INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
COMPUTATIONAL METHODS FOR INTEGRATING DIFFERENT ANATOMICAL DATA SETS OF THE HUMAN TONGUE

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Chao-Min Wu, B.S., M.S.

* * * * *

The Ohio State University

1996

Dissertation Committee:
Osamu Fujimura
Herman R. Weed
John A. Negulesco

Approved by

Adviser
Biomedical Engineering
To my wife, Ling-Chu, and my parents
ACKNOWLEDGEMENTS

There are many people whom I would like to thank for. I want to thank my mentors, Osamu Fujimura and Reiner Wilhelms-Tricarico. I am very grateful to my advisor, Osamu Fujimura, who shows me the world of phonetics and speech science. His lab. and seminar class have been a wonderful place to accumulate my knowledge in speech science. I am deeply indebted to Reiner Wilhelms-Tricarico for his constant guidance as teacher and friend. From him, I have learned more than what I can imagine. Thank you both, for your patience, encouragement, and constant support throughout my Ph.D. program.

I would like to thank Professor Herman R. Weed for being a wonderful teacher and inspiring me to the field of biological control system. I also want to thank Professor John Negulesco for his help on the study of tongue anatomy and helpful comments as my dissertation committee member. I want to thank Cynthia Roberts for her constant help on academic advice through years. My special thanks to Donna Erickson for her encouragement and proofreading a late draft of my dissertation. I would like to express my thanks to Larry Feth who supported me for one quarter while working on other project. I would also like to express my gratitude to Catita Williams, Jianwu Dang, Masashi Sawada, and Klaus Mueller for their friendship and exchange of ideas.
I want to express my deep appreciation to my parents for their love and moral support. I wish my grandfather would have lived to see this day and this document. Finally, but the most important, I thank my wife, Ling-Chu, for your love, support and companion throughout my graduate study.

I would like to thank Kiyoshi Honda, at ATR Human Information Processing Research Laboratories in Japan, for providing me with MRI data and valuable discussions. I want to acknowledge the funding sources throughout my Ph.D. program. This work has been partially supported by National Science Foundation under grant RF 729550 to Osamu Fujimura, ATR Human Information Processing Research Laboratories during my stay in Japan (Apr.- Jun., 1994), a gift from ATR Interpretating Telecommunications Research Laboratories to Osamu Fujimura, and Whitaker Fundation under grant RF 725500 to Reiner Wilhelms-Tricarico. Early stage of my Ph.D. study was supported by The Ohio State University under Seed Grant ID20010308 to Reiner Wilhelms-Tricarico, and support provided to me during Winter 1996 by ATR Media Integration & Communications Research Laboratories.
VITA

June 22, 1960 ...................... Born - Pei-Kang, Taiwan, R.O.C.

June, 1982 ...................... B.S. Biomedical Engineering
Chung-Yuan Christian University
Chung-Li, Taiwan, R.O.C.

1982-1984 ...................... Second Lieutenant, Chinese Army
Taiwan, R.O.C.

Aug., 1984- Dec., 1985 .............. Teaching Assistant
Dept. of Biomedical Engineering
Chung-Yuan Christian University
Chung-Li, Taiwan, R.O.C.

August, 1987 ...................... M.S. Bioengineering
The University of Wyoming
Laramie, Wyoming

Aug., 1987- Aug., 1988 .............. Graduate Research Associate
Bioengineering
The Pennsylvania State University
State College, Pennsylvania

Sep., 1988, Aug., 1990 .............. R & D Supervisor and Engineer
Suntex Instrument Inc.
Taipei, Taiwan, R.O.C.

Oct., 1990-present .............. Graduate Research Associate
Dept. of Speech and Hearing Science
The Ohio State University
Columbus, Ohio

Apr., 1994- June, 1994 .............. Visiting Researcher
ATR Human Information Processing
Research Laboratories
Kyoto, Japan
Publications

Research Publications


Fields of Study

Major Field: Biomedical Engineering

Studies in:

- Biological Control System and Modeling
  - Prof. H. R. Weed
- Speech Science
  - Prof. O. Fujimura
- Bioinstrumentation and Sensory Physiology
  - Prof. C. D. Ferris
# Table of Contents

DEDICATION ................................................................. ii

ACKNOWLEDGEMENTS .................................................. iii

VITA .................................................................................. v

LIST OF TABLES ................................................................. xi

LIST OF FIGURES ............................................................... xiii

CHAPTER PAGE

I Introduction ..................................................................................................... 1

1.1 Background and Motivation ............................................................... 1

1.1.1 Organization of Speech Utterances and Theoretical Phonetic Motivation of this Work 1

1.1.2 Articulatory Movements and their Control .................................... 6

1.2 Scope of this Work .............................................................................. 11

1.3 Significance ........................................................................................... 13

1.4 Organization of this Dissertation ..................................................... 15

II Literature Review ............................................................................................ 16

2.1 Anatomical Data .................................................................................. 16

2.2 Instrumental Methods for Tongue Movement Observation ........ 18

2.2.1 X-ray Cineradiography ............................................................. 18

2.2.2 X-ray Microbeam System .......................................................... 19

2.2.3 Magnetic Methods ................................................................... 21

2.2.4 Imaging Techniques ................................................................ 23

2.2.5 Electrical Methods ................................................................... 26
2.2.6 Remarks on Useful Recording Techniques ...... 28
2.3 Tongue Modeling .................................................. 29
  2.3.1 2-d Model ...................................................... 30
  2.3.2 3-d Model ...................................................... 31
  2.3.3 Dynamic 3-d Model ......................................... 32
2.4 Anatomical Mapping Methods ................................. 34
  2.4.1 Polynomial Mapping Methods ......................... 35
  2.4.2 Thin-plate Spline Mapping Method .................... 37
2.5 Linguistic Relevance: Existing Theories/Models of Temporal Organization of Speech ............. 39
  2.5.1 Phoneme Concatenation-Coarticulation Model .... 39
  2.5.2 Articulatory Phonology ................................. 41
  2.5.3 Converter/Distributor Model (C/D Model) .......... 43

III Methods .................................................................... 46
  3.1 Introduction ....................................................... 46
  3.2 Data Description ................................................ 47
    3.2.1 Magnetic Resonance Image (MRI) Data .......... 47
    3.2.2 Visible Man Data ........................................ 48
  3.3 Image Analysis .................................................. 49
  3.4 Mathematical Methods of Anatomical Description .... 53
    3.4.1 The Thin-plate Spline Mapping Technique ....... 53
    3.4.2 An Example of 2-d Thin-plate Spline Mapping .... 57
    3.4.3 3-d Matching of Anatomical Drawings and Visible Man Data ........................................ 61

IV Landmarks for 3-d Tongues ........................................ 65
  4.1 Introduction ....................................................... 65
  4.2 Types of Landmarks ............................................ 65
  4.3 Landmark Identification for Tongue Image Data .... 68
    4.3.1 Strategy for Landmark Selection ................. 68
    4.3.2 Locating Landmarks ................................. 72
  4.4 Summary ......................................................... 79

