes it is shown that the large anthropometric features of the multi-scale and single scale template shapes are similar. However, because the two template ears differ in finer detail their corresponding acoustics are different.

In light of the above, the contributions of this chapter can be summarized as follows:

1. The construction of the single and multi-scale template ear shape Sec. 5.1.2
2. The construction of the template head and torso shapes Sec. 5.2.2
3. Validating the template ear shapes by comparing the anthropometric features
of the template ear shapes to the SYMARE and CIPIC population Sec. 5.2.3.1 - Sec. 5.2.3.2.

4. Validating the template ear shape orientations (angle and scale factor) by statistically comparing the angle and scale factor of the template ear shapes using rigid matching to the SYMARE population Sec. 5.2.4.

5. Displaying the differences in the multi-scale and single scale template ears by morphoacoustically comparing the multi-scale and single-scale template shapes using the morphological and acoustical comparison tools previously developed Sec. 5.2.5.

This chapter is composed of a methods and results section. The methods section describes the computation of the bary-centre shape in Sec. 5.1.1. After calculating the barycentre shape, the procedure for the single and multi-scale template estimation are discussed in Sec. 5.1.2 and Sec. 5.1.3 respectively. In the results section the template ear shapes using the single and multi-scale template estimation procedure are presented Sec. 5.2.1. Further, in the same section Sec. 5.2.1, the evolution of a selected population of ear shapes using the multi-scale template ear procedure are also shown. In Sec. 5.2.2 the template head and torso shapes are given. In Sec. 5.2.3, statistical analysis on 2D Sec. 5.2.3.1 and 3D Sec. 5.2.3.2 shapes are provided. An analysis on ear shape angles is given in Sec. 5.2.4. Finally, in Sec. 5.2.5 a morphoacoustic comparison between the single and multi-scale template ear shapes is presented.

5.1 Method

5.1.1 Barycentre

Template estimation for a population of shapes in the LDDMM framework is a lengthy and iterative process requiring intense computations. In order to accelerate the computational process, it is advantageous to start with an initial template that is well-centred within the population. One method to rapidly estimate a centred initial
shape is to use the barycentre or iterative centroid algorithm described in [Cury et al., 2014]. The barycentre algorithm is described in Algorithm 4 and closely resembles the running average procedure used in data statistics. The algorithm begins by setting shape $S_1$ as the first estimate of the centred template shape, $B$. It then maps shape $B$ to the second shape in the population, $S_2$, to obtain momentum vectors. These momentum vectors are used to flow shape $B$ diffeomorphically halfway towards shape $S_2$ by setting $t_{\text{end}} = \frac{1}{2}$ in the flow function. The newly obtained shape provides a second estimate of the centred template shape and replaces the old $B$. The process is then repeated for the remaining shapes in the population with $t_{\text{end}} = \frac{1}{r}$ for iteration $r$, so that the flow distance is inversely proportional to the number of shapes contributing to the current running estimate of the centred template shape. In this way we obtain a initial template shape that is well-centred within the population of shapes.

Algorithm 4 Barycentre

inputs: \{\text{S}_1, \ldots, \text{S}_R\}, \sigma_V, \sigma_W.
outputs: B

1: \text{B} \leftarrow \text{S}_1
2: \text{for} \ 2 \leq r \leq R \ \text{do}
3: \ \\ \{\alpha^r(t)\} \leftarrow \mathcal{M}(\text{B}, \text{S}_r, \sigma_V, \sigma_W)
4: \ \text{B} \leftarrow \mathcal{F}(\text{S}_r, \{\alpha^r(t)\}, \sigma_V, 0, \frac{1}{r})
5: \ \text{end for}

5.1.2 Single Scale Template Estimation

The template estimation procedure presented here utilizes the properties of the initial momentums within the larger LDDMM mapping framework (Miller et al., 2006). The Fréchet mean, or ideal template, $T$, for the population \{\text{S}_1, \ldots, \text{S}_R\} of shapes is defined as the shape $T$, where the sum of geodesic LDDMM distances to all other shapes in the dataset is minimal:

$$ T = \arg\min_U \sum_{r=1}^R \int_0^1 \|v_r(t)\|^2_{\nu} dt $$ (5.1)
where \( v_r(t) \) provides exact mappings from the source shape \( U \) to target shapes \( S_r \) in the dataset for every \( r, 1 \leq r \leq R \). Due to the complicated and high dimensional structure of the space of shapes, obtaining the Fréchet mean can be very hard, particularly given that the number of iterations in the gradient descend algorithm is often limited due to time constraints.

In this thesis, the template shape is estimated using a method similar to the template estimation technique detailed in [Vaillant et al. (2004)]. This method uses landmarks for comparing shapes when defining the LDDMM cost function \( J \). However, in the scheme presented here the framework of currents is used. By utilizing the framework of currents all elements (i.e. faces) in the source and target meshes for comparing and ultimately mapping two surfaces is used. Further in the template estimation procedure detailed here we make use of a barycentre shape \( B \) as the initial template shape ([Zolfaghari et al.] 2014). The template shape is defined to be the shape for which the squared norm of the initial momentum vectors from the template shape to all other shapes is zero.

\[
T = U \quad (5.2)
\]

where \( \sum_{r=1}^{R} \alpha_r(0) = 0 \quad (5.3) \)

and \( \{ \alpha_r(t) \} = M(U, S_r, \sigma_V, \sigma_W) \). (5.4)

In practice the value of \( \sum_{r=1}^{R} \alpha_r(0) \) will be small or close to zero due to the limited number of iterations involved.

Figure 5.1 shows the relationship between the barycentre mean (shown in red), the Freche mean (shown in dark blue), and three random points on a sphere (shown in green). The surface of the sphere represents a simple Riemannian space. The three green stars on the surface of the sphere represent points in a Riemannian space. The black lines connecting the points are the optimal geodesic paths connecting the points in this Riemannian setting. It is to be noted that other geodesic paths connecting the three points also exist but are not optimal.
As discussed previously the template estimation procedure can be initiated by first computing a barycentre, $B$, for the population of shapes. The template estimation procedure is an iterative procedure and is composed of 3 major steps:

- The template at the current iterations is mapped to each shape, $S_r$, in the population to obtain the initial momentums, that is Step 11 in Alg. 5.
- The initial momentums are then averaged to form a mean initial momentum vector $\{\tilde{\alpha}\}$ that is Step 12 in Alg. 5.
- a new estimate of the template shape is then obtained by shooting (using Equation 2.31) from the current template shape $T$ and the average initial momentums, that is Step 14 in Alg. 5.

The major steps (1)-(3) detailed above are iterated several times until a given convergence criterion is reached. In Algorithm 5 the number of iterations is denoted by the letter $I$. It is to be emphasized that the barycentre shape, $B$, is usually used as the initial template shape (initializer) for the first iteration of the template estimation procedure. However, if the barycentre shape is not available, then the first shape $S_1$ can be used as the starting shape for the template estimation procedure.
Algorithm 5 Template Estimation

inputs: \( \{S_1, \ldots, S_R\}, \sigma_V, \sigma_W, \) optional: \( B, I. \)
outputs: \( T. \)
1: if \( B \) is not provided then
2: \( T \leftarrow S_1 \)
3: else
4: \( T \leftarrow B \)
5: end if
6: if \( I \) is not provided then
7: \( I \leftarrow 20 \)
8: end if
9: for \( 1 \leq i \leq I \) do
10: \( \text{for } 1 \leq r \leq R \) do
11: \( \{a^r(t)\} \leftarrow \mathcal{M}(T, S_r, \sigma_V, \sigma_W) \)
12: \( \{\tilde{a}\} = \frac{1}{r+1}\{\tilde{a}\} + \frac{r}{r+1}\{a^r(0)\} \)
13: end for
14: \( T = S(\{\tilde{a}\}, T, \sigma_V) \)
15: end for

5.1.3 Sequential Multi Scale Template Estimation

There were two guiding motivations for computing the multi-scale template estimation procedure. Broadly, the first motivating factor related to obtaining a template ear by utilizing different LDDMM metrics, and the second motivating factor was concerned with the shape of the template ear itself.

More specifically, given that LDDMM induces a metric in the highly non-linear Riemannian space of shapes, the first motivating factor utilizes the metric induced by the LDDMM scale parameters, \( \sigma_V \) and \( \sigma_W \), at successively smaller scales (i.e. values) to compute a template shape. It is noted that LDDMM scale parameters, \( \sigma_V \) and \( \sigma_W \), are physical scale parameters and different values of \( \sigma_V \) and \( \sigma_W \) induce a different LDDMM metric in the space of shapes. Further, the LDDMM scale parameters have a direct impact on the deformations obtained between the ear shapes such that large values of \( \sigma_V \) imply a larger coupling between neighbouring vertices in the source shape, at the time of deformation while smaller values of \( \sigma_V \) imply a smaller coupling between the vertices. Also, when performing LDDMM between two ear shapes, larger values of \( \sigma_W \) implies a coarser comparison of two surfaces, while a
smaller $\sigma_W$ value implies a finer comparison between two surfaces using the framework of currents. The multi-scale approach to the template estimation effectively incorporates multiple $\sigma_V$ and $\sigma_W$ LDDMM scales and metrics for the construction of the template ear. In other words the template ear shape is progressively refined when moving from larger scales, hence metrics, to smaller scale values.

The second motivating factor for the use of the multi-scale template estimation procedure was to obtain a template ear shape which is a result of smooth and natural LDDMM deformations. The deformations start from a high level, which resemble a rotation, translation and scale operation and then progress to finer and more detailed deformations of the template shape. An example of smooth and natural deformations can be seen when mapping two ear shapes using SMS-LDDMM, which is shown in Fig. 4.30.

The SMS approach to template estimation is described in Algorithm 6. Essentially repeats the template estimation procedure described previously for successively smaller scales. The objective of the SMS template estimation is to converge to a template that satisfies (or approximately satisfies) Eq. (5.1) at every scale. To this end, after a template, $T^l$, is generated for a scale, $l$, we map and then diffeomorphically flow the population of shapes $\{S_{l-1}^1, \ldots, S_{l-1}^R\}$ to this template to obtain a new population of shapes $\{S^l_1, \ldots, S^l_R\}$ (refer to lines 6 and 7 of Algorithm 6). This new population of shapes then becomes the starting population for the next iteration of the SMS template estimation algorithm. This is repeated until we reach the final scale $L$. 
Algorithm 6 SMS Template Estimation

inputs: \( \{S_1, \ldots, S_R\}, \ [\sigma_V(1), \ldots, \sigma_V(L)], \ [\sigma_W(1), \ldots, \sigma_W(L)]. \)

outputs: \( T^L. \)

1: \( S^0_r = S_r \) for \( r = 1, \ldots, R \)

2: for \( 1 \leq l \leq L \) do

3: \( B^l \leftarrow \text{barycentre}(S^{l-1}_1, \ldots, S^{l-1}_R, \sigma_V(l), \sigma_W(l)) \)

4: \( T^l \leftarrow \text{TempEstim}(S^{l-1}_1, \ldots, S^{l-1}_R, \sigma_V(l), \sigma_W(l), [T_0]) \)

5: for \( 1 \leq r \leq R \) do

6: \( \{\alpha^r(t)\} \leftarrow \mathcal{M} (S^{l-1}_r, T^l, \sigma_V(l), \sigma_W(l)) \)

7: \( S^l_r \leftarrow \mathcal{F} (S^{l-1}_r, \{\alpha^r(t)\}, \sigma_V(l)) \)

8: end for

9: end for

5.2 Results

5.2.1 Single and multi-scale template ear shapes

Single and multi-scale template estimation algorithms were conducted for the left and mirrored right ear shape populations in the SYMARE database. A total of 118 ear shapes belonging to 62 subjects were used for the computation of the average ear shapes. A few ears were omitted from the computation of the average shape due to abnormalities on the surface meshes. The right ear shapes were mirrored by simply negating the "y" coordinate of the points. The "y" coordinate corresponds to the interaural axis passing through the left and right ear shapes. The single scale template estimation procedure was conducted in the manner described in Algorithm 5. For the single scale template estimation procedure the optional barycentre shape \( B \) was also estimated and used as an initializer to the template estimation. The LDDMM scale values used for the estimation of the template were \( \sigma_V = 10 \) and \( \sigma_W = 1.25 \). Fig. 5.2 shows the barycentre shape \( B \) and the template shape \( E_{TSS} \) calculated at a single scale. Further Fig. 5.3 shows the sum of the geodesic distance between the template shape at every iteration of the template estimation procedure (i.e. Algorithm 5) and all shapes involved in the computation of the template ear shape. Note that the geodesic distance or the energy of the deformation was previously discussed in Sec. 2.8.1 and more specifically in Eq. (2.25).
5.2. Results

Figure 5.2: Template shapes from single scale template estimation procedure computed at LDDMM scale values of \( \sigma_V = 10 \), \( \sigma_W(1) = 1.25 \).

Figure 5.3: The sum of geodesic distances from the single scale template shape to all shapes used in the computation of the average ear. The template shape at the first iteration is the barycentre shape.

The multi-scale template ear shape denoted by \( E_{T_{MS}} \) was also estimated using the procedure detailed in Algorithm 6 on the full population of left and mirrored right ear shapes as in the single scale template estimation procedure. The LDDMM scale parameters used for scales 1-4 were \( \sigma_V = [160, 40, 10, 2.5] \) and \( \sigma_W = [20, 5, 1.25, 0.3125] \).

Fig. 5.4 displays a sample of the starting population of ear shapes from the SYMARE database and Fig. 5.5-Fig. 5.7 presents the evolution of the sample ear shapes from large to small LDDMM scale values in the multi-scale template estimation procedure. The evolution for many other ear shapes can be seen in Appendix B. The evolution of the barycentre shape \( B \) and the estimated template shape \( T^l \) are shown in Fig. 5.9. Fig. 5.8 shows the sum of the geodesic distance between the template shape at every iteration of the template estimation procedure and all shapes involved in the computation of the average ear and for every scale of the multi-scale template estimation.
5.2. Results

Figure 5.4: Original and starting population of ears for template \( T^1 \) computed at scale: \( \sigma_V(1) = 160, \sigma_W(1) = 20 \).

Figure 5.5: Starting population of ears for template \( T^2 \) computed at scale: \( \sigma_V(2) = 40, \sigma_W(2) = 5 \).

Figure 5.6: Starting population of ears for template \( T^3 \) computed at scale: \( \sigma_V(3) = 10, \sigma_W(3) = 1.25 \).

Figure 5.7: Starting population of ears for template \( T^4 \) computed at scale: \( \sigma_V(3) = 2.5, \sigma_W(3) = 0.3125 \).
5.2. Results

(a) (b) (c) (d)

Figure 5.8: The sum of geodesic distances from the template shape to all shapes used in the computation of the average ear shape. The template shape at the first iteration is the barycentre shape. (a) for the first scale corresponding to $\sigma_V = 160$ (b) for the second scale corresponding to $\sigma_V = 40$ (c) for the third scale corresponding to $\sigma_V = 10$, (d) for the fourth scale corresponding to $\sigma_V = 10$.

5.2.2 Template head, torso and ear shape

This section describes the production of a high resolution template head and torso shape from low resolution meshes by utilizing the LDDMM framework. The procedure for estimating the high resolution head, torso and ear shapes consisted of three stages. In the first stage, a low resolution template head torso and ear is estimated from low resolution meshes for the head, torso and ears. In the second stage by utilizing the inverse flow $\mathcal{F}^{-1}$ Eq. (4.6) in the LDDMM framework, a high resolution template head and torso was estimated. Finally, in the third stage the high resolution ear, which was previously estimated Sec. 5.2.1 was attached to the template head and torso shape. Now each of these three stages will be described in more detail.

Given the previous, the barycentre head and torso shape was computed using Algorithm 4 from low resolution meshes of the ear, head and torso in SYMARE.
5.2. Results

The reason for computing the head and torso template using low resolution meshes was because features in the head and torso shape do not contribute significantly to the overall acoustical properties of an ear, head and torso shape (Takemoto et al., 2012b; Tao et al., 2003a). For this reason, the head and torso template shapes were computed on low resolution meshes that were suitable for BEM simulations up to a critical frequency of $f_c = 8kHz$. Using the barycentre shape, the low resolution template and head and torso shape $HTE_{T,LR}$ was estimated using Algorithm 5.

In the second stage, in order to obtain a high resolution template head and torso shape, $HTE_T$, first a mapping operation was performed between the low resolution template head and torso shape to a low resolution target shape $HTE_{x,LR}$. Subsequently, an inverse flow operation $F^{-1}$ Eq. (4.6) was performed from the high resolution target head and torso shape $HTE_{x}$ in SYMARE towards the template shape by using the deformation data from matching the template $HTE_{T,LR}$ to the shape $HTE_{x,LR}$. It is reiterated that the deformation information is computed at a low resolution and applied to a high resolution mesh shape in order to obtain a high-resolution.
5.2. Results

Figure 5.10: low resolution mesh of the template head torso and ear shape. Please zoom in if the shape cannot be seen (The above image is embedded as 3D and it can be zoomed and rotated)

resolution template head and torso Eq. (5.5). The deformation from the template shape to the first subject $x = 1$ was chosen for the inverse flow operation.

$$
HTE_T = \mathcal{F}^{-1}(\{a(t)\}^HTE_{T,LR} \rightarrow HTE_{T,LR}, HTE_T, \sigma_V = 10)
$$

$$
= F(\{a(t)\}^HTE_{T,LR} \rightarrow HTE_{T,LR}, HTE_T, \sigma_V = 10, t_s = 1, t_e = 0)
$$

Finally, in the third stage the software Geomagic, was used to attach the high resolution ear shape (estimated previously) to the high resolution template shape $HTE_T$. In order to do this, both the left and right ears in the existing template shape $HTE_T$ were removed using Geomagic and the high resolution single scale $E_{TSS}$ and multi-scale $E_{TMS}$ template ear shapes, were attached as both the left and right ear shapes to the template head and torso shape. Fig. 5.12 shows the high resolution head, torso and ear shape with the high resolution single scale template ears attached, and this is denoted as $HTE_{TSS}$. Further, Fig. 5.11 shows the high resolution template head, torso shape with the high resolution multi-scale template ear shape attached, denoted as $HTE_{TMS}$. 
5.2. Results

Figure 5.11: Template head, torso and ear shape with the high resolution multi-scale ear template attached $HTE_{T+G}$. Please zoom in if the shape cannot be seen (The above image is embedded as 3D and it can be zoomed and rotated).

5.2.3 Comparing anthropometric features in the template head and ear shape to the SYMARE population

The CIPIC database (Algazi et al., 2001) is a well known large database of anthropometric and HRTF measurements. It contains anthropometric measurements from 45 male and female subjects. Nine anthropometric measurements are available for the pinna region, and 17 anthropometric measurements are available for the head and torso region. The measurements for the head and pinna are shown in Fig. 5.13.

This section describes the analysis and procedure for obtaining anthropometric measurements from the population head and ear shapes. It also validates the template shapes by comparing the anthropometric features in the template shape to that of the population. Both the single and multi-scale template ear shapes are examined in this section.

Reasoning intuitively, the template head, torso and ear shapes should have anthropometric features resembling the average anthropometric features of the population they were computed on. For example, the KEMAR mannequin (Burkhard...
5.2. Results

Figure 5.12: Final high resolution template head, torso and ear shape with the high resolution single scale ear template attached as both the left and right ear shapes $HTE_{TS}$. Please zoom in if the shape cannot be seen (The above image is embedded as 3D and it can be zoomed and rotated)

and Sachs, 1975) was constructed based on average ear and facial measurements from a large population of human subjects.

In the work presented here three sets of analysis and measurements were performed on the template and SYMARE ear shapes as follows:

- A set of measurements and analysis was conducted using 2D images of the left and right ear shapes in SYMARE. These measurements are further compared to the measurements in the CIPIC database as well.

- A set of analysis and measurements using 3D points on surface meshes are performed.

- An analysis was done on ear shape angles and scale factors with respect to the template ear shape using the rigid matching technique (Sec. 5.2.4.1).

5.2.3.1 2D Anthropometric Measurements

Anthropometric measurements for the ear shapes using 2D images were taken for the template shape and shapes in SYMARE using a measurement scheme similar to the CIPIC database (Algazi et al., 2001). Measurements were taken from 118 ear
shapes that were used for the computation of the template shapes. Using MATLAB, snapshot of ear shapes were obtained using a predefined view angle, zoom factor, camera position, and camera target. Using a ruler, measurements were taken for the ear shapes similar to measurements defined in the CIPIC database. Fig. 5.13 shows the measurements defined in the CIPIC database for the ear shape.

![Diagram of ear measurements](image)

Figure 5.13: The above figure displays the measurements for the pinna, head and torso shapes taken for the CIPIC database. Image taken from Algazi et al. (2001)

Only five CIPIC distances were measured for the ear shapes. These are $d_1, d_2, d_3, d_5, d_6$ which, in turn, labelled $d_1^{2D}, d_2^{2D}, d_3^{2D}, d_5^{2D}$ and $d_6^{2D}$. Distances $d_1^{2D} + d_2^{2D}$ correspond to the Concha height, distance $d_3^{2D}$ measures the Concha width and distance $d_5^{2D}$ and $d_6^{2D}$ correspond to the pinna height and width respectively. Fig. 5.14 shows a scanned 2D image of the template ear shape where measurement were taken using a ruler. Measurements for ear shapes in the SYMARE database were taken in a similar manner and are shown in Appendix C. Fig. 5.15 shows histogram plots for each of the measured distances. Each of the stem lines in the plot are described below the plot, further Table 5.1 summarises and compares 2D measurements for the ear shapes.
Figure 5.14: The above figure shows an example (Template) ear shape where measurements similar to the ones defined in the CIPIC database were taken.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Description</th>
<th>CIPIC (Mean cm)</th>
<th>CIPIC (std)</th>
<th>SYMARE (Mean cm)</th>
<th>SYMARE (std)</th>
<th>$s_1$</th>
<th>$s_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_{11}^1 + d_{11}^2$</td>
<td>Concha height</td>
<td>2.5478</td>
<td>0.2058</td>
<td>2.4140</td>
<td>0.2411</td>
<td>2.5000</td>
<td>2.5000</td>
</tr>
<tr>
<td>$d_{21}^1$</td>
<td>Concha width</td>
<td>1.5484</td>
<td>0.2547</td>
<td>1.5863</td>
<td>0.2833</td>
<td>1.5500</td>
<td>1.5500</td>
</tr>
<tr>
<td>$d_{22}^1$</td>
<td>Pinna height</td>
<td>6.3946</td>
<td>0.5290</td>
<td>6.3449</td>
<td>0.5414</td>
<td>6.3500</td>
<td>6.3500</td>
</tr>
<tr>
<td>$d_{23}^1$</td>
<td>Pinna width</td>
<td>2.8862</td>
<td>0.2797</td>
<td>2.9714</td>
<td>0.2945</td>
<td>2.9000</td>
<td>2.9000</td>
</tr>
</tbody>
</table>

Table 5.1: Statistics for the pinna measurements from the CIPIC database and SYMARE database using 2D images.
Figure 5.15: Statistics on anthropometric measurements from ear shapes in the SYMARE database and the single and multi-scale template ear shapes. Measurements adhere to the CIPIC database measurements (a) \(d_1^{2D} + d_2^{2D}\), (b) \(d_3^{2D}\), (c) \(d_5^{2D}\), (d) \(d_6^{2D}\). Multiscale template shape (black line), Single-scale template shape (cyan), Population mean (red line), Population standard deviation (yellow lines), Population median (green line), CIPIC mean (blue line), CIPIC standard deviation (magenta lines).
5.2. Results

5.2.3.2 3D Anthropometric Measurements

Anthropometric measurements were also conducted by using the Geomagic software on 3D mesh shapes. Landmark points similar to the points shown in [Jin et al., 2000] were selected on the template head and ear shape and also on the corresponding head and ear population in the SYMARE database. By using these landmark points, anthropometric measurements were computed from the template and the SYMARE population head and ear shapes. Thirteen points were measured for each of the left and right ears, and a total of six points were measured on the head shape. Points $p_1$-$p_{13}$ are for the left ear shape, points $p_{14}$-$p_{26}$ for the right ear, and points $p_{27}$-$p_{33}$ are for the head shape. Fig. 5.16 shows measurement points (red dots) on the template head and ear shape. Measurement points for all shapes in the SYMARE database can be found in Appendix D.

The first set of distances were between point $p_1$ and points $p_{2}$-$p_{13}$ with a total of 12 distances. These distances are termed $d_{3D}^{1}$-$d_{3D}^{12}$ and correspond to measurements for the left ear shape. The second set of distances between point $p_{14}$ and points $p_{15}$-$p_{26}$, which correspond to measurements for the right ear shapes. The distances for the left and right ear shapes were concatenated for plotting. Histogram plots for distances $d_{3D}^{5}$, $d_{3D}^{7}$, $d_{3D}^{8}$, $d_{3D}^{9}$, $d_{3D}^{11}$, and $d_{3D}^{12}$, for both the left and right ear shapes, are shown in Fig. 5.17. Further, Table 5.2 tabulates the statistics for 3D measurements on SYMARE and the two template shapes. Distances $d_5$ and $d_8$ correspond to the pinna size, $d_7^{3D}$ to the size of the Lobe, and distances $d_9^{3D}$, $d_1^{3D}$, and $d_2^{3D}$ correspond to the Concha size.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Description</th>
<th>SYMARE (mean mm)</th>
<th>SYMARE (std)</th>
<th>$E_{50}$</th>
<th>$E_{90}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_{3D}^{5}$</td>
<td>Pinna size</td>
<td>48.3938</td>
<td>4.0277</td>
<td>47.0130</td>
<td>47.4125</td>
</tr>
<tr>
<td>$d_{3D}^{7}$</td>
<td>Lobe Height</td>
<td>19.0261</td>
<td>2.6113</td>
<td>18.3512</td>
<td>18.6874</td>
</tr>
<tr>
<td>$d_{3D}^{8}$</td>
<td>Pinna size</td>
<td>44.3901</td>
<td>4.1115</td>
<td>44.1009</td>
<td>41.9823</td>
</tr>
<tr>
<td>$d_{3D}^{9}$</td>
<td>Concha size</td>
<td>26.4928</td>
<td>2.6963</td>
<td>26.7447</td>
<td>26.2141</td>
</tr>
<tr>
<td>$d_{3D}^{11}$</td>
<td>Concha size</td>
<td>24.5301</td>
<td>2.4869</td>
<td>24.9915</td>
<td>22.6232</td>
</tr>
<tr>
<td>$d_{3D}^{12}$</td>
<td>Concha size</td>
<td>11.0812</td>
<td>1.5768</td>
<td>10.9830</td>
<td>10.8334</td>
</tr>
</tbody>
</table>

Table 5.2: Statistics for the pinna measurements from the SYMARE database using 3D shapes.
Figure 5.16: Land mark points on the template ear and head shape used for anthropometric measurements. Similar points were measured for subjects in the SYMARE database. Points $p_1$-$p_{13}$ are for the left ear shape, points $p_{14}$-$p_{26}$ are for the right ear shape, and points $p_{27}$-$p_{33}$ are for the head shape.
Table 5.3: Statistics for the head shape measurements from the SYMARE database using 3D surfaces.

The third set of anthropometric measurements were obtained for the head shape by calculating the distance between point $p_{27}$ and points $p_{28}$-$p_{31}$ on the template and SYMARE heads. These distances are labelled $d_{13}^{3D}$ . . . $d_{18}^{3D}$. Fig. 5.18 shows a histogram plot of the statistics, while Table 5.3 tabulates the statistics for 3D measurements on SYMARE heads and the template head shape.
5.2. Results

Figure 5.17: Statistics obtained by anthropometric measurements on 3D meshes for combined left and right ear shapes on the template head shapes and the SYMARE population. Multiscale template ear shape $E_{TMS}$ (black stem line), Single-scale template ear shape $E_{TSS}$ (cyan stem line), Population mean (red stem line), Population median (green stem line), +1 and -1 standard deviations from the population mean (yellow stem lines). (a) $d_{3D}^5$ (b) $d_{3D}^7$ (c) $d_{3D}^8$ (d) $d_{3D}^9$ (e) $d_{3D}^{11}$ (f) $d_{3D}^{12}$
Figure 5.18: Statistics obtained by anthropometric measurements on 3D meshes for the head shapes on the template and the SYMARE population. Template head shape $HE_T$ (black stem line), Population mean (red stem line), Population median (green stem line), +1 and -1 standard deviations from the population mean (yellow stem lines). (a) $d_{13}^{3D}$ (b) $d_{14}^{3D}$ (c) $d_{15}^{3D}$ (d) $d_{16}^{3D}$. 
5.2.4 Angle and scale factor measurements

This section presents an analysis in which the scale factor and rotation of the ear shapes is analysed with respect to the single and multi-scale template ear shapes. In order to define the angles of rotation with respect to the template ear shape we first define the standard coordinate system. The standard coordinate system is defined as the "x-axis" passing through the tip of the nose and the back of the head, the "y-axis" being the inter-aural axis passing through both the left and right ears, and the "z-axis" the vertical line through the centre of the head shape. The axis and the rotation angles are shown in Fig. 5.19 to provide the reader with an intuition. The Tait-Bryan angles for each ear with respect to the template ear shapes were calculated using the defined coordinate system by means of rigid matching two measures which represent the ear shapes. The following section discusses rigid matching in more detail.

Figure 5.19: The axis for which the rotation angles are computed is shown above the rotation about the $x,y$ and $z$ axis are denoted by $\theta_x, \theta_y$ and $\theta_z$. It is noted that the rotations are intrinsic, which means that the coordinate system changes by each rotation.
5.2. Results

5.2.4.1 Rigidly Aligning Ear Shapes Using Measures

Matching distributions and specifically measures is discussed in detail in Glaunès et al. (2004). In the current work this scheme was adapted to rigidly align ear shapes in the dataset to the template ear shape. First, we define the measure for the template and target ear shapes, the vertices for the template ear shape, $E_T$, are denoted as $x_i$ and the vertices for the target ear shape, $S_l$ are denoted as $y^l_j$. Here, the superscript $l$ indicates the subject number and the subscript $j$ is the index number for the vertex. Subsequently, the measures for each of the ear shapes are defined as:

$$
\mu = \frac{1}{m} \sum_{i=1}^{m} \delta x_i \quad (5.6)
$$

$$
\nu^l = \frac{1}{n_l} \sum_{j=1}^{n_l} \delta y^l_j
$$

In order to describe rigid matching between the two measures, $\mu$ and $\nu^l$, representing the two ear shapes, we define two spaces, whose detail can be found in Glaunès et al. (2004). The first space denoted by $I$ is a space of functions and has a norm of $||.||_I$. The second space is the dual space to $I$ and is denoted by $I^*$, the norm on the dual space is denoted as $||.||_{I^*}$. Rigid matching the two measures, $\mu$ and $\nu^l$, is obtained by minimizing a cost function $J(M, b)$, which is defined using the Hilbert norm of $||.||_{I^*}$. In Eq. (5.8), $M$ is a rotation and scale matrix and $b$ is a translation vector. The cost function $J(M, b)$, shown in Eq. (5.8), minimizes the norm between the rigidly transformed measure $\nu^l$ and the measure $\mu$. The rotation for the measure $\mu$ is defined as:

$$
\psi_{M,b}(\nu^l) = My^l_j + b, y^l_j \in R^3
$$

$$
J(M, b) = ||\psi_{M,b}(\nu^l) - \mu||_{I^*}^2, \quad (5.8)
$$

Before rigid matching the ear shapes in SYMARE with respect to the template ear shapes, all ears, included the template ear shapes were translated to the centre of
the coordinate system by means of a translation vector $t_l$. The translation vector $t_l$ is equal to the average of all the vertices in the given ear shape.

$$t_l = \frac{1}{V(l)} \sum_{j=1}^{V(l)} y_j$$  \hspace{1cm} (5.9)

Once the ear shapes were translated to the centre of the coordinate system using the translation vector $t_l$, the cost function $J(M_l, b_l)$ was performed for each of the shapes $l$ in the database to obtain rotation and scale matrices $M_l$ and translation vectors $b_l$ with respect to the template ear shapes. The rotation matrix $R_l$ with no scaling from the template ear shape to all the shapes is then obtained by $R_l = \left( \frac{M_l}{\det(M_l)} \right)^{-1}$. The Tait-Bryan angles, $\theta_x$, $\theta_y$, and $\theta_z$, can be calculated from the rotation matrices $R_l$ using Eq. 5.10

$$\theta_x = \arctan \left( \frac{R_l(3, 2)}{R_l(3, 3)} \right)$$

$$\theta_y = \arctan \left( \frac{-R_l(3, 1)}{\sqrt{R_l(3, 2)^2 + R_l(3, 3)^2}} \right)$$

$$\theta_z = \arctan \left( \frac{R_l(2, 1)}{R_l(1, 1)} \right)$$  \hspace{1cm} (5.10)

In Eq. 5.10, the angles are calculated according to the $z$-axis then the $y$-axis and the $x$-axis$^1$. Fig. 5.20 - Fig. 5.21 shows histogram plots for variations in $\theta_x$, $\theta_y$ and $\theta_z$ ear angles and the scale factor with respect to the multi-scale and single scale template ear shapes. Further, Table 5.4 and Table 5.5 tabulate the results for the ear angles for the single and multi-scale ear shapes.

### 5.2.5 Morphoacoustic Comparison of the Single and Multi-Scale Template Shapes

A comparison between the morphology and the acoustics of the single $E_{TSS}$ and multi-scale $E_{TMS}$ template ear shapes was conducted using the shape comparison

---

$^1$The angles can also be calculated using a different order such as with respect to the $x$ or $y$ axis first and this will result in different rotation angles with respect to each of the axis
5.2. Results

Figure 5.20: Statistics on ear shape Tait-Bryan angles obtained from rotation matrices $R_s$ with respect to the multi-scale template ear shapes. Template shape (black stem line), Population mean (red stem line), Population median (green stem line), +1 and -1 standard deviations from the population mean (yellow stem lines). (a) $\theta_x$, (b) $\theta_y$, (c) $\theta_z$, (d) scale factor

More specifically, BEM simulations were conducted on the template head and torso shape and head shape only with the single and multi-scale template shapes attached. Fig. 5.24 shows the difference in the morphology of the template ear shapes using the CSDA measure and, Fig. 5.24 compares the single and multi-scale template ear shapes acoustically when located on the template head and template head and torso shapes.
5.2. Results

Figure 5.21: Statistics on ear shape Tait-Bryan angles obtained from rotation matrices $R_s$ with respect to the single scale template ear shapes. Template shape (black stem line), Population mean (red stem line), Population median (green stem line), +1 and -1 standard deviations from the population mean (yellow stem lines). (a) $\theta_x$, (b) $\theta_y$, (c) $\theta_z$, (d) scale factor

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>SYMARE(mean)</th>
<th>SYMARE(std)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_x$</td>
<td>-1.4996</td>
<td>4.6282</td>
</tr>
<tr>
<td>$\theta_y$</td>
<td>-1.4704</td>
<td>5.9300</td>
</tr>
<tr>
<td>$\theta_z$</td>
<td>1.5435</td>
<td>7.3369</td>
</tr>
<tr>
<td>Scale Factor</td>
<td>0.9891</td>
<td>0.0575</td>
</tr>
</tbody>
</table>

Table 5.4: Statistics on the rotation of the ear shapes using the Tait-Bryan angles and the scale factor with respect to the single scale template ear shape $E_{TSS}$.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>SYMARE(mean)</th>
<th>SYMARE(std)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_x$</td>
<td>-1.8097</td>
<td>4.7084</td>
</tr>
<tr>
<td>$\theta_y$</td>
<td>-1.6966</td>
<td>6.2819</td>
</tr>
<tr>
<td>$\theta_z$</td>
<td>0.3437</td>
<td>7.2975</td>
</tr>
<tr>
<td>Scale Factor</td>
<td>0.9754</td>
<td>0.0609</td>
</tr>
</tbody>
</table>

Table 5.5: Statistics on the rotation of the ear shapes using the Tait-Bryan angles and scale factor with respect to the multi-scale template ear shape $E_{TMS}$. 
5.2. Results

Figure 5.22: Morphological comparison between the single and multi-scale template ear shapes. The regions of differences are highlighted by means of the CSDA technique.

Figure 5.23: Acoustical comparison between the single and multi-scale template ear shapes located on the template head and torso shape.
Figure 5.24: Acoustical comparison between the single and multi-scale template ear shapes located on the template head only shape.
5.3 Conclusion

In this section the new multi-scale template estimation procedure was introduced. It was explained that the new template estimation procedure develops the template at multiple physical scales using a different LDDMM metric at each scale. The procedure for constructing the template ear shape at a single scale was also explained. It was further detailed how this procedure can be made more efficient by the use of a barycentre shape. The production of the barycentre shape was further detailed.

Both the single and multi-scale template ear shapes were validated using anthropometric measurements. The anthropometric measurements from 2D images and 3D surface meshes indicated that important features in the template shapes, such as the size of the pinna and the size of the Concha, are close to the population average.

A further analysis was conducted to examine the rotation and scale factors of the ear shapes with respect to the template ear shapes using rigid matching of distributions. By using the template shapes as the reference shape, it was shown that the orientation of the population of ear shapes is centred around the template ear shapes and the scale factor for the template ear shapes is also close to the average scale factor of the population.

This chapter also detailed the procedure for developing the high resolution template head shape using low resolution meshes. The template head shape was validated by means of 3D anthropometric measurements similar to the ear shapes. It was shown that the template head contained features that were close to the mean of the population.

Finally, the multi-scale and single scale template ear shapes were compared morphologically and acoustically using the tools developed in the previous chapters. Based on the anthropometric measurements both template shapes contained large scale features that were similar. However, based on the acoustical simulations for the template ear shapes located on the template head and template head and torso shape, it was shown that the two ear shapes have different acoustical responses. In
particular it was identified that the two ear shapes have different acoustical responses in the range of 8kHz to 12kHz. Further, other differences were also highlighted in the GAAF figure. Finally, the morphology of the multi-scale template ear shape appears to be much smoother than the single-scale template ear shape. This and other differences using the CSDA measure were further highlighted in the shape comparison figure.

In view of the above, at this point, either the single or multi-scale template ear shapes can be used for further analysis. In the analysis presented in this thesis the multi-scale template ear shape has been chosen for simulations and analysis.
Chapter 6

EXAMINING DEFORMATIONS FROM THE TEMPLATE SHAPE

It may be recalled that the previous chapter detailed the procedure for estimating a template ear shape by the use of the LDDMM framework. It may be recalled that Chapter 4 showed how LDDMM and FM-BEM can be used for the study of the morphology and acoustics in this research project. Also it may be recalled that Chapter 4 described the procedure for mapping a source head and ear shape to a target head and ear shape and explained how the head and ear mesh is processed and made suitable for FM-BEM simulations. It also detailed how the FM-BEM simulations are performed on high resolution surfaces of the head and ear shape.

Using the procedures described in the previous chapters, this chapter studies the quality of the deformations from the template ear shape both morphologically and acoustically with respect to target ears in the SYMARE database by using the tools developed previously. Examining deformations from the template ear shape is important because the template ear shape and deformations from the template ear shape can be used for the modelling of target ears that exist in a database. However, this will be explained in more depth in later chapters (i.e. Sec. 7.7). At this point it

1For this chapter only the template shape used in the analysis, consists of the template head shape and the template ear with a small transformation. This was done by Meshlab to better position the template ear shape on the head shape. The template shape used here is shown in Fig. A.2
is more important to obtain an understanding on the accuracy of a model that is built using deformations from the template ear shape. Fig. 6.1 shows the deformed template shape, the template shape, and the target shape.

**Problem statement:** It is not clear how well the acoustic characteristics of the deformed template ear shape match that of the target ear shape when the template ear shape is deformed towards the target ear shape. There are several potential reasons for the inaccuracies that can occur for the deformations. Firstly, the ear shapes that exist in the SYMARE database are triangulated surfaces and every ear shape has a finite number of vertices and elements. Moreover, every ear shape is represented using different number of vertices and elements. Modelling and mapping a target ear shape with a limited number of vertices available both on the source template shape and target shape needs to be better understood. Secondly, computing the momentum vectors by minimizing the cost function $J$ (Eq. (2.22)) numerically can introduce small errors in the deformation. Thirdly, the use of the LDDMM scale parameters $\sigma_V$ and $\sigma_W$ greatly influences the quality of the deformations.

Deformations from the template ear shape to target ear shapes conducted at various values for the scale parameters were shown in Sec. 4.4. However, in the previous section, no acoustical study was conducted on the quality of the deformations. In this section we analyse the quality of the deformations acoustically which can be
valuable when building a model for ears based on the LDDMM framework.

Contributions of this chapter: As identified previously we need to analyse the similarity of shapes and also their corresponding acoustics. In this chapter the tools developed in Sec. 2.8.2 and Sec. 4.8 are used to conduct morphoacoustic comparison between the deformed template shape and the target shape. The morphoacoustic comparison refers to both the quality of the morphed template ear shape with respect to the target ear shape, and also the quality for the morphed template ear shape DTFs with respect to the target ear shape DTFs. An interesting result from the morphoacoustic analysis is to firstly identify morphological regions within the matched and target ear shapes that have the largest errors. Also, by the use of the global method of comparing DTFs, frequencies and spatial locations with large errors are identified and DTF’s belonging to these locations are shown.

Another contribution of this chapter is that ear shapes can be studied in multiple physical scales. Each physical scale corresponds to deformations that is conducted at a specific LDDMM scale value for $\sigma_V$ and $\sigma_W$. In the first and larger physical scale, mappings from the template ear shape to target ear shapes is conducted using $\sigma_V = 10$ and $1 \leq \sigma_W < 2$. The selection of these values was explained in Sec. 4.4. In the present chapter by acoustically examining the quality of the deformations it is shown that the morphoacoustic quality of mappings is satisfactory for only some of the ear shapes and that this can be improved for other ear shapes by the application of a second order mapping. More specifically, the second order mapping is carried out at a finer LDDMM scale parameter of $\sigma_V = 1$ and $1 \leq \sigma_W < 2$.

The final contribution of this chapter is an interesting analysis between the dissimilarity in the morphology and the acoustics of deformed ear shapes along the geodesic path from the template ear to some target ears in the database. More specifically plots showing the change in the morphology as measured by the LDDMM distance, Hausdorff distance and CSDA distance versus the acoustics dissimilarity is presented. As an example Fig. 6.17 shows the relationship between the mean CSDA error (Sec. 4.3.2) versus the mean SDS error (Sec. 4.8.4).
In summary, this chapter contains the following contributions:

1. A new technique for comparing shapes using the framework of currents, Sec. (4.3).

2. A comprehensive approach to comparing DTFs with respect to each other, Sec. (4.8).

3. A morphoacoustic study for first and second order transformation of ear shapes, Sec. (6.1).

4. A morphoacoustic investigation of the transformations of ear shapes along the LDDMM geodesic path connecting the template ear to target ear shapes, Sec. (6.2).

Chapter organization: This chapter consists of two main sections:

Section. (6.1) explains the morphoacoustic analysis for first and second order transformations for ear shapes.

Section. (6.2) articulates the morphoacoustic analysis of ear shapes along the geodesic
path connecting the source ear shape to the target ear shape.

6.1 Single-Scale and Multi-Scale LDDMM Morphoacoustic Analysis of Ear Shapes

The previous sections of this chapter described the development of tools that can be used for comparing the morphology of ears and their corresponding acoustics. In this section the tools developed in the previous section are used to study ear shape deformations. More specifically, this section shows a series of experiments that examine the deformed template head and ear shape with respect to the target head and ear shapes in the SYMARE database. The study presented here is useful for when modelling ear shapes using the LDDMM framework as it can provide a better insight into the scope and limitations of the morphable model based on the LDDMM framework (Sec. 7.7).

In this section the morphology of the deformed template ear shape is compared to the target ear shape using the CSDA technique. The acoustical quality of the mapping is compared using GAAF. Only the quality of the ear mapping is examined here. The quality of the head mapping is also important, however errors in the mapping of the ear shapes contribute more significantly to the errors seen in the acoustical responses. Results presented in this section will show how the previous statement is further confirmed.

In this study both single-scale and multi-scale deformations (Sec. 4.9) from the template ear shape are examined. By considering both the acoustical and morphological quality of the deformation, it will be shown that for some target ears single-scale deformations of the template ear does not provide an adequate match to the target ear shapes. Therefore, to further improve the acoustical response of the deformed template shape, a second deformation is applied at a successive and smaller scale (hence multi-scale) so as to enhance the quality of the mapping.
Fig. 6.4—Fig. 6.7 shows the acoustical and morphological quality of single-scale deformations from the template ear shape towards selected target ear shapes using the SA and GAAF figures. A comprehensive set of SA and GAAF figures analysing the deformations from the template shape to target shapes in the SYMARE database can be seen in Appendix A. The first deformation applied to the template head and ear shape was conducted in a manner described in Sec. 4.2. The scale parameters used for the deformation are \( \sigma_V = 10 \) and \( 1 \leq \sigma_W < 2 \) for the ear mappings\(^2\) and for the head mapping it was chosen proportional to the maximum element size in the head mesh\(^3\).

By analysing the shape difference and GAAF figures shown in Fig. 6.4—Fig. 6.7 and Appendix A the following observations can be made:

1. For some ears the match between the acoustical responses is inaccurate. For example, the GAAF figures for subject S05 (Fig. 6.5) seen in this section and subjects S01, S18, S28 shown in Appendix A indicate a good match in the acoustical responses.\(^2\) The ear shapes had on average mean edge length of 0.9(mm) so selecting a \( \sigma_W \) value smaller than 2mm is a reasonable choice.\(^3\) The meshes used for the mapping of the head were close to uniform with the maximum edge length being smaller than 2 times the mean edge length so selecting the maximum edge length was a reasonable choice for the \( \sigma_W \) value.
tical responses as well as the morphological mapping of the ear shape. One the other, hand the GAAF figures in this section show that the acoustical responses between the morphed template shape and the target shape are different for subjects S41 and S29 (Fig. 6.7-Fig. 6.6). Many other subjects with a poor quality of matching in the morphology and acoustic response can be found in Appendix A.

2. Generally, the acoustical differences are most visible in the location and depth of the notches because they are the most salient features of the DTFs. For example as seen in Fig. 6.7 for subject S41 the DTF at azimuth 6° and elevation 38° contains a deep notch at around 13kHz in the matched shape (red) where as this notch is very shallow around the same frequency for the target shape (blue). The same phenomenon can also be observed for the DTFs at azimuth 90° and elevation -45°.

3. The acoustic quality of mapping is good for up to 6kHz for all mappings indicating that the quality of the head and torso mapping is particularly adequate. The head and torso shape are known to influence the DTF spectrum at low frequencies to about 5kHz. (Insert a reference here)

4. By observing the shape comparison figures (CSDA) for the deformed template shape, it can be seen that poor quality mappings in the morphology can occur more often in the Helix, Anti-Helix (Stem, Interior), Cymba Concha and Scaphoid-Fossa regions. The ear shape and the different anatomical sections of the ear are shown in Fig. 6.3. Poor quality of mapping in the Scaphoid-Fossa region usually causes minor discrepancies between the acoustical responses of the deformed template shape and target shape, while differences in the Concha and Anti-Helix (Interior) region can cause more sever discrepancies between the acoustical responses.
Figure 6.4: CSDA and GAAF figures comparing the morphological and acoustical quality of deformation from the template shape to the target subject S_{03} at the first scale of $\sigma^E_{W}(1) = 10$, $1 \leq \sigma^E_{W}(1) < 2$ and $\sigma^H_{W} = 10, \sigma^H_{W} = 7.02$
6.1. Single-Scale and Multi-Scale LDDMM Morphoacoustic Analysis of Ear Shapes

Figure 6.5: CSDA and GAAF figures comparing the morphological and acoustical quality of deformation from the template shape to the target subject $S_{05}$ at the first scale of $\sigma_E^V(1) = 10$ and $1 \leq \sigma_W^E(1) < 2$ and $\sigma_H^I = 10, \sigma_W^H = 6.96$
Figure 6.6: CSDA and GAAF figures comparing the morphological and acoustical quality of deformation from the template shape to the target subject $S_{29}$ at the first scale of $\sigma_E V(1) = 10$ and $1 \leq \sigma_W(1) < 2$ and $\sigma_H^H = 10, \sigma_H^W = 6.94$
6.1. Single-Scale and Multi-Scale LDDMM Morphoacoustic Analysis of Ear Shapes

Figure 6.7: CSDA and GAAF figures comparing the morphological and acoustical quality of deformation from the template shape to the target subject $S_{41}$ at the first scale of $\sigma_E^V(1) = 10$ and $1 \leq \sigma_W^E(1) < 2$ and $\sigma_H^I = 10, \sigma_W^I = 6.8$. 

Target

Matched

Front

Back

(a) Matched SFRS 11156 Hz
(b) Target SFRS 11156 Hz
(c) SFRS Correlation
(d) SFRS Error 11156 Hz
(e) Standard deviation in Spectral Difference

$S_{41}$
6.1. Single-Scale and Multi-Scale LDDMM Morphoacoustic Analysis of Ear Shapes

### 6.1.0.1 second-scale deformations

In order to obtain a better quality of mapping in the morphology and the acoustical responses, a multi-scale deformation is proposed for the ear shape only. The second deformation is applied at a smaller deformation scale of $\sigma_V$ on the ear shape while the shape comparison scale parameter was set to $1 \leq \sigma_W < 2$. The algorithm described here has many steps in common with the algorithm for deforming the head and ears described in Sec. 4.2 and, also, the algorithm for multi-scale mapping of the ear shapes detailed in Sec. 4.9. The reader can refer to these sections for more details on each of the algorithms. To describe the multi-scale deformation of the head and ear shape the following five surfaces are used. The notation here is the same as Sec. 4.2.

1. The high resolution source head and ear mesh: $HE_T$
2. The low resolution source head and ear mesh: $HE_{T,LR}$
3. The low resolution head and ear target mesh: $HE_{x,LR}$
4. The high resolution source ear mesh $E_T$
5. The high resolution target ear mesh: $E_x$

Where the subscript $T$ denotes the source shape and the subscript $x$ denotes some target ear shape. The complete set of steps for deforming the head and ear shapes at $n$ sequential scales is shown in Algorithm 7. The text that that follows explains in more detail the deformation of the head shape using only two scales.

**One A LDDMM mapping** $M(HE_{T,LR}, HE_{x,LR}, \sigma^H_V, \sigma^H_W)$ is performed between the low resolution source head and ear shape $HE_{T,LR}$ and low resolution target head and ear shape $HE_{x,LR}$ at a deformation scale of $\sigma^H_V$ and shape comparison scale of $\sigma^H_W$ to obtain a set of time dependant momentum vectors:

$$\{a(t)\}^{HE_{T,LR} \rightarrow HE_{x,LR}} \leftarrow M(HE_{T,LR}, HE_{x,LR}, \sigma^H_V, \sigma^H_W)$$  \hspace{1cm} (6.1)
Algorithm 7 Deforming the head and ears using a multi-scale approach

inputs: \( HE_T, HE_{T,LR}, HE_{x,LR}, ET, EX, [\sigma_V^E(1), \ldots, \sigma_V^E(L)], [\sigma_W^E(1), \ldots, \sigma_W^E(L)], \sigma_H^E, \sigma_H^W \).

outputs: \( HE_{Tm} \)

1. \( \{a(t)\}^{HE_{T,LR} \rightarrow HE_{E,LR}} \leftarrow M(HE_{T,LR}, HE_{x,LR}, \sigma_H^E, \sigma_H^W) \)
2. \( ET_1 \leftarrow F(\{a(t)\}^{HE_{T,LR} \rightarrow HE_{E,LR}}) \)
3. \( HE_{Tm} \leftarrow F(\{a(t)\}^{HE_{T,LR} \rightarrow HE_{E,LR}}, HE_T, \sigma_H^E) \)
4. for \( 2 \leq i \leq L \) do
5. \( \{a(t)\}^{ET_1 \rightarrow EX} \leftarrow M(EX_{(i-1)}, EX, \sigma_V^E(i-1), \sigma_W^E(i-1)) \)
6. \( ET_i \leftarrow F(\{a(t)\}^{ET_1 \rightarrow EX}, ET_{(i-1)}, \sigma_V^E(i-1)) \)
7. \( HE_{Tm} \leftarrow F(\{a(t)\}^{ET_1 \rightarrow EX}, HE_{Tm}, \sigma_V^E(i-1)) \)
8. end for

Two, Using the set of time dependant momentum vectors (i.e \( \{a(t)\}^{HE_{T,LR} \rightarrow HE_{E,LR}} \)) obtained in the first step an LDDMM flow function \( F \) is then performed on the high resolution source ear shape \( ET \) at a deformation scale of \( \sigma_H^E \) to obtain an intermediate ear shape \( ET_1 \):

\[
ET_1 \leftarrow F(\{a(t)\}^{HE_{T,LR} \rightarrow HE_{E,LR}}, ET, \sigma_H^E) \quad (6.2)
\]

Three, A mapping \( M \) between the intermediate high resolution ear shape \( ET_1 \) and the target ear shape \( EX \) is performed at scale values of \( \sigma_V^E(1) \) and \( \sigma_W^E \) to obtain a set of time dependant momentum vectors \( \{a(t)\}^{ET_1 \rightarrow EX} \) for the ear shapes.

\[
\{a(t)\}^{ET_1 \rightarrow EX} \leftarrow M(ET_1, EX, \sigma_V^E(1), \sigma_W^E(1)) \quad (6.3)
\]

Four, An LDDMM flow \( F \) is performed on the high resolution intermediate ear shape \( ET_i \) using the momentum vectors \( \{a(t)\}^{ET_1 \rightarrow EX} \) and deformation scale parameter \( \sigma_V^E(1) \) in order to obtain a second intermediate high resolution ear shape \( ET_2 \)

\[
ET_2 \leftarrow F(\{a(t)\}^{ET_1 \rightarrow EX}, ET_1, \sigma_V^E(1)) \quad (6.4)
\]

Five, A mapping \( M \) is performed between the second intermediate high resolution ear shape \( ET_2 \) and the target ear shape \( EX \) at scale values of \( \sigma_V^E(2) \) and \( \sigma_W^E(2) \) and a set
of time dependent momentum vectors is obtained \( \{ \alpha(t) \}_{E_{T2} \rightarrow E_x} \) for the ear shapes.

\[
\{ \alpha(t) \}_{E_{T2} \rightarrow E_x} \leftarrow M(E_{T2}, E_x, \sigma^F_V(2), \sigma^F_W(2)) 
\]

(6.5)

**Six**, A LDDMM flow \( \mathcal{F} \) is performed using the momentum vectors \( \{ \alpha(t) \}_{HE_{T,LR} \rightarrow HE_{E,LR}} \) on the high resolution source shape \( HE_T \) and deformation scale of \( \sigma^H_{E,V} \) in order to obtain an intermediate high resolution shape \( HE_{Tm} \)

\[
HE_{T1} \leftarrow \mathcal{F}(\{ \alpha(t) \}_{HE_{T,LR} \rightarrow HE_{E,LR}}, HE_T, \sigma^H_V) 
\]

(6.6)

**Seven**, A LDDMM flow \( \mathcal{F} \) is performed using the momentum vectors \( \{ \alpha(t) \}_{E_{T1} \rightarrow E_x} \) on the intermediate high resolution shape \( HE_{Tm} \) to obtain the high resolution mapped head and ear shape at the first physical scale.

\[
HE_{Tm} \leftarrow \mathcal{F}(\{ \alpha(t) \}_{E_{T1} \rightarrow E_x}, HE_{Tm}, \sigma^F_V(1)) 
\]

(6.7)

**Eight**, A LDDMM flow \( \mathcal{F} \) is performed using the momentum vectors \( \{ \alpha(t) \}_{E_{T2} \rightarrow E_x} \) on the intermediate high resolution shape \( HE_{Tm} \) to obtain the high resolution mapped head and ear shape at the second physical scale.

\[
HE_{Tm} \leftarrow \mathcal{F}(\{ \alpha(t) \}_{E_{T2} \rightarrow E_x}, HE_{Tm}, \sigma^F_V(2)) 
\]

(6.8)

Further, if the ear shapes need to be mapped using smaller scales, then steps Five and Four can be repeated.

**Fig. 6.8-Fig. 6.11** show matchings conducted at a second scale for the same pair of ear shapes shown in **Fig. 6.4-Fig. 6.7**. Further, as previously indicated, for the simulations conducted here \( L = 2 \) and \( \sigma^H_{E,V} = \sigma^F_{E,V}(1) \) and \( 1 \leq \sigma^F_{E,W}(1), \sigma^F_{E,W}(2) < 2 \) and \( \sigma^H_{E,W} \) is proportional to the maximum element size in the mesh. It can be seen that both the morphological quality of mapping and the quality for the acoustical responses for the ear shapes have improved. This suggests that ear shapes can be examined using
deformations at multiple physical scales. When analysing ear shapes at multiple
physical scales both the deformation scale parameter $\sigma_V$ and shape comparison
parameter $\sigma_W$ can be changed. However the results presented in this section suggest
that changing the deformation scale $\sigma_V$ has a greater impact when deforming ear
shapes once the shape comparison parameter $\sigma_W$ is set appropriately. This was
further discussed in Sec. 4.4. The remainder of the work related to this research
conducts and examines only single scale deformations of the ears. However, if there
is a need for further accuracy, the option to examine the ear shapes at more than one
physical scale exists.
6.1. Single-Scale and Multi-Scale LDDMM Morphoacoustic Analysis of Ear Shapes

![Figure 6.8: SA and GAAF figures comparing the morphological and acoustical quality of deformation from the template shape to the target subject $S_3$ at the second scale of $\sigma^T/V(2) = 1$ and $1 \leq \sigma^V/E(2) < 2.$](image)
6.1. Single-Scale and Multi-Scale LDDMM Morphoacoustic Analysis of Ear Shapes

Figure 6.9: SA and GAAF figures comparing the morphological and acoustical quality of deformation from the template shape to the target subject $S_5$ at the second scale of $\sigma_E^V(2) = 1$ and $1 \leq \sigma_E^W(2) < 2$. 
6.1. Single-Scale and Multi-Scale LDDMM Morphoacoustic Analysis of Ear Shapes

Figure 6.10: SA and GAAF figures comparing the morphological and acoustical quality of deformation from the template shape to the target subject $S_{29}$ at the second scale of $\sigma_E^2(2) = 1$ and $1 \leq \sigma_E^2(2) < 2$. 
Figure 6.11: SA and GAAF figures comparing the morphological and acoustical quality of deformation from the template shape to the target subject $S_{41}$ at the second scale of $\sigma_E^V(2) = 1$ and $1 \leq \sigma_E^W(2) < 2$. 
6.2 Morphoacoustic Distance Analysis Along the Deformation Path

This section details a morphoacoustic distance analysis along the deformation path from the source template ear shape to some select target ear shapes in the SYMARE database.

LDDMM ensures smooth and invertible deformations between two pair of ears at each step along the deformation path. The deformation path, which is also a geodesic path, signifies that an optimized path in the space of deformations is obtained between the source template ear shape and the target ear shape. An example of the flow of an ear shape along the deformation path was shown previously in Fig. 2.24.

Like the previous section (Sec. 6.1), the morphoacoustic analysis in this section includes the analysis of the morphology of the ear shapes along the deformation path and, also, the corresponding acoustical analysis. The deformation from the template shape examined are deformations conducted at a single physical scale. The scale values that are used for deforming the template ear shape is \( \sigma_E^T = 10 \) and \( 1 \leq \sigma_{EW}^T < 2 \).

It may be recalled that the mapping operator \( \mathcal{M} \) defined in Eq. (4.1) produces a complete set of time dependant momentum vectors \( \{ \alpha(t) \}_{0 \leq t \leq 1}^{S_1 \rightarrow S_2} \) that parametrizes the deformation path between the source and target shapes \( S_1 \) and \( S_2 \). The artificial time parameter \( t \), which is in the range of \( [0 \ldots 1] \), is used to point to a specific location in the deformation path, with \( t = 0 \) being the location of the starting source shape, and \( t = 1 \) being the location of the matched target shape. Further, it may be recalled that the flow function \( \mathcal{F} \) defined in Sec. 4.4 uses the momentum vectors obtained from the mapping and the source shape and two time parameters, \( t_s \) and \( t_e \), signifying the start and finish time for the deformation to deform the source shape.

Given the previous it is appreciated that ear shapes along the deformations path can easily be obtained by appropriately setting the \( t_e \) parameter in the function \( \mathcal{F} \). Algorithm lists the steps involved in the production of shapes for various locations.
along the geodesic path using the source shape \( S_1 \), momentum vectors \( \{\alpha(t)\}_0^1 \) and locations along the geodesic path indicated by a time vector \([t_1, t_2 \ldots t_n]\):

**Algorithm 8 Generating Shapes Along The Deformation Path**

*inputs:* \( E_1, E_2, \sigma_V, \sigma_W, [t_1, t_2 \ldots t_n] \).

*outputs:* \( E_1(t_1), E_1(t_2) \ldots E_1(t_n) \).

1. \( \{\alpha(t)\}^E_{1 \rightarrow 2} \leftarrow \mathcal{M}(E_1, E_2, \sigma_V, \sigma_W) \),
2. \( \text{for } 1 \leq i \leq n \text{ do} \)
3. \( E_1(t_n) \leftarrow \mathcal{F}(\{\alpha(t)\}^{S_1 \rightarrow S_2}, S_1, \sigma_V, 0, t_i) \)
4. \( \text{end for} \)

The previous paragraphs explained how to obtain ear shapes along the geodesic path. We now describe how to obtain head shapes with deformed ear shapes that belong to evenly spaced locations along the geodesic path between the source and target ears. FM-BEM simulations can be conducted on these head and ear shapes to obtain corresponding acoustical responses. The procedure detailed here is very similar to the procedures detailed in Sec. 6.1.0.1 and Sec. 4.2 albeit with some differences that will be explained. We only consider deformations of the left ear. Five mesh surfaces are required for the deformation process these are:

1. The high resolution source head and ear mesh: \( H_E \)
2. The low resolution source head mesh without the left ear: \( HENL_{T,LR} \)
3. The low resolution target head mesh without the left ear: \( HENL_{x,LR} \)
4. The high resolution source ear mesh \( E \)
5. The high resolution target ear mesh: \( E_x \)

It is to be noted that low resolution matching of the head shapes is conducted on head shapes that do not contain the left ear. The reason for this will be explained later in this chapter. The procedure to obtain head meshes with ears along the deformation path is given in Algorithm 9. The algorithm takes in the five surface meshes listed above, a list of \( n \) time parameters signifying the locations along the
Algorithm 9 Generating Head Meshes With Ear Shapes Along The Deformation Path
inputs: $HENL_{T,LR}, HENL_{x,LR}, HE_T, E_T, E_x, [t_1, t_2, \ldots, t_L], \sigma_V, \sigma_W$.
outputs: $S_{g1}, S_{g2}, \ldots, S_{gL}$
1: $\{\alpha(t)\}^{HENL_{T,LR}}_{HENL_{x,LR}} \leftarrow \mathcal{M}(HENL_{T,LR}, HENL_{x,LR}, \sigma_V, \sigma_W)$
2: $HE_{T1} \leftarrow \mathcal{F}(\{\alpha(t)\}^{HENL_{T,LR}}_{HENL_{x,LR}}, HE_T, \sigma_V)$
3: $E_{T1} \leftarrow \mathcal{F}(\{\alpha(t)\}^{HENL_{T,LR}}_{HENL_{x,LR}}, E_T, \sigma_V)$
4: $\{\alpha(t)\}^{E_{T1}}_{E_x} \leftarrow \mathcal{M}(E_{T1}, E_x, \sigma_V, \sigma_W)$
5: for $1 \leq i \leq L$ do
6: $S_{gi} \leftarrow \mathcal{F}(\{\alpha(t)\}^{E_{T1}}_{E_x}, HE_{T1}, \sigma_V, 0, t_i)$
7: end for

geodesic path and produces $n$ head surfaces with ear shapes corresponding to the locations along the geodesic path, these surfaces are denoted as $S_{g1}, S_{g2}, \ldots, S_{gN}$.

The reason for removing the left ear shape when mapping low resolution head shapes is now explained. This is also one of the chief differences of the current algorithm with the algorithms for deforming head and ear shapes in previous sections. It may be recalled that step 4 in Algorithm I performs a mapping between the low resolution head and ear shapes and further it may be recalled that step 2 in the same algorithm performs a flow function on the source ear shape. The flow from the low resolution mapping can deform the source ear shape, which is not desirable for the purposes of the current experiment. This can be seen more clearly in the second step of Fig. 4.3. A picture displaying the template head shape with no left ears is shown in Fig. 6.12.

Morphological deformations were applied from the template ear shape $E_T$ to three target ears $E_1, E_4$ and $E_{11}$ for locations along the geodesic path corresponding to $T = [0, \frac{1}{3}, \frac{2}{3}, 1]$. Fig. 6.13–Fig. 6.15 shows the results for the morphoacoustic analysis of the three ear shapes along the deformation path. The first column in the figures is the target ear shape with regions of large differences highlighted in red color. The second column in the figures is the deformed template ear shape and shows the evolution of the template ear along the deformation path, again regions of large morphological differences are highlighted in red color. The third column in the figures is the SDS map (function $V$) which was discussed in Sec. 4.8.4. The function $V$ between two
Figure 6.12: The above figure shows the template head shape without the left ear that is used in deformations in the current section.
shapes provides information on the quality of the acoustical mapping between the two shapes for all frequencies and for computed azimuth and elevation angles. The third and fourth columns in the figures show DTFs belonging to directions in space where the discrepancy between the acoustical responses is large. The locations where the DTFs correspond to are highlighted with a green cross on the SDS map.

Figure 6.13: Morphoacoustic analysis of ear shapes along the LDDMM deformation path for the template ear shape to target ear shape $E_1$. 

$E_1$
Figure 6.14: Morphoacoustic analysis of ear shapes along the LDDMM deformation path for the template ear shape to target ear shape $E_4$
6.2. Morphoacoustic Distance Analysis Along the Deformation Path

Figure 6.15: Morphoacoustic analysis of ear shapes along the LDDMM deformation path for the template ear shape to target ear shape $E_{11}$
6.2.1 Morphoacoustic Distance Analysis

An analysis was conducted to quantify the relationship between the morphological dissimilarities in the ear shapes with respect to the acoustical dissimilarities in the ears. The morphological discrepancies examined between the ear \( E_T(t_i) \) and target ear shapes \( E_x \) were measured using:

1. The mean CSDA (Sec. 4.3.2) value denoted by \( \bar{d}(E_T(t_i), E_x) \).

2. The geodesic distance between the two shapes induced by the LDDMM framework \( d_{De}f(E_T(t_i), E_x) \) (refer to Sec. 2.8.1 for further details).

The acoustical differences was calculated using the mean SDS measure (Sec. 4.8.4) \( \bar{V}_{E_T(t_i), E_x} \). In the previous \( t_i \in T \) corresponds to the time locations and \( x \in [1, 4, 11] \) corresponds to the three target ear shapes under examination. Fig. 6.16 shows a plot of the geodesic distance \( d_{De}f \) (Eq. 2.25) versus the acoustical dissimilarity while Fig. 6.17 shows a plot of \( \bar{d} \) versus \( \bar{V} \). Each of the colors in the plots belong to a specific target ear shape and the black line in the plots signifies the average of the three plots.
6.2. Morphoacoustic Distance Analysis Along the Deformation Path

Figure 6.16: The above shows a plot of the geodesic distance $d_{Def}$ versus acoustical discrepancy measured by the function $\bar{V}$. Each of the red, blue and red lines separately shows the relationship for a particular subject. The black line is the average value for the three subjects.

Figure 6.17: The above shows a plot of morphological discrepancy measured by the function $\bar{d}$ versus acoustical discrepancy measured by the function $\bar{V}$. Each of the red, blue and red lines shows the relationship for a particular ear. The black line is the average value for the three ears.
6.3 Conclusion

In this chapter deformations from the template ear shape to target ears in SYMARE were examined both morphologically and acoustically. The CSDA tool developed in a previous chapter was used to highlight and measure differences in ear shapes using the framework of currents. The tool was shown to be effective in highlighting differences in ear shapes that cause acoustic differences. Further, GAAF was used to compare the acoustics of the two ear shapes globally both in the frequency domain and spatially.

Further, this section showed that ear shapes can be analysed more accurately in two physical scales. It was shown that in the first and larger physical scale important morphological and acoustical features are captured and mapped. Some of the morphological features that did not match accurately after transformation at the first scale were seen in the Interior-Crus Anti Helix, Concha and Scapha region. Acoustic features that did not match were mainly related to the depth and location of the notches seen in the DTF spectrum. For the ear shapes that required a more accurate mapping a second and finer LDDMM mapping was then performed. This mapping fixed any inaccurate morphological mappings and further reduced significantly the errors between the acoustic spectrum’s for the two ear shapes.

This section also showed the morphological and acoustic evolution of the template ear to target ear shapes. In particular, plots showing the spectral differences between DTFs using the SDS measure and DTF’s belonging to directions in space where the acoustic error was large is shown. Among other transformations it could be observed from the plots that as the template ear shape evolved the location of the notches and peaks of the deformed template ear shape evolved towards the target ear shape. Finally, this section, contained a distance analysis between the morphology of the ears and the corresponding acoustics. The distances between the ear shapes were measured using the mean CSDA value, and also the geodesic distance calculated by LDDMM. The acoustical distances were calculated using the mean SDS value.
The plots showed that on average the morphological and acoustic distance exhibit a piecewise linear (or close to linear) relation, however, more experiments are needed to confirm this claim.
Chapter 6 PCA Analysis
Chapter 7

KERNEL PRINCIPAL COMPONENT ANALYSIS OF EAR MORPHOLOGY

This chapter describes a process for statistically analysing variations in ear morphology and their impact on acoustics. The aim of this statistical analysis is to obtain a better understanding of the significant and most common changes in the morphology of the ear shapes across a population of ears. The statistical analysis conducted over the space of ear shapes uses the kernel principal component analysis (KPCA). Further, it utilizes the framework of large deformation diffeomorphic metric mapping (LDDMM) and the vector space that is constructed over the space of initial momentums. The initial momentum’s describe the diffeomorphic transformations from the reference template ear shape to ears in the database. The population of ear shapes examined by the KPCA is a total of 124 left and right ear shapes from the SYMARE database that were aligned to the template (population average) ear using affine transformation\(^1\).

There are multiple benefits in studying ear shapes using KPCA. Firstly, in the KPCA technique we are analysing variations with respect to the mean or template

\(^1\) Affine transformation include scaling, rotation and translation among other collinear movements
ear shape. The template ear was shown in the previous section to have features that resembled the average of the population. The requirement of obtaining a mean or average point is important when performing any PCA analysis. **For the analysis in this thesis and more particularly the KPCA the multi-scale template ear shape is used.**

Secondly, similar to the PCA analysis, the KPCA provides us with an orthogonal set of bases vectors wherein each of these bases vectors captures some form of the variation in the ear shapes. Importantly, each of these bases vectors can be changed by applying a multiplication by a scalar value. Using the LDDMM framework, we can then generate ear shapes using each of the changed kernel principal components and in this manner we can analyse the variation captured by each of the kernel principal components. Changes made to the kernel principal components is an integral part of the research conducted in this chapter and more details will be given in the relevant sections. It is to be noted that the ear shapes that are analysed in the following sections using KPCA are affine transformed to the template ear shape. In other words, KPCA was performed on affine transformed ear shapes because it was desirable for the kernel principal components to only capture important structural differences in ears and not the rotational, scaling or translational variations.

The complete process of analysing ear shapes using the KPCA technique, like projection onto the orthogonal bases and reconstruction of ear shapes is not simple and includes solving many differential equations. This process will be described in later sections. At this point it might be useful for the reader to familiarize himself/herself with the standard PCA technique that was explained in the Background chapter of this thesis (refer Sec. 2.7). Sec. 2.7 will be referred to when explaining concepts in this chapter. For clarity this chapter contains PCA analysis on both shape data and acoustic data. The analysis performed on shape data will be referred to as shape PCA and the associated kernel principal components will be referred to as shape principal components. Further, the analysis performed on acoustic data will be referred to as acoustic PCA and the principal components obtained will be
referred to as acoustic principal components.

Presently, the research questions that are addressed in this chapter will now be detailed. A total of six research questions are addressed in this chapter. The first research question that is addressed in this chapter investigates how much of the acoustic variation in the ear shapes is captured by the affine transformation of the ear shapes to the template ear shape? This is an important issue as the shape KPCA analysis presented in this chapter is performed on ear shapes that are affine transformed to the template ear shape. It is shown that the rigid alignment captures approximately thirteen percent (13%) of the shape variation.

The second research question relates to identifying how many kernel principal components are required for reconstructing a target ear shape in order for the acoustic response of the reconstructed ear to closely match that of the target ear? The analysis of this chapter suggests that matching the target shape acoustics to a very fine level is not efficient because it requires the use of a large number of kernel principal components (i.e. > 50). The results show that on average for 20 shape principal components approximately 47% of the acoustic variations and 44% of the morphological variation is captured. Further, using the shape comparison tool (i.e CSDA) that was developed in Sec. 6, this chapter also provides a quantitative measure of how the distance between ear shapes varies (according to CSDA) as the number of kernel principle components changes.

The third research question examines what kind of variations in ear shape are captured by the first few shape principal components. The sections that follow display both the morphological variations captured by the kernel principal components and also the corresponding acoustic variations. For simplicity, this chapter focuses on the first five kernel principal components which capture the largest amount of variation in the data. The appendix to this thesis further plots the variations captured by the remaining 20 kernel principal components. It is shown that the first five principal components capture large structural changes in the ear shape, which includes the shape and size of the Concha, Triangular Fossa, Helix and Lobe of the ear to name a
few features.

The fourth research question investigates whether there are identifiable features in the DTF spectrum that change systematically with changes made to a single shape kernel principal component. In order to answer this question PCA was performed on the HRTF acoustic data. It is shown that variations in the weight for a single shape kernel principle component corresponds to variations in the weights for two acoustic principle components. The variation for DTF data is shown against variation for the shape data.

The fifth research question addressed here asks whether there are identifiable features in the acoustic spectrum that change coherently for two or more angles, given that the ear shape is systematically modified by changing a single shape principal component? This analysis also entails the application of a PCA onto the DTF data. Using this analysis, it is shown that features exists in the form of peaks and notches in the DTF spectrum belonging to the various angles that do indeed change coherently.

The sixth research question examines the accuracy and efficiency of a morphable-model based on the SYMARE database to represent new unseen ears outside of the database. The model is derived from the KPCA data on a population of 58 left ear shapes from the SYMARE database. The strength of the model when reconstructing test ear shapes which have not been included in the KPCA is examined. It is shown that the morphable model is promising and is able to reconstruct some ear shapes with higher accuracy than others. The CSDA measure is used to measure the level of accuracy in the reconstruction of the shapes.