V Integration of Different Data Sets of the Human Tongue ................................................... 81
  5.1 Introduction ....................................................... 81

viii
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2</td>
<td>Alignment Mapping</td>
<td>81</td>
</tr>
<tr>
<td>5.3</td>
<td>3-d Thin-plate Spline Mapping</td>
<td>91</td>
</tr>
<tr>
<td>5.4</td>
<td>Evaluation of the Accuracy of Mapping</td>
<td>93</td>
</tr>
<tr>
<td>5.5</td>
<td>Summary</td>
<td>105</td>
</tr>
<tr>
<td>VI</td>
<td>Discussion</td>
<td>107</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>107</td>
</tr>
<tr>
<td>6.2</td>
<td>Limitation of the Current Data</td>
<td>108</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Miyawaki's Data</td>
<td>108</td>
</tr>
<tr>
<td>6.2.2</td>
<td>MRI Data</td>
<td>114</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Visible Man Data</td>
<td>114</td>
</tr>
<tr>
<td>6.3</td>
<td>Implications of this Work</td>
<td>115</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Application for Morphological Studies of Soft Tissues</td>
<td>115</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Applications to Phonetic Studies</td>
<td>116</td>
</tr>
<tr>
<td>6.3.3</td>
<td>Application to Biomechanical System Modeling</td>
<td>118</td>
</tr>
<tr>
<td>6.4</td>
<td>Future Work</td>
<td>118</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td></td>
<td>122</td>
</tr>
<tr>
<td>APPENDICES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>Program Description and Listing for 2-d Thin-plate Spline Mapping</td>
<td>136</td>
</tr>
<tr>
<td>A.1</td>
<td>Overview</td>
<td>136</td>
</tr>
<tr>
<td>A.2</td>
<td>Program Description</td>
<td>137</td>
</tr>
<tr>
<td>A.2.1</td>
<td>Program Description for drawcolorscl.m</td>
<td>137</td>
</tr>
<tr>
<td>A.2.2</td>
<td>Program Description for tpmp2dsc.m</td>
<td>140</td>
</tr>
<tr>
<td>A.3</td>
<td>Program Listing</td>
<td>142</td>
</tr>
<tr>
<td>A.3.1</td>
<td>Program Listing for drawcolorscl.m</td>
<td>142</td>
</tr>
<tr>
<td>A.3.2</td>
<td>Program Listing for tpmp2dsc.m</td>
<td>150</td>
</tr>
<tr>
<td>B</td>
<td>Program Description and Listing for 3-d Thin-plate Spline Mapping</td>
<td>155</td>
</tr>
<tr>
<td>B.1</td>
<td>Overview</td>
<td>155</td>
</tr>
<tr>
<td>B.2</td>
<td>Program Description</td>
<td>157</td>
</tr>
<tr>
<td>B.2.1</td>
<td>Program Description for Newtracetongue.m</td>
<td>157</td>
</tr>
<tr>
<td>B.2.2</td>
<td>Program Description for New2Colorcut.m</td>
<td>159</td>
</tr>
<tr>
<td>B.3</td>
<td>Program Listing</td>
<td>162</td>
</tr>
</tbody>
</table>
B.3.1 Program Listing for Newtracetongue.m ............... 162
B.3.2 Program Listing for New2Colorcut.m .................. 170
B.3.3 Program Listing for other programs .................... 176

C  Technical Details for Program Implementation ............... 185

C.1 Computer Environment ........................................... 185
  C.1.1 Getting Started ............................................... 185
  C.1.2 Setting MATLABPATH ........................................ 186
C.2 Program Implementation ........................................ 187
  C.2.1 Programs for 2-d Thin-plate Spline Mapping ........... 187
  C.2.2 Programs for 3-d Thin-plate Spline Mapping ........... 187
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The specifications of the MRI data</td>
</tr>
<tr>
<td>2</td>
<td>Landmarks for 3-d tongues, as shown in Figure 16</td>
</tr>
<tr>
<td>3</td>
<td>Parameters of the alignment mapping for each transverse slice</td>
</tr>
<tr>
<td>4</td>
<td>Coordinates (in pixel) of the control points from Figure 17</td>
</tr>
<tr>
<td>5</td>
<td>Coordinates of the control points for each transverse slice of Miyawaki's data</td>
</tr>
<tr>
<td>6</td>
<td>Translation of the z-coordinate for each transverse slice</td>
</tr>
<tr>
<td>7</td>
<td>Slice-wise correspondence between the Miyawaki data and the Visible Man data</td>
</tr>
<tr>
<td>8</td>
<td>Summary of initially evaluated correspondence in vertical position (slice number) for each landmark pair (identified by landmark number to the left) that were used for mapping shown in Figure 25</td>
</tr>
<tr>
<td>9</td>
<td>Revised coordinates of the landmark pairs for the Visible Man data and the Miyawaki data</td>
</tr>
<tr>
<td>10</td>
<td>The distance from the midsagittal plane to the closest point of the contour of the hyoglossus muscle in transverse slices of the Miyawaki data</td>
</tr>
<tr>
<td>11</td>
<td>Lengths of the transverse slices of the Miyawaki data</td>
</tr>
</tbody>
</table>
Normalized Lengths (L3) of the transverse slices of the Miyawaki data
# List of Figures

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Display examples of the Visible Man data on three cutting planes. The top-left panel is a display on the sagittal (S) plane (viewed from the right side; the bottom-left panel is a display on the transverse (T) plane (viewed from below, the upper half is the right side; the top-right panel is a display on the coronal (C) plane (viewed from the front). Cross lines indicate the cutting locations for the other two displays and the number shows the location in pixel value.</td>
</tr>
<tr>
<td>2</td>
<td>(a) Enlarged oral portion of the MRI transverse section at a level of the third vertebra from the Japanese vowel /a/. (b) Full tongue sketch of transverse section at a comparable level of Figure 2(a).</td>
</tr>
<tr>
<td>3</td>
<td>(a) Bold polygon represents the tongue contour of MRI data in Figure 2(a). (b) Closed polygons (in colors) circumscribe regions of different muscles and line segments for muscle fiber directions in Figure 2(b).</td>
</tr>
<tr>
<td>4</td>
<td>Demonstration of the mapping with 2-d thin-plate splines. (Hand-drawn polygons were used in this example to show the concept.) (a) A typical sketch of a tongue section as in Miyawaki's tongue drawings. (b) Outline of tongue as obtained from MRI. (c) Specification of corresponding landmarks (here 27 landmark pairs for A and B). (d) Panel A mapped onto panel B with thin-plate spline to make all landmarks coincide. (e) - (f) Demonstration of the mapping by using a grid.</td>
</tr>
<tr>
<td>5</td>
<td>(a) 2-d thin-plate spline mapping from tongue sketch polygons onto MRI tongue contour with 5 landmark points; and (b) with 7 landmark points.</td>
</tr>
</tbody>
</table>
6 The left panel: the new scanned-in transverse slice number 9 of Miyawaki’s data. The right panel: the extracted polygon file of the left panel (V: Verticalis; T: Transversus; SL: Superior Longitudinalis; GG: Genioglossus muscle). .......................................................... 62