This chapter presents a significant number of acoustic simulations performed on ear shapes. In order to obtain more insight into the features seen in the DTF spectrum, the DTF spectrum’s are presented using different kind of plots in the frequency and spatial domains. For simplicity, the figures mainly display the DTF spectrum only at angles corresponding to the median plane.

The organization of the rest of this chapter is as follows: Sec. 7.1 will describe the
theory for the kernel based principal component analysis. Sec. 7.3 will examine the question of how many shape principal components are required to match a target shape acoustically to a prescribed level of accuracy. Sec. 7.4 examines the changes in ear shape described by the first few shape kernel principal components. Sec. 7.5 examines the DTFs for the median plane and identifies the important features in the DTF that change when the ear shapes change. Sec. 7.6 performs an analysis to investigate how coherently the acoustic features belonging to a few angles in the median plane change. Finally, Sec. 7.7 develops a morphable model for ear shapes and examines its ability to represent new ear shapes.

## 7.1 Kernel Principal Component Analysis (KPCA)

This section details the KPCA technique for analysing the variations of ear shapes. The KPCA technique is chosen over the standard PCA technique because it entails an inner product between observations that is more suitable for comparing shapes. The inner product that is used in KPCA will be described in detail later in this section. Further the KPCA analysis conducted in this section is applied only to single-scale deformations from the template ear shape to target ear shapes.

It may be recalled from the analysis in Sec. 6.1 that single-scale deformations captured important structural and acoustic information on the ear shapes. Further it was detailed that second scale transformations can be applied then onwards if more accuracy in the mapping of the morphology or acoustics is required.

The KPCA analysis over the space of deformations defined by mapping the template ear shape to all the other ear shapes in the SYMARE database is made possible by utilizing the initial momentum vectors that describe the geodesic path for each of the shape deformations. The initial momentum vectors form a linear vector space and can be added, subtracted and or multiplied by a scalar value. The kernel PCA is different from the standard PCA technique in that the inner products between observations are calculated via a kernel function.
7.1. Kernel Principal Component Analysis (KPCA)

It may be recalled from Sec. 2.7 that the standard principal component analysis (PCA) performs an orthogonal linear transformation on a given set of data. This orthogonal transformation produces a set of orthogonal bases vectors such that by linearly combining these bases vectors any point within the dataset can be reconstructed. The KPCA (Cury, 2015; Cury et al., 2015; Vaillant et al., 2004) over the space of deformations is similar, that is, a set of orthogonal bases vectors are formed with respect to the kernel matrix (see below) and shape deformations are described using a linear combination of the kernel principal components. However it is noted that shapes reside in a non-linear Riemannian space which is more complicated than a set of points in an Euclidean space and performing the KPCA and reconstructing ear shapes from the kernel principal components involves extra steps such as solving the shooting differential equations (Sec. 2.31).

7.1.1 KPCA On Ear Shapes

The KPCA analysis that follows is conducted on ear shapes that are rotated, translated and scaled towards the template ear. The technique for the affine transformation of the ears to the template ear shape was previously discussed in Sec. 5.2.4.1. Assuming that $L$ ear shapes are taken from the SYMARE database the first step in the KPCA analysis is to calculate the initial momentum vectors for deformations from the source template shape $T$ to every ear, $S_l$, in the population of $L$ ears, as follows:

$$\{a_n^{(l)}(t)\}_{0 \leq t \leq 1}^{0 < i \leq 1} = \mathcal{M}(T, S_l, \sigma_V, \sigma_W) \quad (7.1)$$

In order to calculate the principal components, the covariance matrix, $C$, which expresses the mutual correlation of the different ear shapes in the space of deformations is calculated. To compute this matrix, first a data matrix $A \in \mathbb{R}^{3N \times L}$ which contains the initial momentum vectors for the entire population of ears is constructed:

$$A = [a_1, a_2, \ldots, a_L]_{3N \times L} \quad (7.2)$$
where \( \mathbf{a}_l \) denotes the column vector containing all the initial momentum vector scores for shape \( S_l \). Subsequently the data is centred by subtracting the population average momentum vectors. The centred data matrix, \( \hat{\mathbf{A}} \), is given by:

\[
\hat{\mathbf{A}} = [\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \ldots, \hat{\mathbf{a}}_L]_{3N \times L}
\]  

(7.3)

where \( \hat{\mathbf{a}}_l \) is the vector of the centred momentum vectors for the \( l \)-th shape:

\[
\hat{\mathbf{a}}_l = \mathbf{a}_l - \bar{\mathbf{a}} \quad \text{with} \quad \bar{\mathbf{a}} = \frac{1}{L} \sum_{i=1}^{L} \mathbf{a}_i .
\]  

(7.4)

The kernel matrix, \( \mathbf{K} \), which contains the values of the kernel function for every pair of vertex positions that comprise the vertices, \( \mathbf{X} \), of the template shape \( T \) is also formed:

\[
\mathbf{K} = 
\begin{bmatrix}
\mathbf{K}_{11} & \mathbf{K}_{12} & \ldots & \mathbf{K}_{1N} \\
\mathbf{K}_{21} & \mathbf{K}_{22} & \ldots & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{K}_{N1} & \ldots & \ldots & \mathbf{K}_{NN}
\end{bmatrix},
\]  

(7.5)

where \( \mathbf{I}_{3 \times 3} \) denotes the \( 3 \times 3 \) identity matrix.

The correlation between two shapes is calculated as the inner product of the initial momentum vectors in the Hilbert space of deformations, \( \mathcal{V} \). The correlation between shapes \( S_i \) and \( S_j \) is given by:

\[
c_{ij} = \frac{1}{L-1} \left\langle \{ \alpha_n^{(i)}(0), \{ \alpha_n^{(j)}(0) \} \} \right\rangle_{\mathcal{V}} = \hat{\mathbf{a}}_i^T \mathbf{K} \hat{\mathbf{a}}_j ,
\]  

(7.6)

where \( (\cdot)^T \) denotes the transpose of a vector or matrix. Thus, the covariance matrix for the entire population of ears, \( \mathbf{C} \), is given by:

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

(7.7)
In order to calculate the principal components, as well as the coordinates of the ears in the basis of the principal components, the singular value decomposition of the covariance matrix $C$ is performed:

$$C = VDV^T. \quad (7.8)$$

The matrix of the principal components, $U$, can be then calculated as:

$$U = \hat{A}VD^{-\frac{1}{2}}. \quad (7.9)$$

Note that the principal components are orthogonal in the Hilbert space of deformations, i.e., $U^TKU = I$. It follows from Eq. (7.9) that $\hat{A} = UD\frac{1}{2}V^T$ and, therefore, $D\frac{1}{2}V^T$ provides the principle components scores which are the coordinates of the different ear shapes in the basis of the principal components. Each ear can thus be reconstructed by:

1. computing $a_l = \bar{a} + UD\frac{1}{2}v_l$ ( $v_l$ is the $l$-th column of $V^T$); and
2. shooting from the template in the $a_l$ direction, i.e.,

$$\{a_n^l(t)\}_{0\leq l\leq 1}^{1\leq n\leq N} = S(E_{TMs}, \{a_l\}_{1<n<N}) \quad (7.10)$$

3. applying a diffeomorphic flow using the momentum vectors:

$$S_l = F(E_{TMs}, \{a_n^l(t)\}_{1\leq n\leq N} \{0\leq l\leq 1\}) \quad (7.11)$$

Please note that the notation used here for representing the initial momentums indicates both the range for the time parameter and the number of vertices, this is slightly different from what was presented in Sec. 4.1. Using the KPCA framework just described, each ear shape in the population is described by $L$ parameters, where $L$ is the size of the population of ears. It may be noted that the dimension of the model can be further reduced at the cost of reduced shape reconstruction accuracy.
by keeping only the first \( K (K \leq L) \) principal components. Generating morphable-models for ear shapes is further described in Sec. 7.7 of this chapter.

### 7.2 Quantifying the Reduction in Acoustic Variation Obtained With Alignment of the Ear Shapes to The Template Ear

This section establishes that the acoustic variation between ears is primarily attributable to changes in ear shape and not affine transformations. More specifically, this section provides a quantitative estimate of the reduction in acoustic variation that occurs when the ear shapes are aligned\(^2\) to the template ear shape compared to when they are not aligned. As detailed previously the KPCA analysis over the space of ear shape deformations are for ears that have been affine transformed to the template ear shape\(^3\).

Six ear shapes were selected from the SYMARE database and were paired up using all combinations of 2. The acoustic distances for every pair were calculated and compared when aligned to the template shape and when not aligned. More specifically we require the, acoustic responses for head, torso and ear shapes for a population of normal ear shapes (without alignment) and the same ears aligned to the multi-scale template ear shape \( E_{T_{MS}} \). Algorithm\(^{10}\) details the steps for obtaining two ear, head, and torso meshes \( HTE_{l_{TS}} \) and \( HTE_{l} \) for an ear shape \( E_{l} \) belonging to subject \( l \) in the SYMARE database. The shape \( HTE_{l_{TS}} \) is the template head and torso with the ear shape \( E_{l} \) aligned to the template ear and the shape \( HTE_{l} \) is the template head and torso shape with the ear shape \( E_{l} \) translated to the template shape. The translation is done in order to align the ear shape better with the template head shape, a similar alignment was performed in the paper by [Zolfaghari et al. (2014)](https://www.mdpi.com/2072-666X/7/11/158). A total 12 head, torso and ear meshes were created with the proper ear shapes attached.

---

\(^2\) Alignment means they are affine transformed to the template ear shape

\(^3\) please view Sec. 5.2.4.1 to see how ears are affine transformed to the template ear shape.
7.2. Quantifying the Reduction in Acoustic Variation Obtained With Alignment of the Ear Shapes to The Template Ear

**Algorithm 10** Obtain Head and Ears For Acoustic Distance Analysis

**Inputs:** $E_{TMS}, E_l, HTE_{TMS}$.

1. $E_l \leftarrow \text{Translate ear } E_l \text{ towards } E_{TMS} \text{ using centre of mass (see Eq. (5.9))}.$
2. \{a(t)\}$_{0 \leq t \leq 1}^{E_{TMS} \rightarrow E_l} \leftarrow \mathcal{M}(E_{TMS}, E_l, \sigma_V = 10, \sigma_W = 1.25)$
3. $HTE_l \leftarrow \mathcal{F}(\{a(t)\}^{E_{TMS} \rightarrow E_l}, HTE_{TMS}, \sigma_V = 10, 0, 1)$
4. $E_{RTS} \leftarrow \text{Rigidly align ear } E_l \text{ towards } E_{TMS} \text{ (see Sec. 5.2.4.1)}.$
5. \{a(t)\}$_{0 \leq t \leq 1}^{E_{TMS} \rightarrow E_{RTS}} \leftarrow \mathcal{M}(E_{TMS}, E_{RTS}, \sigma_V = 10, \sigma_W = 1.25)$
6. $HTE_{RTS} \leftarrow \mathcal{F}(\{a(t)\}^{E_{TMS} \rightarrow E_{RTS}}, HTE_{TMS}, \sigma_V = 10, 0, 1)$
7. **return** $HTE_{RTS}, HTE_l$

Next, FM-BEM simulations were performed on the 12 head and torso meshes. The procedure for conducting the FM-BEM simulations are as presented in Sec. 4.6

DTF distances using the function $\bar{V}$ (see Sec. 4.8.4) between shapes $HTE_{i_{RTS}}$ and $HTE_{j_{RTS}}$ and shapes $HTE_i$ and $HTE_j$ for $1 \leq i, j \leq 6$ were calculated. Given six ear shapes a total of 15 unique pair of ears when aligned to the template ear and in normal orientation were investigated.

Fig. 7.1 shows the morphology of the six normal ear shapes, denoted by $E_l$ and the corresponding ear shapes that are aligned to the template ear denoted by $E_{RTS}$ using affine transformation. Further Fig. 7.2 plot (a-c) show the acoustic analysis for the 15 pairs of ears. Plot (a) shows the acoustic distances calculated using the function $\bar{V}$ between the shapes $HTE_{i_{RTS}}, HTE_{j_{RTS}}$ and $1 \leq i, j \leq 6$ (i.e. Normally oriented ear shapes). Plot (b) shows the acoustic distances also calculated using $\bar{V}$ for the affine transformed ears, on the shapes $HTE_{i_{RTS}}, HTE_{j_{RTS}}$ and $1 \leq i, j \leq 6$ and plot (c) shows the ratio of plot (a) over plot (b).

The value in plot (c) in Fig. 7.2 signifies that ear shapes that are aligned to the template shape on average have more similar acoustic responses as measured by the function $\bar{V}$ compared to the same pair of ears when not-aligned to the template ear shape. This is because as observed in plot(c) of Fig. 7.2 for almost every pair of ear shapes $E_{RTS}$ and $E_{RTS}$ the ratio of the acoustic distances is less than 1.

$\frac{1}{2}C_2 = \binom{6}{2} = 15$, which means choose a combination of 2 ears from a set of six ear shapes.
shapes the ratio is below 1. However there is an exception for the pair \( \{E_4, E_5\} \) where alignment did not reduce the acoustic function \( \bar{V} \). It is noted that the same pair of ears had the lowest acoustic dissimilarity when aligned or in normal orientation when measured using the function \( \bar{V} \). Further plot (c) in Fig. 7.2 contains a red stem bar which signifies the average ratio. The red stem bar is located at approximately 0.87 which shows that on average the aligned ear shapes have an acoustic response
that is 13% more similar to each other compared to when the same pair of ears is not aligned.

![Graphs showing acoustic distances](image)

**Figure 7.2:** The above plots shows the acoustic distances calculated using the function $\bar{V}$ between (a) every pair of normal ears on shapes $HTE_i$ and $HTE_j$ for $1 \leq i, j \leq 6$ and (b) every pair of affine transformed ears to the multi-scale template ear shape on shapes $HTE_{iRTS}$ and $HTE_{jRTS}$ for $1 \leq i, j \leq 6 E_{TS}$. (c) shows the ratio of plot (a) over plot (b). The red stem bar on plot (c) shows the mean ratio and is located at approximately 0.87.

### 7.3 Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Sec. 7.1 describes how to compute the kernel principal components and their associated scores (also referred to as weights) for each ear shape. More specifically, a procedure is detailed (7.1.1) in which target ear shapes in the dataset could be fully
reconstructed using the complete set of scores obtained by the KPCA technique and
the template ear. The analysis conducted in this section aims to provide more insight
into the effects and significance of the kernel principal components when modelling
the morphology and acoustics of the ear shapes.

There are three aims relating to this section of the research. The first aim is to
obtain an estimate on the number of principle components required to reconstruct
a target ear shape from a source ear shape, such that the acoustic response of the
reconstructed ear shape matches the acoustic response of the target ear. Target
ears are ear shapes that are affine transformed to the template shape which can be
reconstructed using the full set of scores and principle components. The second aim
was to obtain an estimate on the significance of changes and or perturbations made
to the kernel principal components on the acoustic responses of the ear shapes. To be
more specific the SYMARE database contains 124 left and right ear shapes, and by
using the KPCA analysis an ear shape will require 124 scores and associated kernel
principal components for complete and accurate reconstruction. At this point it was
not clear as to how many scores and kernel principal components out of the 124
are required for modelling ears in SYMARE. The third aim relating to this study is
to obtain an estimate on the impact of changing the 50th or 100th kernel principal
component on the morphology and acoustics of the reconstructed ear shape.

### 7.3.1 Choosing The Ear Shapes in SYMARE

The analysis in this section is conducted between pairs of ears that have significantly
different acoustic responses. Differences in the DTFs were measured using the
function $\bar{C}$, named the spatial correlation metric (SCM), which was described in
Eq. (4.38), and the function $\bar{V}$, named standard deviation on spectral difference
(SDS), which was described in Eq. (4.40). The DTFs were obtained from the BEM’s
available in the SYMARE database. These BEM’s are valid up until 16kHz and
provide good information on the acoustic responses for the ear shape.

In order to choose ear shapes that had significantly different acoustic responses,
the DTFs for every pair of ear shape from the SYMARE database was examined using two distance matrices, $M(V)$ and $M(C)$. These matrices were computed using the measures $C$ and $\bar{V}$. Each element $M_{ij}(V)$ is a measure of the difference in the acoustic response between ear $i$ and ear $j$ in SYMARE using $\bar{V}$ and element $M_{ij}(C)$ is a measure for the difference between the same pair of ears using the measure $\bar{C}$. When examining the acoustic differences only DTFs belonging to the left ear shapes were considered. Only ipsilateral responses for the ear shapes were taken into consideration and the frequency range considered in the DTFs was between 5kHz and 16kHz. Using these two matrices pairs of ears which had large DTF distances were selected and used for the experiments. Histogram plots for the matrices $M(V)$ and $M(C)$ are shown in Fig. 7.3.

The ear pairs $S_{55}, S_5$ and $S_{13}, S_{23}$ available in the SYMARE database appeared consistently in the top 1% of ears that had large DTF distances with each other as measured by the functions $V$ and $C$. These ear shapes are shown as the source and target ears in Fig. 7.6-Fig. 7.7.
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Figure 7.3: (a) DTF distances between pairs of left ear shapes in the SYMARE database calculated using the function $\bar{C}$ (SCM measure). Higher values indicate a larger distance between the DTFs. (b) DTF distances between pairs of left ear shapes in the SYMARE database calculated using the function $\bar{V}$ (SDS measure). Higher values indicate a larger distance between the DTFs. (c) Histogram plot of the normalized distribution between all pairs of ear shapes in the SYMARE database. The black line indicates the mean of the population, the yellow lines are at -2 and +2 standard deviation and the red line is a pair of ears chosen for the simulation in this section.

7.3.2 Methodology

Now a description of how the experiments were set-up and conducted is given. It may be recalled that Sec. 7.1 described how to compute the kernel principal components in the form of a matrix $U$ and further showed how the scores for all ear shapes in the database are computed and stored in the matrix $V$. For example, the scores representing ear shape $l$ are in the column vector $v_l$ within the matrix $V^T$. Further, it may be recalled that the ear shape $E_l$ can be generated by its initial
momentum vectors located on the template shape as follows:

\[ a_l = \bar{a} + UD^2 v_l \]  
\[ \{a(t)\}_{1 \leq n \leq N}^{0 < t < 1} = S(E_{TMS}, a_l, \sigma_V) \]  
\[ E_l = F(E_{TMS}, \{a(t)\}_{1 \leq n \leq N}^{0 < t < 1}, \sigma_V) \]

where in the above \( \bar{a} \) is the mean momentum vector across the population of ear shapes (Eq. (7.4)). Note that the KPCA is done using the template ear shape and the initial momentum vectors obtained are located on the template ear. Once the momentum vectors for the particular ear shape is obtained, the geodesic shooting equations can be used to obtain the full set of momentum vectors between the template ear shape and the specified target ear shapes.

In order to find the number of kernel principal components that are required when trying to obtain the acoustic features of the target shape, first two or more ear shapes that have different DTF spectrum’s were selected. This was described in Sec. 7.3.1. Once the ear shapes were selected then Algorithm [11] is used to generate ears between the source and target ear shapes. Let's denote the scores for the source ear shape by \( v_s \), and the scores representing the target ear shape as \( v_l \) for some \( l, s \leq L \), where \( L \) is the number of principal components and also the number of ears in the dataset. Note that both these ear shapes have to exist in the dataset in which the KPCA was computed. Further, let’s denote the \( j^{th} \) row of the column vector \( v_s \) and \( v_l \) as, \( v_s(j) \) and \( v_l(j) \) respectively, then we can generate an ear shape between the source and target ears by taking the first \( P \) scores from the KPCA scores belonging to the target ear shape and the remaining \( L - P \) scores from the KPCA scores belonging to the source ear shape. Algorithm [11] detailed how an ear shape that is between a source and target ear is obtained using the KPCA scores. In addition to the previously mentioned parameters the algorithm also takes in the kernel principal components \( U \) the eigenvalues \( D \) and the multi-scale template ear shape \( E_{TMS} \). For the work presented here ear shapes were generated for values of
0 ≤ P ≤ 124. It is to be noted that when \( P = 0 \) the source ear shape is generated and when \( P = 124 \) the target ear shape is attained.

**Algorithm 11 Generate An Ear Using Source And Target KPCA Weights**

**Inputs:** \( U,D,v_s,v_l,P,E_{TMS},\sigma_V \).

1: for \( 1 \leq n \leq P \) do
2: \( v_{tmp}(n) = v_l(n) \)
3: end for
4: for \( P + 1 \leq n \leq L \) do
5: \( v_{tmp}(n) = v_s(n) \)
6: end for
7: \( a = \bar{a} + UD_1^{-1}v_{tmp} \)
8: \( \{a(t)\}_{0<\frac{n}{N} \leq 1} = S(E_{TMS},a,\sigma_V) \)
9: \( S = \mathcal{F}(E_{TMS},\{a(t)\}_{0<\frac{n}{N} \leq 1},\sigma_V) \)
10: return \( S,\{a(t)\}_{0<\frac{n}{N} \leq 1} \)

To perform acoustic analysis and also conduct FM-BEM simulations on the ear shapes, the ear shapes have to be attached to the template head and torso shapes. Algorithm 12 lists the additional steps involved in generating head and torso meshes with appropriate ear shapes attached for the purposes of FM-BEM simulations. Additional parameters required for the algorithm are the template head, torso and ear mesh \( HTE_T \) and the various values of \( P \) to generate the ear shapes.
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Algorithm 12 Generate Full Head And Torso Mesh Using Source And Target Ear KPCA Weights

\textbf{Inputs:} $U,D,v_{s},v_{t},\{P_{1} \ldots P_{M}\},\sigma_{V}, E_{T_{MS}}, HTE_{T}$.

1: \textbf{for} $1 \leq y \leq M$ \textbf{do}
2: \quad $\{\alpha^{s,l,y}(t)\}_{0 \leq t \leq 1} \leftarrow \text{GenerateAnEarUsingSourceAndTargetKPCAWeights}(U,D,v_{s},v_{t},P_{y},E_{T_{MS}},\sigma_{V})$\newline
3: \quad $HTE_{s,l,y} = F(HTE_{T_{MS}}, \{\alpha^{s,l,y}(t)\}_{1 \leq n \leq N}, \sigma_{V})$
4: \textbf{end for}
5: \textbf{return} $HTE_{s,l,y}$ for $1 \leq y \leq M$

7.3.3 Analysis Of Results

Two pair of ears were used for experimentation. The first pair, $S_{5}, S_{55}$ and the second pair, $S_{13}, S_{23}$. Fig. 7.4 and Fig. 7.5 show scatter plots of the absolute value for the kernel principal component scores of the evolving ear shapes (i.e. $S_{5}$ and $S_{13}$) and also a scatter plot of the empirical standard deviation of the kernel principal component scores belonging to the full population of ears. This plot is useful in order to obtain an understanding on how much change is applied to the kernel principal component as the ear shape is evolving.

Fig. 7.6 and Fig. 7.7 shows the morphology of the ears generated when moving between the source ear $S_{5}$ to target ear $S_{55}$, and when moving between the source ear $S_{13}$ and target ear $S_{23}$. The number of scores $P$ that belong to the target ear shape is indicated above each ear. The ear shape with zero scores is the source ear shape, while the ear shape with 124 scores is the target ear shape. Further, Fig. 7.6 highlights the differences between the given ear shape and the target ear shape using the CSDA technique (Sec. 4.3).
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Figure 7.4: The above figures displays the absolute value for the KPCA scores of the evolving ear shape $S_5$ with respect to the empirical standard deviation of the scores calculated for the population of 124 ear shapes.

Figure 7.5: The above figure displays the absolute value for the scores of the evolving ear shape $S_{13}$ with respect to the empirical standard deviation of the scores calculated for the population of 124 ear shapes.
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Figure 7.6: The above figure shows how the source ear shape $S_5$ evolves towards the target ear shape $S_55$ as more kernel principal component scores belonging to the target ear shape is used for the generation of the ear. The numbers $x$ shown above the ear shapes indicate that scores 1 to $x$ belong to the target ear shape while scores $x + 1$ to 124 belong to the source ear shape. The colours in the shapes highlight regions of difference between the given ear and the ear where all 124 scores are used for the construction of the ear shape.
Figure 7.7: The above figure shows how the source ear shape $S_{13}$ evolves towards the target ear shape $S_{23}$ as more kernel principal component scores belonging to the target ear shape is used for the generation of the ear. The numbers $x$ shown above the ear shapes indicate that scores 1 to $x$ belong to the target ear shape while scores $x + 1$ to 124 belong to the source ear shape. The colours in the shapes highlight regions of difference between the given ear and the ear where all 124 scores are used for the construction of the ear shape.
Fig. 7.8 - Fig. 7.9 and Fig. 7.10 - Fig. 7.11 show the acoustical responses in the median plane for the ears generated between $S_5, S_{55}$ and $S_{13}, S_{23}$. The numbers beside each figure indicates the number $P$ of KPCA scores for the target ear shape used when generating the ears. Another set of plots were generated in order to obtain further insight into the similarity or differences for the features in the DTF spectrum. Fig. 7.12 - Fig. 7.15 displays plots of the ears and a few DTFs belonging to the median plane at the specified angles. The black colour curve in the plot shows the DTF at the specified angle for the ear when $P$ scores of the target ear shape is used for generating the shape. The blue colour curve on the plots shows the DTF at the specified angle for the target ear shape.

The figures show that the acoustic responses for the ear shapes are evolving as the value for $P$ increases. Specifically, for the two pairs it can be observed that for $P \leq 10$ the features in the DTF spectrum do not resemble that of the target ear shape, while for values of $P \geq 10$ general features in the DTF spectrum in the form of the location of the peaks and notches resemble that of the target ear shape. It is to be remembered that the aim in this study was not to accurately obtain the DTF spectrum, rather it was aimed to estimate the number of scores needed in order for the general features in the DTF spectrum to be similar. The results further indicate that acoustic modelling of the ear shapes can be conducted in multiple scale. In the larger scale the interest is to obtain and match the broad features in the acoustic spectrum, and results show that around 10 to 20 kernel principal components are required to obtain such a matching. Finer levels of matching in the acoustic spectrum can be obtained if more kernel principal components are utilized. Particularly, it can be seen from Fig. 7.12 - Fig. 7.15 that finer tunings will involve adjusting the depth of a notch or a shift and or scaling in the DTF spectrum.
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Figure 7.8: The above figure shows an image of the acoustic response for the median plane plus a 3D surface plot for the response also in the median plane. The plots show how the acoustic responses change as the source ear shape $S_5$ evolves towards the target ear shape $S_{55}$ as more kernel principal component scores belonging to the target ear shape is used for the generation of the ear. The numbers $x$ shown next to the plots is the value $P$, that is the ear shape was generated by assigning scores 1 to $x$ from the target ear shape while scores $x + 1$ to 124 from the source ear shape.
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Figure 7.9: The above figure shows an image of the acoustic response for the median plane plus a 3D surface plot for the response also in the median plane. The plots show how the acoustic responses change as the source ear shape $S_5$ evolves towards the target ear shape $S_{55}$ as more kernel principal component scores belonging to the target ear shape is used for the generation of the ear. The numbers $x$ shown next to the plots is the value $P^*$, that is the ear shape was generated by assigning scores 1 to $x$ from the target ear shape while scores $x + 1$ to 124 from the source ear shape.
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Figure 7.10: The above figure shows an image of the acoustic response for the median plane plus a 3D surface plot for the response also in the median plane. The plots show how the acoustic responses change as the source ear shape $S_{13}$ evolves towards the target ear shape $S_{23}$ as more kernel principal component scores belonging to the target ear shape is used for the generation of the ear. The numbers $x$ shown next to the plots is the value $P$, that is the ear shape was generated by assigning scores 1 to $x$ from the target ear shape while scores $x + 1$ to 124 from the source ear shape.
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Figure 7.11: The above figure shows an image of the acoustic response for the median plane plus a 3D surface plot for the response also in the median plane. The plots show how the acoustic responses change as the source ear shape $S_{13}$ evolves towards the target ear shape $S_{23}$ as more kernel principal component scores belonging to the target ear shape is used for the generation of the ear. The numbers $x$ shown next to the plots is the value $P$, that is the ear shape was generated by assigning scores 1 to $x$ from the target ear shape while scores $x + 1$ to 124 from the source ear shape.
Figure 7.12: A plot of the source ear shape and the DTFs belonging to the median plane are shown around it. The source subject is $S_5$ and target subject is $S_{55}$. The black curve on the DTF plots is the DTF belonging to the displayed ear shape and the blue curve is the DTF belonging to the target ear shape. The angle for the DTFs are indicated on the plot of the shape. Positive angles correspond to the front of the subject while negative angles correspond to the back of the subject. Further, the angle zero corresponds to the location directly above the head.
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Figure 7.13: A plot of the ear shape where the first $P = 5$ scores belong to the target ear shape. The source subject is $S_5$ and target subject is $S_{S5}$. Further, the DTFs belonging to the median plane are shown around it. The black curve on the DTF plots is the DTF belonging to the displayed ear shape and the blue curve is the DTF belonging to the target ear shape. The angle for the DTFs are indicated above each DTF. Positive angles correspond to the front of the subject while negative angles correspond to the back of the subject. Further, the angle zero corresponds to the location directly above the head.
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Figure 7.14: A plot of the ear shape where the first $P = 10$ scores belong to the target ear shape. The source subject is $S_5$ and target subject is $S_{55}$. Further the DTFs belonging to the median plane are shown around it. The black curve on the DTF plots is the DTF belonging to the displayed ear shape and the blue curve is the DTF belonging to the target ear shape. The angle for the DTFs are indicated above each DTF. Positive angles correspond to the front of the subject while negative angles correspond to the back of the subject. Further the angle zero corresponds to the location directly above the head.
Figure 7.15: A plot of the ear shape where the first $P = 20$ scores belong to the target ear shape. The source subject is $S_5$ and target subject is $S_{55}$. Further the DTFs belonging to the median plane are shown around it. The black curve on the DTF plots is the DTF belonging to the displayed ear shape and the blue curve is the DTF belonging to the target ear shape. The angle for the DTFs are indicated above each DTF. Positive angles correspond to the front of the subject while negative angles correspond to the back of the subject. Further, the angle zero corresponds to the location directly above the head.
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Figure 7.16: A plot the source ear shape and the DTFs belonging to the median plane are shown around it. The source subject is $S_{13}$ and target subject is $S_{23}$. The black curve on the DTF plots is the DTF belonging to the displayed ear shape and the blue curve is the DTF belonging to the target ear shape. The angle for the DTFs are indicated on the plot of the shape. Positive angles correspond to the front of the subject while negative angles correspond to the back of the subject. Further, the angle zero corresponds to the location directly above the head.
Figure 7.17: A plot of the ear shape where the first $P = 5$ scores belong to the target ear shape. The source subject is $S_{13}$ and target subject is $S_{23}$. Further the DTFs belonging to the median plane are shown around it. The black curve on the DTF plots is the DTF belonging to the displayed ear shape and the blue curve is the DTF belonging to the target ear shape. The angle for the DTFs are indicated above each DTF, positive angles correspond to the front of the subject while negative angles correspond to the back of the subject. Further the angle zero corresponds to the location directly above the head.
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Figure 7.18: A plot of the ear shape where the first $P = 10$ scores belong to the target ear shape. The source subject is $S_{13}$ and target subject is $S_{23}$. Further the DTFs belonging to the median plane are shown around it. The black curve on the DTF plots is the DTF belonging to the displayed ear shape and the blue curve is the DTF belonging to the target ear shape. The angle for the DTFs are indicated above each DTF, positive angles correspond to the front of the subject while negative angles correspond to the back of the subject. Further the angle zero corresponds to the location directly above the head.
Figure 7.19: A plot of the ear shape where the first $P = 20$ scores belong to the target ear shape. The source subject is $S_{13}$ and target subject is $S_{23}$. Further the DTFs belonging to the median plane are shown around it. The black curve on the DTF plots is the DTF belonging to the displayed ear shape and the blue curve is the DTF belonging to the target ear shape. The angle for the DTFs are indicated above each DTF, positive angles correspond to the front of the subject while negative angles correspond to the back of the subject. Further the angle zero corresponds to the location directly above the head.
7.3.4 Distance Analysis Between Morphology and Acoustics

A distance analysis was carried out for the morphology and acoustics between the shapes $HTE_{s,l,y}$ and $HTE_{s,l,124}$ where $y \in \{0, 5, 10, 20, 40, 50, 60, 80, 100, 124\}$ for the given pairs of the source and the target subjects (i.e. $s \in [5, 13]$ and $l \in 55, 23$). The acoustic distances were calculated using function $V$, namely, the SDS measure described in Sec. 4.8.4 and the morphological distances were calculated using three different measures:

1. Frobenius norm between the initial momentum vector $\{\alpha_{s,l,y}(0)\}_{1 \leq n \leq N}$ and $\{\alpha_{s,l,124}(0)\}_{1 \leq n \leq N}$
2. The Hausdorff distance between the ear shapes $E_{s,l,y}$ and $E_{s,l,124}$ (Sec. 4.3.1)
3. the mean CSDA distances using the function $\bar{d}(E_{s,l,y}, E_{s,l,124})$ (Sec. 4.3.2) between the ear shapes $E_{s,l,y}$ and $E_{s,l,124}$

Fig. 7.20 shows the acoustic distance calculated using the function $V$ and Fig. 7.21 - Fig. 7.22 shows the morphological distances between the source and target ear shapes using the previously discussed measures. It can be observed that while the morphological distances decrease monotonically, however the same is not true for the acoustic differences. Although the acoustic differences also converge to zero when the number of principal components used is large, however the decrease in the acoustic differences was non-monotonic (i.e. it is not always decreasing). In particular, for the source and target shapes $S_5$ and $S_{55}$ it can be observed that the acoustic differences decrease monotonically between 0 and 40 principal components, however, there is a small increase in the acoustic differences between 40 and 80 principal components.
7.3. Significance of the Kernel Principal Components for the Acoustic Study on Ear Shapes

Figure 7.20: The above figure shows the acoustic differences measured using the function $V$, between source ear shapes $E_{s,l,y}$ and target ear $E_{s,l,124}$, where $y \in \{0, 5, 10, 20, 40, 50, 60, 80, 100, 124\}$ and $s \in \{5, 13\}$ and $l \in \{55, 23\}$. The legend shows the pair numbers used. The black line shows the average of the two curves.
Figure 7.21: The above figure shows the morphological differences measured using (1) the Frobenius norm (2) Hausdorff distance (3) the mean CSDA value $\bar{d}$ and (4) the morphological distance $\bar{d}$ versus the acoustic distance $\bar{V}$ between source ear shapes $E_{5,55,y}$ and target ear $E_{5,55,124}$, where $y \in [0,5,10,20,40,60,80,100,124]$. 
Figure 7.22: The above figure shows the morphological differences measured using (1) the Frobenius norm (2) Hausdorff distance (3) the mean CSDA value \( \bar{d} \) and (4) the morphological distance \( \bar{d} \) versus the acoustic distance \( \bar{V} \) between source ear shapes \( E_{13,23,y} \) and target ear \( E_{13,23,124} \), where \( y \in \{0, 5, 10, 20, 40, 50, 80, 100, 124\} \).
7.4 Examining the Kernel Principal Components

In this section the kernel principal components belonging to a population of ear shapes that are aligned to the multi-scale template ear shape using affine transformation are analysed. This study provides further insight into the most common and important features in the ear shape that change across a population of ears and explores the corresponding changes in ear acoustics.

It may be recalled that in Sec. 7.1 it was shown how a KPCA over a given population of ear shapes is conducted and how the kernel principal components in the form of a matrix $U$ are obtained. For a given population of ears each of the columns in the KPCA bases matrix $U$, here denoted as $u_p$, encodes some of the variations that exists in the initial momentum vectors and ultimately captures some of the variations in the ear shapes across the population. The subscript $p$, signifies the $p^{th}$ column of the kernel principal component matrix. The aim of this section is to examine both the morphology and acoustics for the variations captured by the first few kernel principal components individually. The benefit of such an analysis is that it can provide an overall insight into the variations seen in the ear shapes and ultimately their respective acoustic spectrum. The ear shapes that are examined in this section consist of 62 left ear shapes and 62 mirrored right ear shapes from the SYMARE database, which add up to $L = 124$ ears in total. As indicated in the introduction to this chapter, the ear shapes that are examined here are aligned to the template ear shape using an affine transformation. It may be recalled that the method for aligning the ear shapes to the template ear shape was explained previously in Sec. 5.2.4.1.

7.4.1 Methodology

In this section the procedure for analysing variations in the ear morphology that is captured by changing a single kernel principal component is described. It also details how to compute the acoustic responses for the generated ear shape that
belongs to the changed kernel principal component.

By changing the kernel principal components only the ear shapes are changed, however, in order to obtain the acoustic responses for the ear shapes using FM-BEM the generated ear shapes are attached to the template head and torso shape. In order to obtain deformed ear shapes located on a head and torso surface, techniques similar to Sec. 4.2 for deforming of the head and ear shape is used. Algorithm 13 describes the procedure to obtain a full head, torso and ear mesh such that the ear shapes are obtained by changing a single kernel principal component. The inputs to the algorithm are:

1. The p\textsuperscript{th} kernel principal component vector \( u_p \).
2. The mean momentum vector \( \bar{a} \) corresponding to the population of shapes.
3. A scalar value \( c_{u_p} \) that multiplies the given kernel principal component.
4. The multi-scale template ear shape \( E_{TMS} \) and the template head, torso shape \( HTE_{TMS} \). (The procedure for making these template shapes was discussed in Sec. 5.1.3 and Sec. 5.2.2)

**Algorithm 13** Obtain full head, torso and ear mesh shapes by changing the kernel principal components

**Inputs:** \( u_p, \bar{a}, c_{u_p}, E_{TMS}, HTE_{TMS} \).

1: \( \alpha^{(p,c_{u_p})} = c_{u_p} u_p \)
2: \( \{\alpha^{(p,c_{u_p})}_n(t)\}_{0 < t < N} = S(E_{TMS}, \alpha^{(p,c_{u_p})} + \bar{a}, \sigma_V) \)
3: \( S = F(HTE_{TMS}, \{\alpha^{(p,c_{u_p})}_n(t)\}_{0 < t < N}, \sigma_V) \)
4: return \( S \)

Importantly, step 3 in Algorithm 13 applies the flow corresponding to the changed principal component to the template head and torso and ear shape, \( HTE_{TMS} \). The value of \( c_{u_p} \) that is used in the algorithm is based upon statistical information on the scores corresponding to the kernel principal components and is further described below. The distribution of the scores for the first five principal components is shown.
in Fig. 7.23. Recall that the matrix $\mathbf{D}^{\frac{1}{2}} \mathbf{V}^T$, which was computed in Sec. 7.1, contains all

the scores for the kernel principal components belonging to all ears in SYMARE. The
KPCA computes the kernel principal components in a manner that the variance of the scores across the whole population decreases as the order of the kernel principal components increase. Fig. [7.24] shows a plot of the empirical standard deviation for the scores versus the principal component number for the left and right population of ear shapes in the SYMARE database.

![Standard Deviation in the Scores](image)

Figure 7.24: The standard deviation for the scores belonging to each kernel principal component. The scores get multiplied with the kernel principal components. There is a total of 124 kernel principal components.

In the experiments discussed in this section, the scalar value, \( c_{u_p} \), multiplies the kernel principal component, \( u_p \), replacing the corresponding score that would normally be multiplied when reconstructing an ear shape. The values chosen for \( c_{u_p} \) in our experimentation is computed using the following equation:

\[
 c_{u_p} = mD_{pp}^{1/2} 
\]  

(7.15)

Where \( m \in \mathbb{R} \) is some real number. For example, when \( m = 2 \) the value of \( c_{u_p} \) will be two standard deviation away from the mean score for the \( p^{th} \) kernel principal component, and when \( m = 0 \) the kernel principal component is not used, that is, the mean initial momentum vector is used only. The mean value was previously denoted...
as $\tilde{a}$ in Sec. 7.1. In the text that follows $c_{u_p}$ is referred to as simply $c_p$. Fig. 7.25 shows the value of $c_p$ for the first five kernel principal components when $-5 \leq m \leq 5$. The red, green, blue, cyan and magenta crosses are for the first, second, third, fourth and fifth principal components respectively. It is noted that as the value of $|m|$ increases deformations applied to the template ear shape will be larger. The section to follow will detail the experimental set-up and the range of $m$ used.

### 7.4.2 Experimental Setup

Due to the extensive computational time required for obtaining the acoustic responses for the ear shapes using FM-BEM, only the first five kernel principal component vectors were examined. As indicated in the introduction to this chapter, selecting the first five principal components was a logical choice as they would capture the largest variations in the momentum vectors and, consequently, the ear shapes. More specifically, acoustic and morphological simulations were performed on ear shapes when Algorithm 13 was applied where $1 \leq p \leq 5$ and for values of $m$
in the set $B_1$:

$$B_1 = \{ \pm 2.5, \pm 2, \pm 1.5, \pm 1, \pm 0.5, 0 \}$$

(7.16)

Fig. 7.23 shows the probability distribution of the scores for the first five kernel principle components using histogram plots.

It will be seen in the results section for values of $|m| \leq 2.5$ the changes in the ear morphology caused by changing the specified kernel principal component might not be very visible. For this reason ears shapes were also generated for values of $m = \pm 5$ and $m = \pm 7$ in order to magnify the changes caused by the kernel principal components. The morphological data is shown for $m$ in the set $B_2$:

$$B_2 = \{ \pm 7, \pm 5, \pm 2.5, \pm 2, \pm 1.5, \pm 1, \pm 0.5, 0 \}$$

(7.17)

For the analysis to follow the ear shapes generated by the values of $m$ are denoted as $E^{(p,m)}_{T_{MS}}$, which indicates that the ear shape $E^{(p,m)}_{T_{MS}}$ was generated from the multi-scale template ear shape $E_{T_{MS}}$ by changing the kernel principal component, $p$, by a magnitude of $m$.

### 7.4.3 Results

The previous section explained the set-up of the experiments, and detailed the range of $m$ used for generating the ear morphology and acoustic data. The results presented in this section use four kinds of figures to show changes in the morphological and acoustic data.

- The first set of figures show the morphological changes in the ear shapes. An example of such a figure is Fig. 7.27. These figures show how the morphology of the ear shapes change for various values of $m$. The value of $m$ is shown above each ear shape. Further Appendix E to this thesis plots the variations captured by some other kernel principal components.

- The second set of figures indicate the acoustic changes in the median plane
as an image. An example of such plots are seen in Fig. 7.28. In Fig. 7.28 the numbers appearing in the left hand column indicate the value of $m$. These plots use different colours to show how the gain in the median plane changes as a function of elevation and frequency. As discussed previously, the front section of the head is represented by positive angles and the back of the head is represented by negative angles. Further, the angle directly below the head and torso is either $\pm \pi$ and the angle directly above the head is zero degrees.

- The third set of figures present the DTF spectrum in the median plane as a surface plot, which are used in addition to the images to provide a global view of the DTF spectrum in the median plane. An example of such plots are seen in Fig. 7.28. In particular, by the use of these plots notches and peaks can then be easily identified.

- The fourth set of figures present the morphology of the ear shape and also the DTF spectrum for six angles in the median plane. In these plots the value of $m$ is given above the ear shape, while the angle for the DTF spectrum is shown above the DTF spectrum. An example of such plots are seen in Fig. 7.30. Observing individual DTF spectrum’s can be very useful because features in the DTF spectrum in the form of peaks and notches can be easily compared for different ear shapes and, also, for different angles in the median plane.

Figure 7.26: The acoustic response of the template ear shape for the median plane.

It is noted that the acoustic and morphological variation of ears GUI (AMVE-GUI)
detailed in Appendix [H] was also used to analyse the variations in the morphology and acoustics of the ears obtained from the KPCA analysis.
Figure 7.27: Variations in the ear morphology by changing the first principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2.5, -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5, 5, 7$. These numbers are shown above the ear shape. Further, the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure 7.28: Variations in the acoustics by changing the first principal component and for the median plane only is shown as an image and surface plot. For both the image and surface plots positive angles indicate directions in the front of the head, while negative angles indicate directions at the back of the head. The value for $m$ corresponding to the acoustic plots is shown on the left hand side of the plot.
Figure 7.29: Variations in the acoustics by changing the first principal component and for the median plane only is shown as an image and surface plot. For both the image and surface plots positive angles indicate directions in the front of the head, while negative angles indicate directions at the back of the head. The value for $m$ corresponding to the acoustic plots is shown on the left hand side of the plot.
7.4. Examining the Kernel Principal Components

Figure 7.30: Plots of the ear shape and some DTFs for angles in the median plane are shown above. The ear shape is generated by changing the first kernel principal component. The value $m$ used for the generation of the ear shape is given above the ear, while the angle of the DTF is shown above the DTF spectrum.
Figure 7.31: Plots of the ear shape and some DTFs for angles in the median plane are shown above. The ear shape is generated by changing the first kernel principal component. The value $m$ used for the generation of the ear shape is given above the ear, while the angle of the DTF is shown above the DTF spectrum.
Figure 7.32: Variations in the ear morphology by changing the second principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2.5, -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5, 5, 7$. These numbers are shown above the ear shape. Further, the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure 7.33: Variations in the acoustics by changing the second principal component and for the median plane only is shown as an image and surface plot. For both the image and surface plots positive angles indicate directions in the front of the head, while negative angles indicate directions at the back of the head. The value for $m$ corresponding to the acoustic plots is shown on the left hand side of the plot.
Figure 7.34: Variations in the acoustics by changing the second principal component and for the median plane only is shown as an image and surface plot. For both the image and surface plots positive angles indicate directions in the front of the head, while negative angles indicate directions at the back of the head. The value for \( m \) corresponding to the acoustic plots is shown on the left hand side of the plot.
7.4. Examining the Kernel Principal Components

Figure 7.35: Plots of the ear shape and some DTFs for angles in the median plane are shown above. The ear shape is generated by changing the second kernel principal component. The value \( m \) used for the generation of the ear shape is given above the ear, while the angle of the DTF is shown above the DTF spectrum.
Figure 7.36: Plots of the ear shape and some DTFs for angles in the median plane are shown above. The ear shape is generated by changing the second kernel principal component. The value $\sigma$ used for the generation of the ear shape is given above the ear, while the angle of the DTF is shown above the DTF spectrum.
The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2.5, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5, 5, 7$. These numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure 7.38: Variations in the acoustics by changing the third principal component and for the median plane only is shown as an image and surface plot. For both the image and surface plots positive angles indicate directions in the front of the head, while negative angles indicate directions at the back of the head. The value for $m$ corresponding to the acoustic plots is shown on the left hand side of the plot.
7.4. Examining the Kernel Principal Components

Figure 7.39: Variations in the acoustics by changing the third principal component and for the median plane only is shown as an image and surface plot. For both the image and surface plots positive angles indicate directions in the front of the head, while negative angles indicate directions at the back of the head. The value for $m$ corresponding to the acoustic plots is shown on the left hand side of the plot.
Figure 7.40: Plots of the ear shape and some DTFs for angles in the median plane are shown above. The ear shape is generated by changing the third kernel principal component. The value $m$ used for the generation of the ear shape is given above the ear, while the angle of the DTF is shown above the DTF spectrum.
Figure 7.41: Plots of the ear shape and some DTFs for angles in the median plane are shown above. The ear shape is generated by changing the third kernel principal component. The value $m$ used for the generation of the ear shape is given above the ear while the angle of the DTF is shown above the DTF spectrum.
Figure 7.42: Variations in the ear morphology by changing the fourth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2.5, -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5, 5, 7$. These numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure 7.43: Variations in the acoustics by changing the fourth principal component and for the median plane only is shown as an image and surface plot. For both the image and surface plots positive angles indicate directions in the front of the head, while negative angles indicate directions at the back of the head. The value for $m$ corresponding to the acoustic plots is shown on the left hand side of the plot.
Figure 7.44: Variations in the acoustics by changing the fourth principal component and for the median plane only is shown as an image and surface plot. For both the image and surface plots positive angles indicate directions in the front of the head, while negative angles indicate directions at the back of the head. The value for $m$ corresponding to the acoustic plots is shown on the left hand side of the plot.
Figure 7.45: Plots of the ear shape and some DTFs for angles in the median plane are shown above. The ear shape is generated by changing the fourth kernel principal component. The value \( m \) used for the generation of the ear shape is given above the ear, while the angle of the DTF is shown above the DTF spectrum.
Figure 7.46: Plots of the ear shape and some DTFs for angles in the median plane are shown above. The ear shape is generated by changing the fourth kernel principal component. The value \( m \) used for the generation of the ear shape is given above the ear while the angle of the DTF is shown above the DTF spectrum.
Figure 7.47: Variations in the ear morphology by changing the fifth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2.5, -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5, 5, 7$. These numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
7.4. Examining the Kernel Principal Components

Figure 7.48: Variations in the acoustics by changing the fifth principal component and for the median plane only is shown as an image and surface plot. For both the image and surface plots positive angles indicate directions in the front of the head, while negative angles indicate directions at the back of the head. The value for $m$ corresponding to the acoustic plots is shown on the left hand side of the plot.
Figure 7.49: Variations in the acoustics by changing the fifth principal component and for the median plane only is shown as an image and surface plot. For both the image and surface plots positive angles indicate directions in the front of the head, while negative angles indicate directions at the back of the head. The value for \( m \) corresponding to the acoustic plots is shown on the left hand side of the plot.
Figure 7.50: Plots of the ear shape and some DTFs for angles in the median plane are shown above. The ear shape is generated by changing the fifth kernel principal component. The value $m$ used for the generation of the ear shape is given above the ear, while the angle of the DTF is shown above the DTF spectrum.
Figure 7.51: Plots of the ear shape and some DTFs for angles in the median plane are shown above. The ear shape is generated by changing the fifth kernel principal component. The value $m$ used for the generation of the ear shape is given above the ear, while the angle of the DTF is shown above the DTF spectrum.
7.4.4 Analysis of Morphological Results

In this section, a qualitative analysis of the variation of the ear shapes with respect to the reference template ear shape when changing the kernel principal components is presented. Performing a qualitative analysis on the changes occurring to the ear shapes is difficult, because changes occurring to the ear shapes are complicated. In the LDDMM framework a change or deformation in a given section of the ear, such as enlarging the Cavum-Concha or Crus-Anti-Helix, will also cause deformations to the surrounding regions. Naturally the changes that happen in the ear shapes is gradual as one moves away from the template ear shape and these changes increase when $|m|$ gets larger. The external ear shape and the different anatomical regions with numbering is shown in Fig. 7.52. These anatomical numberings will be used when referring to ear shape changes. The numbers corresponding to the anatomical regions is also listed below:

1. Helix
2. Scaphoid-Fossa
3. Cymba-Concha
4. Scapha
5. Anti-Helix-Stem
6. Lobule/Lobe
7. Anti-Tragus
8. Intertragic notch/Incisura
9. Cavum-Concha
10. Tragus
11. Crus-Helix
12. Interior-Crus-Anti-Helix

13. Triangular-Fossa

14. Superior-Crus-Anti-Helix

Two kinds of observations are made on the generated ear shapes. The first set of observations on the ear shapes are made for $|m| < 2.5$. For this range of $m$ large structural differences in the ear shapes are not easily visible. However, changes to different sections of the ear can be observed utilizing the colours produced by the CSDA, which highlights shape differences with respect to the template ear shape. Table 7.1 lists regions of the ear shape that have changed with respect to the template ear shape using their anatomical numbers for when $|m| \leq 2.5$. Generally the ears exhibit different kind of changes for when $m > 0$ compared to the case when $m < 0$ for a given principal component.

![Figure 7.52: The external ear shape and the regions of the ear plus the anatomical names for the regions. The numbers used to indicated the anatomical region will be used extensively in this section to refer to the regions of the ear that has changed. The red lines indicate the vertical and horizontal axis which are also used for describing the changes in the ear shapes.](image-url)
### Morphological Changes In Ears, $|m| \leq 2.5$

<table>
<thead>
<tr>
<th>KPC Number</th>
<th>Changes for $m &lt; 0$ (Anatom Sec Num)</th>
<th>Changes for $m &gt; 0$ (Anatom Sec Num)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>• 2, 6, 9, 10</td>
<td>• 5, 6, 9, 10</td>
</tr>
<tr>
<td>2</td>
<td>• 4, 5, 6, 12</td>
<td>• 1, 3, 6, 9, 10, 13.</td>
</tr>
<tr>
<td>3</td>
<td>• 1, 4, 5, 6</td>
<td>• 1, 5, 6, 9, 13.</td>
</tr>
<tr>
<td>4</td>
<td>• 1, 3, 6</td>
<td>• 3, 4, 6, 10</td>
</tr>
<tr>
<td>5</td>
<td>• 1, 4, 6, 11</td>
<td>• 1, 5, 6, 7, 10.</td>
</tr>
</tbody>
</table>

Table 7.1: Summary of small morphological changes seen in the ear shapes captured by the kernel principal components.
On the other hand, when $|m| > 2.5$ the changes made to the kernel principal component are amplified and this made analysing the structural differences in ear shapes with respect to the template ear easier. When considering the ear shapes for $|m| > 2.5$ the large scale changes in the ear shapes will be explained using five basic operations. These operations are folding, stretching, compression and pulling. These operations take place with respect to a given axis, a specific point or region of the ear shape. The vertical axis is the axis that runs from the top of the Helix to the bottom of the Lobe, while the horizontal axis connects the Anti-helix to the Tragus. These axis are also shown in Fig. 7.52. Table 7.2 summarizes the observations made on the large scale morphological changes that occur to the ear shape with respect to the template ear based on the ears for when $|m| > 2.5$. 
Large Scale Morphological Changes In Ears, $|m| > 2.5$

<table>
<thead>
<tr>
<th>KPC Number</th>
<th>changes for $m &lt; 0$</th>
<th>changes for $m &gt; 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>• The ear appears to be wider in width with a bigger Concha and Superior-Crus-Anti-Helix regions.</td>
<td>• The ear appears to be folded inwards and compressed with respect to the horizontal axis. • Narrower ear.</td>
</tr>
<tr>
<td>2</td>
<td>• Larger Superior-Crus-Anti-Helix. • Anti-Helix-stem is moving out of the ear. • Flatter Lobe.</td>
<td>• The opening of the Concha is very wide, making the Anti-Helix region smaller and pushing down the Anti-Tragus. • The Helix of the ear appears to be larger. • Helix and the Lobe are ticker and moving out of the ear.</td>
</tr>
<tr>
<td>3</td>
<td>• The lobe is stretching and is further folding towards inside the head. This causes the Incisura to be pulled along the lobe as well. • There appear to be some folding with respect to the vertical axis.</td>
<td>• The Anti-Helix is folded towards inside the head. • The Lobe is folded out of the ear shape. • The Helix tail is forming a sharp angle. • The Helix is moving outwards.</td>
</tr>
<tr>
<td>4</td>
<td>• Folding in the ear shape with respect to the horizontal axis. • A pointy Helix in the top which could also indicate the ear is compressed with respect to the horizontal axis. • Ticker Lobe. Ticker Helix tail.</td>
<td>• Tragus being pulled towards the front of the head. The Concha appears to be stretched. • The Lobe moving downwards.</td>
</tr>
<tr>
<td>5</td>
<td>• Folding of the Helix towards inside of the ear shape.</td>
<td>• Widened Helix tail, making a sharp angle at the bottom. Anti-Helix and Anti-Tragus moving outwards.</td>
</tr>
</tbody>
</table>

Table 7.2: Summary of large morphological changes seen in the ear shapes that are captured by the kernel principal components.
7.4.5 Morphology and Acoustic Distance Analysis

This section presents a quantitative analysis on the morphology and acoustics of the ear shapes generated by changing the first five kernel principal components. Distances with respect to the template ear are calculated for the morphology and acoustics of the ears using the tools developed in Sec. 4.3.2 and Sec. 4.8.4. The acoustic analysis presented in this section will be focused on the DTF spectrum of the median plane. To remind the reader a plot of the median plane and the convention for representing sound source directions in the median plane is shown in Fig. 7.53.

![Figure 7.53: The convention used when representing angles in the median plane.](image)

A distance analysis for the morphology between the generated ear shapes, $E_{T_{MS}}^{(p,m)}$, and the template ear shape, $E_{T_{MS}}$, was performed and is shown in Fig. 7.54. The CSDA measure (discussed in Sec. 4.3.2) was used to calculate the overall distances...
between the ear shapes. More specifically, the function $\bar{d}(E_{T_{MS}}^{(p,m)}, E_{T_{MS}})$ (given in Eq. (4.19)) was used to compute the overall dissimilarity between the surfaces. The plots in Fig. 7.54 show the mean CSDA value between $E_{T_{MS}}^{(p,m)}$ and $E_{T_{MS}}$ for $1 \leq p \leq 5$ and $-5 \leq m \leq 5$.

Figure 7.54: Morphology distance between the ear $E_{T_{MS}}^{(p,m)}$ and the template ear $E_{T_{MS}}$ for $1 \leq p \leq 5$ and $-5 \leq m \leq 5$, measured using the CSDA technique. First principal component (blue line), second principal component (red line), third principal component (green line), fourth principal component (black line), fifth principal component (purple line).

More over an acoustic distance analysis was performed on the generated ear shapes $E_{T_{MS}}^{(p,m)}$ with respect to the template ear $E_{T_{MS}}$ on the median plane. Recall that in Sec. 4.7, the log magnitude of the DTFs in the median plane were denoted by $W_{E_{T_{MS}}^{(p,m)}}(f, \phi)$ where $\phi$ varies between $-\pi \leq \phi \leq +\pi$. Fig. 7.55 plots the acoustic distance between $E_{T_{MS}}^{(p,m)}$ and $E_{T_{MS}}$ using the function $\bar{V}_{E_{T_{MS}}^{(p,m)}, E_{T_{MS}}}(w)$ (for further details see Eq. (4.43)) and Fig. 7.56 shows a plot of the acoustical distance $\bar{V}_{E_{T_{MS}}^{(p,m)}, E_{T_{MS}}}$ versus the morphological distance $\bar{d}(E_{T_{MS}}^{(p,m)}, E_{T_{MS}})$. It is observed from the above plots that as $|m| \to 0$ both the morphological and acoustical distances converge monotonically to zero (i.e. $\bar{d}(E_{T_{MS}}^{(p,m)}, E_{T_{MS}}) \to 0$ and $\bar{V}_{E_{T_{MS}}^{(p,m)}, E_{T_{MS}}} \to 0$ when $|m| \leq 2.5$).
7.4. Examining the Kernel Principal Components

Figure 7.55: Acoustic distance between the ear $E_{T_{MS}}^{(p,m)}$ and the template ear $E_{T_{MS}}$ for $1 \leq p \leq 5$ and $-2.5 \leq m \leq 2.5$, measured using the function $\bar{V}$. First principal component (blue line), second principal component (red line), third principal component (green line), fourth principal component (black line), fifth principal component (purple line).

Figure 7.56: Shape distance versus acoustical distance calculated using the functions $\bar{V}_{E_{T_{MS}}^{(p,m)} E_{T_{MS}}}$ and $\bar{d}_{E_{T_{MS}}^{(p,m)} E_{T_{MS}}}$ respectively. First principal component (blue line), second principal component (red line), third principal component (green line), fourth principal component (black line), fifth principal component (purple line).

### 7.4.6 Acoustic Variations in the Median Plane

We now present an analysis that shows regions in the median plane where changing the kernel principal components had the biggest impact on the acoustics. For the
7.4. Examining the Kernel Principal Components

In this section, the median plane DTFs are denoted similar to the previous section. Fig. 7.57 shows the variations of the gain in the DTF, measured using the function $Y_p(f, \phi)$:

$$Y_p(f, \theta) = \sqrt{\sum_{m=1}^{11} (W_{E_p,m}^{TMS}(f, \phi) - W_{E_p,m}^{TMS}(f, \phi))^2}$$  \hspace{1cm} (7.18)

$$W_{E_p,m}^{TMS}(f, \phi) = \sum_{m=1}^{11} W_{E_p,m}^{TMS}(f, \phi)$$  \hspace{1cm} (7.19)

$Y_p(f, \phi)$ is a function of frequency and elevation angle and the subscript $p$ in $Y_p$ indicates the principal component number which is in the range of $1 \leq p \leq 5$.

In Fig. 7.57, regions with brighter colour show large variations in the gain as the value of $m$ changes. The following summarizes some of the observations made:

- At low frequencies, approximately below 5kHz, the variation in the gain with respect to $m$ is very small as indicated by the dark regions.

- As the frequency increases the variation in the DTF spectrum increases in certain regions of space.

- Regions of large and small variation appear in bands at high frequencies (i.e. above 5kHz).
Figure 7.57: The above figure shows the variation of the gain using the function $Y_p(f, \theta)$ where $1 \leq p \leq 5$ and $-\pi \leq \phi \leq +\pi$. The value of $p$ (principal components) are indicated above the plots. The bright areas in each plot indicate regions of high variation in the acoustics, while darker regions indicate regions of no or small variation in the acoustics as the value of $m$ changes.
7.5 Mechanisms That Generate Changes In The DTF Spectrum

The question that is investigated in this section is whether one or more features in the DTF spectrum, in the form of peaks or notches, change systematically as the ear shapes change? This is a very challenging question and has been the focus of much research in the scientific community (Mokhtari et al., 2010, 2011; Takemoto et al., 2012b). The approach to answering the above question is conducted in two stages. In the first stage, the information contained in the DTF spectrum is modelled using PCA. Utilizing the obtained model of the DTF’s by means of the PCA, the features in the principal components are analysed to understand the underlying mechanisms that cause the changes seen in the DTF spectrum for the generated ear shapes.

It is worth mentioning that the analysis presented in this section can greatly assist in building a simple model that links the shape morphology to the acoustic response. To be more specific such a model can help in predicting the acoustic responses for a given angle in the median plane for a set of ear shapes without the need of acoustic simulations or measurements. The analysis and the model presented here is based on limited data and can be used to model the acoustic responses of a small set of ear shapes and their corresponding acoustics. Nevertheless, it does show the usefulness of KPCA towards the understanding and modelling of the acoustic responses of ear shapes.

In the next section a simple analysis of the acoustic gain is presented for a few angles and frequencies in the median plane. This analysis provides further insight to the changes in the DTF spectrum and justifies the use of PCA which is a more global approach to the analysis of the DTF data.

7.5.1 Gain Analysis On Selected Frequencies and Angles

An analysis was performed on the variation of the gain of the DTF spectrum over a range of \( m, -2.5 \leq m \leq 2.5 \) (see Eq. 7.15), and a few frequencies and angles in the
median plane. The aim of this analysis was to obtain a better understanding on how the gain of the DTF spectrum changes for the given angles and frequencies.

Fig. 7.58 shows the changes in the gain of the DTF spectrum as a function of $m$ for angles $a = \left[\frac{3\pi}{4}, \frac{\pi}{2}, \frac{\pi}{4}\right]$ and frequency components $f = \{7kHz, 11kHz, 13kHz, 15kHz\}$ for the first five principal components. The curves with different colour correspond to different principal components.

The results show that the gain exhibits a complex relationship with respect to $m$. In particular the the pattern of change for the gain is different at different frequencies and angles. Further a simple modelling of the gain might be both inefficient complicated. The next section will show how a simple model of the acoustic responses can be achieved using PCA.

### 7.5.2 Methodology: Modelling the DTFs in the Median Plane Using PCA

The modelling of the DTF data is now described using PCA. Further, a new error metric is proposed to examine the strength of the PCA model for the DTF data. Modelling the DTF spectrum directly is a very challenging task because the DTF spectrum can be very complicated in shape and, among other features, can contain many peaks and notches. To obtain a better appreciation on the complexity of the DTFs, the DTFs belonging to specific angles in the median plane can be seen in Fig. 7.51.

In order to model the DTF spectrum the information contained here is simplified by applying the PCA. The method for conducting the PCA was described previously in Sec. 2.7. At this point there are two options for performing the PCA on the acoustic data in the median plane. The option that is chosen will also effect the modelling of the acoustic data. The options are:

- The PCA can be performed for all angles in the median plane for the given ear shapes.
7.5. Mechanisms That Generate Changes In The DTF Spectrum

- The PCA can be performed for a specific angle in the median plane for the given ear shapes.

Each of the above options will be examined in more detail in Sec. 7.5.3, however at first the following section will describe a quantitative measure to examine the

Figure 7.58: Plots of the variation of the gain with respect to the value $m$ for the given angles and frequencies. The angles are at $\frac{3\pi}{4}$, $\frac{\pi}{2}$, $\frac{\pi}{4}$. The frequencies are at: $f = [7kHz, 11kHz, 15kHz]$ and the value of $m$ varies in the range of $[-2.5, -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5]$
quality for the reconstruction of the DTFs.

### 7.5.2.1 Iteratively Reconstructing a DTF and Calculating the Reconstruction Error

This section describes a quantitative measure to examine the quality for the reconstruction of the DTFs. To begin the discussion the notation introduced in Sec. 4.7 is used to represent the DTFs in the median plane. In the first step one or more DTFs that belong to single and or multiple ear shapes are concatenated into a data matrix that is denoted by $X$:

$$X = \begin{bmatrix} W_{E_{1,-2.5}} \\ \vdots \\ W_{E_{1,2.5}} \end{bmatrix}_{R \times M} \quad (7.20)$$

In the above the data matrix is of size $R \times M$, where $N$ is the number of DTFs in the median plane for a given ear shape, $M$ is the number of frequency bins in the DTFs and $r = 11$ which corresponds to 11 ear shapes that are generated for values of $m \in B_1$ for the first kernel principal component. Next, a PCA analysis is performed on the data matrix $X$ as described in Sec. 2.7 where by the principal components are in columns of $F \in R^{M \times M}$, and the score matrix $Q \in R^{N \times M}$ and mean vector $\bar{m} \in R^M$ is obtained. An accurate and complete reconstruction of a specific DTF identified by row $b$ in the data matrix $X$ can be made by multiplying its weights with the principal components:

$$\bar{m} + q_b F \quad (7.21)$$

where $q_b$ is the $b^{th}$ column of $Q^T$. In order to get a better understanding on how many acoustic principal components are required for reconstructing the DTFs, the error for reconstructing a DTF using the function $\tilde{V}$ (Eq. (4.40)), namely, SDS which was introduced previously in Sec. 4.8.4 was employed. The process of reconstructing a DTF is done iteratively by adding more weights and principal components to the analysis. Algorithm [14] shows how the error for reconstructing a particular DTF
denoted by $x_b$ is computed. In the algorithm $f_j$ denotes the $j^{th}$ column of the matrix $F$ and the DTF $x_b$ is the $b^{th}$ column in the matrix $X^T$.

Algorithm 14  Compute Error when Reconstructing a DTF

**Inputs:** $F, Q, \bar{m}, x_b$.

1. for $i = 1$ to $M$ do  
2. $\hat{x}_b = \bar{m}$  
3. for $j = 1$ to $i$ do  
4. $\hat{x}_b = \hat{x}_b + f_j(Q_{bj})$  
5. end for  
6. $e(i) = \bar{V}(\hat{x}_b, x_b)$  
7. end for

8. return $e \in \mathbb{R}^{1 \times M}$

Algorithm 14 calculates the error for a single DTF in the data matrix $X$ as more principal components are progressively added. The output to Algorithm 14 is a row vector $e \in \mathbb{R}^{1 \times M}$, Algorithm 15 extends Algorithm 14 to calculate the reconstruction error for every DTF in the data matrix $X$.

Algorithm 15  Compute Error when Reconstructing a DTF

**Inputs:** $F, Q, \bar{m}, X$.

1. $E \leftarrow []$  
2. for $b = 1$ to $N$ do  
3. $x \leftarrow$ row $b$ of $X$  
4. $e \leftarrow$ ComputeErrorwhenReconstructingaDTF($F, Q, \bar{m}, x$)  
5. $E = \begin{pmatrix} E \\ e \end{pmatrix}$  
6. end for

7. return $E \in \mathbb{R}^{N \times M}$
7.5.3 Results: Modelling the DTFs In The Median Plane Using PCA

As noted previously in Sec. 7.5.2 the DTFs can be modelled using PCA in two different schemes. In the first scheme a data matrix is formed by concatenating all the DTFs for all the 11 ear shapes and for a given shape principal component, \( p \), in a manner detailed in Eq. (7.20) into a data matrix \( X_1 \in \mathbb{R}^{1342 \times 342} \). That is:

\[
X_1 = \begin{bmatrix}
W_{E, -2.5} \\
\vdots \\
W_{E, 2.5}
\end{bmatrix}_{1342 \times 342}
\tag{7.22}
\]

Fig. 7.59 shows how the average error in reconstructing the DTFs in the median plane (i.e. averaged for all the DTFs in the data matrix) changes as more principal components are added. Further, Fig. 7.60 illustrates the reconstruction of a DTF directly in the front when 10 and 20 acoustic principal components are used.

![Graph showing average error vs. principle component](image)

Figure 7.59: The average error using the SDS measure is plotted versus the number of principal components used for the reconstruction of all DTFs in the median plane.

The second type of PCA was conducted by concatenating all the DTFs for a single
7.5. Mechanisms That Generate Changes In The DTF Spectrum

Figure 7.60: The reconstruction of a DTF (red) directly located at the front of the head, i.e. $\phi = \frac{\pi}{2}$ using 10 and 20 principal components is shown with respect to the actual DTF for the angle (blue).

-angle $\phi_i$ across all the 11 ear shapes to form a data matrix $X_2 \in \mathbb{R}^{11 \times 342}$:

$$X_2 = \begin{bmatrix}
    W_{E_{\phi,-2.5}}(\phi_i) \\
    \vdots \\
    W_{E_{\phi,2.5}}(\phi_i)
\end{bmatrix}_{11 \times 342} \quad (7.23)$$

The above was done for each of the 122 angles in the median plane. In other words, a total of 122 PCA’s were computed for every angle. The plot of the average error when reconstructing a DTF using the second kind of PCA analysis is displayed in Fig. 7.61. The errors are computed using Algorithm 15 for a single angle and then averaged for all the 122 angles.

Fig. 7.62 - Fig. 7.64 show the reconstruction of the DTF spectrum using 1 and 2 principal components for angles in the median plane corresponding to 135, 90 and 45 degrees respectively.

The results indicate that the DTFs for a single angle in the median plane can be successfully modelled with 2 scores and corresponding acoustical principle components. On the other hand, if the full set of DTFs in the median plane is to be modelled using PCA, the results indicate around 20 scores are required in order to get a similar level of reconstruction accuracy.

To summarize the discussion in this section the DTFs $W_{E_{\phi,m}}(\phi_i = \frac{\pi}{2})$ for ear
7.5. Mechanisms That Generate Changes In The DTF Spectrum

Figure 7.61: The average error using the SDS measure is plotted versus the number of principal components used for the reconstruction of all DTFs in the median plane.

Figure 7.62: The reconstruction of a DTF (red) located at $\phi = \frac{3\pi}{2}$ using 1 and 2 acoustical principal components is shown with respect to the actual DTF for the angle (blue).

Shapes for a given $p$ and $m$ belonging to a specific angle $\frac{\pi}{2}$ can be modelled as:

$$\hat{W}_{E,p,m}(\phi_i = \frac{\pi}{2}) = Q_{r_1}^{(p,\phi_i)} f_1^{(p,\phi_i)} + Q_{r_2}^{(p,\phi_i)} f_2^{(p,\phi_i)} + \bar{m}^{(p,\phi_i)}$$

$$(7.24)$$

$$\hat{W}_{E,p,m}(\phi_i = \frac{\pi}{2}) \approx W_{E,p,m}(\phi = \frac{\pi}{2})$$

$$(7.25)$$

In Eq. (7.25) the index, $r$, is chosen appropriately depending on the value of $m$. Further $f_1^{(p,\phi_i)}$ signifies the first acoustic principal component calculated for the shape.
7.5. Mechanisms That Generate Changes In The DTF Spectrum

7.5.4 Examining The Relationship Between Shape and Acoustic Models

In the previous section the modelling of the DTF spectrum in the median plane was examined using two PCA schemes. Moreover, it was illustrated that the PCA model is much more simpler when only a single angle is considered. Further, it was shown that 2 acoustic principal components provided good reconstruction for the DTFs. A model for the DTF based on two principal components was presented in Eq. (7.25).
In this section, the morphological and acoustic data are examined together. More specifically, here it would be shown how the scores belonging to the acoustic principal components change as the scores for the shape principal component change. In addition, important features in the acoustic principal components will be shown and, further, it will be identified how changes in the weighted acoustic principal components (the score multiplied by the principle component) result in changes to the DTF spectrum when the ear shapes change.

Fig. 7.65 and Fig. 7.66 display how the scores for the first and second acoustic principal component change with respect to the first five shape principal components for elevation angle of \( \phi = \frac{\pi}{2} \) and \( \phi = \frac{\pi}{4} \) degrees in the median plane. The important features that cause changes to the DTF spectrum for the given angles can be identified within the acoustic principal components. These features are in the form of peaks and notches in the acoustic principal components. When the appropriate score factors for different ear shapes are multiplied by the acoustic principal components, the prominent features in the acoustic principal components will change causing changes to the overall DTF spectrum. Fig. 7.68 shows features for the first acoustic principal component that were generated by the first shape principal component (i.e. \( p = 1 \)).

From the discussion in Sec. 7.5.3 it is understood that for modelling the DTF for a single angle only two acoustical principle components is required. There are three stages involved when modelling the DTFs using the two acoustic principal components. In the first stage the first score factor is multiplied by the first acoustic principal component. In the second stage the second score factor is multiplied by the second acoustic principal component, and in the third stage the mean of the DTF spectrum is added to the weighted first and second principal components. The following section will show results for each of these three stages.
7.5.5 Results: Examining the Changes in the Acoustic Principal Components as the Ear Shape Changes

Analysis on the acoustic data was carried out for the first two shape principal components \( p = 1, p = 2 \) and for two angles \( \phi = \frac{\pi}{2} \) and \( \phi = \frac{\pi}{4} \) in the median plane. Fig. 7.67 shows the changes in the ear morphology for the first shape principal component \( p = 1 \) as \( m \) changes from -2.5 to +2.5. Further, Fig. 7.68 displays the prominent features in the first two acoustic principal components belonging to elevation angle \( \phi = \frac{\pi}{2} \) in the median plane. The features appear as a notch or a peak in the principal component spectrum. It can be observed that there are in total of 10 prominent features in the acoustic principal components, four features, F1-F4, belong to the first acoustic principal component and six features, F1-F6, belong to the second acoustic principal component. As the weight factors change the peaks could change to become a notch or could disappear. On the other hand, a feature that appears as a notch could change to a peak or disappear. Fig. 7.71 - Fig. 7.72 shows the first and second weighted principal components (i.e. when the PCA score factor is multiplied by the principal component), the weighted principal components plus the added mean and target DTF for different values of \( m \).