7 The three types of landmarks. Type 1: Juxtapositions of tissues. Type 2: Maxima of curvature. Type 3: Extremal points given a selected coordinate system. (adaptation from Figure 3.3.1 in Bookstein, [10]) .................................................................................................. 67

8 The left panel: landmark #1 is identified by a filled (red) circle at mid right portion of the tongue (V: Verticalis; T: Transversus; SL: Superior Longitudinalis; GG: Genioglossus; HG: Hyoglossus; SG: Styloglossus muscle). The downward protrusion of the lower (left) half of the tongue was caused by a missing tooth. The right panel: the transverse slice shown is number 9. The filled circle at the upper left extreme is the corresponding landmark #1 in Miyawaki’s data. 73

9 The vertical stripes are the projected boundaries of the styloglossus muscle and the horizontal stripes are the projected boundaries of the hyoglossus muscle. The bold polynomial curves are the fitted boundaries of the styloglossus (green) and the hyoglossus (blue) muscles. The filled (red) circle is the landmark #2, the SG-HG point. The anterior of the tongue is on the right-hand side and the coordinate system is in pixel number. ...................................................... 75

10 The transverse slice shown here is number 15. The filled circle is the corresponding landmark #2 in Miyawaki’s data. ........................ 76

11 The left panel: landmark #3 is shown in filled (red) circle (right of IL)(V: Verticalis; T: Transversus; IL: Inferior Longitudinalis; GG: Genioglossus; HG: Hyoglossus; SG: Styloglossus muscle). The right panel: the transverse slice shown here is number 12. The filled circle is the corresponding landmark #3 in Miyawaki’s data. ............... 76
12 The left panel: landmark #4 is shown in filled (red) circle (to the left of IL) (V: Verticalis; T: Transversus; IL: Inferior Longitudinalis; GG: Genioglossus; HG: Hyoglossus; SG: Stylloglossus muscle). The right panel: the transverse slice shown here is number 11. The filled circle is the corresponding landmark #4 in Miyawaki's data.  

13 The left panel: landmark #5 is shown in filled (red) circle (above V & T) (V: Verticalis; T: Transversus; SL: Superior Longitudinalis; GG: Genioglossus; HG: Hyoglossus; SG: Stylloglossus muscle). The protrusion of the lower (left) half of the tongue was caused by a missing tooth. The right panel: the transverse slice shown here is number 7. The filled circle is the corresponding landmark #5 in Miyawaki's data.  

14 The left panel: landmarks #6 and #7 are shown in filled (red) circles (to the right of SL & right of V & T) (V: Verticalis; T: Transversus; SL: Superior Longitudinalis muscle). The right panel: the transverse slice shown here is number 5. The filled circles are the corresponding landmarks #6 and #7 in Miyawaki's data.  

15 The left panel: landmark #8 is shown in filled (red) circle (at SL) (V: Verticalis; T: Transversus; SL: Superior Longitudinalis; GG: Genioglossus muscle). The right panel: the transverse slice shown here is number 3. The filled circle on the midsagittal (upper) edge is the corresponding landmark #8 in Miyawaki's data.  

16 The left panel: a 3-d display of all landmarks open (red) circles identified from the Visible Man data. The right panel: the corresponding landmarks are displayed in cross (green) marks for Miyawaki's data. The landmark numbers are given in both panel to show their correspondences.  

17 The outline of a tongue was viewed from the side. The parallel line shows where the transverse section was made. The number on the right side is the slice number (adaptation from Figure 5-i in Miyawaki, [106]).
The definition of the control points based on hypothesis (a). They are shown in filled (red) circles identified by numbers 1 and 2. Midsag. Pl. stands for the midsagittal plane.

The definition of the control points based on hypothesis (b). (cf. Figure 18.)

Graphic illustration of the alignment transformation.

The result of the alignment mapping for the polygon files of the transverse slices of Miyawaki's data (lateral view). The filled (red) circles indicate the locations of the control points. The numbers indicate the landmark numbers for Miyawaki's data.

The result of the 3-d thin-plate spline mapping for the polygon files of the transverse slices of Miyawaki's data (lateral view). The filled (red) circles indicate the locations of the control points and closed points with numbers indicate the landmark numbers for the Visible Man data.

The left panel: the mapped result of transverse slice number 4 of Miyawaki's data. The right panel: the original contour of the same Miyawaki slice superimposed on the left panel image after 3-d thin-plate spline mapping (V: Verticalis; T: Transversus; SL: Superior Longitudinalis muscle).

The left panel: the scanned-in transverse slice number 4 of Miyawaki's data. The right panel: the corresponding slice in the Visible Man data to the left panel (V: Verticalis; T: Transversus; SL: Superior Longitudinalis muscle).

The left panel: the mapped result of transverse slice number 17 of Miyawaki's data. The right panel: the contour of transverse slice number 17 of Miyawaki's data superimposed on the left panel image (V: Verticalis; T: Transversus; GG: Genioglossus; HG: Hyoglossus; MY: Mylohyoid muscle).
The left panel: the scanned-in transverse slice number 17 of Miyawaki's data. The right panel: the corresponding slice in the Visible Man data to the left panel (V: Verticalis; T: Transversus; GG: Genioglossus; HG: Hyoglossus; MY: Mylohyoid muscle). ........................................ 100

The transverse slice shown here is number 16. The filled circle is the new landmark #2 in Miyawaki's data (T: Transversus; GG: Genioglossus; HG: Hyoglossus; SG: Styloglossus muscle). ..................... 102

The left panel: another Visible Man slice corresponding to transverse slice number 4 of Miyawaki's data. The right panel: A similar Visible Man slice corresponding to transverse slice number 17 of Miyawaki's data (V: Verticalis; T: Transversus; GG: Genioglossus; HG: Hyoglossus; MY: Mylohyoid; SL: superior longitudinalis muscle). 103

The mapped result of transverse slice number 9 of Miyawaki's data (V: Verticalis; T: Transversus; GG: Genioglossus muscle). ..................... 104

The mapped result of transverse slice number 12 of Miyawaki's data (V: Verticalis; T: Transversus; GG: Genioglossus; IL: Inferior Longitudinalis; MY: Mylohyoid; HG: Hyoglossus; SG: Styloglossus muscle). 105

The left panel: top view of the hyoglossus muscle under a uniform scaling. The right panel: top view of the hyoglossus muscle under a scaling in x direction only. The slice # 10-18 (as shown) goes from top to bottom along the vertical dimension. .................... 111

A display produced by the program drawcolorsc.m. ..................... 138

A display produced by the program tpmp2dsc.m. ............................ 141

A display produced by the program Newtracetongue.m. ........................ 157

A display produced by the program New2Colorcut.m. .......................... 159
CHAPTER I

Introduction

1.1 Background and Motivation

This dissertation is concerned with a general computational method for evaluating quantitative relations among the shapes and internal muscular structures of the tongue observed by various methods.