A similar analysis was conducted for the elevation angle \( \frac{\pi}{4} \) in the median plane for the first shape principal component. Fig. 7.71 gives the prominent features in the acoustic principal components spectrum for this elevation angle. At this angle a total of nine prominent features were detected on the acoustic principal components, which are marked on the plots. Further, Fig. 7.72 - Fig. 7.73 show the first and second weighted principal components (when the weight is multiplied by the principal component), the weighted principal components plus the added mean and target DTF for different values of \( m \). It is noted that due to the different weight factors between the first and second acoustic principal components not all of these features effect the final DTF spectrum equally.
Figure 7.65: The above figure shows how the weights of the acoustic principal components (vertical axis) change with respect to the weights of the shape principal components (horizontal) for the elevation angle $\phi = \frac{\pi}{2}$ is plotted for the first five shape principal components.
7.5. Mechanisms That Generate Changes In The DTF Spectrum

Figure 7.66: The above figure shows how the weights of the acoustic principal components (vertical axis) change with respect to the weights of the shape principal components (horizontal) for the elevation angle $\phi = \frac{\pi}{4}$ is plotted for the first five shape principal components.
Figure 7.67: The ear shapes generated by changing the first shape principal component is shown. The value $m$ used for the generation of the ear shapes is indicated above the shapes.
Figure 7.68: The above figure shows the prominent features marked by the letter Fx, where x is the feature number for the first and second acoustic principal components. These features belong to elevation angle of $\phi = \frac{\pi}{2}$ in the median plane for when the ear shapes generated by changing the first shape principal component change. When weight factors corresponding to different ear shapes are multiplied to the above principal components these features in the forms of sharp peaks and notches change, which ultimately cause changes in the modelled DTF spectrum.
Figure 7.69: The above figure shows the changes in the weighted first (first column from left) and second (second column from left) acoustic principal components for elevation angle of $\phi = \frac{\pi}{2}$ in the median plane. Each row in the figure corresponds to different ear shapes that are obtained by changing the first shape principal component. The first row corresponds to the ear generated when $m = -2.5$ and last row corresponds to the ear generated for $m = 0$. The mean is then added to the weighted first and second principal components (third column from left). The target DTF is shown on the fourth column from the left. (The rest of the shapes are continued on the next page)
7.5. Mechanisms That Generate Changes In The DTF Spectrum

Figure 7.70: The above figure shows the changes in the weighted first (first column from left) and second (second column from left) principal components for elevation angle of $\phi = \pi \over 2$ in the median plane. Each row in the figure corresponds to different ear shapes that are obtained by changing the first shape principal component. The first row corresponds to the ear generated when $m = 0.5$ and last row corresponds to the ear generated for $m = 2.5$. The mean is then added to the weighted first and second principal components (third column from left). The target DTF is shown on the fourth column from the left.
7.5. Mechanisms That Generate Changes In The DTF Spectrum

Figure 7.71: The above figure shows the prominent features marked by the letter Fx, where x is the feature number for the first and second acoustic principal components. These features belong to elevation angle of $\phi = \frac{\pi}{4}$ in the median plane for the ear shapes generated by changing the first shape principal components $p = 1$. When weight factors corresponding to different ear shapes are multiplied to the above principal components these features in the forms of sharp peaks and notches change, which ultimately cause changes in the modelled DTF spectrum.
Figure 7.72: The above figure shows the changes in the weighted first (first column from left) and second (second column from left) principal components for elevation angle of $\phi = \frac{\pi}{4}$ in the median plane. Each row in the figure corresponds to different ear shapes that are obtained by changing the first shape principal component. The first row corresponds to the ear generated when $m = -2.5$ and last row corresponds to the ear generated for $m = 0$. The mean is then added to the weighted first and second principal components (third column from left). The target DTF is shown on the fourth column from the left. (The rest of the shapes are continued on the next page)
Figure 7.73: The above figure shows the changes in the weighted first (first column from left) and second (second column from left) principal components for elevation angle of $\phi = \frac{\pi}{4}$ in the median plane. Each row in the figure corresponds to different ear shapes that are obtained by changing the first shape principal component. The first row corresponds to the ear generated when $m = 0.5$ and last row corresponds to the ear generated for $m = 2.5$. The mean is then added to the weighted first and second principal components (third column from left). The target DTF is shown on the fourth column from the left.
A similar analysis was conducted using the second shape principal component \( p = 2 \). Fig. 7.74 shows the changes in the morphology of the ear shapes for when \( p = 2 \). The elevation angles examined here are \( \phi = \frac{\pi}{2} \) and \( \phi = \frac{\pi}{4} \). Fig. 7.75 shows the prominent features detected in the first two acoustic principal components when modelling the elevation angle of \( \phi = \frac{\pi}{2} \). At this angle a total of nine prominent features were identified on the two principal components which are marked on the figure. Further, Fig. 7.76 - Fig. 7.77 shows the first and second weighted principal components (i.e. when the score factor is multiplied by the principal component), the weighted principal components plus the added mean and target DTF for different values of \( m \). Finally, Fig. 7.78 - Fig. 7.80 show a similar analysis for the second shape principal component \( p = 2 \) when the elevation angle is at \( \phi = \frac{\pi}{4} \) in the median plane.

Figure 7.74: The ear shapes generated by changing the first shape principal component is shown. The value \( m \) used for the generation of the ear shapes is indicated above the shapes.
7.5. Mechanisms That Generate Changes In The DTF Spectrum

Figure 7.75: The above figure shows the prominent features marked by the letter Fx, where x is the feature number for the first and second principal components. These features belong to elevation angle of $\phi = \frac{\pi}{2}$ in the median plane for when the ear shapes generated by changing the second shape principal component change. When weight factors corresponding to different ear shapes are multiplied to the above principal components these features in the forms of sharp peaks and notches change, which ultimately cause changes in the modelled DTF spectrum.
7.5. Mechanisms That Generate Changes In The DTF Spectrum

Figure 7.76: The above figure shows the changes in the weighted first (first column from left) and second (second column from left) principal components for elevation angle of $\phi = \frac{\pi}{2}$ in the median plane. Each row in the figure corresponds to different ear shapes that are obtained by changing the second shape principal component. The first row corresponds to the ear generated when $m = -2.5$ and last row corresponds to the ear generated for $m = 0$. The mean is then added to the weighted first and second principal components (third column from left). The target DTF is shown on the fourth column from the left. (The rest of the shapes are continued on the next page)
7.5. Mechanisms That Generate Changes In The DTF Spectrum

Figure 7.77: The above figure shows the changes in the weighted first (first column from left) and second (second column from left) principal components for elevation angle of $\phi = \frac{\pi}{2}$ in the median plane. Each row in the figure corresponds to different ear shapes that are obtained by changing the second shape principal component. The first row corresponds to the ear generated when $m = 0.5$ and last row corresponds to the ear generated for $m = 2.5$. The mean is then added to the weighted first and second principal components (third column from left). The target DTF is shown on the fourth column from the left.
Figure 7.78: The above figure shows the prominent features marked by the letter Fx, where x is the feature number for the first and second principal components. These features belong to elevation angle of $\phi = \frac{\pi}{4}$ in the median plane for when the ear shapes generated by changing the second shape principal component change. When weight factors corresponding to different ear shapes are multiplied to the above principal components these features in the forms of sharp peaks and notches change, which ultimately cause changes in the modelled DTF spectrum.
### 7.5. Mechanisms That Generate Changes In The DTF Spectrum

In the analysis of the DTF spectrum, certain mechanisms are observed to generate changes in the spectrum. These mechanisms are associated with the behavior of the principal components. The target principle components, when combined with the mean, result in the observed changes in the spectrum.

#### Figure 7.79: Changes in Principal Components

The above figure illustrates the changes in the weighted first (first column from left) and second (second column from left) principal components for an elevation angle of $\phi = \frac{3}{4}$ in the median plane. Each row in the figure corresponds to different ear shapes that are obtained by changing the second shape principal component. The first row corresponds to the ear generated when $m = -2.5$ and the last row corresponds to the ear generated for $m = 0$. The mean is then added to the weighted first and second principal components (third column from left). The target DTF is shown on the fourth column from the left. (The rest of the shapes are continued on the next page.)
Figure 7.80: The above figure shows the changes in the weighted first (first column from left) and second (second column from left) principal components for elevation angle of $\phi = \frac{\pi}{4}$ in the median plane. Each row in the figure corresponds to different ear shapes that are obtained by changing the first shape principal component. The first row corresponds to the ear generated when $m = 0.5$ and the last row corresponds to the ear generated for $m = 2.5$. The mean is then added to the weighted first and second principal components (third column from left). The target DTF is shown on the fourth column from the left.
7.6 How Coherently Do DTFs Change In The Median Plane

In the previous section (Sec. 7.5) the underlying mechanisms that caused changes in the DTF spectrum as the ear shape changed for a given shape principal component was identified. In this section, the acoustic properties of the ear shapes that are obtained by changing a given shape principal component are further examined. More specifically, this section aims to provide a better understanding on the inter-relationship changes in the DTF spectrum for a few angles in the median plane. The motivation for asking this question is along the same lines as that which was raised in the previous section, namely, given that the ear shapes produced have been generated by changing a single shape principal component, then do the DTFs for a few given angles change in a coherent manner?

The approach in investigating this problem is very similar to Sec. 7.5 albeit an important difference. In Sec. 7.5 it may be recalled that every observation consisted of a single DTF that belonged to a particular direction in space. However, here, as it would be explained throughout this section, every observation in the data matrix consists of DTFs belonging to a few angles.

The overall approach is to first build a simple model for the DTF data using PCA. The aim is not to accurately represent the DTF data using the built model. However, the built model needs to contain important features found in the real DTF spectrum. Once the DTF model is constructed, then important features using the principal components that form a bases for the DTF data are analysed.

At first the number of principal components that are required for the reconstruction of the data is examined, here each of the observations consists of multiple DTF spectrums which belong to multiple angles in the median plane. In Sec. 4.7 and Sec. 7.5 much of the notation for representing the DTFs was presented and this notation is further used here.

First a data matrix $X$ is formed by concatenating the DTFs for the three angles, in
7.6. How Coherently Do DTFs Change In The Median Plane

\[ \phi = 135, \phi = 45, \phi = -45 \]

Figure 7.81: The above figure shows the reconstruction of the DTFs for three elevation angles of \( \phi = 135, \phi = 45 \) and \( \phi = -45 \) degrees using only one acoustic principal component.

\[ \phi = 135, \phi = 45, \phi = -45 \]

Figure 7.82: The above figure shows the reconstruction of the DTFs for three elevation angles of \( \phi = 135, \phi = 45 \) and \( \phi = -45 \) degrees using two acoustic principal component.

\[
X = \begin{bmatrix}
W_{EP,-25}(\phi_1) & W_{EP,-25}(\phi_2) & W_{EP,-25}(\phi_3) \\
\vdots & \vdots & \vdots \\
\end{bmatrix} \in \mathbb{R}^{11 \times 1026}
\]

(7.26)

In the above, \( p \), can range from 1-5, and \( \phi_1, \phi_2 \) and \( \phi_3 \) indicate three distinct angles in the median plane. For the present discussion the three angles are \( \phi_1 = \frac{3\pi}{2}, \phi_2 = \frac{\pi}{2} \) and \( \phi_3 = \frac{\pi}{4} \), however, this procedure can easily be adjusted if there is a need to analyse the changes for fewer or more angles. The notation used here is adopted from Sec. 4.9 to represent the DTF data in the median plane. Given that there are 11 ear shapes for a given shape principal component the size of the data matrix is \( X \in \mathbb{R}^{11 \times 1026} \). Once the data matrix \( X \) is formed the rest of the PCA procedure is very similar to Sec. 7.5. By performing a PCA the principal components in the
matrix $F \in \mathbb{R}^{1026 \times 1026}$ and the weight matrix $Q^T \in \mathbb{R}^{11 \times 1026}$ and the mean vector $\bar{m}$ is obtained. The data matrix $X$ can be reconstructed by performing the following matrix multiplication:

$$X = (FQ + \bar{m})^T$$ (7.27)

Next, the number of principal components needed for the reconstruction of our observations in the data matrix $X$ is examined. The number of principal components used will determine the modelling of our DTF data.

Fig. 7.81 - Fig. 7.82 shows the reconstruction of $[W_{E1-25}(\phi_1), W_{E1-25}(\phi_2), W_{E1-25}(\phi_3)]$ when one and two acoustic principal components are used. Further, an analysis on the error of the reconstructed DTF spectrum was performed using Algorithms 14 and 15. Fig. 7.83 shows the average error for reconstructing the DTFs as more principal components are added. Note that the error seen for the reconstruction of the DTFs is very similar to what was observed previously Fig. 7.61.

The results obtained indicate that using two principal components is adequate for modelling the DTFs for the given angles. Fig. 7.84 shows the changes of the
7.6. How Coherently Do DTFs Change In The Median Plane

weights for the first and second acoustic principal components with respect to $m$. It is recalled that $m$ changes from -2.5 to +2.5 and indicates the amount of change applied to the shape principal components when generating the ear shapes. Given the previous argument a single observation from the data matrix $X$ which consists of three DTF spectrum’s can be modelled as:

$$
\begin{align*}
[W_{EP=1,m}(\phi_1), W_{EP=1,m}(\phi_2), W_{EP=1,m}(\phi_3)]^T &= \\
&= Q_1^{(p,\phi_1,\phi_2,\phi_3)} f_1^{(p=1,\phi_1,\phi_2,\phi_3)} + Q_2^{(p,\phi_1,\phi_2,\phi_3)} f_2^{(p=1,\phi_1,\phi_2,\phi_3)} + \hat{m} \\
&\approx [\hat{W}_{EP=1,m}(\phi_1), \hat{W}_{EP=1,m}(\phi_2), \hat{W}_{EP=1,m}(\phi_3)]^T
\end{align*}
$$

(7.28)

(7.29)

where $f_1^{(p,\phi_1,\phi_2,\phi_3)}$ and $f_2^{(p,\phi_1,\phi_2,\phi_3)}$ indicate the first and second acoustic principal component that belong to the first shape principal component $p = 1$ computed for the angles $(\phi_1, \phi_2, \phi_3)$. Further $Q_1^{(p,\phi_1,\phi_2,\phi_3)}$ and $Q_2^{(p,\phi_1,\phi_2,\phi_3)}$ are the appropriate first and second weight factors that multiply the principal components for the given DTFs.

In the next step, the acoustic principal components for the combined angles and for the 11 ear shapes are examined. As indicated previously the acoustic principal components form a common basis for reconstructing the DTFs for the three angles.

Figure 7.84: The above figure shows the weights for the first and second acoustic principle as they change with respect to the weights for the shape principal components for the data matrix obtained by concatenating the three angles.
in the median plane. By examining the features found in the acoustic principal components, the changes in the DTF spectrum for the three angles can be simultaneously examined. Fig. 7.86 - Fig. 7.87 illustrates the first principal component that has been divided into three separate plots corresponding to the three angles under consideration. The plot also shows the prominent features that have been marked by "FxAy", where "x" is the feature number and "y" is the index to the angle. To elaborate let's consider the following features that are located on the first acoustic principal component:

1. F1A1, located at 10kHz for the DTF at elevation angle \( \phi = \frac{3\pi}{2} \), and
2. F2A2, located around 12kHz for the DTF at elevation angle \( \phi = \frac{\pi}{4} \) and
3. F1A3, located at 11kHz for the DTF at elevation angle of \( \phi = -\frac{\pi}{4} \)

Then, as the score factor that multiplies this (first) principal component decreases and/or increases these features will all increase or decrease with each other. Similarly let's consider the features located on the second acoustic principal component (Fig. 7.87):

1. F1A1, located around 7.5kHz for the DTF at elevation angle \( \phi = \frac{3\pi}{2} \),
2. F1A2, located around 7.5kHz for the DTF at elevation angle \( \phi = \frac{\pi}{4} \) and
3. F2A3, located near 11kHz for the DTF at elevation angle of \( \phi = -\frac{\pi}{4} \)

Then, as the score factor that multiplies the first acoustic principal component decreases and/or increases these features will all increase or decrease with each other. Fig. 7.88 - Fig. 7.89 show the sum of the weighted first and second acoustic principal components for these three angles when the score factors displayed in Fig. 7.84 are applied to the principal components appropriately. To clarify, the following equation details the plots for each angle shown in Fig. 7.88 - Fig. 7.89:

\[
Q_1 r_1^{p_1 \phi_1 \phi_2 \phi_3} + Q_2 r_2^{p_2 \phi_1 \phi_2 \phi_3}, \quad 1 \leq r \leq 11
\]  

(7.30)
7.6. How Coherently Do DTFs Change In The Median Plane

![Diagram of ear shapes generated by changing the first shape principal component.](image1)

Figure 7.85: The ear shapes generated by changing the first shape principal component is shown for convenience. The value $m$ used for the generation of the ear shapes is indicated above the shapes.

![Graphs of the first acoustic principal component for different elevation angles.](image2)

Figure 7.86: The above three plots show the first acoustic principal component that represent the common bases for the three elevation angles $\phi = \frac{3\pi}{2}, \phi = \frac{\pi}{4}$ and $\phi = -\frac{\pi}{4}$ which are also indicated above the plots. Further the acoustic PCA is computed for the first shape principal component $p = 1$. Prominent features are also labelled. As different weight factors is multiplied to this principal component these features will change in a coherent manner ultimately changing the shape of the model DTF spectrum.

![Graphs of the second acoustic principal component for different elevation angles.](image3)

Figure 7.87: The above three plots show the second acoustic principal component that represent the common bases for the three elevation angles $\phi = \frac{3\pi}{2}, \phi = \frac{\pi}{4}$ and $\phi = -\frac{\pi}{4}$ which are also indicated above the plots. Further the acoustic PCA is computed for the first shape principal component $p = 1$. Prominent features are also labelled. As different weight factors is multiplied to this principal component these features will change in a coherent manner ultimately changing the shape of the model DTF spectrum.
Figure 7.88: The above figure shows how the sum of the weighted first and second acoustic principal components change as the ear shape changes for the first shape principal component $p = 1$. The number $m$ indicating the amount of change exerted on the shape principal component is shown next to the plots. The angles are shown above the plots. Notice that the prominent features that were shown in a previous plot cause the majority of the changes seen in the above spectrum's. (Continued on next figure)
Figure 7.89: The above figure shows how the sum of the weighted first and second acoustic principal components change as the ear shape changes for the first shape principal component $p = 1$. The number $m$ indicating the amount of change exerted on the shape principal component is shown next to the plots. The angles are shown above the plots. Notice that the prominent features that were shown in a previous plot cause the majority of the changes seen in the above spectrum's.
7.7 Morphable Model of Ears

This section describes a morphable model for ear shapes and examines its ability to reconstruct new ear shapes, i.e., ear shapes outside of the database used for constructing the model (Zolfaghari et al., 2016). The significance of the morphable model is its ability to compress the representation of 3D ear shapes to a set of parameters typically obtained by projection onto a set of orthogonal bases functions (Blanz and Vetter, 2003; Cashman and Fitzgibbon, 2013). This parametrisation of ear shapes, using a morphable model greatly aids in the study of morphoaoustics (Jin et al., 2000; Mokhtari et al., 2011; Tew et al., 2012; Zotkin et al., 2003), where the goal is to understand the link between variations in the shape of an ear and their effect on the corresponding set of 3D audio filter functions, referred to as HRIRs.

Modelling ear shapes is a challenging task and ear shape deformations are arguably best described using a Riemannian space. The morphable model presented here is based on single scale LDDMM transformations. In the work presented in this section the LDDMM framework combined with the kernel based principal component analysis (KPCA) technique Sec. 7.1 is used to construct a morphable model for ears. In particular, the morphable model uses the concept of the linear space of initial momemtums (Miller et al., 2006) within the framework of LDDMM and a set of coupled differential equations known as the “shooting equations” (Eq. (2.31)) to construct and model ears. The model is derived from a statistical analysis of a population of 58 left ears from the SYMARE database.

7.7.1 Method

This section will detail a procedure on how a new ear shape, $S_p$, that was not included in the computation of the principal components, can be described using the KPCA data (Sec. 7.1). The computation of the model parameters for, $S_p$, can be divided into three steps. First, compute the initial momentum vectors, $a_p$, corresponding to the morphing of the template, $E_{TM}$, into shape, $S_p$. Second, the population average
momentum vectors are subtracted from $a_p$ to obtain the centred initial momentum vectors, $\hat{a}_p$. Third, the centred momentum vectors are projected onto the principal components to obtain the model parameters, $\tilde{v}_p$. The procedure is summarized below:

**Algorithm 16** Computation of the Model Parameters for an Ear

**Inputs:** $U$, $\bar{a}$, $S_p$.

1. $a_p = M(E_{TMS}, S_p, \sigma_V, \sigma_W)$
2. $\hat{a}_p = a_p - \bar{a}$
3. $\tilde{v}_p = U^T K \hat{a}_p$
4. return $\tilde{v}_p$

The reconstruction of shape, $S_p$, from the model parameters is performed in two steps. First, the initial momentum vectors for shape $S_p$ are estimated by combining the principal components according to the model parameters. Second, the shooting operation is used to morph the template into $\tilde{S}_p$, an approximation of shape $S_p$. The shape reconstruction operation is summarized below:

**Algorithm 17** Reconstruction of an Ear from the Model Parameters

**Inputs:** $E_{TMS}$, $U$, $\tilde{v}_p$.

1. $\tilde{a}_p = \bar{a} + U \tilde{v}_p$
2. $\{\tilde{a}_p\}_{0 < t < 1}^{1 \leq n \leq N} = S(E_{TMS}, \tilde{a}_p, \sigma_V)$
3. $\tilde{S}_p = F(E_{TMS}, \{\tilde{a}_p\}_{0 < t < 1}^{1 \leq n \leq N}, \sigma_V)$
4. return $\tilde{S}_p$

Note that shape $\tilde{S}_p$ is an approximation of $S_p$ because: (1) the LDDMM operation $M(E_{TMS}, S_p, \sigma_V, \sigma_W)$ does not match shapes perfectly; and (2) the principal components may not enable perfect reconstruction of the initial momentum vectors for shape $S_p$. 


7.7.2 Experimental Setup

A morphable model of ear shapes was created based on 58 different ear shapes from the SYMARE database. The ears used in the morphable model were in their normal orientation, that is they were not aligned to the template ear shape using an affine transformation. While the SYMARE database is the largest database of its kind, 58 ears is not a large number considering the human population and it is unclear how well the morphable model can describe an arbitrary ear not included within the database. In order to address this issue, one ear shape, $S_i$, was repeatedly left out of the dataset of 58 ears and formed a morphable model based on the remaining 57 ears. Then the ability of the morphable model to reconstruct the ear that was left out was examined. In other words, for each shape $S_i$ in the dataset, the KPCA analysis described in Sec. 7.1 using 57 ear shapes was applied (i.e., leaving $S_i$ out). Then, using the method described in Sec. 7.7.1 an approximate ear shape $\tilde{S}_i$ was reconstructed. We then examined how accurately the approximation $\tilde{S}_i$ matches the original shape $S_i$ using a shape difference analysis based on currents (refer to Sec. 4.3.2 for details). Further, we examined the shape reconstruction accuracy as a function of the number of principal components used to reconstruct the ear shape. Note that in order to exclude the mismatch caused by the LDDMM matching procedure, we actually compared the $\tilde{S}_i$ shapes to the shapes $Z_i$ obtained by matching the template $E_{T_{MS}}$ to $S_i$ and shooting using the true initial momentum vectors.

7.7.3 Results

Results are summarized in Fig. 7.91. As expected the accuracy of the model improves as the number of principal components increases. However, there is very little difference between the results obtained with 50 and 57 principal components, which indicates that the last 7 principal components have very little influence on the accuracy of the model. Interestingly, the quality of the reconstruction strongly depends on the ear considered. Some ears are reconstructed with greater accuracy using
Figure 7.90: This figure compares ear shapes reconstructed using 50 principal components, $\tilde{S}_{i}^{(50)}$, to the corresponding reference shapes $Z_{i}$. Colors (constant luminance, so examine online) on the shape indicate local shape mismatch calculated using the measure $\hat{d}(\tilde{S}_{i}^{(50)}, Z_{i}, f)$. Please refer to equation Eq. 4.18.

relatively few principal components, while others are poorly reconstructed using the full basis of principal components. This is illustrated in Fig. F.2 where examples of reconstructed ears are compared to the corresponding reference shapes. The quality
for the reconstruction of all the 58 ear shapes can be seen in Appendix F. It can be observed that shapes \( Z_1 \) and \( Z_2 \) were reconstructed with no apparent mismatch, while there is clear mismatch for shapes \( Z_3, Z_4 \) and \( Z_5 \). In summary, these results indicate that the morphable model is promising, but it requires a larger population of ears to enable the model to morph into any possible ear shape.

### 7.8 Conclusion and Summary Of Chapter

This chapter was the culmination of the work conducted in this thesis. Overall, this chapter presented the KPCA technique to examine the variations in the ear shapes and their acoustics. The KPCA technique was also used for generating a morphable
model of ear shapes. Further, the standard PCA technique was used to model DTFs in the median plane, and, also to obtain a better understanding on how the DTF spectrum changes as the ear shapes change. The findings discussed in this chapter are now summarized below:

- The first question that was investigated in this chapter related to quantifying the magnitude of acoustic variations that is lost when aligning the ear shapes to the template ear using an affine transformation. For this purpose six ear shapes were considered for this examination. The function $\bar{V}$ that was formulated in Eq. (4.43) was used to measure the acoustic similarity/dissimilarity for the ear shapes when aligned versus when in normal orientation. It was shown that on average when aligning the ear shapes to the template ear there is on average a 13% reduction in the acoustic variation of the ears with respect to each other.

- The second question that was examined in this chapter was the number of shape principal components that were required to represent the morphology and the acoustic properties of the ear shapes. This research question also investigated the significance of lower order principal components (i.e. principal components which capture a smaller variation in the population) on the acoustic responses of the ear shapes. The results indicated that between 10 and 20 shape principal components were needed to capture on average 47% of the acoustic variations and 44% of the morphological variation. It was argued that exact matching of the acoustics was not the aim of this research question, but rather the emphasize was on identifying and replicating the general features in the HRTF for the angles considered.

- Subsequently a qualitative and quantitative analysis on the changes for the ear shapes was presented when a given shape principal component was changed. The acoustic changes for the ears on the median plane was shown. These results were presented using images and surface plots. Further, observations were made on the structural changes that occur by changing each of the shape
7.8. Conclusion and Summary Of Chapter

principal components which were tabulated in Table 7.1 and Table 7.2. As a quantitative analysis acoustical and morphological distances were computed and plotted from the generated ear shapes to the template ear (using the measures developed in previous chapters). Using the measures it was shown that both the morphology and acoustics of the ears approached the template shape monotonically.

- The next research question investigated the mechanisms that cause changes in the DTF spectrum for specific angles in the median plane. It was shown that by performing a PCA across the ear shapes for a single angle in the median plane a reasonable model of the DTFs can be obtained. In the same section a measure was devised using the function \( \overline{V} \) to quantify the error when reconstructing the DTFs using progressively more principle components (see Algorithm 14). It was shown that when reconstructing DTFs an average SDS error between 1 and 0.5 indicated a reasonable quality of reconstruction. Further, it was shown that around 2 principal components were sufficient to model the acoustic data. Subsequently the prominent features in the acoustic principal components that influenced changes in the DTF spectrum were identified. When the first shape principal component was considered for elevation angle \( \frac{3\pi}{2} \), it was shown that the changes in 10 prominent features in the first and second acoustical principal components influenced the changes in the DTF spectrum. Using the same analysis and for the first shape principal component it was shown that 9 prominent features in the acoustic principal components influenced the changes in the DTF spectrum at the elevation angle \( \frac{\pi}{4} \).

- The research also examined how coherently the DTFs change for a few angles in the median plane. The angles that were investigated are elevation angle \( \frac{3\pi}{4} \), elevation angle \( \frac{\pi}{4} \), and elevation angle \( -\frac{\pi}{4} \). For these angles a total of 11 prominent features were identified. It was found that the gain at the location of these features either all increase in magnitude or for
some the gain increased and for others it decreased.

- Finally, this chapter concluded by presenting a morphable model using the KPCA analysis on ear shapes. The ears that were considered were not aligned to the template ear shape (i.e., they had their normal orientation). The morphable model was able to construct a few ear shapes with good accuracy while for others it was not very accurate. This suggested that more ear shapes were required in order to increase the strength of the current morphable model used in this research.
Chapter 8

CONCLUSION AND FUTURE WORK

In this thesis a systematic study was undertaken in order to analyse the variations in the ear morphology and their corresponding acoustics using the large deformation diffeomorphic metric mapping (LDDMM) and the fast multipole boundary element method (FM-BEM). Based on a review of the prevailing literature in this field, LDDMM was found to be a new and promising method for analysing ear shapes. This is because it enabled ears to be analysed using large diffeomorphic deformations within a convenient metric space. In contrast, the previous methods analysed ears by means of small perturbation and deformations (Mokhtari et al., 2010; TEW et al., 2012) without having the convenience of a diffeomorphic metric space (refer Sec. 3.4). A metric over the space of ear shapes enables morphological distances to be analysed over the space of ears. In conjunction with a suitable metric over the space of acoustics it can enable a quantitative study of morphoacoustics. This can assist in the improved modelling of individualized HRIR filter functions which in turn necessary for the production of VAS. A preliminary study of the morphological distances versus acoustic distances was conducted in Sec. 6.2 and this is summarized in the next section (i.e. Sec. 8.1).

Essentially, the present research culminated in the novel analysis of the morpho-
logical and acoustic variations for ear shapes using the LDDMM framework and the kernel principal component analysis (KPCA) over the space deformations from the reference/template ear shape to target ears in the SYMARE database. More specifically, the SYMARE database of ears was utilized by combining the left and right ear shapes (totalling a 124 ears) for 62 subjects in order to obtain a larger sample of ears to perform the statistical analysis. The template ear shape was computed by means of a new procedure that was developed using the LDDMM framework. The research also necessitated the construction of novel tools to examine both shape and acoustic differences between ear morphologies. These tools were subsequently used to examine the deformations from the template ear. As a result of analysing the deformations from the template ear shape it was understood that ear shapes can be more accurately analysed at multiple physical scales. In this thesis the variations in the ear shapes was examined at a single physical scale. However, if more accuracy is required for analysing ear shapes a multi-scale approach needs to be used for more exact deformations (refer to Sec. 6.1).

The following section provides a brief overview of the research questions and summarizes some of the main findings related to each question.

8.1 Questions and Answers

**Question 1:** Given that the LDDMM framework had not previously been employed for the analysis of ear shapes, the first question that was posed for this investigation was: (a) how do the LDDMM parameters used in the LDDMM cost function influence the quality of the ear mappings? (b) what would be the appropriate values for these parameters in the study of ear shapes?

Given the above question a study was undertaken to identify the most important parameters when performing ear mappings (Sec. 4.4). These parameters were the deformation scale parameter $\sigma_V$, the shape comparison parameters $\sigma_W$, and the

---

$^1$The SYMARE database contains 61 subjects however for the KPCA analysis an extra subject was used.
optimization parameter $\gamma$. It was shown that $\sigma_V$ and $\sigma_W$ are physical scale parameters and greatly influence the quality of deformations when conducting a mapping between a source and a target ear shape. Based on a series of simulations it was concluded that there is a range of values that could be used for these parameters. More specifically, it was clarified that for the deformation scale parameter $\sigma_V$ small values will produce meaningless deformations, while large values will produce inaccurate mappings (refer Fig. 4.11). A range of values was examined for both the $\sigma_V$ and $\sigma_W$ parameters and by means of visual inspection for the deformation scale parameter a value of $\sigma_E^V = 10$, and for the shape comparison scale parameter a value of $1 \leq \sigma_E^W < 5$ was identified to be adequate based on the resolution of the ear meshes\textsuperscript{2}. Further, Sec. 4.4.3 argued that the optimization parameter $\gamma$ had a large range of values (i.e. $\gamma \leq 1$), which produced very similar deformations on the ear shapes. Based on this finding the $\gamma$ parameters was set to a small value of 0.0005.

**Question 2:** The next question that was asked in this research related to obtaining a template shape that represented the population average at multiple scales and metrics. This question was inspired by the SMS-LDDMM approach for mapping the ear shapes which mapped two ears at successive and sequentially smaller LDDMM scales using smooth and natural transformations (refer Sec. 4.9). It also involved investigating how a template ear shape obtained by mappings conducted at multiple LDDMM scales and metrics compared to a template ear shape that was obtained at a single LDDMM scale. Further, it inquired whether these template shapes were valid when considering the size and orientation for features in the ear population.

In order to tackle this question a new multi-scale template estimation procedure using the LDDMM framework was developed in Sec. 5.1.3. The multi-scale template ear shape was estimated utilizing four LDDMM scales and metrics and a template ear based on single-scale deformations was also estimated. In order to both verify

\textsuperscript{2}The $\sigma_E^V$ and $\sigma_E^W$ parameters are used for ear shape deformations, for the head and torso a similar values were used for the deformation scale parameter, for the shape comparison parameter the value dependent on the mesh edge length properties.
and compare these template ear shapes, a study on the ear shape features was conducted using 2D images and also 3D surfaces of the ear. Distances similar to the ones defined in the CIPIC database \cite{Algazi2001} were measured on the template ear shapes and all ears in SYMARE. The distances were the pinna width and height and the Concha width and height. By performing a statistical analysis on the measured distances it was identified that the template ear shapes were within ±1 standard deviations from the population mean. Moreover, a similar study was conducted over 3D shapes using distances similar to the ones defined in \cite{Jin2000} and it was demonstrated there that the distances in both the multi-scale and single scale template ear shapes were close to the population mean and well within the range of ±1 standard deviation of the population mean. Once the two template ear shapes were verified they were further compared using the shape and acoustic comparison tools previously developed for this research. Using these tools it was shown that even though at a relatively larger scale the template shapes have similar features, nevertheless, they have different acoustic characteristics and differ morphologically at a finer scale too. Both their morphological and acoustic differences were highlighted using comprehensive plots.

In addition, the orientation (size and rotation) of the template ear shapes was also compared to the SYMARE population (Sec. 5.2.4) using a rigid matching procedure (refer Sec. 5.2.4.1). By conducting a statistical analysis it was shown that both the scale and orientation of the template ear shapes was close to the population of ears in SYMARE.

**Question 3:** The next question that was explored was to identify the accuracy of the acoustic and morphological quality of deformations from the template shape to target shapes using the chosen LDDMM parameters. Although this question was partially investigated by virtue of the preceding question (refer Sec. 4.4), the acoustic quality of the mappings was further investigated here. Also the morphological quality of mappings was examined more comprehensively using a large population
of ear shapes.

In order to answer this question Sec. 6 described new tools for comparing ear shapes and their corresponding acoustics. In particular, it was shown that the new shape comparison tool (refer Sec. 4.3.2), which is based on the metric defined by the framework of currents, was superior to methods based on the Euclidean metric (Sec. 4.3.1) when comparing the ear morphology. The strength of the new tool developed here was particularly useful when highlighting morphological differences that contributed to acoustic differences.

Further, by analysing the deformations from the template ear shape it was shown that the acoustic quality of mapping is adequate only for some ear shapes (Sec. 6.1). These findings guided the rationale for the construction of a new multi-scale technique for analysing ear shapes (Sec. 6.1.0.1). In the multi-scale approach the ear shapes can be accurately studied at multiple, sequential LDDMM scales. Single scale and two scale deformations from the template ear shape were examined for a couple of ear shapes and it was demonstrated how multi-scale deformations can greatly improve the accuracy of the mapping from an acoustic and morphological standpoint (Sec. 6.1.0.1).

Finally, a distance analysis on the morphology and acoustics for deformations from the template ear shape to select ears in SYMARE was performed (Sec. 6.2). The distances computed were based on the newly developed techniques that were especially conceptualized and constructed for this research. It was shown that on an average the acoustic distances with respect to the morphological distances was close to linear for the given ear shapes.

**Question 4:** Given that the template/reference ear shape was estimated as explained above, the next question aimed at estimating the variation in the population of ears using the reference/template shape by means of the kernel principal component analysis (KPCA). KPCA was chosen to analyse the deformations from the template ear shape as it induced a suitable inner product over the space of initial
momentums. At this stage of the investigation it was desirable to delimit our study to important structural differences only in the ear shapes and towards this end the population of ears in SYMARE was rigidly aligned to the template ear. This part of the research involved investigating a number of related sub-questions, namely:

1. How much of the acoustic variation in the ear shapes is lost due to rigidly aligning the ears to the template ear shape (Sec. 7.2)? Based on acoustic simulations on six ear shapes in normal positions and when rigidly aligned to the template ear shape it was found that on an average 13% of the acoustic variation in the ear shape was lost when rigidly aligning the ear shapes.

2. The second sub-section question examined the number of kernel principal components that were required for modelling the morphology and acoustic properties of ear shapes (Sec. 7.3). The experiment involved gradually changing a source ear shape to a target ear shape by means of changing the KPCA coefficients. A distance analysis for the acoustics and the morphology was conducted between the intermediate shapes starting from the source shape and ending with the target shape. It was shown that both the shape distances and acoustic responses converged to that of the target ear shape. It was found that for an accurate acoustic match nearly all of the 124 kernel principal components were required. Further, while the shapes distances converged monotonically the same could not be said about the acoustic responses. It was shown that 20 kernel principal components captured on average 47% of the acoustic variation of the target shape and around 44% of the morphological variation of the target ear shape. Based on the current findings Sec. 8.2 will highlight future research avenues that can assist in the modelling of ear shapes both morphologically and acoustically.

3. The third sub-question inquired into the significant and most common variations that exist in the population of ears and examined the variations captured by the kernel principal components (Sec. 7.4). Due to the extensive time re-
quired for computing the HRIRs through FM-BEM simulations only the first five kernel principal components were analysed. More specifically, each of the kernel principal components were changed based on the scores obtained from the KPCA analysis. The kernel principal components were altered using weight factors ranging from -2.5 to +2.5 standard deviations of the scores in the population of ear shapes. Both the nature of changes in the morphology and the corresponding acoustics were shown (Sec. 7.4.3) and analysed (Sec. 7.4.4). Further, in order to obtain a better understanding in regard to variations that each kernel principal component captured, the principal components were varied at ±5 and ±7 standard deviations and subsequently the corresponding ear shapes were examined.

4. The fourth question asked whether features in the acoustic spectrum change systematically as the ear morphology changes for a given kernel principal component? To answer this question the DTF spectrum was modelled using the PCA technique. Modelling was done for the DTF spectrum for both single directions in the median plane as well as for all the directions in the median plane combined. It was shown that the DTF model was much more simpler (i.e. requiring fewer principal components) for a single direction in the median plane as compared to when all angles in the median plane were considered. In particular, it was identified that only two principal components and corresponding coefficients were sufficient for capturing the majority of the variations in the acoustic spectrum. Using plots it was shown that a functional relationship exists between the acoustic scores belonging to the acoustic principal components to the scores belonging to the kernel (shape) principal components. Finally, features in the acoustic principal components were identified that changed systematically when the ear shapes changed for a given shape principal component. As an example, for the ear shapes generated by changing the first kernel principal component and for directions directly in the front (i.e. $\phi = \frac{\pi}{2}$ in the median plane), 10 features were identified in the
acoustic principal components that changed systematically as the ear shapes changed.

5. The fifth sub-question examined the coherence in the changes to the acoustic spectrum belonging to two or more directions in space. By combining the DTF spectrum for a few angles, and by means of applying a PCA on the acoustic data, features in the acoustic principal component spectrum belonging to the angles under consideration were identified that changed coherently with each other. For example, for the ear shapes generated by the first shape principal component and when considering angles: \( \phi = \frac{3\pi}{4}, \phi = \frac{\pi}{4}, \phi = -\frac{\pi}{4}, \) it was identified that a feature in the acoustic spectrum at around 12kHz for \( \phi = \frac{\pi}{4} \) was negatively correlated with a feature at 15kHz for the angle \( \phi = -\frac{\pi}{4}. \)

6. The sixth sub-question examined the strength of a morphable model based on a sample of 58 normal left ear shapes drawn from SYMARE database (refer Sec. 7.7). It involved ears that had not been rigidly aligned to the template ear shape. The strength of the morphable-model was examined on ear shapes that were not included in the data set when computing the KPCA. It was shown that the morphable model was able to construct some ear shapes with a high level of accuracy while for other ear shapes the accuracy was poor. Based on these finding avenues for improving the morphable model are further discussed in the next section (refer Sec. 8.2).

### 8.2 Future Work

The analysis of ear shapes by means of the LDDMM framework is a novel and exciting area of research with significant implications for the better understanding of ear shape morphology and human HRIR functions. Importantly, this will have applications towards the production of high quality VAS. Arguably, the most crucial and immediate work that can assist any morphoacoustic study of ear shapes is by improving the efficiency of BEM or other numerical techniques for obtaining
8.2. Future Work

the HRIR. Currently, the FM-BEM computations on high resolution ear, head and torso meshes require several hours for completion, which greatly limits any large scale morphoacoustic analysis on ear shapes. Techniques have been proposed to coarsen the mesh geometry, particularly on the contra lateral side of the ear, while preserving the quality of the computed HRIRs through numerical simulations. However, these techniques typically involve manual intervention which could make them unsuitable for large scale morphoacoustic analysis on ear shapes. In short, although BEM simulations are now much more rapid with the use of the fast multipole methods, computations of the HRIRs numerically are not efficient for large scale morphoacoustic study of ear shapes at this point of time.

In chapter 4 of this document a distance analysis was carried out between the morphology of the shapes and the corresponding acoustics by means of the tools and functions developed for this research. While the present work has been foundational, significantly more simulations are needed to obtain a better understanding on the relationship between the acoustic and morphological distances. These simulations will likely pave the way to measure the differences in the acoustics and the morphology of the ear shapes more accurately.

While a significant initial step was taken using KPCA to explore the space of ears and their corresponding acoustics using the first five kernel principal components (Sec. 7.4), the results indicated that this is not adequate. Firstly, a linear combination of the various kernel (shape) principal components is required to more comprehensively explore the space of deformations from the template shape. Secondly, as it was shown in Sec. 7.3, the morphology and acoustic properties for ear shapes cannot be modelled well by a combination of the first five kernel principal components. Deriving a satisfactory model for the ear morphology and acoustics would have required a significant amount of simulations. Unfortunately, this was not feasible within the time frame of the current research. However, this would be an important area for future research.

In this thesis a preliminary morphable model for ear shapes was examined based
on a sample of 58 left ear shapes with normal orientation. It was shown that some of the ear shapes were not modelled adequately and that the strength of the morphable model can be improved by increasing the sample size. For future research one possible avenue for increasing the sample size of ears is by means of generating synthetic ear shapes. Fortunately the LDDMM framework enables synthetic ear shapes to be generated relatively easily utilizing the geodesic shooting equations (e.g. Fig. 2.29). This option can be explored to strengthen the morphable model on ear shapes.

Finally, it would be valuable for future research to incorporate the modelling of the ear shapes or the acoustics with appropriate psychoacoustic tests. By the use of numerical simulations the present research showed that around 10-20 kernel principal components can capture a large amount of acoustic and morphological variation for the ear shapes (see Sec. 7.3). However this finding needs to be accompanied by acoustic experiments on real subjects.
Bibliography


Partial differential equations in physics. (cited page 21)


Blauert, J. (2013). The technology of binaural listening. Springer. (cited pages 13 and 30)


Duda, R. O., V. R. Algazi, and D. M. Thompson (2002). The use of head-and-


Gumerov, N., R. Duraiswami, D. N. Zotkin, et al. (2007). Fast multipole accelerated boundary elements for numerical computation of the head related transfer func-


Vaillant, M., A. Qiu, J. A. Glaunés, and M. I. Miller (2007). Diffeomorphic metric


Appendices
Appendix A

Morphoacoustics analysis of
Deformations from the
Template Ear to SYMARE Ears

This appendix presents figures which examine the quality of the morphoacoustic mapping from the template ear shape to ears in the SYMARE database. The template ear shown and discussed in Chapter 5 is mapped to SYMARE ear shapes and the quality of mappings at a single scale is shown here. When comparing the acoustic properties of the mapped and target shapes using GAAF, two shapes are used as the source shape when conducting LDDMM mappings 1- the template ear attached to the head of subject 2 of the SYMARE database shown in Fig. A.1 and 2- the template ear attached to the head and torso shape computed and discussed in Chapter 5 show in Fig. A.2. CSDA and GAAF figures presented in Sec. 4.3 and Sec. 4.8 are used to examine the quality of the deformations morphologically and acoustically. When mapping the template head and ear shape the $\sigma^E_V = \sigma^H_V = 10$ for all mapping,

---

1Two intermediate shapes were constructed with the template ear before the construction of the final template shape shown in Fig. 5.11. In the first intermediate shape the template ear was attached to subject number 2 in the SYMARE database and in the second intermediate shape the template ear was attached to the template head and torso shape detailed in Sec. 5.2.2 albeit with a minor rotation. The small rotation in the ear shapes were done using mesh processing software to position the template ear more conveniently on the template head and torso shape. It is noted that the final template shape consisting of the template ear plus the template head and torso shape are discussed in Chapter 5.
Figure A.1: First intermediate template shape used in the acoustic study of ears. The template ear is attached to a head shape from SYMARE.

however the $\sigma^E_W, \sigma^H_W$ value chosen for the mappings is different between the two template shapes and is proportional to the edge length of the mesh.
Figure A.2: Second intermediate template shape used for the acoustic study of ears. The template ear is attached to the template head.
Target Matched

Front

Back

Target Matched
Target Matched
Front
Back

Target Matched
Front
Back

S21
Target Matched

Front

Back

Matched
Target Matched
Front
Back

S37
(a) (b) (c)
(d) (e) (f)
(g) (h) (i)

Target Matched
Front
Back

S60

Target Matched
Front
Back

S60
Appendix B

Evolving Ear Population When Generating the Multi scale Template Ear
Figure B.1: The evolution of the ear shapes and the starting population at each scale of the multi-scale template ear shape.
Figure B.2: The evolution of the ear shapes and the starting population at each scale of the multi-scale template ear shape.
Figure B.3: The evolution of the ear shapes and the starting population at each scale of the multi-scale template ear shape.
Figure B.4: The evolution of the ear shapes and the starting population at each scale of the multi-scale template ear shape.
Figure B.5: The evolution of the ear shapes and the starting population at each scale of the multi-scale template ear shape.
Appendix C

Anthropometric Measurements of Ears from 2D Images
$E_1 d_1 + d_2 = 2.3 \, d_3 = 1.9 \, d_5 = 6.1 \, d_6 = 3.2$

$E_2 d_1 + d_2 = 2.3 \, d_3 = 1.5 \, d_5 = 7.1 \, d_6 = 3.1$

$E_3 d_1 + d_2 = 2.7 \, d_3 = 1.9 \, d_5 = 6.7 \, d_6 = 3.4$

$E_4 d_1 + d_2 = 2.1 \, d_3 = 2.0 \, d_5 = 6.3 \, d_6 = 3.2$

$E_5 d_1 + d_2 = 2.9 \, d_3 = 2.1 \, d_5 = 7.2 \, d_6 = 3.3$

$E_6 d_1 + d_2 = 2.8 \, d_3 = 1.7 \, d_5 = 6.0 \, d_6 = 2.7$
\[ E_7, d_1 + d_2 = 2.3 \quad d_3 = 1.4 \quad d_5 = 5.9 \quad d_6 = 2.5 \]

\[ E_8, d_1 + d_2 = 1.9 \quad d_3 = 1.3 \quad d_5 = 6.1 \quad d_6 = 2.8 \]

\[ E_9, d_1 + d_2 = 2.2 \quad d_3 = 1.6 \quad d_5 = 6.1 \quad d_6 = 3.0 \]

\[ E_{10}, d_1 + d_2 = 2.7 \quad d_3 = 1.6 \quad d_5 = 6.5 \quad d_6 = 3.0 \]

\[ E_{11}, d_1 + d_2 = 2.0 \quad d_3 = 1.4 \quad d_5 = 6.4 \quad d_6 = 3.1 \]

\[ E_{12}, d_1 + d_2 = 2.4 \quad d_3 = 1.4 \quad d_5 = 7.1 \quad d_6 = 3.2 \]
$E_{13}: d_1 + d_2 = 2.5\quad d_3 = \text{NaN} \quad d_5 = 6.4\quad d_6 = \text{NaN}$

$E_{14}: d_1 + d_2 = 2.3\quad d_3 = 1.4\quad d_5 = 6.5\quad d_6 = 2.8$

$E_{16}: d_1 + d_2 = 2.0\quad d_3 = 1.9\quad d_5 = 6.2\quad d_6 = 3.4$

$E_{17}: d_1 + d_2 = 2.5\quad d_3 = 1.0\quad d_5 = 6.9\quad d_6 = 2.9$

$E_{18}: d_1 + d_2 = 2.2\quad d_3 = 1.7\quad d_5 = 5.9\quad d_6 = 2.9$
\[ E_{19}, d_1 + d_2 = 2.4 \quad d_3 = 1.8 \quad d_5 = 6.3 \quad d_6 = 3.1 \]

\[ E_{20}, d_1 + d_2 = 2.9 \quad d_3 = 1.1 \quad d_5 = 6.4 \quad d_6 = 2.6 \]

\[ E_{21}, d_1 + d_2 = 2.2 \quad d_3 = 1.4 \quad d_5 = 6.0 \quad d_6 = 2.7 \]

\[ E_{22}, d_1 + d_2 = 2.6 \quad d_3 = 1.7 \quad d_5 = 6.9 \quad d_6 = 3.3 \]

\[ E_{24}, d_1 + d_2 = 2.4 \quad d_3 = 1.3 \quad d_5 = 6.3 \quad d_6 = 3.0 \]
\begin{align*}
E_{25}, & 
\quad d_1 + d_2 = 2.1 \quad d_3 = 1.6 \quad d_5 = 6.5 \quad d_6 = 3.0 \\
E_{26}, & 
\quad d_1 + d_2 = 2.1 \quad d_3 = 1.4 \quad d_5 = 6.3 \quad d_6 = 3.4 \\
E_{27}, & 
\quad d_1 + d_2 = 2.7 \quad d_3 = 1.7 \quad d_5 = 5.8 \quad d_6 = 2.7 \\
E_{28}, & 
\quad d_1 + d_2 = 2.1 \quad d_3 = 1.2 \quad d_5 = 6.1 \quad d_6 = 2.5 \\
E_{29}, & 
\quad d_1 + d_2 = 2.3 \quad d_3 = 1.2 \quad d_5 = 6.2 \quad d_6 = 2.6 \\
E_{30}, & 
\quad d_1 + d_2 = 2.2 \quad d_3 = 1.6 \quad d_5 = 6.4 \quad d_6 = 3.2
\end{align*}
\[ E_{31}, d_1 + d_2 = 2.4 \ d_3 = 1.7 \ d_5 = 6.2 \ d_6 = 2.9 \]

\[ E_{32}, d_1 + d_2 = 2.0 \ d_3 = 1.2 \ d_5 = 6.0 \ d_6 = 3.0 \]

\[ E_{33}, d_1 + d_2 = 2.0 \ d_3 = 1.5 \ d_5 = 6.5 \ d_6 = 2.7 \]

\[ E_{34}, d_1 + d_2 = 2.2 \ d_3 = 1.6 \ d_5 = 5.9 \ d_6 = 2.8 \]

\[ E_{35}, d_1 + d_2 = 2.6 \ d_3 = 1.5 \ d_5 = 6.9 \ d_6 = 2.9 \]

\[ E_{36}, d_1 + d_2 = 2.4 \ d_3 = 1.6 \ d_5 = 6.1 \ d_6 = 3.2 \]
$E_{37}, d_1 + d_2 = 2.2, d_3 = 1.3, d_5 = 5.8, d_6 = 2.8$

$E_{39}, d_1 + d_2 = 2.2, d_3 = 1.8, d_5 = 6.0, d_6 = 2.9$

$E_{40}, d_1 + d_2 = 2.4, d_3 = 1.6, d_5 = 5.8, d_6 = 3.1$

$E_{41}, d_1 + d_2 = 2.8, d_3 = 1.4, d_5 = 6.6, d_6 = 2.8$

$E_{42}, d_1 + d_2 = 2.4, d_3 = 1.4, d_5 = 6.9, d_6 = 3.0$
$E_{43}, d_1 + d_2 = 2.4, d_3 = 1.1, d_5 = 6.4, d_6 = 3.4$

$E_{44}, d_1 + d_2 = 2.4, d_3 = 1.9, d_5 = 6.3, d_6 = 3.4$

$E_{45}, d_1 + d_2 = 2.1, d_3 = 1.9, d_5 = 6.3, d_6 = 3.3$

$E_{46}, d_1 + d_2 = 2.4, d_3 = 1.9, d_5 = 6.8, d_6 = 3.3$

$E_{47}, d_1 + d_2 = 2.6, d_3 = 2.1, d_5 = 7.1, d_6 = 3.3$

$E_{48}, d_1 + d_2 = 2.5, d_3 = 1.5, d_5 = 6.6, d_6 = 2.8$
\[ E_{49} d_1 + d_2 = 2.7 \quad d_3 = 2.1 \quad d_5 = 6.8 \quad d_6 = 3.3 \]

\[ E_{50} d_1 + d_2 = 2.3 \quad d_3 = 1.3 \quad d_5 = 6.1 \quad d_6 = 2.6 \]

\[ E_{52} d_1 + d_2 = 2.2 \quad d_3 = 1.7 \quad d_5 = 5.8 \quad d_6 = 3.1 \]

\[ E_{53} d_1 + d_2 = 2.4 \quad d_3 = 1.4 \quad d_5 = 6.0 \quad d_6 = 2.9 \]

\[ E_{54} d_1 + d_2 = 2.2 \quad d_3 = 1.3 \quad d_5 = 5.8 \quad d_6 = 3.1 \]
\[ E_{55}, d_1 + d_2 = 2.4 \quad d_3 = 1.6 \quad d_5 = 5.9 \quad d_6 = 2.7 \]

\[ E_{56}, d_1 + d_2 = 2.3 \quad d_3 = 1.8 \quad d_5 = 6.0 \quad d_6 = 3.3 \]

\[ E_{57}, d_1 + d_2 = 2.2 \quad d_3 = 1.7 \quad d_5 = 5.9 \quad d_6 = 3.4 \]

\[ E_{59}, d_1 + d_2 = 2.6 \quad d_3 = 1.6 \quad d_5 = 7.0 \quad d_6 = 3.3 \]

\[ E_{60}, d_1 + d_2 = 2.5 \quad d_3 = 1.4 \quad d_5 = 6.4 \quad d_6 = 2.9 \]
\[ E_{63}, d_1 + d_2 = 2.3 \ d_3 = 1.2 \ d_5 = 5.7 \ d_6 = 2.8 \]
\[ E_{62}, d_1 + d_2 = 2.4 \ d_3 = 1.6 \ d_5 = 6.7 \ d_6 = 3.0 \]
\[ E_{63}, d_1 + d_2 = 2.3 \ d_3 = 1.7 \ d_5 = 6.2 \ d_6 = 3.2 \]
\[ E_{64}, d_1 + d_2 = 2.5 \ d_3 = 1.4 \ d_5 = 6.9 \ d_6 = 2.9 \]
\[ E_{65}, d_1 + d_2 = 2.6 \ d_3 = 1.5 \ d_5 = 6.7 \ d_6 = 2.5 \]
\[ E_{66}, d_1 + d_2 = 2.3 \ d_3 = 1.7 \ d_5 = 5.7 \ d_6 = 2.9 \]
$E_{67}, d_1 + d_2 = 3, d_3 = 2, d_5 = 7, d_6 = 3.4$

$E_{68}, d_1 + d_2 = 2.9, d_3 = 1.6, d_5 = 6.9, d_6 = 2.6$

$E_{69}, d_1 + d_2 = 2.2, d_3 = 1.6, d_5 = 6.1, d_6 = 2.5$

$E_{70}, d_1 + d_2 = 2.4, d_3 = 1.8, d_5 = 6.4, d_6 = 2.9$

$E_{71}, d_1 + d_2 = 2.5, d_3 = 1.6, d_5 = 6.5, d_6 = 2.9$

$E_{72}, d_1 + d_2 = 2.6, d_3 = 1.4, d_5 = 7.0, d_6 = 2.9$
\[ E_{73} d_1 + d_2 = 2.3 \quad d_3 = 1.5 \quad d_5 = 5.7 \quad d_6 = 2.8 \]

\[ E_{74} d_1 + d_2 = 2.5 \quad d_3 = 1.6 \quad d_5 = 7.4 \quad d_6 = 3.5 \]

\[ E_{75} d_1 + d_2 = 2.6 \quad d_3 = 1.4 \quad d_5 = 6.6 \quad d_6 = 2.6 \]

\[ E_{76} d_1 + d_2 = 2.9 \quad d_3 = 1.6 \quad d_5 = 6.7 \quad d_6 = 2.9 \]

\[ E_{77} d_1 + d_2 = 2.25 \quad d_3 = 1.75 \quad d_5 = 6.2 \quad d_6 = 2.75 \]

\[ E_{78} d_1 + d_2 = 2.6 \quad d_3 = 1.0 \quad d_5 = 6.0 \quad d_6 = 3.0 \]
\[ E_{79}, d_1 + d_2 = 2.6 \quad d_3 = 1.5 \quad d_5 = 7.0 \quad d_6 = 3.0 \]

\[ E_{80}, d_1 + d_2 = 2.4 \quad d_3 = 1.4 \quad d_5 = 6.0 \quad d_6 = 2.8 \]

\[ E_{81}, d_1 + d_2 = 1.9 \quad d_3 = 1.8 \quad d_5 = 6.3 \quad d_6 = 3.2 \]

\[ E_{82}, d_1 + d_2 = 2.5 \quad d_3 = 1.4 \quad d_5 = 6.3 \quad d_6 = 2.6 \]

\[ E_{83}, d_1 + d_2 = 2.4 \quad d_3 = 1.5 \quad d_5 = 5.9 \quad d_6 = 2.5 \]

\[ E_{84}, d_1 + d_2 = 2.7 \quad d_3 = 1.6 \quad d_5 = 6.8 \quad d_6 = 3.2 \]
\[ E_{86}, d_1 + d_2 = 2.3 \quad d_3 = 1.5 \quad d_5 = 6.6 \quad d_6 = 3.0 \]

\[ E_{87}, d_1 + d_2 = 2.4 \quad d_3 = 1.6 \quad d_5 = 6.6 \quad d_6 = 2.6 \]

\[ E_{88}, d_1 + d_2 = 2.3 \quad d_3 = 1.4 \quad d_5 = 6.1 \quad d_6 = 3.2 \]

\[ E_{89}, d_1 + d_2 = 2.5 \quad d_3 = 1.7 \quad d_5 = 5.7 \quad d_6 = 2.8 \]

\[ E_{90}, d_1 + d_2 = 2.3 \quad d_3 = 1.9 \quad d_5 = 6.1 \quad d_6 = 2.6 \]
\( E_{91} d_1 + d_2 = 2.4 \ d_3 = 1.4 \ d_5 = 6.0 \ d_6 = 2.6 \)

\( E_{92} d_1 + d_2 = 2.3 \ d_3 = 1.9 \ d_5 = 6.6 \ d_6 = 3.3 \)

\( E_{93} d_1 + d_2 = 2.4 \ d_3 = 1.6 \ d_5 = 6.4 \ d_6 = 2.7 \)

\( E_{94} d_1 + d_2 = 2.3 \ d_3 = 1.0 \ d_5 = 6.2 \ d_6 = 3.0 \)

\( E_{95} d_1 + d_2 = 2.7 \ d_3 = 1.6 \ d_5 = 6.6 \ d_6 = 2.5 \)

\( E_{96} d_1 + d_2 = 2.3 \ d_3 = 1.6 \ d_5 = 6.4 \ d_6 = 3.0 \)
\[ E_{97}, d_1 + d_2 = 2.6 \quad d_3 = 1.6 \quad d_5 = 6.8 \quad d_6 = 3.0 \]

\[ E_{98}, d_1 + d_2 = 2.4 \quad d_3 = 1.4 \quad d_5 = 6.8 \quad d_6 = 3.3 \]

\[ E_{99}, d_1 + d_2 = 2.2 \quad d_3 = 1.3 \quad d_5 = 6.1 \quad d_6 = 2.5 \]

\[ E_{100}, d_1 + d_2 = 2.6 \quad d_3 = 2.1 \quad d_5 = 7.3 \quad d_6 = 3.7 \]

\[ E_{101}, d_1 + d_2 = 2.1 \quad d_3 = 2.0 \quad d_5 = 5.7 \quad d_6 = 2.8 \]

\[ E_{102}, d_1 + d_2 = 2.5 \quad d_3 = 1.55 \quad d_5 = 5.75 \quad d_6 = 2.75 \]
\[ E_{103} d_1 + d_2 = 2.8 \quad d_3 = 1.5 \quad d_5 = 6.7 \quad d_6 = 3.7 \]

\[ E_{104} d_1 + d_2 = 2.9 \quad d_3 = 1.3 \quad d_5 = 6.9 \quad d_6 = 2.8 \]

\[ E_{105} d_1 + d_2 = 2.4 \quad d_3 = 1.2 \quad d_5 = 6.3 \quad d_6 = 3.3 \]

\[ E_{106} d_1 + d_2 = 2.6 \quad d_3 = 1.6 \quad d_5 = 6.7 \quad d_6 = 3.3 \]

\[ E_{107} d_1 + d_2 = 2.3 \quad d_3 = 1.6 \quad d_5 = 5.9 \quad d_6 = 3.3 \]

\[ E_{108} d_1 + d_2 = 2.5 \quad d_3 = 2.2 \quad d_5 = 6.6 \quad d_6 = 3.3 \]
\[ E_{109}: d_1 + d_2 = 2.5 \quad d_3 = 1.85 \quad d_5 = 7 \quad d_6 = 3.45 \]

\[ E_{110}: d_1 + d_2 = 2.6 \quad d_3 = 1.3 \quad d_5 = 6.7 \quad d_6 = 2.9 \]

\[ E_{111}: d_1 + d_2 = 2.6 \quad d_3 = 1.8 \quad d_5 = 6.7 \quad d_6 = 3.1 \]

\[ E_{112}: d_1 + d_2 = 2.4 \quad d_3 = 1.2 \quad d_5 = 6.2 \quad d_6 = 2.6 \]

\[ E_{113}: d_1 + d_2 = 3.0 \quad d_3 = 1.0 \quad d_5 = 7.8 \quad d_6 = 3.4 \]

\[ E_{114}: d_1 + d_2 = 2.3 \quad d_3 = 1.7 \quad d_5 = 5.9 \quad d_6 = 2.9 \]
$E_{115} d_1 + d_2 = 2.4 \ d_3 = 1.9 \ d_5 = 6.3 \ d_6 = 2.7$

$E_{116} d_1 + d_2 = 2.3 \ d_3 = 1.4 \ d_5 = 6.1 \ d_6 = 2.7$

$E_{117} d_1 + d_2 = 2.3 \ d_3 = 1.7 \ d_5 = 5.9 \ d_6 = 2.8$

$E_{118} d_1 + d_2 = 2.5 \ d_3 = 2.1 \ d_5 = 6.1 \ d_6 = 2.9$

$E_{119} d_1 + d_2 = 2.2 \ d_3 = 1.5 \ d_5 = 5.7 \ d_6 = 2.7$
\[ E_{121}, d_1 + d_2 = 2.7, d_3 = 1.8, d_5 = 6.7, d_6 = 3.5 \]

\[ E_{123}, d_1 + d_2 = 2.9, d_3 = 1.9, d_5 = 5.8, d_6 = 2.7 \]

\[ E_{124}, d_1 + d_2 = 2.0, d_3 = 1.5, d_5 = 6.6, d_6 = 3.3 \]
Appendix D

Anthropometric Measurements of the Head Shape from 3D Surfaces

This appendix displays the 3D measurement points for the SYMARE population. The red dots on the images show the points of measurements. Note that not all points on the head shape were used in the analysis. Further it also provides boxplots for the head and ear shape statistics for the SYMARE database. For 3D head and ear measurements and further discussion on their statistics please refer to section Sec. 5.2.3.2 of Chapter 5.
Appendix E

KPCA Analyses of Ear Morphology
Figure E.1: Variations in the ear morphology by changing the sixth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, -0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.2: Variations in the ear morphology by changing the seventh principal component is shown. The value \( m \) which indicates the magnitude of deformation from the template ear shape are \(-7, -5, -2, -1, -0, 1, 2, 5, 7\), these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.3: Variations in the ear morphology by changing the eight principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, -0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.4: Variations in the ear morphology by changing the ninth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, 0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.5: Variations in the ear morphology by changing the tenth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, 0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.6: Variations in the ear morphology by changing the eleventh principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, 0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.7: Variations in the ear morphology by changing the twelfth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, 0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.8: Variations in the ear morphology by changing the thirteenth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, -0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.9: Variations in the ear morphology by changing the fourteenth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, -0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.10: Variations in the ear morphology by changing the fifteenth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, 0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.11: Variations in the ear morphology by changing the sixteenth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, -0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.12: Variations in the ear morphology by changing the seventeenth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, 0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.13: Variations in the ear morphology by changing the eighteenth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, -0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.14: Variations in the ear morphology by changing the nineteenth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, 0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Figure E.15: Variations in the ear morphology by changing the twentieth principal component is shown. The value $m$ which indicates the magnitude of deformation from the template ear shape are $-7, -5, -2, -1, -0, 1, 2, 5, 7$, these numbers are shown above the ear shape. Further the ear shapes are coloured using the CSDA measure, regions of higher intensity show larger differences with respect to the template ear shape.
Appendix F

Reconstructing Ears Using the Morphable Model
Figure F.1: This figure compares ear shapes reconstructed using 50 principal components, \( \tilde{S}_i^{(50)} \), to the corresponding reference shapes \( Z_i \). Colors (constant luminance, so examine online) on the shape indicate local shape mismatch calculated using the measure \( \hat{d}(\tilde{S}_i^{(50)}, Z_i, f) \).
Figure F.2: This figure compares ear shapes reconstructed using 50 principal components, $\tilde{S}^{(50)}_i$, to the corresponding reference shapes $Z_i$. Colors (constant luminance, so examine online) on the shape indicate local shape mismatch calculated using the measure $\hat{d}(\tilde{S}^{(50)}_i, Z_i, f)$. 
Appendix G

Coustyx Configuration File

... 

// NUMBER OF SENSORS
var NUM_SENSORS=[
  2562
];

// OUTPUT FILE NAME
var SENSORS_FILE_NAME=[
  "THTE_P1Sigma0_v2-26253Hz.dat"
];

function Run(){
  try{
    SetOptionNumThreadsToMax();
    SetOptionFormulationType("BURTON_MILLER_GALERKIN");
    SetOptionUseFMM(TRUE);
    SetOptionFMMPrecomputeNearFieldMatrices(TRUE);
    SetOptionFMMTransitionMethod("SPEED");
    SetOptionFMMNumLevels(10);
    SetOptionUseGMRES(TRUE);
    SetOptionGMRESPreconditioner("NEARFIELD");
SetOptionGMRESConvergenceCriterion("RESIDUAL");
SetOptionGMRESMaxResidual(0.00500000000000);
SetOptionGMRESInitialGuess("PREVIOUS_FREQ");
SetOptionGMRESNVectorsKrylovSubspaceAtRestart(1000);
SetOptionIntegrationOrderTriangleElem(3);
SetOptionIntegrationOrderQuadrilateralElem(3);
...
};
Appendix H

AMVE-GUI, A Tool for Analysing Acoustic and Morphological Variations in Ear Shapes

This appendix describes the design and the functionality of a graphical user interface (GUI) named the AMVE-GUI. The AMVE-GUI was designed in order to assist the analyses of the acoustic and morphological variations of the ears using the KPCA technique which was developed in Chapter 7. It is also used for analysing the morphoacoustics of rigidly aligned ears to the template shape versus normal ears. A snapshot of the AMVE-GUI is shown in Fig. H.1.

The AMVE-GUI contains five main functionality that will be explained below:

1. Plot(a) of the AMVE-GUI presents the user with a high resolution SFRS map (refer to Sec. 4.7) for the given frequency. The frequency can be adjusted using ”Slider 1” and the frequency value is also indicated above the Slider 1. The SFRS map corresponds to the ear morphology shown in Plot(b).

2. Plot(b) of the AMVE-GUI shows the selected ear morphology. Corresponding acoustic data for this ear shape are shown in Plot(a) and Plot(c)-Plot(e) of the AMVE-GUI. The ear morphology can be chosen for the first five kernel
Figure H.1: A screenshot of the AMVE-GUI for analysing the acoustic variations of ear shape morphology using KPCA. The GUI is composed of a control section where the user can choose the ear shape and its corresponding acoustic data. The morphological information that can be selected includes the kernel principal component number and the value of \( m \) which was previously explained. The acoustic information that can be selected includes, frequency for plotting the SFRS map and azimuth and elevation angles for plotting the DTF. There are a total of five plots in the AMVE-GUI these are, from top-left in clockwise direction (a) SFRS map for the ear shown in plot(b) and the frequency selected by the controls. (b) the shape of the ear for the given kernel principal component and value of \( m \). (c) the median plane DTFs. (d) the horizontal plane DTFs. (e) the HRTF for the given ear shape at the specified Azimuth and Elevation angles.

principal components and for a range of \( m \) (An explanation for the quantity \( m \) was given in Eq. (7.15) in Sec. 7.4). Details about the parameter \( m \) can be obtained from Sec. 7.4

3. Plot(c) of the AMVE-GUI shows high resolution plots of the median plane DTFs. The convention used for the elevation angle is shown in Fig. 4.20. Elevation angle \( \phi = \pi \) and \( \phi = -\pi \) are for a source originating directly below the head, elevation angles \( \phi = \frac{\pi}{2} \) and \( \phi = -\frac{\pi}{2} \) are for sources directly at the front or back of the head and elevation angle 0 is directly at the top of the head.

4. Plot(d) of the AMVE-GUI shows high resolution plots of the horizontal plane DTFs. The angular convention used for the azimuth angle is shown in Fig. 4.19. The azimuthal angle of \( \theta = \pi \) and \( \theta = -\pi \) is for a source originating directly at the back of the head, azimuthal angle of \( \theta = 0 \) is directly at the front of the
head and azimuthal angle of \( \theta = \frac{\pi}{2} \) is directly to the left of the head.

5. Plot(e) shows the DTF corresponding to the azimuthal and elevation angles entered by the user in the Azimuth and Elevation text boxes. A cross on the SFRS map shown in Plot(a) further visually shows the direction of the DTF.

The following explains the "User Control" section of the AMVE-GUI.

1. The radio button "KPCA" highlighted in Fig. [H.2] loads morphological and acoustic data for ear shapes that are obtained by changing the kernel principal components.

2. The radio button "Translated Ears to Template" shown in Fig. [H.3] loads morphological and acoustic data for the translated (only) ears towards the template. The ears are translated to align with the size of the template head shape. This mode can be used when analysing ear shapes that are rigidly aligned to the template versus when they are in normal orientation.

3. The radio button "Rigidly Aligned Ears To Template" shown in Fig. [H.4] loads data for a few rigidly aligned ear shapes to the template ear. This mode can be
Figure H.3: The AMVE-GUI with the "Translated Ears to Template" radio button highlighted.

used when analysing ear shapes that are rigidly aligned to the template versus when they are in normal orientation.

Figure H.4: The AMVE-GUI with the "Rigidly Aligned Ears To Template" radio button highlighted.

4. The drop-down list box "PCA Component Number" highlighted in Fig. H.5 lists the kernel principal component numbers for which both acoustic and morphological data is available. By selecting different values from the drop
The drop-down list box "Ear,Ear Pair Num" is used in conjunction with "Translated Ears To Template" or "Rigidly Aligned Ears To Template" radio buttons. It lists the pair of ears for which simulation data is available. 

"Slider 1" shown in Fig. [H.7] enables the user to select and plot the SFRS maps for a large range of frequencies starting from 1kHz going up to 16kHz for the selected data. Currently the step size for this slider is 1kHz, however this can be adjusted in the program.

"Slider 2" highlighted in Fig. [H.9] is a multifunction slider. When the radio button "KPCA" is selected the slider enables the user to adjust the value of $m$ as detailed previously. When this slider is moved morphological and acoustic data for the given value of $m$ and for the specific kernel principal component is loaded and displayed. For morphological data the valid range is between $-5 \leq m \leq 5$, and for the acoustic data the valid range is $-2.5 \leq m \leq 2.5$.

---

1. This functionality is available due to a continued and larger study of ear shapes for which parts of this study was conducted in Sec. 7.3.
When values of $|m| > 2.5$ are selected only morphological data is displayed and the acoustic plots are empty (refer to Fig. H.8). On the other hand when the radio button "Translated Ears To Template" or "Rigidly Aligned Ears To Template" is selected the slider shows either the source or target ear shape depending on the position of the slider. At the extreme left the source shape
Figure H.8: The AMVE-GUI displaying only the morphology data when the value of $|m| > 2.5$

data is displayed and at the extreme right the target shape data is displayed.

Figure H.9: The AMVE-GUI with "Slider 2" is highlighted. If the radio button "KPCA" is selected this slider enables acoustic and morphological information displayed for various values of $m$. On the other hand if the "Translated Ears to Template" or "Rigidly Aligned Ears To Template" is selected then the acoustics and morphology corresponding to that ear is displayed.

8. Input fields "Azimuth" and "Elevation" highlighted in Fig. H.10 enable the user

\footnote{As detailed previously this is useful for the study of ear shapes similar to the discussion in Sec. 7.3}
to plot the DTF for the given azimuth and elevation angle for the displayed ear shape.

Figure H.10: The AMVE-GUI with the "Text boxes" highlighted. DTFs corresponding to the direction indicated in the text boxes will be highlighted.
Appendix I

Publications
ABSTRACT

This paper describes how Large Deformation Diffeomorphic Metric Mapping (LDDMM) can be coupled with a Fast Multipole (FM) Boundary Element Method (BEM) to investigate the relationship between morphological changes in the head, torso, and outer ears and their acoustic filtering (described by Head Related Transfer Functions, HRTFs). The LDDMM technique provides the ability to study and implement morphological changes in ear, head and torso shapes. The FM-BEM technique provides numerical simulations of the acoustic properties of an individual’s head, torso, and outer ears. This paper describes the first application of LDDMM to the study of the relationship between a listener’s morphology and a listener’s HRTFs. To demonstrate some of the new capabilities provided by the coupling of these powerful tools, we morph the shape of a listener’s ear, while keeping the torso and head shape essentially constant, and show changes in the acoustics. We validate the methodological framework by mapping the complete morphology of one listener to a target listener and obtaining the target listener’s HRTFs. This work utilizes the data provided by the Sydney York Morphological and Acoustic Recordings of Ears (SYMARE) database.

Index Terms— SYMARE, LDDMM, HRTF, Ear Morphology, Binaural Hearing

1. INTRODUCTION

Morphoacoustic signal processing is a relatively new term in binaural acoustics [1]. It refers to signal processing in which the relationship between a listener’s morphology (the shape of the torso, head, and ears) and the listener’s individualized acoustic filtering properties plays a prominent role. The acoustic filters required for synthesizing high-fidelity 3D audio for an individual listener are referred to as head-related impulse response (HRIR) filters. These filters describe the acoustic filtering properties of the torso, head, and outer ears and how they transform the sound waves that arrive from some location in space, interact with the ear, and ultimately reach the tympanic membrane. This acoustic filtering imparts a signature to the incoming sound that the human auditory system perceptually decodes as spatial information. The Fourier transform of the HRIR filters are referred to as head-related transfer functions (HRTFs).