1.1.1 Organization of Speech Utterances and Theoretical Phonetic Motivation of this Work

The current work provides a substantive methodological preparation for future research, in the form of computational tools and example data, to investigate the phonetic characteristics of various articulations. A typical application is comparison of different vowel gestures by the same tongue (such as the same speaker) and by different tongues (i.e. different speakers) in stationary (quasistatic) articulatory positions. The present dissertation constitutes a significant step toward establishing a systematic and quantitative approach for anatomically accurate speaker normalization of articulatory data as well as integration of different sources of information (i.e. preserved cadaver and live articulating tongues). The computational method is also applicable to describing other biomedical structures,
particularly those comprising highly deformable organs such as the heart. Phonetic and most other muscular studies require understanding non-stationary dynamics. The method explored here forms a basis for dealing with dynamic simulations in articulatory research of speech.

This investigation attempts to contribute to our understanding of the communicative functions of the tongue, and to clarify those irrelevant factors observed in the data that obscure communicative functions. Clarification of what is relevant from a particular functional point of view (in this case verbal communication) about a given biological system (in this case the tongue) is the broad goal of this dissertation. In particular, a general model of tongue shape representation is a crucial element of a recently proposed phonetic theory, called the C/D model (see below). In this model, much of the seemingly complicated diversity of phonetic phenomena corresponding to the same phonological unit (in our case a vowel) is assumed to be ascribed to the complex physical nature of the articulatory organ (the tongue) itself, as opposed to the complexity of the linguistic control that uses the organ. From this point of view, it is crucial to be able to compute, quantitatively and meaningfully, how a tongue shape in one situation is related to the corresponding but quite different shape of the same biological organ in a different phonetic situation.

Speech signals are traditionally viewed as a concrete manifestation of a linear string of phonemic segments (consonants and vowels). Phonemes are abstract functional units by the choice of which, in a given language, meaningfully different
words are distinguished from each other. In the early 1950's, Jakobson, Fant & Halle, ([68], 3rd ed.) proposed a further analysis of these segments, and described each phoneme as a simultaneous bundle of distinctive features, which constitute the minimal units of speech sounds for contrasting functions. In their theory, phonemes as simultaneous bundles of distinctive features were assumed to be integral speech segments with physical (i.e. phonetic) properties, which may be modified according to certain simple principles when used in a specific phonetic environment (context), and may manifest themselves differently depending on the particular language.

Based on the assumption of distinctive features as the minimal ingredients, Chomsky & Halle ([19]) published an elaborate description of generative phonology of English, in which applications of a linearly ordered set of rewrite rules alter a linearly arranged string of segments to produce observed phonetic forms. Later, in the 1970's, students of generative phonology revised this linear theory of phonology by assuming multiple dimensions, called tiers or planes, to function more or less independently from each other in the process of phonological derivation. To account for the phonological derivation, at the same time, they questioned the validity of having a completely ordered set of rules (see e.g. Halle & Vergnaud, [53]; Clements & Keyser, [21]; Goldsmith, [45]; Halle & Idsardi, [52]; Idsardi, [67]). The new framework resulting from these discussions is called nonlinear phonology. The nonlinearity means that the rules are not linearly ordered (see Halle & Vergnaud, [53]), but it also means that the representation is multi-dimensional and does not strictly form a linear string of segments.
One important and common characteristic of the current nonlinear phonology is underspecification, the lack of complete specifications of distinctive feature values for each segment, which would uniquely identify the phonetic property of the sound. There are different ways of skipping and restoring (from a segmental point of view) redundant feature values which are necessary for actual physical forms of speech, and some properties may remain unspecified up to the final form of the phonological representation to be handed over to a phonetic implementation process for the given language. This concept of underspecification is important for this dissertation work, because it is related to the expectation that certain aspects of the tongue gestures are more variable than others depending on the situation. Such a nonuniform distribution of the functional load, from an information-theoretical point of view, makes it necessary to pay particular attention to selected factors (most probably related to the use of particular muscles as opposed to others from a physiological point of view) that determine the observed physical shape of the given organ when it is used in the given language.

Neither the phoneme nor the distinctive feature can be concretely uttered by itself. The building blocks of uttered speech are generally identified as syllables, despite some difficulty in clarifying where the boundaries are among contiguous syllables in the phonetic signals (see Fujimura, [39]). Prevailing nonlinear phonological theories use syllables as organizational units, technically by linking syllable nodes to individual segmental feature trees representing (only partially identified) phonemic segments (for the first treatment of this kind in generative phonology, see
Kahn, [71]). In other words, syllables play a critical role in the linguistic description of speech signals, but are not the theoretical primitives in most phonological descriptions; they are treated as "prosodic" units formed out of the organization of features, where phonemic segments (or roots) serve as the basic units for forming a multi-dimensional time series, elements of which underlie phonetic events. In addition, there is a certain deviation from the classical identification of phonemic segments (roots) as the unit of concatenation, considering functional independence of subsegments (see e.g. Steriade, [141]).

Despite such basic changes over the past two decades in phonological theory, the process of converting the phonological representations to speech signals has not been revisited comprehensively, apart from few exceptions (particularly Browman & Goldstein's articulatory phonology which claims to eliminate the distinction between phonology and phonetics (see further discussion in Section 2.5.2, below) and Fujimura's C/D model (see Section 2.5.3)). A common characteristic of these two theories is that both are based on multi-dimensional organization of articulatory events. Most speech engineers currently use the classical concept of phoneme concatenation and smoothing (called coarticulation) as the basic principle of phonetic implementation, either explicitly or implicitly, in speech synthesis as well as in automatic recognition (see Klatt, [83]). Speech pathologists and experts in language acquisition tend to prefer this classical picture of speech organization as their basis of conceptualizing the mental processes and analyzing observed phonetic phenomena, while many of them are aware of its inadequacy.
1.1.2 Articulatory Movements and their Control

In order to investigate the open question about the segmental structure of speech, the process of converting abstract phonological structure to concrete speech signals has to be understood as exactly as possible. In particular, the process that relates the speech organization principle to the articulatory processes must be quantitatively described. Therefore, observations of articulatory movements have often been taken into consideration together with acoustic recordings for systematically acquiring utterance data, to interpret the complex relations between the acoustic signal properties and the articulatory process. Direct observations of articulatory movements have been made possible by recent technical progress, such as computer-controlled X-ray microbeam system (Fujimura et al., [42]; Nadler et al., [110]), ultrasound and magnetic resonance imaging (MRI) (Watkin and Rubin, [148]; Stone, [142]; Kaburagi and Honda, [69]; Baer et al., [3], [2]), magnetic methods (Sonoda and Kiritani, [136]; Schönle et al., [132]; Perkell et al., [127]), and Tagging Snapshot MRI techniques (Kumada et al., [89], [88]; Niitsu et al., [112]). A review of various techniques can be found in Fujimura ([39]).