The study of the relationship between morphology and acoustic filtering has been on-going for over a decade now. For example, [2] applied regression and principle component analysis to establish a linear relation between morphological and acoustic data. In another study, [3] proposed a “best fit” HRTF based on a similarity score between individual anthropometric measurements and data available in the CIPIC database. More recently [1] has characterized the effect of small perturbations on the magnitude of the notches and peaks appearing in the frequency spectrum of a single HRTF. Similarly, [4] describes the effect of ear morphology on the magnitude of the N1 notch in the median plane.

We now briefly introduce the tools that are applied in this study, beginning with the FM-BEM. Traditionally, individual HRTFs are acoustically measured in an anechoic chamber that incorporates a robotic arm to move a sound source in space. Microphones are placed in the ears of the listener and impulse responses are recorded for a number of directions in space. This clearly requires specialized equipment and is time consuming for the listener. With the increasing computational power of personal computers and also the improvement in numerical methods for boundary element method simulations, such as the fast multipole method [5], a listener’s HRTFs can now be numerically derived using acoustic simulations conducted using high-resolution surface meshes of the listener. Considerable research has been done to establish the validity of numerical simulations of HRTFs using the BEM [5, 6, 7]. These research show that high resolution surface meshes of the torso, head and ears are sufficient to provide accurate numerical simulation of HRTFs that provide a reasonable match to the acoustically-measured HRTFs.

We now introduce a shape analysis tool referred to as Large Deformation Diffeomorphic Metric Mapping (LDDMM) [8, 9] that will be used to manipulate and study ear shapes. LDDMM models the mapping of one surface, $C$, to another surface, $S$,
as a dynamic flow of diffeomorphisms of the ambient space, \( \mathbb{R}^3 \), in which the surfaces are embedded. The computations for LDDMM are not performed directly in the space of diffeomorphisms; instead, the method works with time dependent vector fields, \( v(t) : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \) for \( t \in [0, 1] \), which model the infinitesimal flows of the flow. The flow of diffeomorphisms, \( \phi^t(\cdot, X) \), operating on a subset \( X \subset \mathbb{R}^3 \) is defined via the partial differential equation:

\[
\frac{\partial \phi^t(\cdot, X)}{\partial t} = v(t) \circ \phi^t(\cdot, X),
\]

where \( \circ \) denotes the composition of functions. Note that the superscript \( v \) on \( \phi^t(\cdot, X) \) has no significance except to denote that the flow of diffeomorphisms is defined for a particular set of time dependent vector fields \( v(t) \).

The diffeomorphism at time \( t = 0 \), is the identity diffeomorphism: \( \phi^0(0, C) = C \). The result of the flow of diffeomorphisms at time \( t = 1 \) maps \( C \) to \( S \); we write this as \( \phi^1(t, C)|_{t:0 \rightarrow 1} = S \). The time dependent vector fields, \( v(t) \), belong to a Hilbert space of regular vector fields that has a kernel, \( k_V \), and a norm \( \| \cdot \|_V \) that models the infinitesimal cost of the flow. In the LDDMM framework, we solve an inexact matching problem which minimizes the cost function, \( J_{C,S} \), defined as:

\[
J_{c,s} \left( v(t)_{t \in [0,1]} \right) = \gamma \int_0^1 ||v(t)||_V^2 dt + E \left( \phi^t(t, C)|_{t:0 \rightarrow 1}, S \right),
\]

where \( E \) is a norm-squared cost measuring the degree of matching between \( \phi^t(t, C)|_{t:0 \rightarrow 1} \) and \( S \). In this work, we use the Hilbert space of currents \([10,11,12]\) to compute \( E \) because it is easier and more natural than using landmarks.

In order to make the concept of the flow of diffeomorphisms explicit, consider Fig. 1. In this figure, we transform the surface mesh for Ear A to that for Ear B. We show the flow of diffeomorphisms \( \phi^t(t, \text{Ear A})|_{t:i \rightarrow ti} \) for times \( t_i = \{0,0.2,0.4,0.6,0.8,1.0\} \).

In order to apply the tools of FM-BEM and LDDMM, we use the SYMARE database \([6]\) which provides high-resolution surface meshes of the head and torso both with and without ears. The database also provides high-resolution surface meshes of the ears alone. The basic idea behind this paper is to manipulate shapes using LDDMM and to examine the resulting acoustics using FM-BEM. The problem that we consider is the classical question of: “what does it mean to listen through another individual’s outer ears.” In examining this question, we demonstrate and validate our approach to combining the tools of LDDMM and FM-BEM. We also reveal an interesting finding regarding the influence of the head and torso on HRTFs. In Section 2 we describe the methods; Section 3 describes the results and Section 4 concludes.

2. METHOD

This study starts with high-resolution surface meshes (approximately 800,000 triangular elements) of the torso, head, and ears for two subjects in the SYMARE database, referred to as S1 and S2 (top row of Fig. 2). The surface meshes are triangulated meshes composed of a collection of vertices, \( x_n, 1 \leq i \leq N \), such that \( x_n \in \mathbb{R}^3 \) and a set of connectivity information for these points, \( f_m, 1 \leq m \leq M \), which constitutes the triangular faces of the mesh. Note that for simplicity, we use S1 and S2 to designate both the subjects and their corresponding surface meshes. Using the tools of LDDMM, we apply two shape transformations to the surface mesh for S1. In one transformation, we transform the left ear of S1 such that it is similar to the left ear of S2. The resulting surface mesh is referred to as S12ear−only. In a second transformation, we transform the torso, head, and left ear (but not the right ear) of S1 to be similar to the torso, head, and left ear of S2. The resulting surface mesh is referred to as S12all. The two shape transformations described above are shown in the bottom row of Fig. 2. We also perform similar shape transformations to S2, but in the reverse direction from S2 to S1, to obtain S21ear−only and S21all. All together, we have six different surface meshes consisting of the two original surface meshes and the four transformed surface meshes.

In order to describe how the shape transformations are im-
implemented using LDDMM, we define three mathematical operations – translate: $T$; match: $M$; and flow: $F$. The mathematical operation $T(S_1, S_2)$ applies a translation to the surface mesh $S_1$ such that it best matches surface mesh $S_2$. The mathematical operation $M(S_1, S_2)$ finds the momentum vectors, $\alpha_n(t)$, that minimize the cost function $J_{S_1, S_2}(v(t)_{t \in [0, 1]})$. The momentum vectors describe the change to a point, $x(t)$, in the ambient space, at a given time step via the following differential equation:

$$
\frac{dx(t)}{dt} = \sum_{n=1}^{N} k_V(x_n(t), x)\alpha_n(t), \quad (3)
$$

where $k_V(x, y)$ is the Cauchy kernel defined by:

$$
k_V(x, y) = \frac{1}{1 + \frac{\|x-y\|^2}{\sigma_V^2}}. \quad (4)
$$

The parameter $\sigma_V$ determines the range of influence of the momentum vectors $\alpha_n(t)$. The mathematical operation $F(S_1, \{\alpha\})$ applies the flow of diffeomorphisms defined by a set of momentum vectors, $\{\alpha\}$, to the surface mesh $S_1$. For simplicity in this notation, we have not made explicit the time dependence of the momentum vectors, $\{\alpha\}$.

With the definitions of the three mathematical operations ($T$, $M$, and $F$) in hand, it is straightforward to describe the methods to determine $S_{12_{\text{ear-only}}}$ and $S_{12_{\text{all}}}$ from $S_1$. Fig. 3 shows and describes the five steps required to synthesize $S_{12_{\text{ear-only}}}$ from $S_1$. Fig. 4 shows and describes the three steps required to synthesize $S_{12_{\text{all}}}$ from $S_1$. In order to follow the procedures listed in Figs. 3 and 4, it is important to understand that the SYMARE database provides high-resolution surface meshes of the torso and head without ears. So, for example, in Fig. 3 we first learn the momentum vectors required to match $HT1$ (the torso and head surface mesh of $S_1$ without ears) to $HT2$ (the torso and head surface mesh of $S_2$ without ears). We then apply this flow of diffeomorphisms to $LE1$ (the left ear of $S_1$) to obtain the intermediate left ear $LE3$. We then learn the momentum vectors matching $LE3$ to $LE2$ (the left ear of $S_2$). We then apply both flows of diffeomorphisms sequentially to $S_1$ to obtain $S_{12_{\text{all}}}$. Synthesizing $S_{12_{\text{all}}}$ is a control condition in the sense that the HRTFs for $S_{12_{\text{all}}}$ should be identical to that for $S_2$. In Fig. 4 we show the steps to obtain $S_{12_{\text{ear-only}}}$. Because the head diameter of $S_1$ and $S_2$ are not identical, we translate $LE2$ to match $LE1$. We then learn the momentum vectors matching $LE1$ to the translated
version of LE2 and apply the flow of diffeomorphisms to S1 to obtain S12ear−only.

In order to obtain the HRIRs corresponding to the six surface meshes (S1, S2, S12ear−only, S12all, S21ear−only, S21all) we apply FM-BEM simulations. In this work we used the Coustyx software by Ansol. The simulations were performed by the FM-BEM solver using the Burton-Miller Boundary Integral Equation (BIE) method. Using the acoustic reciprocity principle, a single simulation is used to determine all of the HRIRs in one go by placing a source on a surface mesh element that forms part of the blocked ear canal and then setting a uniform normal velocity boundary condition on this surface element. A post-processing step was used to refine the meshes prior to the FM-BEM simulation using the open-source software ACVD. The criteria the meshes need to meet during the mesh refinement are described in [6].

3. RESULTS

We now compare the FM-BEM simulated HRIR data for S2, S12ear−only, S12all. To make these comparisons, we plot the spatial frequency response surfaces (SFRS) corresponding to the HRTF data. An SFRS plot (see [13] for details) shows the magnitude gain of the HRTF for a single frequency as a function of direction in space. Fig.5 shows the SFRS plots for S2, S12ear−only, and S12all at 2 kHz and 10 kHz. The SFRS plots for S2 and S12all are practically identical, thus validating the methodological framework. On the other hand, the SFRS plot for S12ear−only shows differences that can be attributed to the different torso and head. The spatial correlation between the SFRS’s for S2 and S12all and between the SFRS’s for S2 and S12ear−only was calculated as a function of frequency and are shown in Fig 6. Morphological differences in the torso and head cause the spatial correlation to dip around 2 kHz and, surprisingly, around 9 kHz.

Fig. 5: SFRS plots for S2 (a,d); S12all (b,e); and S12ear−only (c,f) are shown at two frequencies, 2 kHz (top row) and 10 kHz (bottom row).

Fig. 6: The spatial correlation between various SFRS’s are shown as a function of frequency. The spatial correlation is shown for the following pairing: (S2 and S12ear−only) and (S2 and S12all) – solid line; (S1 and S21ear−only) and (S1 and S21all) – dotted line.

4. CONCLUSIONS

We have demonstrated the first application combining the tools of LDDMM and FM-BEM to gain insights into binaural acoustics. We have shown how these tools can be applied to study the influence of morphological changes on acoustics. LDDMM provides a powerful and flexible tool to study, characterize and manipulate ear shapes. In future work, we will statistically characterize the distribution of torso, head and ear shapes and their relationship to binaural acoustics.
5. REFERENCES


A Multiscale LDDMM Template Algorithm for
Studying Ear Shape Variations

Reza Zolfaghari, Nicolas Epain, Craig T. Jin
School of Electrical and
Information Engineering
University of Sydney
Sydney, Australia 2006
reza.zolfaghari@sydney.edu.au
craig.jin@sydney.edu.au

Anthony Tew
Department of Electronics
University of York
Heslington, York, UK YO10 5DD
tony.tew@york.ac.uk

Joan Glaunès
MAP5
Université Paris 5-René Descartes
Paris, France 75006
alexis.glaunes@mi.parisdescartes.fr

Abstract—This paper describes a method to establish an
average human ear shape across a population of ears by se-
quentially applying the Large Deformation Diffeomorphic Metric
Mapping (LDDMM) framework at successively smaller physical
scales. Determining such a population average ear shape, also
referred to here as a template ear, is an essential step in studying
the statistics of ear shapes because it allows the variations in
ears to be studied relative to a common template shape. Our
interest in the statistics of ear shapes stems from our desire to
understand the relationship between ear morphology and the
head-related impulse response (HRIR) filters that are essential
for rendering 3D audio over headphones. The shape of the ear
varies among listeners and is as individualized as a fingerprint.
Because the acoustic filtering properties of the ears depend on
their shape, the HRIR filters required for rendering 3D audio
are also individualized. The contribution of this work is the
demonstration of a sequential multiscale approach to creating
a population template ear shape using the LDDMM framework.
In particular we apply our sequential multiscale algorithm to a small
population of synthetic ears in order to analyse its performance
given a known reference ear shape.

I. INTRODUCTION

Morphoacoustic signal processing refers to signal process-
ing in which the relationship between a listener’s morphology
(the shape of the torso, head, and ears) and the listener’s
individualized acoustic filtering properties plays a prominent
role [1]. In this work, we focus on ear morphology and describe
a sequential multiscale approach to calculate a population av-
erage ear shape, which we refer to as a template ear shape. Our
work uses the Large Deformation Diffeomorphic Metric Map-
ing (LDDMM) framework [2]–[4]. This framework enables
one to compute spatial transformations between shapes that are
diffeomorphic, i.e., continuous, invertible, and differentiable.
These spatial transformations also enable one to define a metric
distance in shape space. The value of being able to calculate
the metric distance between two shapes is that it enables one to
define and calculate a population template shape: i.e., a shape
for which the sum of the squared metric distance between it
and each shape in the population is minimum.

There are several methods for estimating a template shape
using the LDDMM framework [5]–[9]. Generally these meth-
ods work at a single, fixed physical scale which is usually
chosen to be small relative to the shape features of interest.
In other words, the metric distance is defined for shape
transformations that operate on a single, fine physical scale.
In this work, we consider metric distances defined at multiple
physical scales. The motivation for studying metric distances at
multiple physical scales originates from considerations of the
acoustic properties of ears. For example, when considering ear
acoustics it is important for us to understand the influence of
a change in the size, orientation, and position of the ear on
its acoustic properties as defined by the head-related impulse
response (HRIR) filters. These physical changes likely result
in gross spectral changes in the frequency spectrum of the
HRIR filters. Similarly, it is also important to consider how
the acoustic properties of the ear change when the morphology
changes on a smaller physical scale.

In order to accommodate metric distances defined at multi-
ple physical scales, we introduce a sequential multiscale (SMS)
approach to shape deformation using LDDMM. The standard
LDDMM framework provides a method to smoothly deform
one shape, the source, into another shape, the target. The shape
deformation is performed at a fixed physical scale. In the SMS
approach, we iteratively apply the standard LDDMM shape
deformation process at successively smaller physical scales so
that the final deformation is a composition of deformations at
different scales. We refer to this method as the SMS-LDDMM
method. The SMS-LDDMM method enables one to robustly
deform a source shape into a target shape with high precision.
In this work, we also apply the SMS approach to LDDMM
template estimation. In this case, the estimated template shape
is sequentially refined at increasingly finer physical scales. For
this work, we use the template estimation algorithm that is
described in [6], [8], [9]. This template estimation algorithm
relies on the theory of currents from geometric measure theory.
The advantage of using currents to represent surfaces is that
one does not need to define pairs of corresponding points
(landmarks) on two surfaces in order to calculate their spatial
similarity. As well, the space of currents is a vector space
and one can linearly combine currents to represent a single
template shape.

Our paper is organized as follows, in Section II we
briefly review LDDMM and its mathematical framework and
describe the SMS-LDDMM shape deformation algorithm. In
Section III, we describe the template estimation algorithm and
the SMS-LDDMM approach to template estimation. In Sec-
tion IV, we describe the results of applying the SMS-LDDMM
template estimation algorithm to a small synthetic population of ear shapes generated using the SYMARE database [10]. We then conclude the paper.

II. METHOD

A. LDDMM Framework

The LDDMM framework is presented in [2], [11] and specifically developed and applied to shapes in [3], [4]. It is based on theories from functional analysis, variational analysis and reproducible kernel Hilbert spaces. In order to describe surface matching using LDDMM, we define two surfaces, \( S_1(X) \) and \( S_2(Y) \), where \( X \subset \mathbb{R}^{3 \times N} \) and \( Y \subset \mathbb{R}^{3 \times M} \) specify the vertices of the two surfaces, respectively. LDDMM models the mapping of \( S_1(X) \) to \( S_2(Y) \) as a dynamic flow of diffeomorphisms of the ambient space, \( \mathbb{R}^3 \), in which the surfaces are embedded (refer to Fig. 1). The computations for LDDMM are not performed directly in the space of diffeomorphisms; instead, the method works with time dependent vector fields, \( \mathbf{v}(t) : \mathbb{R}^3 \rightarrow \mathbb{R}^3 \) for \( t \in [0,1] \), which model the infinitesimal efforts of the flow. The flow of diffeomorphisms is defined via the partial differential equation:

\[
\frac{\partial \phi^y(t, X)}{\partial t} = \mathbf{v}(t) \circ \phi^y(t, X),
\]

where \( \circ \) denotes composition. The initial condition for (1) is \( \phi^y(0, X) = X \). Note that the superscript \( y \) on \( \phi^y(t, X) \) simply denotes that the flow of diffeomorphisms is defined for a particular set of time dependent vector fields \( \mathbf{v}(t) \). The result of the flow of diffeomorphisms from time \( t = 0 \) to \( t = 1 \) maps \( S_1 \) to \( S_2 \) as shown below:

\[
S_1(\phi^y(t, X)|_{t=0 \rightarrow 1}) \equiv S_2(Y) + \mathcal{E},
\]

where \( \equiv \) denotes geometrical equivalence between surfaces despite differing number of vertices and \( \mathcal{E} \) represents the error in the matching. To simplify the notation, we henceforth assume that all diffeomorphic flows start from time \( t = 0 \).

The time dependent vector fields, \( \mathbf{v}(t) \), belong to a Hilbert space of regular vector fields equipped with a kernel, \( k_V \), and a norm \( \| \cdot \|_V \) that models the infinitesimal cost of the flow. In the LDDMM framework, we determine \( \mathbf{v}(t) \) by minimizing the cost function, \( J_{S_1,S_2} \):

\[
J_{S_1,S_2}(\mathbf{v}(t)) = \gamma \int_0^1 \| \mathbf{v}(t) \|^2 dt + E(S_1(\phi^y(1, X)), S_2(Y)),
\]

where \( E \) is a norm-squared cost measuring the degree of matching between \( \phi^y(1, S_1) \) and \( S_2 \), and \( \gamma \) is a parameter that sets the relative weight of the two terms in the cost function. In this work, \( \gamma = 1 \). It can be shown that the time dependent vector fields, \( \mathbf{v}(t) \), can be expressed as a sum of momentum vectors, \( \mathbf{\alpha}_n(t) \), defined for the vertices \( X \):

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

where in this work we use the Cauchy kernel defined by:

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

for \( x \) and \( y \in \mathbb{R}^3 \). The \( \sigma_V \) parameter is a scale parameter, that determines through the kernel, \( k_V \), the range of influence of the momentum vectors \( \mathbf{\alpha}_n(t) \). Setting \( \sigma_V \) to a larger value increases the coupling in the motion of vertices that are further apart.

We now consider the two terms of the cost function, \( J_{S_1,S_2} \). The first term \( \| \mathbf{v}(t) \|^2 \) is a kinetic energy term that measures geodesic distance in the Riemannian space of diffeomorphic deformations. It can be written in expanded form using the kernel function, \( k_v \), as:

\[
\| \mathbf{v}(t) \|^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \mathbf{\alpha}_i(t), k_V(\mathbf{x}_i(t), \mathbf{x}_j(t)) \mathbf{\alpha}_j(t) \rangle,
\]

where \( \langle \cdot, \cdot \rangle \) denotes the usual Euclidean inner product. The second term of the cost function, \( E(S_1(\phi^y(1, X)), S_2(Y)) \), measures the difference in the surface geometry of \( S_1(\phi^y(1, X)) \) and \( S_2(Y) \) and is based on the theory of currents [3]. Currents are linear functionals on the space of differential 2-forms and can be used to represent surfaces. The intuition behind using currents to represent surfaces is that differential 2-forms can be integrated over a surface to give a real value and when two surfaces have similar geometry, the difference in the value of the surface integrals for the two surfaces should be close to zero. We use the notation \( |S| \) to denote the current representing the surface \( S \). In the discrete setting, a surface, \( S \), is approximated by a triangular mesh in \( \mathbb{R}^3 \). Given a face \( f \) of \( S \), let \( c_S(f) \) denote the center of the face and \( \mathbf{n}_S(f) \) denote the normal vector to the face with a length equal to the area

![Fig. 1](image-url) The result of the flow of diffeomorphisms for several time steps is shown for the matching of Ear A to Ear B. The color indicates cumulative displacement. For clarity, a constant luminance colormap has been used and the figure must be examined in color.
of the face. We can then express $E(S_1, S_2)$ as follows:

$$E(S_1, S_2) = \sum_{f, g} (n_{S_1}(f), k_W(c_{S_1}(g), c_{S_1}(f)) n_{S_1}(g))$$

$$- 2 \sum_{f, q} (n_{S_2}(f), k_W(c_{S_2}(q), c_{S_1}(f)) n_{S_1}(q))$$

$$+ \sum_{p, q} (n_{S_2}(p), k_W(c_{S_2}(p), c_{S_2}(q)) n_{S_2}(q)),$$  \tag{7}

where $f, g$ index the faces of $S_1$ and $p, q$ index the faces of $S_2$. In (7), the kernel function, $k_W$, is the kernel of the reproducing kernel Hilbert space associated with the space of currents and in this work we use the Cauchy kernel:

$$k_W(x, y) = \frac{1}{1 + ||x-y||^2_{\sigma_W}}.$$

where $\sigma_W$ is a second physical scale parameter. Small values for $\sigma_W$ result in a high pass filter. The $\sigma_V$ and $\sigma_W$ parameters described in the previous section to vary the physical scale at which the shape matching algorithm operates. In the SMS approach, we consider shape matching that is performed sequentially at successively smaller scales. In order to describe the SMS-LDDMM algorithm we extend two operators, $\mathcal{M}$ and $\mathcal{F}$, that were first introduced in [12]. The mapping operation $\mathcal{M}(S_1, S_2, \sigma_V, \sigma_W)$ determines the momentum vectors, $\alpha(t)$, that minimize the cost function $J_{S_1, S_2}(v(t))$. The flow operation $\mathcal{F}(S, \{\alpha(t)\}, \sigma_V, t_{end})$ applies the diffeomorphic transformation defined by $\{\alpha(t)\}$ to the surface mesh $S$ from time $t = 0$ to $t = t_{end}$.

The SMS-LDDMM algorithm is described in Algorithm 1. The algorithm starts by mapping the source shape to the target shape at the scale defined by $\sigma_V(1)$ and $\sigma_W(1)$. The momentum vectors obtained are then used to flow the source shape diffeomorphically to the target shape resulting in a shape $T_1$ that matches the target shape at the first scale. $T_1$ then becomes the source shape for the mapping applied at the second scale. These operations are repeated for each successive scale specified by $\sigma_V(l)$ and $\sigma_W(l)$, where $\sigma_V(l + 1) < \sigma_V(l)$ and $\sigma_W(l + 1) < \sigma_W(l)$.

**Algorithm 1 SMS-LDDMM**

**inputs:** $S_1, S_2, [\sigma_V(1), ..., \sigma_V(L)], [\sigma_W(1), ..., \sigma_W(L)]$.  
**outputs:** $\{\alpha_i(t)\}, T_l$, for $l = 1 \ldots L$.

1: $T_0 \leftarrow S_1$
2: for $1 \leq l \leq L$ do
3:  $\{\alpha_i(t)\} \leftarrow \mathcal{M}(T_{l-1}, S_2, \sigma_V(l), \sigma_W(l))$
4:  $T_l \leftarrow \mathcal{F}(T_{l-1}, \{\alpha_i(t)\}, \sigma_V(l), t_{end})$
5: end for

Fig. 2 illustrates the SMS-LDDMM algorithm applied to match two distinct ears in the SYMARE database using four successively smaller scales. The scale parameters for the first scale, $\sigma_V(1)$ and $\sigma_W(1)$, are relatively large and have been chosen to capture differences in scale, rotation and translation. The scale parameters for the second scale, $\sigma_V(2)$ and $\sigma_W(2)$, are a factor of four times smaller than the first scale and have been chosen to capture important structural differences between the shapes. For the following two scales, we continue to divide the scale parameters by a factor of four to capture increasingly finer details in the target shape. Note that the final shape deformation, $\phi_{\text{SMS}}(t, X)$, can be described by the composition of the diffeomorphic flows at the different scales, i.e., $\phi_{\text{SMS}}(t, X) = \phi^{X_4} \circ \phi^{X_3} \circ \phi^{X_2} \circ \phi^{X_1}(t, X)$.

### III. SMS Template Estimation

This section describes the SMS approach to estimating a template for a population of shapes. In this work, we follow the template estimation methods described in [6].

**A. Barycenter**

Template estimation for a population of shapes in the LDDMM framework is a lengthy and iterative process requiring intensive computation. In order to accelerate the computational process, it is advantageous to start with an initial template that is well-centered within the population. One method to rapidly estimate a centered template shape is to use the barycenter or iterative centroid algorithm described in [9]. The barycenter algorithm is described in Algorithm 2 and closely resembles the running average procedure used in data statistics. The algorithm begins by setting shape $S_1$ as the first estimate.
of the centered template shape, $B$. It then maps shape $B$ to the second shape in the population, $S_2$, to obtain momentum vectors. These momentum vectors are used to flow shape $B$ diffeomorphically halfway towards shape $S_2$ by setting $t_{\text{end}} = \frac{1}{3}$ in the flow function. The newly obtained shape provides a second estimate of the centered template shape and replaces the old $B$. The process is then repeated for the remaining shapes in the population with $t_{\text{end}} = \frac{1}{3}$ for iteration $r$ so that the flow distance is inversely proportional to the number of shapes contributing to the current running estimate of the centered template shape. In this way, we obtain a template shape that is well-centered within the population of shapes.

Algorithm 2 Barycenter

**inputs:** \{${S}_1, \ldots, {S}_R$\}, $\sigma_V$, $\sigma_W$.

**outputs:** $B$

1: $T \leftarrow {S}_1$
2: for $2 \leq r \leq R$ do
3: \{${\alpha}_r(t)$\} $\leftarrow \mathcal{M}(B, {S}_r, \sigma_V, \sigma_W)$
4: $B \leftarrow \mathcal{F}(S_r, \{\alpha_r(t)\}, \sigma_V, \frac{1}{r})$
5: end for

The template estimation algorithm relies on the use of currents to represent shapes. To begin, as mentioned in the previous section, we first usually compute a barycenter, $B$, for the population of shapes. We then initialize the template estimation as described in Algorithm 3. Each shape, $S_r$, in the population is mapped to the barycenter and then diffeomorphically flowed nine-tenths of the way toward the barycenter to obtain a shape $S'_r$. The factor nine-tenths is not critical and has been chosen to move all of the shapes in the population much closer toward the barycenter while preserving the relative distribution of the population of shapes. This is to say, the relative distance of the shapes from the barycenter in the Riemannian space is preserved. The utility in performing such an operation is that we can then compute an initial representation of the template as a current [$T_0$]:

$$T = \arg\min_{\mathcal{U}} \sum_{r=1}^{R} \int_{0}^{1} \|v_r(t)||^2 dt$$

where $v_r(t) = \arg\min_{\omega_r(t)} J_{S_r, U}(\omega_r(t))$. (9)

The template estimation algorithm relies on the use of currents to represent shapes. To begin, as mentioned in the previous section, we first usually compute a barycenter, $B$, for the population of shapes. We then initialize the template estimation as described in Algorithm 3. Each shape, $S_r$, in the population is mapped to the barycenter and then diffeomorphically flowed nine-tenths of the way toward the barycenter to obtain a shape $S'_r$. The factor nine-tenths is not critical and has been chosen to move all of the shapes in the population much closer toward the barycenter while preserving the relative distribution of the population of shapes. This is to say, the relative distance of the shapes from the barycenter in the Riemannian space is preserved. The utility in performing such an operation is that we can then compute an initial representation of the template as a current [$T_0$]:

$$[T_0] = \frac{1}{R} \sum_{r=1}^{R} [S'_r].$$

Algorithm 3 Template Estimation Initialization

**inputs:** \{${S}_1,\ldots, {S}_R$\}, $\sigma_V$, $\sigma_W$. **optional:** $B$.

**outputs:** [$T_0$].

1: if $B$ is not provided then
2: [$T_0] \leftarrow \frac{1}{R} \sum_{r=1}^{R} [S_r]$ 
3: else
4: for $1 \leq r \leq R$ do
5: \{${\alpha}_r(t)$\} $\leftarrow \mathcal{M}(S_r, B, \sigma_V, \sigma_W)$
6: $S'_r \leftarrow \mathcal{F}(S_r, \{\alpha_r(t)\}, \sigma_V, 0.9)$
7: end for
8: [$T_0] \leftarrow \frac{1}{R} \sum_{r=1}^{R} [S'_r]$ 
9: end if

Algorithm 4 Template Estimation

**inputs:** \{${S}_1,\ldots, {S}_R$\}, [$T_0$], $\sigma_V$, $\sigma_W$.

**outputs:** $T_r$.

1: for $1 \leq i \leq 1$ do
2: for $1 \leq r \leq R$ do
3: \{${\alpha}_r(t)$\} $\leftarrow \mathcal{M}(S_r, [T_0], \sigma_V, \sigma_W)$
4: $S'_r \leftarrow \mathcal{F}(S_r, \{\alpha_r(t)\}, \sigma_V, 1)$
5: end for
6: [$T_0] \leftarrow \frac{1}{R} \sum_{r=1}^{R} [S'_r]$ 
7: end for
8: $T_{r} \leftarrow S'_r$

It can be appreciated that the ability to combine, that is add currents in the vector space of currents as in Equation 10, and then perform a mapping operation, $\mathcal{M}$, from a shape, $S_r$, to this sum of currents is the critical and advantageous aspect of the template estimation operation. Intuitively, it may be helpful to think of the current representation of the template as a union of shapes that are becoming more and more similar as the iterations proceed. The objective of the SMS template estimation is to converge to a template that satisfies (or approximately satisfies) Equation 9 at every scale. To this end, after a template, $T^i$, is generated for a scale, $i$, we map and then diffeomorphically flow the population of shapes \{${S}_1,\ldots, {S}_R$\} to this template to obtain a new population of shapes \{${S}_1,\ldots, {S}_R$\} (refer to lines 7 and 8 of Algorithm 5). This new population of shapes

C. Sequential Multiscale Template Estimation

The SMS approach to template estimation is described in Algorithm 5 and essentially repeats the template estimation procedure described previously for successively smaller scales. The objective of the SMS template estimation is to converge to a template that satisfies (or approximately satisfies) Equation 9 at every scale. To this end, after a template, $T^i$, is generated for a scale, $i$, we map and then diffeomorphically flow the population of shapes \{${S}_1,\ldots, {S}_R$\} to this template to obtain a new population of shapes \{${S}_1,\ldots, {S}_R$\} (refer to lines 7 and 8 of Algorithm 5). This new population of shapes
then becomes the starting population for the next iteration of
the SMS template estimation algorithm. This is repeated until
we reach the final scale $L$.

IV. RESULTS

In this section, we demonstrate the performance of the
SMS template estimation algorithm using the synthetic pop-
ulation of ears shown in Fig. 3 for which there is a hidden
reference. This synthetic population of ears was generated
from two surface meshes of ears randomly selected from the
SYMARE database [10]. We refer to these surface meshes
as $E_1$ and $E_2$. Surface mesh $E_1$ was mapped to surface
mesh $E_2$ using a scale: $\sigma_V = 12$ and $\sigma_W = 1.25$. We
then generated eight ear shapes evenly distributed along
the geodesic flow between $E_1$ and $E_2$. These eight ears
were then rotated about the interaural axis by the angles
$[-20^\circ, -13.3^\circ, -10^\circ, -8^\circ, 8^\circ, 10^\circ, 13.3^\circ, 20^\circ]$ and scaled by
a factor of $[0.8, 0.85, 0.9, 0.95, 1.05, 1.1, 1.15, 1.2]$, re-
spectively. The eight ears were remeshed using the software
ACVD [13] to create meshes with differing numbers of vertices
and variable mesh connectivity. The number of vertices were
randomly generated from a Gaussian distribution with a mean

Fig. 3. Original and starting population of ears for template $T^1$ computed at scale: $\sigma_V(1) = 160, \sigma_W(1) = 20$.

Fig. 4. Starting population of ears for template $T^2$ computed at scale: $\sigma_V(2) = 40, \sigma_W(2) = 5$.

Fig. 5. Starting population of ears for template $T^3$ computed at scale: $\sigma_V(3) = 10, \sigma_W(3) = 1.25$.

Fig. 6. Comparisons between the reference ear and template ears, $T^1, T^2, T^3$, obtained at scales 1 to 3.
Algorithm 5 SMS Template Estimation

**inputs:** \( \{S_1, \ldots, S_R\}, [\sigma_V(1), \ldots, \sigma_V(L)], [\sigma_W(1), \ldots, \sigma_W(L)] \)

**outputs:** \( T^L \)

1. \( S_0^r = S_r \) for \( r = 1, \ldots, R \)
2. for \( 1 \leq l \leq L \) do
3. \( B^l \leftarrow \text{Barycenter}(S_1^{-l}, \ldots, S_R^{-l}, \sigma_V(l), \sigma_W(l)) \)
4. \( [T_0] \leftarrow \text{TempInit}(S_1^{-l}, \ldots, S_R^{-l}, \sigma_V(l), \sigma_W(l), B^l) \)
5. \( T^l \leftarrow \text{TempEstim}(S_1^{-l}, \ldots, S_R^{-l}, [T_0], \sigma_V(l), \sigma_W(l)) \)
6. for \( 1 \leq r \leq R \) do
7. \( \{\alpha_r(t)\} \leftarrow \mathcal{M}(S_1^{-l}, T^l, \sigma_V(l), \sigma_W(l)) \)
8. \( S_r^l \leftarrow F(S_1^{-l}, \{\alpha_r(t)\}, \sigma_V(l), 1) \)
9. end for
10. end for

equal to the number of vertices of \( E_1 \) (8102) and a standard deviation of 300. The hidden reference surface mesh was generated as the surface mesh located in the middle of the geodesic flow between \( E_1 \) and \( E_2 \).

The SMS template estimation algorithm was applied to the synthetic population of eight ear surface meshes for the three scales: \( \sigma_V = [160, 20, 10]; \sigma_W = [20, 5, 1.25] \). The starting population of ears for each scale are shown in Figs. 3-5. The resulting templates for each scale are shown and contrasted with the hidden reference in Fig. 6. Vertices of the surface meshes are colored according to the minimum of the distances between the given vertex and all vertices in the hidden reference mesh. A constant luminance colormap is used so the figures must be examined in color. The templates more closely match the hidden reference as the scale is successively decreased.

In order to examine the issue of whether the final template, \( i.e., \) the template at scale level 3, satisfies Equation 9 at all scales, we computed the ratio \( \mu(l) \):

\[
\mu(l) = \frac{d^2(S_l^1, S_l^3)}{\max_{r=1, \ldots, 8} d^2(S_r^l, S_r^3)}, \tag{11}
\]

where \( l \) indicates the scale level and \( d^2(\cdot) = \int_0^1 \|v(t)\|^2 dt \) indicates the geodesic distance between shapes computed at scale level \( l \). The numerator of \( \mu(l) \) computes the squared distance of the final template to the template at scale level \( l \) and the denominator of \( \mu(l) \) computes the maximum distance between the surfaces comprising the current which represents the template at scale level \( l \). Small values for the ratio \( \mu(l) \) for all \( l \) indicate that Equation 9 is approximately satisfied at all scales and we empirically found \( \mu(1) = 0.275 \) and \( \mu(2) = 0.082 \).

V. Conclusion

We have introduced a new sequential multiscale approach to template estimation in the LDDMM framework which we refer to as SMS template estimation. The method calculates templates at successively smaller scales. We have demonstrated the application of the method to a synthetic population of surface meshes of ears with a hidden reference and found that the estimated template reasonably approximates the hidden reference. In future work, we will be testing the SMS template estimation method on a large population database of ears. We also intend to compare the SMS template estimation approach with a single scale approach to template estimation.

REFERENCES


GENERATING A MORPHABLE MODEL OF EARS

Reza Zolfaghari, Nicolas Epain, Craig T. Jin, Joan Glaunès, Anthony Tew

ABSTRACT

This paper describes the generation of a morphable model for external ear shapes. The aim for the morphable model is to characterize an ear shape using only a few parameters in order to assist the study of morphoacoustics. The model is derived from a statistical analysis of a population of 58 ears from the SYMARE database. It is based upon the framework of large deformation diffeomorphic metric mapping (LDDMM) and the vector space that is constructed over the space of initial momentums describing the diffeomorphic transformations. To develop a morphable model using the LDDMM framework, the initial momentums are analyzed using a kernel based principal component analysis. In this paper, we examine the ability of our morphable model to construct test ear shapes not included in the principal component analysis.