These techniques make it possible, to a limited extent, to provide a basis for understanding the speech organization patterns with recorded articulatory data. Fragments of information obtained through different means of measurements can provide useful data for a comprehensive simulation. For example, vocal tract configurations were obtained by extracting vocal tract cross-sectional areas from series of 2-d slices of MR images for sustained vowel production (Baer et al., [2]). These
vocal tract configurations were then used to study the acoustic characteristics of vowels. The transformation from a vocal tract configuration to acoustic output is computed through the area function of an articulatory model (Mermelstein, [105]; Maeda, [97]; Carré and Mrayati, [18]).

The purpose of the computational simulation is to develop a model of the real system so that we can manipulate model parameters to observe consequences in the output signal. A model should be able to explain observed data from a real system (i.e. articulatory movements) and also may predict critical variables that can be evaluated systematically and extensively against observable data, thus suggesting direction for future research. An articulatory model uses the physiological parameters, such as the tongue-tip position in the midsagittal plane and lip opening, to control the vocal tract configuration. Although physiological parameters were used as control signals in early articulatory modeling (e.g. Coker and Fujimura, [23]), an effectively accurate simulation of speech production processes via vocal tract shapes is still difficult to obtain. For this reason, the dependence on data only in the midsagittal plane persists in the research literature, and the vocal tract area function is derived from the midsagittal dimensions in prevailing methods, including Browman & Goldstein ([14]).

Such existing methods for mapping from the vocal tract area functions to acoustic signals (via formants) do not correctly reflect the multi-dimensional control of the articulatory system and resulting dynamic behaviors. In order to understand the basic principle of speech production, a physiologically-based articulatory
modeling is necessary. The consideration of 3-d structure, as opposed to the commonly adopted two-dimensional (2-d) description pertaining to midsagittal configurations of articulatory organs, is critical for understanding many basic properties of speech sounds. It is easy to see that the 3-d consideration is crucial for the lateral consonants (as in 'lap'), which can be similar to stop consonants (as in 'tap') as far as the midsagittal vocal tract shape is concerned. One crucial difference between the lateral and stop consonants is that the sides of the tongue blade are left open for the lateral consonants, allowing the area function in this area to remain nonzero. The lateral openings not only characterize this vocal tract shape in three dimensions for the laterals, distinctly from stops, but also make the corresponding acoustic output adequate for /l/ with additional pole-zero pairs as well as differences in formant frequencies. These differences will not be obtained if only the 2-d midsagittal vocal tract is considered directly.

There are many other cases including vowels, for which the essence of the articulatory strategy is totally missed unless the 3-d characteristics of the articulatory gesture is considered. For example, the formation of the vowel /i/ is achieved largely by the combination of the anterior and posterior genioglossus (gg) muscles. The anterior gg (apparently counteracting the elevation of the tongue blade in the midsagittal image of the tongue shape by pulling it down) is necessary for the narrow groove near the midsagittal plane to be maintained without collapsing, as in the stop articulation, while the sides of this groove are elevated by the contraction of the posterior gg.
Three-dimensional (3-d) tongue models have been implemented using the finite element (FE) method (Zienkiewics, [153]; Hughes, [66]) by several workers (Kiritani et al., [80]; Fujimura, [37]; Hashimoto & Suga, [58]) (see also Perkell, [123] for an earlier attempt using a similar computational method), and have been used for explaining some basic properties of the relation between articulatory control of vowels and acoustic properties (formant patterns) (Fujimura & Kakita, [34]; Kakita et al., [72]).

A limitation of all previous 3-d studies, however, is that the models only dealt with static gestures without accounting for the effects of inertia. The traditional cineradiographic studies of the tongue movement in speech (see e.g. Houde, [65]; Perkell, [121]), optical tracking of mandible and lip movements (see e.g. Kelso & Tuller, [73]; Munhall, [109]; Ostry & Munhall, [118]; Krakow, [84]; Cohn, [22]), and even earlier motion picture studies of lip movements (Fujimura, [35]) have revealed peculiar characteristics of articulatory movement patterns that could not be accounted for by ignoring dynamics of the articulators. For example, Fujimura ([35]) revealed a damped oscillation of the lower lip following the explosion from the stop closure, which cannot be there unless the moving system has both acceleration-related inertia and position-related restoring force. More data of rapid movements in speech articulation are being collected and studied by the microbeam (Fujimura et al., [42]; Kiritani et al., [78]; Nadler et al., [110]) and new magnetic methods (Perkell et al., [127]), in order to elucidate the phonetically crucial nature of speech dynamics (see Fujimura, [38], [39]).
Existing 3-d tongue models also lack anatomical accuracy. They adopt oversimplified mathematical methods in using linear infinitesimal deformation theories, and ignore detailed properties of tongue muscles. Thus, an additional limitation of previous studies using the finite-element methods and other similar computational simulations was that, due to the linear approximation adopted for simplicity of the mathematical algorithms, realistic movement patterns in large deformation of the tongue as a massive and complex muscular body have not been accounted for. Virtually all these limitations of the computational simulation studies of tongue movements are being removed to a large extent in the method proposed by Wilhelms-Tricarico. The basic computational algorithm and associated tools including graphics, data structure, etc. have been provided (Wilhelms-Tricarico, [149]), and have been improved and implemented in part in the current dissertation work.

In this course of study, the most immediately needed information is the anatomical and physiological data, which provide accurate factual basis for the simulation work. The main objective of this dissertation work is to develop a computational scheme to obtain reliable and detailed anatomical data which can be used for the new dynamic 3-d tongue model, by combining the existing insufficient MRI and sectional anatomical data. It also prepares for future studies by providing tools for processing more detailed and massive MRI data, based on emerging improvement of MRI methods. By supplying realistic anatomical data for the tongue, based on static tongue shapes and knowledge of muscular structures of the tongue, an
effective dynamic model of the tongue can be constructed with sufficient accuracy, according to Wihelms-Tricarico's computational simulation method. Dynamic movements of such a realistic model can be evaluated by comparing computed time functions of sample flesh points on such a model with actual time functions of corresponding flesh points on the tongue as observed by the microbeam or magnetic methods discussed above. This seems to be the only feasible quantitative and empirical approach for understanding the complex articulatory process used in realistic speech production.

1.2 Scope of this Work

The general task of this dissertation work is to develop a computational method for integrating different 3-d anatomical data of the human tongue for their use in a 3-d tongue model. This work introduces a strategy for defining a set of landmark points in reference to a general 3-d tongue structure representation. Landmark-based morphometrics has been used for the human skulls and brain tissues in studies such as shape averaging (Cutting et al., [24]), variability (Bookstein, [11]), and matching one shape to another by transformation (Bookstein, [10]). Landmarks define locations that have anatomical significance and have identifiable geometric coordinates. They constitute common reference points for mathematical mapping between any pair of data sets and define the anatomical correspondence or homology between the data set. The integration of different data sets of the human tongue is based on this correspondence, and it is achieved by a method of
nonuniform geometrical transformation called thin-plate spline mapping. At the same time, in this dissertation, the thin-plate spline mapping is used to evaluate the sufficiency and robustness of the landmark sample set for comparing 3-d tongue shapes.

Our specific tasks are:

(1) to develop guidelines for defining a set of landmark points between two different anatomical data sets of the human tongue, such as detailed anatomical drawings using histological dying technique from one preserved cadaver sample, published atlas images from other deep-frozen samples, or MRI head scans of living subjects using different articulatory gestures of vowels; and

(2) to integrate different data sets of the tongues supplementing tongue shape data with tissue boundary information of the internal organization of the tongue. The latter information is obtained from cadaver data pertaining to internal properties such as fiber directions of muscles, which is largely missing in observations of more natural tongue shapes. The integration is performed by thin-plate spline mapping. The adequacy of the selected landmarks is tested by comparing the original image data before the mapping with the image obtained after another similar mapping back to the original image.

Software tools have been developed to implement tasks (1) and (2). The result of this dissertation work using different sources of information ranging from preserved cadaver to live tongues in articulations demonstrates the feasibility of this new approach.
1.3 Significance

Finding landmarks for the tongue, in particular within the tongue body, has interested anatomists (Negulesco, personal communication) in the understanding of specimen-to-specimen variability of the tongue as well as for the production of averaged or normalized images of the tongue. Anatomists and pathologists are interested in finding differences between groups of data sets (i.e., MRI or dissectional pictures) and correlating these differences to age, disease, or treatment. With a highly deformable nature and individual-to-individual variability of the tongue, the selection of robust landmarks between different data sets in the case of the tongue is rather difficult to objectivize and quantify, but this issue is critical for providing a solid basis for application of MRI and other recording techniques in the tongue pathology and speech science research.

Considerable individual anatomical differences for phonetic functions have been noted by several researchers (see, for example, Perkell ([122], [124], [126]); Kuehn & Moll ([87]); also for related issues of articulatory gesture variation see Ladefoged et al., ([90])). Various instrumental methods have been used for tongue movement observation in speech production (Fujimura, [39]). Nevertheless, data collected from these methods are always incomplete due to their inherent constraints with respect to 3-d and dynamic aspects (see review in Section 2.2).

The simulation of tongue deformation with a model adapted to individual anatomy provides well-defined quantitative means to compare different subjects and different articulations based on anatomical structure. It will also provide a
theoretical basis for characterizing subject-to-subject variation of formant patterns via the quantitative parametric description of subject anatomy, thereby normalizing individual phonetic data for subject-to-subject comparison in order to identify differences in terms of phonetic control. The data will be available for use by researchers in phonetics in general, speech science and pathology, child development and applied linguistics as well as speech technology, regardless of the particular theory of phonetics and phonology.

A very concrete application of the data is for the computational model of phonetic implementation, called Converter-Distributor (C/D) model. This model was proposed by Fujimura ([40], [41]). It has four computational modules that are organized in a serial order, namely, the converter, the distributor, a parallel set of actuators, and the signal generator (see Section 2.5.3, below). The signal generator is a physical model that computes movements of the tongue and other articulatory systems of the vocal tract, such as the lips and the mandible, with the multi-dimensionally organized impulse response functions as its input control signals. The computational method, for which this dissertation work provides the method and tools for obtaining design data, is the core of this component of the C/D model. The computational algorithm of the tongue model has been studied, designed, and demonstrated recently in a simplified implementation by Wilhelms-Tricarico ([149]).
1.4 Organization of this Dissertation

This dissertation is organized as follows.

Chapter II reviews previous works on data recording techniques, tongue modeling, thin-plate spline mapping, and phonetic and phonological theory.

Chapter III describes the details of the implementation method used in processing the anatomical drawings and MRI data, and mathematical principles used in the thin-plate spline mapping.

Chapter IV provides guidelines about how the landmarks are selected and shows the selected landmark pairs for example pairs of data.

Chapter V describes the computation method to integrate different sets of 3-d anatomical data of the human tongue and a method to evaluate how good the selection of landmarks is, using a sample pair of data sets from anatomical drawings and dissectional images.

Chapter VI discusses the results of the dissertation work and its possible applications in other research, and gives pointers to future research.
CHAPTER II

Literature Review

This chapter gives a brief review of previous background studies. They are in the fields of oral anatomy, methods and tools for observing tongue movement, measuring and recording the speech production process, and modeling of the tongue shape variation. The application of the anatomical mapping methods, in particular the thin-plate spline mapping, to the description of tongue shape variation is described. In addition, the linguistic relevance of this dissertation work in terms of existing theories of speech organization is discussed.

2.1 Anatomical Data

The tongue has been recognized to be the most important organ for the speech production process. Nevertheless, not as many detailed anatomical studies of the tongue have been reported, as they have been of other speech organs such as the mandible. This is largely due to the difficulty of determining an accurate descriptive framework in relation to variably-shaped anatomical components, which have interwoven muscles without solid bony frames. An early comprehensive study on the morphology of the human tongue was conducted by Abd-el-Malek ([28]) in the
late 1930's. Although the dissecting methods were not clearly described, his obser-
servations covered detailed descriptions of the connective tissue and musculature as well as the nerves and blood vessels of the tongue. However, the description of the complexity of the tongue structure is limited to a few anatomical drawings and seven good microphotographs of selected sections of the tongue.

More recently, Miyawaki's study ([106]) shows drawings of three different excised tongue halves (two from a cadaver of a 39-yr. old Japanese female and one from a cadaver of a 29-yr. old Japanese male), microtomed in slices (2.5-3 mm in thickness) of three roughly perpendicular planes. In each drawing, a projection of tongue muscle fiber directions was marked, giving information for estimating the 3-d fiber directions for most locations in the tongue. Transparent serial sections were prepared with formaline for a period of approximately four years, dehydrated with increasing concentration of alcohol for about two weeks, and then dyed with xyrol to make specimens brownish and muscular fibers transparent. Miyawaki's work reveals muscle tissue directions with unusual detail. The drawings are utilized as a guide in the anatomical interpretation of the current dissertation work.

Due to recent progress in imaging techniques, computed tomography (CT) (Larsson et al., [91]), ultrasound (Shawker et al., [133]) and MR imagings (Lufkin et al., [96]; McKenna et al., [101]) also have been used extensively to study the normal and abnormal soft tissue anatomy of the tongue and the floor of the mouth. Nevertheless, these studies focused more on anatomy and pathology of other types of tissue, such as glands and blood vessels, rather than on the muscle fiber
directions which are of primary interest in this study. Muscle fiber directions are extremely difficult, if even possible, to determine from any of these modern observation techniques.

### 2.2 Instrumental Methods for Tongue Movement Observation

In addition to the X-ray cineradiography (Öhman and Stevens, [116]; Houde, [65]; Perkell, [121]; Kent & Moll, [76]) and the microbeam system (Fujimura et al., [42]; Kiritani et al., [78]; Nadler et al., [110]), several other methods have been used for observing the shapes and movements of the tongue in speech articulation (Fujimura, [39]), such as magnetic measurements (Sonoda and Kiritani, [136]; Schöule et al., [132]; Perkell et al., [127]), ultrasound and magnetic resonance imaging (Stone et al., [143]; Stone, [142]; Baer et al., [3], [2]; Dang et al., [25]), and electrical methods (Fujimura et al., [44]; Kiritani et al., [79]; Hardcastle et al., [55]).