Index Terms— Morphable model, Ears, Currents, Shape analysis

1. INTRODUCTION

This paper describes a morphable model for external ear shapes. Our objective in creating the morphable model is to assist research into the prediction of individualized 3D audio filters for listeners based on the shape of their ears. The significance of the morphable model is its ability to compress the representation of 3D ear shapes to a set of parameters typically obtained by projection onto a set of orthogonal base functions [1, 2]. This parameterisation of ear shapes using a morphable model greatly aids in the study of morphoacoustics [3, 4, 5, 6], where the goal is to understand the link between variations in the shape of an ear and their effect on the corresponding set of 3D audio filter functions, referred to as head related impulse responses (HRIRs). HRIRs vary for each listener because each listener has differently shaped ears. There is an HRIR filter for each ear and each direction in space and these HRIR filters enable the rendering of binaural 3D audio for a listener.

Modeling ear shapes is a challenging task and ear shape deformations are arguably best described using a Riemannian space. In this regard, large deformation diffeomorphic metric mapping (LDDMM) is a framework to perform non-rigid diffeomorphic registration and mapping between images, surfaces, curves and distributions in two and three dimensional space [7, 8, 9, 10, 11, 12]. Diffeomorphic maps provide a smooth, invertible, one-to-one transformation between the source and target shape. In particular, considerable work has been undertaken to formulate an algorithm for mapping 3D triangulated surfaces [13, 14]. In a recent paper [15] we show how LDDMM coupled with fast multipole boundary element method (FM-BEM) simulations can assist with the study of morphoacoustics and in [16] we show how a template ear shape can be estimated using LDDMM. The template ear is a critical element of our morphable model, but we leave the description of its calculation to [16] as it is beyond the focus of this paper. While LDDMM permits a multiscale approach to mapping ear shapes as discussed in [16], the morphable model presented here is based on single scale LDDMM transformations. In this work we use the LDDMM framework combined with a kernel based principal component analysis (KPCA) technique [17, 18, 19] to construct a morphable model for ears. In particular, our morphable model uses the concept of the linear space of initial momentums [20] within the framework of LDDMM and a set of coupled differential equations known as the “shooting equations” to construct and model ears. We use the SYMARE database of ears described in [21] for generating our morphable model. This paper describes the morphable model and examines its ability to reconstruct new ear shapes, i.e., ear shapes outside of the database used for constructing the model.

2. METHODS

2.1. LDDMM Framework

LDDMM [22, 12] is a mathematical framework that can be employed for the registration and morphing of three-dimensional shapes [14, 13]. It is based on theories from functional analysis, variational analysis and reproducible kernel Hilbert spaces. In the LDDMM framework we model a 3D-shape as a mesh with triangular faces, which we refer to as $S(X)$ where $X$ is the matrix specifying the mesh vertices and $S$ represents the mesh connectivity (the triangular faces). We now describe two fundamental LDDMM operations that are at the core of this work. The first operation, referred to as LDDDM mapping, consists in determining the diffeomorphic transformation that morphs an initial shape $S_1(X)$, with $X \in \mathbb{R}^{N \times 3}$, into a target shape $S_2(Y)$ with $Y \in \mathbb{R}^{M \times 3}$. The result of this operation is a set of vectors, $\{\alpha_n(0)\}_{1 \leq n \leq N}$, defined at the vertices $X$ and known as the initial momentum vectors, that characterize the diffeomorphic transformation entirely. The second operation, referred to as geodesic shooting, applies the morphing operation (i.e. the diffeomorphic flow) described by the initial momentum vectors to a given shape.

2.1.1. LDDMM mapping

LDDMM models the mapping or morphing of $S_1(X)$ to $S_2(Y)$ as a dynamic flow of diffeomorphisms of the ambient space, $\mathbb{R}^3$, in which the surfaces are embedded. This flow of diffeomorphisms, $\phi^\nu(t, \cdot)$, is defined via the partial differential equation:

$$
\frac{\partial \phi^\nu(t, X)}{\partial t} = \nu(t) \circ \phi^\nu(t, X),
$$

(1)

where $\nu(t)$ is a time-dependent vector field, $\nu(t) : \mathbb{R}^3 \to \mathbb{R}^3$ for $t \in [0, 1]$, which models the infinitesimal efforts of the flow, and $\circ$ denotes function composition. Note that the superscript $\nu$ on $\phi^\nu(t, X)$ simply denotes that the flow of diffeomorphisms is defined
for a particular time-dependent vector field \( v(t) \). This vector field belongs to a Hilbert space of regular vector fields equipped with a kernel, \( k_V \), and a norm \( \| \cdot \|_V \) that models the infinitesimal cost of the flow. In the LDDMM framework, we determine \( v(t) \) by minimizing the cost function, \( J_{S_1,S_2} \):

\[
J_{S_1,S_2}(v(t)) = \gamma \int_0^1 \|v(t)\|^2 dt + E(S_1(\phi^r(1,X)),S_2(Y)),
\]

where \( E \) is a norm-squared cost measuring the degree of matching between \( S_1(\phi^r(1,X)) \) and \( S_2(Y) \). In this work we use the Hilbert space of currents \([8, 14]\) to compute \( E \) since it is easier and more natural than using landmarks. The parameter \( \gamma \) is a parameter that sets the relative weight of the two terms in the cost function. In this work \( \gamma = 5 \times 10^{-5} \).

It can be shown that the time-dependent vector field, \( v(t) \), can be expressed as a sum of momentum vectors, \( \alpha_n(t) \), with one momentum vector defined for each of the \( N \) vertices in \( X \):

\[
v(t) = \frac{dx(t)}{dt} = \sum_{n=1}^{N} k_V(x_n(t), x(t)) \alpha_n(t),
\]

where in this work we use the Cauchy kernel defined by:

\[
k_V(x,y) = \frac{1}{1 + \|x-y\|^2}, \tag{4}
\]

for \( x \) and \( y \) in \( \mathbb{R}^3 \). The \( \sigma_V \) parameter is a scale parameter that determines through the kernel, \( k_V \), the range of influence of the momentum vectors \( \alpha_n(t) \). Setting \( \sigma_V \) to a larger value increases the coupling in the motion of vertices that are further apart. In this work, \( \sigma_V = 10 \) mm. Further, the initial momentum vectors, \( \alpha_n(0) \), determine the diffeomorphic mapping of \( S_1 \) to \( S_2 \) entirely \([17]\). In other words, \( S_2 \) can be represented as a deformation of \( S_1 \) through the diffeomorphic flow defined by the initial momentum vectors \( \{\alpha_n(0)\}_{1 \leq n \leq N} \). In the following, we refer to the calculation of the initial momentum vectors as the mapping operator, \( \mathcal{M} \):

\[
\{\alpha_n(0)\}_{1 \leq n \leq N} = \mathcal{M}(S_1, S_2). \tag{5}
\]

### 2.2. Kernel Based Principal Component Analysis (KPCA)

In the previous section, we have shown how a given shape could be represented as the deformation of another shape through a flow of diffeomorphisms. In order to build a morphable model of ear shapes, we represent every ear in a given population as a deformation of a unique template shape, \( T \), which represents the population’s average ear. Details regarding the calculation of such a template are described in \([16]\). In this work, we assume the template shape, \( T \), is given. The first step in our analysis is to calculate the initial momentum vectors for every ear, \( S_i \), in the population of \( L \) ears, as follows:

\[
\{\alpha_n^{(i)}(0)\}_{1 \leq n \leq N} = \mathcal{M}(T, S_i). \tag{8}
\]

Together with the template shape, the set of initial momentum vectors form a rudimentary model of ear shape such that the template shape can be morphed into any of the shapes in the population. In order to simplify this model, we apply a kernel based Principal Component Analysis (KPCA) to the deformations represented by the initial momentum vectors. We use the kernel version of PCA because the space of deformations is Riemannian.

In order to calculate the principal components, we calculate the covariance matrix, \( C \), which expresses the mutual correlation of the different ear shapes in the space of deformations. To compute this matrix we first construct a data matrix \( A \in \mathbb{R}^{3N \times L} \) which contains the initial momentum vectors for the entire population of ears:

\[
A = [a_1, a_2, \ldots, a_L]_{3N \times L}, \tag{9}
\]

where \( a_i \) denotes the column vector containing all the initial momentum vector coefficients for shape \( S_i \). We then center the data by subtracting the population average momentum vectors. The centered data matrix, \( \hat{A} \), is given by:

\[
\hat{A} = [\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_L]_{3N \times L}, \tag{10}
\]

where \( \hat{a}_i \) is the vector of the centered momentum vectors for the \( i \)-th shape:

\[
\hat{a}_i = a_i - \bar{a} \quad \text{with} \quad \bar{a} = \frac{1}{L} \sum_{i=1}^{L} a_i. \tag{11}
\]

We also form the kernel matrix, \( K \), which contains the values of the kernel function for every pair of vertex positions that comprise the vertices, \( X \), of the template shape \( T \):

\[
K = \begin{bmatrix}
K_{11} & K_{12} & \cdots & K_{1N} \\
K_{21} & K_{22} & \vdots & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
K_{N1} & \cdots & \cdots & K_{NN}
\end{bmatrix}, \tag{12}
\]

where \( I_{3 \times 3} \) denotes the 3 \( \times \) 3 identity matrix.

The correlation between two shapes is calculated as the inner product of the initial momentum vectors in the Hilbert space of deformations, \( V \). The correlation between shapes \( S_i \) and \( S_j \) is given by:

\[
e_{ij} = \left\langle \{\alpha_n^{(i)}(0)\}, \{\alpha_n^{(j)}(0)\} \right\rangle_V = \hat{a}_i^T K \hat{a}_j, \tag{13}
\]

where \( (\cdot)^T \) denotes the transpose of a vector or matrix. Thus, the covariance matrix for the entire population of ears, \( C \), is given by:

\[
C = \hat{A}^T K \hat{A}. \tag{14}
\]
In order to calculate the principal components, as well as the coordinates of the ears in the basis of the principal components, we perform the singular value decomposition of the covariance matrix $C$:

$$C = VDV^T.$$  \hspace{1cm} (15)

The matrix of the principal components, $U$, can be then calculated as:

$$U = \hat{A}VD^{-\frac{1}{2}}.$$

Note that the principal components are orthogonal in the Hilbert space of deformations, i.e., $U^T KU = I$. It follows from Equation (16) that $\hat{A} = U\tilde{D}^{\frac{1}{2}}V^T$ and therefore $D^{\frac{1}{2}}V^T$ provides the coordinates of the different ear shapes in the basis of the principal components. Each ear can thus be reconstructed by: (1) computing $a_i = \tilde{a} + U\tilde{D}^{\frac{1}{2}}v_i$ (where $v_i$ is the $i$-th column of $V$); (2) shooting from the template in the $a_i$ direction, i.e., $S_i = \mathcal{S}(T, \{a_i\}_{1 \leq i \leq N})$. In other words, we now have a morphable model of ears in which each ear shape in the population is described by $L$ parameters, where $L$ is the size of the population of ears. Note that the dimension of the model can be further reduced at the cost of reduced shape reconstruction accuracy by keeping only the first $Q$($Q \leq L$) principal components.

### 2.3. Morphable Model of Ears

We now describe how a new ear shape $S_p$, that was not included in the computation of the principal components, can be described using the KPCA data. The computation of the model parameters for $S_p$ can be divided into three steps. First, compute the initial momentum vectors, $\alpha_p$, corresponding to the morphing of the template $T$ into shape $S_p$. Second, the population average momentum vectors are subtracted from $\alpha_p$ to obtain the centered initial momentum vectors, $\bar{a}_p$. Third, the centered momentum vectors are projected onto the principal components to obtain the model parameters, $v_p$. The procedure is summarized below:

**Algorithm 1** Computation of the Model Parameters for an Ear

**Inputs:** $U, \bar{a}, S_p$

1. $\{\alpha^{(p)}_i(t)\}_{1 \leq i \leq N} = \mathcal{M}(T, S_p)$
2. $\bar{a}_p = a_p - \bar{a}$
3. $v_p = U^T \hat{K} \bar{a}_p$
4. **return** $v_p$

The reconstruction of shape $S_p$ from the model parameters is performed in two steps. First, the initial momentum vectors for shape $S_p$ are estimated by combining the principal components according to the model parameters. Second, the shooting operation is used to morph the template into $\hat{S}_p$, an approximation of shape $S_p$. The shape reconstruction operation is summarized below:

**Algorithm 2** Reconstruction of an Ear from the Model Parameters

**Inputs:** $T, U, v_p$

1. $\tilde{a}_p = \tilde{a} + Uv_p$
2. $\hat{S}_p = \mathcal{S}(T, \{\tilde{a}_p\}_{1 \leq i \leq N})$
3. **return** $\hat{S}_p$

Note that shape $\hat{S}_p$ is an approximation of $S_p$ because: (1) the LDDMM operation $\mathcal{S}(T, S_p)$ does not match shapes perfectly; and (2) the principal components may not enable perfect reconstruction of the initial momentum vectors for shape $S_p$.

### 3. EXPERIMENTS

#### 3.1. Experimental setup

A morphable model of ear shapes was created based on 58 different ear shapes from the SYMARE database [21]. While the SYMARE database is the largest database of its kind, 58 ears is not a large number considering the human population and it is unclear how well the morphable model can describe an arbitrary ear not included within the database. In order to address this issue, we repeatedly left one of the ears, $S_i$, out of the dataset of 58 ears and formed a morphable model based on the remaining 57 ears. We then examined the ability of the morphable model to reconstruct the ear that was left out. In other words, for each shape $S_i$ in the dataset, we applied the KPCA analysis described in Section 2.2 using 57 ear shapes (i.e., leaving $S_i$ out). Then, using the method described in Section 2.3 an approximate ear shape $\hat{S}_i$ was reconstructed. We then examined how accurately the approximation $\hat{S}_i$ matches the original shape $S_i$ using a shape difference analysis based on currents (please refer to the Appendix). Further, we examined the shape reconstruction accuracy as a function of the number of principal components used to reconstruct the ear shape.

Note that in order to exclude the mismatch caused by the LDDMM matching procedure, we actually compared the $\hat{S}_i$ shapes to the shapes $Z_i$ obtained by matching the template $T$ to $S_i$ and shooting using the true initial momentum vectors.

#### 3.2. Results

Results are summarized in Figure 2. As expected the accuracy of the model improves as the number of principal components increases. However, there is very little difference between the results obtained with 50 and 57 principal components, which indicates that the last 7 principal components have very little influence on the accuracy of the model. Interestingly, the quality of the reconstruction strongly depends on the ear considered. Some ears are reconstructed with great accuracy using relatively few principal components, while others are poorly reconstructed using the full basis of principal components. This is illustrated in Figure 1 where examples of reconstructed ears are compared to the corresponding reference shapes. Observe that shapes $Z_1$ and $Z_2$ were reconstructed with no apparent mismatch, while there is clear mismatch for shapes $Z_3$, $Z_4$ and $Z_5$. In summary, these results indicate that the morphable model is promising, but requires a larger population of ears to enable the model to morph into any possible ear shape.

### 4. CONCLUSION

In this paper we have presented a method for generating a morphable model of ears using the LDDMM framework. The core idea of this method is to apply KPCA to the initial momentum vectors corresponding to morphing a template ear into the different ears in the dataset. We tested the method over a dataset of 58 ear shapes and examined how well each ear in this dataset would be reconstructed using a model formed with the 57 remaining ears. The results indicate that a larger dataset would be required to generate a model that can morph into any ear shape.

### 5. APPENDIX: ANALYSING SHAPE DIFFERENCE USING CURRENTS

In this appendix we describe a method for measuring the local mismatch between two ear shapes. This method is based on the represen-
where \( k \) is a scale parameter and \( \sigma \) is a measure of sensitivity patterns revealing reflecting and diffracting surfaces that are almost equal and \( d(S_1, S_2, f) \) is very small. On the other hand, when the two shapes are very dissimilar (orthogonal or far away from each other), \( \beta_2(f) \) is very small and \( d(S_1, S_2, f) \) is relatively large. In order to enable meaningful comparisons across different triangular faces and different shapes, we normalise \( d(S_1, S_2, f) \) by the absolute value of \( \beta_1(f) \). We also limit the maximum value of \( d(S_1, S_2, f) \) to unity to ensure the measure does not blow up when \( \beta_1(f) \) is very small.

The overall similarity between two ear shapes is calculated as the average similarity measure, \( d(S_1, S_2) \), given by:

\[
d(S_1, S_2) = \frac{1}{F} \sum_{f=1}^{F} d(S_1, S_2, f),
\]

where \( F \) denotes the total number of faces in shape \( S_1 \).

6. REFERENCES


ABSTRACT

This paper describes features in the ear shape that change across a population of ears and explores the corresponding changes in ear acoustics. The statistical analysis conducted over the space of ear shapes uses a kernel principal component analysis (KPCA). Further, it utilizes the framework of large deformation diffeomorphic metric mapping and the vector space that is constructed over the space of initial momentums, which describes the diffeomorphic transformations from the reference template ear shape. The population of ears examined by the KPCA are 124 left and right ear shapes from the SYMARE database that were rigidly aligned to the template (population average) ear. In the work presented here we show the morphological variations captured by the first two kernel principal components, and also show the acoustic transfer functions of the ears which are computed using fast multipole boundary element method simulations.

Index Terms—Morphoacoustics, LDDMM, Kernel principal Component Analysis, Ear shape analysis, FM-BEM

1. INTRODUCTION

This paper describes the most important features in the ear shape that change across a population of ears and explores the corresponding changes in ear acoustics. The work forms part of the study of morphoacoustics [1, 2, 3, 4], where the goal is to understand the link between variations in the shape of an ear and their effect on the corresponding set of 3D audio filter functions, referred to as head related impulse responses (HRIRs). HRIRs vary for each listener because each listener has differently shaped ears. There is an HRIR filter for each ear and each direction in space and these HRIR filters enable the rendering of binaural 3D audio for a listener. The purpose of the study is to assist research into the prediction of individualized 3D audio filters for listeners based on the morphology of their ears.

The outer ear is an intricate shape and examining the non-linear variations in the ear morphology between listeners is a challenging task. We consider ear shape diffeomorphisms as belonging to a Riemannian space. In this regard, large deformation diffeomorphic metric mapping (LDDMM) is a framework to perform non-rigid diffeomorphic registration and mapping between images, surfaces, curves and distributions in two and three dimensional space [5, 6, 7, 8, 9, 10]. Diffeomorphic maps provide a smooth, one-to-one transformation between the source and target shape. In particular, considerable work has been undertaken to formulate an algorithm for mapping 3D triangulated surfaces [11, 12].

In a recent paper [13] we show how LDDMM coupled with fast multipole boundary element method (FM-BEM) simulations can assist with the study of morphoacoustics and in [14] we show how a template or population average ear shape can be estimated using LDDMM. Furthermore, in [15] we show how a morphable model for ear shapes based on the LDDMM framework and the kernel principal component analysis (KPCA) is constructed. The template ear is a critical element of the statistical analysis conducted here, but we leave the description of its calculation to [14, 16] as it is beyond the focus of this paper. While LDDMM permits a multiscale approach to mapping ear shapes as discussed in [14], the statistical analysis of ear shapes presented here is based on single scale LDDMM transformations from the reference template ear to ear shapes that have been aligned to the template ear shape via an affine transformation. In this work we use the LDDMM framework combined with a KPCA technique [16, 17, 18] to perform a statistical analysis of ear morphology. In particular, the statistical analysis conducted here is performed over the linear space of initial momentums [19] within the framework of LDDMM. By utilizing a set of coupled differential equations known as the “shooting equations” we examine the morphological variations seen in the ear shape. We use the population of left and right ear shapes in the SYMARE database [20] to conduct a statistical analysis of ear shapes. This paper shows the variations in the ear morphology captured by the first and second kernel principal component and also shows the associated changes in ear acoustics as determined by FM-BEM numerical acoustic simulations.

2. METHODS

2.1. LDDMM Framework

LDDMM [21, 10] is a mathematical framework that can be employed for the registration and mapping of three-dimensional shapes [12, 11]. It is based on theories from functional analysis, variational analysis and reproducible kernel Hilbert spaces. We model a 3D-shape as a mesh with triangular faces, which we refer to as a 3D-shape model. LDDMM is a framework to perform non-rigid diffeomorphic registration and mapping between images, surfaces, curves and distributions in two and three dimensional space [5, 6, 7, 8, 9, 10]. Diffeomorphic maps provide a smooth, one-to-one transformation between the source and target shape. In particular, LDDMM is a mathematical framework that can be employed for the registration and mapping of three-dimensional shapes [12, 11]. It is based on theories from functional analysis, variational analysis and reproducible kernel Hilbert spaces. We model a 3D-shape as a mesh with triangular faces, which we refer to as a 3D-shape model. LDDMM is a framework to perform non-rigid diffeomorphic registration and mapping between images, surfaces, curves and distributions in two and three dimensional space [5, 6, 7, 8, 9, 10]. Diffeomorphic maps provide a smooth, one-to-one transformation between the source and target shape. In particular, LDDMM is a mathematical framework that can be employed for the registration and mapping of three-dimensional shapes [12, 11]. It is based on theories from functional analysis, variational analysis and reproducible kernel Hilbert spaces. We model a 3D-shape as a mesh with triangular faces, which we refer to as a 3D-shape model. LDDMM is a framework to perform non-rigid diffeomorphic registration and mapping between images, surfaces, curves and distributions in two and three dimensional space [5, 6, 7, 8, 9, 10]. Diffeomorphic maps provide a smooth, one-to-one transformation between the source and target shape. In particular, LDDMM is a mathematical framework that can be employed for the registration and mapping of three-dimensional shapes [12, 11]. It is based on theories from functional analysis, variational analysis and reproducible kernel Hilbert spaces. We model a 3D-shape as a mesh with triangular faces, which we refer to as a 3D-shape model. LDDMM is a framework to perform non-rigid diffeomorphic registration and mapping between images, surfaces, curves and distributions in two and three dimensional space [5, 6, 7, 8, 9, 10]. Diffeomorphic maps provide a smooth, one-to-one transformation between the source and target shape. In particular, LDDMM is a mathematical framework that can be employed for the registration and mapping of three-dimensional shapes [12, 11]. It is based on theories from functional analysis, variational analysis and reproducible kernel Hilbert spaces. We model a 3D-shape as a mesh with triangular faces, which we refer to as a 3D-shape model. LDDMM is a framework to perform non-rigid diffeomorphic registration and mapping between images, surfaces, curves and distributions in two and three dimensional space [5, 6, 7, 8, 9, 10]. Diffeomorphic maps provide a smooth, one-to-one transformation between the source and target shape. In particular, LDDMM is a mathematical framework that can be employed for the registration and mapping of three-dimensional shapes [12, 11]. It is based on theories from functional analysis, variational analysis and reproducible kernel Hilbert spaces. We model a 3D-shape as a mesh with triangular faces, which we refer to as a 3D-shape model. LDDMM is a framework to perform non-rigid diffeomorphic registration and mapping between images, surfaces, curves and distributions in two and three dimensional space [5, 6, 7, 8, 9, 10]. Diffeomorphic maps provide a smooth, one-to-one transformation between the source and target shape. In particular,
LDDMM framework, we determine \( v(t) \) by minimizing the cost function, \( J_{S_1, S_2} \):

\[
J_{S_1, S_2}(v(t)) = \gamma \int_0^1 \|v(t)\|^2 dt + E \left( S_1(\phi^v(1, X)), S_2(Y) \right),
\]

where \( E \) is a norm-squared cost measuring the degree of matching between \( S_1(\phi^v(1, X)) \) and \( S_2(Y) \). In this work we use the Hilbert space of currents \([6, 12]\) to compute \( E \) because it is easier and more natural than using landmarks. The parameter \( \gamma \) is a parameter that sets the relative weight of the two terms in the cost function. In this work \( \gamma = 5 \times 10^{-5} \).

The optimal \( v(t) \) can be expressed as a sum of momentum vectors, \( \alpha_n(t) \), with one momentum vector defined for each of the \( N \) vertices in \( X \):

\[
v(t) = \frac{dx(t)}{dt} = \sum_{n=1}^{N} k_V(x_n(t), x(t)) \alpha_n(t),
\]

where in this work we use the Cauchy kernel defined by:

\[
k_V(x, y) = \frac{1}{1 + \|x-y\|^2},
\]

for \( x \) and \( y \) in \( \mathbb{R}^3 \). The \( \sigma_V \) parameter is a scale parameter that determines through the kernel, \( k_v \), the range of influence of the momentum vectors \( \alpha_n(t) \). Setting \( \sigma_V \) to a larger value increases the coupling in the motion of vertices that are further apart. In this work, \( \sigma_V = 10 \) mm.

We now define three fundamental LDDMM operations that are at the core of this work, 1- LDDMM matching, 2- geodesic shooting and 3- diffeomorphic flow. The first LDDMM operation denoted by \( \mathcal{M} \) refers to the calculation of the momentum vectors that represent the matching between two shapes \( S_1 \) and \( S_2 \):

\[
\{\alpha_n(t)\}_{0 \leq n \leq 1} = \mathcal{M}(S_1, S_2).
\]

The second LDDMM operation denoted by \( \mathcal{S} \) consists in obtaining the deformed shape \( S_2' \) and the time dependent momentum vectors that completely parametrize the deformation between the shapes \( S_1 \) to \( S_2 \) from the initial momentum vectors:

\[
\{S_2', \{\alpha_n(t)\}_{0 \leq n \leq 1}\} = \mathcal{S}(S_1, \{\alpha_n(0)\}_{1 \leq n \leq N}).
\]

The operation \( \mathcal{S} \) is achieved by solving a set of coupled differential equations known as the shooting equations \([15]\). The third LDDMM operation is known as the diffeomorphic flow operation \( \mathcal{F} \) and uses Eq. 3 and the time dependant momentum vectors \( \{\alpha_n(t)\}_{0 \leq n \leq 1} \) to morph the shape \( S_1 \) to shape \( S_2' \):

\[
S_2' = \mathcal{F}(S_1, \{\alpha_n(t)\}_{0 \leq n \leq 1}^{-1}).
\]

In this work, \( S_2' \) is very close to \( S_2 \) but not identical depending on the LDDMM matching process.

### 2.2. Kernel Based Principal Component Analysis (KPCA)

The previous section shows how a given shape can be represented as the deformation of another shape through a flow of diffeomorphisms which is completely parameterized using the initial momentum vectors. In this section, we statistically analyse the deformation from the template ear shape, \( \overline{E} \), to all ears in the dataset, taking for granted that the template shape has already been computed. We use the kernel principal component analysis (KPCA) to statistically analyse the initial momentum vector data corresponding to the deformations. KPCA uses the same inner product as in the computation of the deformation in the LDDMM cost function. The first step in our analysis is to calculate the momentum vectors for every ear, \( S_i \), in the population of \( L \) ears, as follows:

\[
\{\alpha_n^{(l)}(t)\} = \mathcal{M}(\overline{E}, S_i)
\]

In order to calculate the principal components, we calculate the covariance matrix, \( C \), which expresses the mutual correlation of the different ear shapes in the space of deformations. To compute this matrix we first construct a data matrix \( \tilde{A} \in \mathbb{R}^{3N \times L} \) which contains the initial momentum vectors for the entire population of ears:

\[
\tilde{A} = [a_1, a_2, \ldots, a_L]_{3N \times L}
\]

where \( a_l \) denotes the column vector containing all the initial momentum vector coefficients for shape \( S_i \). We then center the data by subtracting the population average momentum vectors. The centred data matrix, \( \hat{A} \), is given by:

\[
\hat{A} = [\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_L]_{3N \times L}
\]

We also form the kernel matrix, \( K \), which contains the values of the kernel function for every pair of vertex positions, \( x_n \), that comprise the vertices, \( X \), of the template shape \( \overline{E} \):

\[
K = \begin{bmatrix}
K_{11} & K_{12} & \cdots & K_{1N} \\
K_{21} & K_{22} & \cdots & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
K_{N1} & \cdots & \cdots & K_{NN}
\end{bmatrix},
\]

\[
K_{mn} = k_V(x_m, x_n) I_{3 \times 3}.
\]

where \( I_{3 \times 3} \) denotes the \( 3 \times 3 \) identity matrix.

The correlation between two shapes is calculated as the inner product of the initial momentum vectors in the Hilbert space of deformations, \( V \). The correlation between shapes \( S_i \) and \( S_j \) is given by:

\[
c_{ij} = \frac{1}{L-1} \left< \{\alpha_n^{(i)}(0)\}, \{\alpha_n^{(j)}(0)\} \right>_V = \frac{1}{L-1} \hat{a}_i^T K \hat{a}_j,
\]

where \( \left< \cdot, \cdot \right>_V \) denotes the transpose of a vector or matrix and \( \frac{1}{L-1} \) is a normalization factor. Thus, the covariance matrix for the entire population of ears, \( C \), is given by:

\[
C = \frac{1}{L-1} \hat{A}^T K \hat{A}
\]

In order to calculate the principal components, as well as the coordinates of the ears in the basis of the principal components, we perform the singular value decomposition of the covariance matrix \( C \):

\[
C = V D V^T.
\]

The matrix of the principal components, \( U \), can be then calculated as:

\[
U = \hat{A} V D^{-\frac{1}{2}}.
\]
Note that the principal components are orthogonal in the Hilbert space of deformations, i.e., \( U^T K U = I \). It follows from Equation (16) that \( \hat{A} = UD \tilde{\mathbf{v}} \) and therefore \( D \tilde{\mathbf{v}} \) provides the coordinates of the different ear shapes in the basis of the principal components. Each ear can thus be reconstructed by: (1) computing \( \mathbf{u}_i = \bar{a} + UD \tilde{\mathbf{v}}_i \) (\( \mathbf{v}_i \) is the i-th column of \( \mathbf{V}^T \)); and (2) shooting from the template in the \( \mathbf{a}_i \) direction, i.e., \( S_i = \mathcal{G}(\bar{E}, \{ \mathbf{u}_i \}) \).

2.3. Examining the Kernel Principal Components

We now describe how the kernel principal components can be used to examine important changes in the ear morphology and their corresponding acoustics. Each kernel principal component (KPC) captures some form of morphological variation seen in the population of shapes. In mathematical notation the kernel principal components are denoted by \( \mathbf{u}_i \), where \( i \) signifies the principal component number and also the column \( n \) in the matrix \( \mathbf{U} \). In order to examine the morphological variations captured by a single principal component, \( \mathbf{u}_i \), two steps are involved. In the first step, \( \mathbf{u}_i \), is multiplied by a suitable weight factor that is chosen to be a scalar multiple of the eigenvalue, \( m \mathbf{D}^i \), for some real value \( m \in \mathbb{R} \). Because of the normalization used in Eq. (13), the eigenvalue \( \mathbf{D}^i \) is equal to the variance of the scores (coordinates of the ear shapes) belonging to the \( i \)-th principal component. In the second step, the morphological changes with respect to the template shape can be observed by constructing the ear shape, \( \bar{E}(i, m) \), by using the shooting operation:

\[
\{ \bar{E}(i, m), \{ \mathbf{a}_n(t, i, m) \} \} = \mathcal{G}(\bar{E}, \bar{a} + m \mathbf{D}^i \mathbf{u}_i)
\] (17)

In order to obtain the acoustic response for the ear shape, the FM-BEM simulations need to be conducted on ears that are attached to the template head and torso shape. However, the template head and torso shape also has the template ear shape, \( \bar{E} \), attached. Fig. 1 shows a picture of the template head and torso shape with the template ear attached (i.e \( \mathcal{HTE} \)). In order to appropriately morph the template ear shape to the modified ear shape, \( \bar{E}(i, m) \), we use the time-dependent momentum vectors obtained in Eq. 17 to perform a flow operation on the template shape, \( \mathcal{HTE} \), to obtain a template head and torso shape with the modified ear:

\[
\mathcal{HTE}(i, m) = \mathcal{F}(\mathcal{HTE}, \{ \mathbf{a}_n(t, i, m) \})
\] (18)

Fig. 1. The template head, torso, and ear shape, \( \mathcal{HTE} \), is shown.

3. EXPERIMENTS

3.1. Experimental setup

We conducted KPCA on ear shapes that were obtained from 62 subjects in the SYMARE database. The right ear shapes of the 62 subjects were reflected to obtain left ear shapes so that we had a total of 124 left ear shapes. In this work, we exclude any scale, rotational or translational variations in ear shapes when conducting the KPCA and thus focus the KPCA solely on structural differences in ear morphology. The structural differences in ear morphology are the most difficult to study. Changes in scale, orientation and position are referred to as affine transformations. Thus, we first optimally align all of the left ear shapes to the template ear, \( \bar{E} \), using affine transformations based on a distribution matching technique described in [5].

In order to observe structural differences in ear shape, KPCA was performed on the 124 left ear shapes that were aligned to the template ear via an affine transformation. The KPCA was performed as detailed in Section 2.2. New ear shapes were generated by varying the weights corresponding to the first and second kernel principal components as described in the previous section. More precisely, new ear shapes were obtained using Eq. 17 for values of \( m \in B_1 = \{ \pm 7, \pm 2, \pm 1, 0 \} \). In order to study the acoustics of the new ear shapes, the template, \( \mathcal{HTTE} \), was modified according to Eq. 18 to obtain \( \mathcal{HTTE}(i, m) \) for \( m \in B_2 = \{ \pm 2, \pm 1, 0 \} \) and \( i \in \{ 1, 2 \} \). The only reason for generating ear morphologies at the large and nonsensical values of \( m = \pm 7 \) was to clearly visualize how the changes in the kernel principal component weights relate to changes in the ear morphology.

HRIRs corresponding to the shapes \( \mathcal{HTTE}(i, m) \) were generated using FM-BEM simulations. For this work the Coustyx software by Ansol was used [23]. The simulations were performed by the FM-BEM solver using the Burton-Miller Boundary Integral Equations (BIE) and the Galerkin implementation. Using the acoustic reciprocity principal, a single simulation is used to determine all of the HRIRs in one go by placing a source on a surface mesh element that forms part of the blocked ear canal and then setting a uniform normal velocity boundary condition on this surface element. A post-processing step was used to refine the meshes prior to the FM-BEM simulation using the open-source software ACVD [24]. The meshes had a critical frequency of 26 kHz for six elements per acoustic wavelength and further, met the FM-BEM mesh criterion detailed in [20, 25].

3.2. Results

Fig. 3 shows the ear shapes that have been generated from the template ear by changing the weights corresponding to the first and second principal components. The differences with respect to the template ear are highlighted in color using a normalized dissimilarity measure based on currents [15]. By changing the weights corresponding to the first and second kernel principal components, there are changes to the size and structure of features in the pinna. Please keep in mind that the overall size and rotation of the ears in the dataset have been aligned to the template, so that it is the structural features of the pinna...
Dissimilarity

Fig. 3. Ear morphologies are shown corresponding to systematic changes in (a) the first and (b) the second kernel principal components. Note that $E(i,m)$ denotes the template ear, $E$, modified with the $i$-th KPC using a weight of $m$ standard deviations. The degree of difference between the given ear shape and the template ear shape are highlighted in color using a normalized dissimilarity measure based on currents [15]. (The colors have constant luminance and so do not appear in grey-scale, instead please view the online version.)

that are changing. For the ensuing discussion, please refer to both Fig. 3 and Fig. 2 which shows the names of anatomical features of the outer ear. For the first KPC, when $m = -7$ the ear appears to be wider in width and have a larger Concha and Superior-Crus-Anti-Helix region. On the other hand, when $m = 7$ the ear appears to fold inwards and become narrower. For the second KPC, when $m = -7$ the Anti-Helix-Stem has moved outwards (even out of the ear) and the Superior-Crus-Anti-Helix region is larger. On the other hand, when $m = 7$ the opening of the Concha is very wide, making the Superior-Crus-Anti-Helix region smaller and pushing down the Anti-Tragus.

Consider now the acoustics of the ears. Fig. 4 shows the log-magnitude of the Directional Transfer Functions (DTFs) [26] in the median plane for several of the ear shapes in Fig. 3. In order to quantitatively evaluate the differences in the median plane DTFs, we used a measure similar to that described in [26]. Assume the log-magnitude spectra are given for two DTFs, so that we have $D_1(f)$ and $D_2(f)$. We then compute a log-magnitude spectral difference, $\sigma$, as:

$$\sigma = \sqrt{\frac{1}{N} \sum_{n=1}^{N} \left[ (D_1(f_n) - \overline{D}_1) - (D_2(f_n) - \overline{D}_2) \right]^2},$$  (19)

where $N$ is the number of frequency bins and $\overline{D}$ is the mean value of $D(f)$. The log-magnitude spectral difference between the median plane DTFs for $E(i,m)$ and $E$ were computed and then averaged across all elevations on the median plane (see Table 1). The average log-magnitude spectral difference for both KPCs is fairly similar and varies between 3 and 5.3 dB.

Table 1. The average log-magnitude spectral difference for the median plane DTFs are shown in dB for the first and second KPC.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$i = 1$</th>
<th>$i = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-7</td>
<td>5.3</td>
<td>4.4</td>
</tr>
<tr>
<td>-1</td>
<td>4.0</td>
<td>3.3</td>
</tr>
<tr>
<td>1</td>
<td>3.1</td>
<td>3.3</td>
</tr>
<tr>
<td>2</td>
<td>4.2</td>
<td>4.7</td>
</tr>
</tbody>
</table>

4. CONCLUSION

This paper shows variations in ear morphology that commonly occur across a population of ears and the associated changes in the ear acoustics. The analysis was performed using KPCA within the LDDMM framework. The morphological and corresponding acoustic variations of the ear shapes within the SYMARE population are shown for the first and second kernel principal components. The work detailed in this paper forms part of ongoing morphoacoustics research. Future studies will further examine the relationship between the kernel principal components for ear shape and the associated changes in ear acoustics.
5. REFERENCES