#### 2.2.1 X-ray Cineradiography

For a long time, X-rays were the only source of direct information about tongue movements (Fang, [29]; Perkell, [121]; Wood, [151]). The only supplemental information available was from optical recording (high-speed motion picture) (Fujimura, [35]) and dynamic electro-palatography, i.e. electrical recordings of lingo-palatal contact patterns. Research based on X-ray film recording provided the primary information for the understanding of the physics and physiology of articulation.
However, repeated exposure to X-ray radiation for the acquisition of articulatory data is a major health risk. Furthermore, the frame-by-frame study of the X-ray film is time-consuming, and the poor boundary resolution between tissue and air is an additional problem. The use of metal pellets fixed on selected flesh points on the tongue surface in the midsagittal plane by speech production researchers such as Houde ([65]) and Kent & Moll ([76]) improved the accuracy of position measurements. The basic problem that remained unsolved, both in terms of radiation dose and optical noise due to scatter photons, stems partly from the simultaneous X-ray exposure covering the entire image field. The use of computer tomography (CT) scanning (Kiritani et al., [82]) solves the problem of boundary contour to a large extent with substantially reduced optical noise and 3-d image convolution, but the high radiation dose and the long exposure duration make this modern method inadequate for most articulatory studies.

2.2.2 X-ray Microbeam System

The X-ray microbeam system uses a deflectable thin beam of X-ray under the control of a digital computer. X-rays are produced at a selected point of an X-ray generating target that is hit by a deflectable electron beam. A portion of the X-rays comes out from a fixed pinhole as a thin deflectable X-ray beam. This X-ray microbeam passes through the object, a subject's head, and the penetrating photons are measured by a Na-I scintillation counter which serves as an X-ray photon detector. The counter output is fed into the computer which controls the beam deflection. The computer then identifies pellet positions in real time through
a local spatial search for the shadow of each pellet sequentially in time. Since the
number of pellet trackings per second is limited, slowly moving pellets (such as head
fixed reference pellets) are usually tracked at a lower sampling rate than pellets
on fast articulators (such as the tongue tip) to optimize the overall performance.
This nonuniform sampling rate also saves the total dose. The microbeam emission
can be stopped by an overdeflection control. When this control is enabled, the
electron beam is deflected far out of the effective area so that the X-ray radiation
through the pinhole does not take place. A reset pulse disables this control. This
mechanism is provided to restrict the exposure under strict control (the over­
deflection is usually not used in practice except at the end of each exposure due to
the limitation of the total power of the currently operating system).

A pilot system with 50-kV electron acceleration voltage was implemented at
the University of Tokyo in the late 1960's (Fujimura et al., [42]). A second version
of the system with a 2-mA electron beam that is accelerated by a voltage of 150 kV
was actually used for articulatory studies since 1974 (see Kiritani et al., [78]). The
implementation of the system at the University of Wisconsin (Nadler et al., [110])
is the third generation. It uses a new X-ray generation scheme of a transmission
(as opposed to the usual reflection) type with a nominal 600-kV maximum electron
acceleration voltage and a nominal 5-mA maximum electron beam current. The
use of high energy photons makes it possible to include experiments with subjects
with dental fillings, and also to reduce the radiation dose required, since photon
interaction with the biological tissue is reduced. While the radiation hazards are
minimized by the use of the microbeam, performing comparable tasks without using ionizing photons at all would be ideal. Furthermore, the cost to implement and maintain such a system is another issue of concern. Instrumentation associated with the Wisconsin system also allows for simultaneous detection/transduction and processing of the following speech production variables: (1) the speech acoustic signal, (2) an accelerometer/throat microphone, (3) multi-channel electromyographs.

2.2.3 Magnetic Methods

In principle, magnetic methods are safer than the X-ray microbeam system. Sonoda ([138]) introduced a new application of the magnetometer to measure the tongue movement in speech production. A small permanent magnet is used as a pellet to be attached on a selected point on the tongue surface. Its position is detected by externally located field detecting coils (Sonoda and Ogata, [139], [114]). The voltage that is induced in the sensing coil is inversely proportional roughly to the cube of the distance between the pellet and the sensing coil. The long axes of the magnetic dipole of the permanent magnet and sensing coil are aligned in parallel. This system has the advantage of not using wires for transmitting electrical signals from the pellets. However, it cannot track more than one sample point simultaneously.

The use of an externally created alternating magnetic field picked up by small detectors in the mouth makes simultaneous tracking of multiple points possible (Perkell et al., [127]; Schönle et al., [132]). In the case of Perkell and Cohen's method (called EMMA system), each detector is a 4 × 4 × 2.5 mm coil (coils of
$2 \times 2 \times 4 \, \text{mm}$ are now being used) wrapped around a ferrite core with a pair of thin wires for external connections. The detector is glued onto the tongue surface. The necessity of attached wires connected to the external measurement system is a disadvantage. Furthermore, careful control of the sensing coil placement and calibrations are required to obtain a position resolution and accuracy between 0.5 and 1.0 mm. Slight deviation from the midsagittal plane or a slight disorientation of the axis of the coil causes measurement inaccuracy. In these respects, the microbeam system is distinctly superior. However, the magnetic method does have the distinct advantage of not being affected by small metal objects in the mouth such as dental fillings and caps. Further, it does not use any ionizing rays.

There is still another experimental magnetic articulatory measurement system designed by Panagos and Strube ([119], [120]) of University of Göttingen, Germany. Unlike previous methods, they use uniform magnetic fields and two specially designed detectors, a magnetic potentiometer (MPM) and a flat flexible coil (FC), to measure jaw and tongue movements during speech production. The MPM (a long thin flexible coil) is used to measure the distance between the upper and lower jaws in the midsagittal plane by fixing its both ends to the upper and lower incisors. The FC is a coil with one or two rectangular windings that is embedded between two flexible plastic sheets. The placement of the FC could be on any part of the tongue surface in the midsagittal plane (the distance from the front edge of the FC to the tongue tip is between 5 and 50 mm). The FC is used to measure the tongue movements (the distance between two short edges of the FC). The disadvantage of
this system is that it is insensitive to parallel displacements of the detecting coils. Thus, the system is used to measure the vectorial distance of the jaw and tongue movements.

Currently two commercial systems for studies of articulatory movements are available. One is the Articulograph AG-100 by Carstens Medizinelektronik GmbH, Göttingen, Germany; the other is the Movetrack system by Botronic AB, Hägersten, Sweden. The former is a commercial product of the Schönle system. The latter uses two transmitters and single-axis transducers without conducting misalignment correction. The modular and analog hardware design of the Movetrack system is simpler than those of the EMMA and the Articulograph AG-100 systems, but its application to articulatory studies has not been reported.

2.2.4 Imaging Techniques

Ultrasound Imaging

Even though the image acquisition time of MRI and X-ray CT techniques could be shortened by technical improvements, the sampling rate would be still too low to obtain the precise shape of the moving vocal tract during speech (Baer et al., [3]; Kiritani et al., [82]). However, continuously rather than discretely sampled surface contour information about articulatory organs, such as the tongue and the larynx, can be obtained using ultrasonic techniques which do not employ pellets for recording position of the organ. With an ultrasound transducer placed under subject’s chin, the ultrasonic measurements allow coronal ultrasound imaging to provide angled cross-sectional images of the tongue during speech (Stone et al.,
A transducer holder has been used to provide an accurate alignment for the ultrasound transducer and to track the movement of the mandible. This type of scan cannot, however, provide information about anterior-posterior tongue movement (Stone et al., ibid.). The jaw and hyoid bone appear as shadows of conical shapes because they refract the sound away from the straight path (Stone, [142]). In addition, the ultrasound system is unable to image the palate because the air in between does not allow transmission of ultrasonic waves. Another limitation is the inability to trace the tongue tip because of the air beneath the tongue tip. To provide the information about anterior-posterior tongue movement, Stone's group (Stone et al., [144], [145]) used b-splines to approximate curves between the points representing the detected tongue surface contours, and bicubic Bézier patches to synthesize these 2-d curves into a 3-d anterior-posterior tongue surface from a series of ultrasound slices. Their approach is to reconstruct a 3-d surface over a grid of splines in order to examine differences between vowels and consonants.

Another approach of ultrasound imaging was proposed by Kaburagi and Honda ([69]). They combined ultrasonic pulsed-echo imaging techniques with pulsed-transmission techniques to visualize lateral tongue shape for monitoring tongue motion. With an ultrasonic sensor placed on the tongue surface, the position of the sensor was tracked to monitor tongue motion. The position of the sensor was determined by its direction relative to the pulsed-transmission transducer (called "sensor angle") and its distance from the transducer in that direction. The sensor angle was determined by finding the ultrasonic pulse emitting angle with the
maximum output signal from the sensor; and the distance from the sensor to the transducer was estimated from the tongue shape on the ultrasonic image. They evaluated the measurement accuracy by investigating the effects of the sensor-transducer distance, tilt angle and lateral displacement of the sensor on the measurement error of the sensor angle. Their results showed that the measurement accuracy is similar to the angle resolution of the pulsed-transmission transducer (1.55 degrees) if the tilt angle of the sensor is less than 20 degrees, and the lateral displacement of the sensor is less than 2.0 mm. Their method provides another approach to track tongue motion during the production of continuous speech utterance.

**Magnetic Resonance Imaging**

The use of MRI techniques has been popular for visualizing the soft-tissue of human body for diagnostic purposes because of its noninvasiveness. It has been applied to anatomical and pathological studies of the tongue and oropharynx (see Section 2.1). Baer et al. ([3], [2]) at Haskins Laboratories and Moore ([108]) at University of Pittsburgh first applied the MRI technique to the analysis of speech production. Vocal tract shapes were obtained by extracting vocal tract areas from series of 2-d slices of MR images for sustained vowel production. These vocal tract shapes were then used to study the acoustic characteristics of vowels. Dang et al. ([25]) at ATR, Japan, used MRI to investigate the effects of nasal cavities and paranasal sinuses to nasal sounds. Combined with acoustic analysis, their data showed that the nasal cavities and paranasal sinuses play an important role in the production
of nasal sounds. Narayanan et al. ([111]) at University of California, Los Angeles, used MRI to study fricative consonants with morphological analyses of the vocal tract and tongue shapes. In addition, custom software for edge detection and 3-d reconstruction for MR images were developed by Matsumura and his colleague ([99], [100]) at Osaka Electro-Communication University, Japan, to measure the area function for vowel, nasal, and fricative sounds. In their studies, a special dental crown plate was used to make the air space and the contact region between tongue tip and the area behind the upper incisors visible during the production of the fricative sound /s/. Recently, a new technique, which is called Tagging Snapshot MRI, was applied to determine the contraction of tongue muscles during the vowel production (Kumada et al., [89], [88]; Niitsu et al., [112], at Tsukuba University and University of Tokyo, Japan). Although the MRI technique has been proven a useful tool for speech production research, dynamic observation using MRI techniques is still infeasible due to the limitation in speed of image acquisition.

2.2.5 Electrical Methods

Electropalatography (EPG)

Computerized palatography makes the traditionally static palatography capable of recording palato-lingual contact patterns as a function of time. Contact signals from the electrodes, which are embedded in the artificial palate, are processed in the EPG unit. The digitized contact signals are then stored in a computer. Tongue-palate contact patterns can be monitored on the screen display in real
time. Since the introduction of electropalatography in the 1970's (Fujii et al., [33]), groups at the University of Tokyo (Fujimura et al., [44]; Miyawaki et al., [107]; Sawashima and Kiritani, [131]) and Edingburg University (Hardcastle et al., [55], [54]) worked independently on studying the palate-tongue contact patterns for different consonant sounds. The group at the University of Tokyo called it dynamic palatography. In addition to its use for speech research, the real time aspect of the display of electropalatographic measurements can be quite useful in speech training for speech disorder therapy. Opto-electrical sensing for measuring tongue dynamics also has been tried (Sonoda et al., [137]; Chuang and Wang, [20]).

**Electromyography (EMG)**

Electromyography is a technique that allows researchers to make physiological measurements of muscle activity via electrodes placed either directly in the muscle or indirectly on the skin surface close to the muscle. Since the introduction of hooked-wire electrodes (Hirano and Ohala, [61]) in intrinsic laryngeal muscles for EMG studies, EMG has become a uniquely informative method for investigating speech control via direct physiological measurements. It has been utilized to investigate the contributions of mandible (Oshima and Gracco, [117]), tongue muscles (Harris et al., [57]; Baer et al., [1]), and laryngeal muscles (Simada et al., [134]; Kiritani et al., [77]; Hirose and Gay, [63]; Hirose, [62]) to speech production and to interpret the motor control mechanisms of these muscles. Currently, the use of surface electrodes is being investigated to replace the invasive electrodes for tongue
muscle measurements (Honda, personal communication). However, combinations with other methods for physical observation are often desirable.

2.2.6 Remarks on Useful Recording Techniques

Efforts to combine two observing schemes, such as EMG with X-ray microbeam system (Kiritani et al., [81]), ultrasound imaging with X-ray microbeam system (Stone, [142]), and ultrasound imaging with magnetometer (Kaburagi and Honda, [70]), have been tested to obtain a better method for observing tongue movements. Based on the data that are collected from the above mentioned methods, computer simulation of the complex articulatory system for studying temporal characteristics of speech is being explored (Wilhelms-Tricarico, [149]). The current dissertation work constitutes a basic contribution to such future research.

We are interested in a computational method for evaluating quantitative relations among the shapes and internal muscular structures of the tongue observed from various methods. In our computational method, we need to select a set of landmarks (see Section 1.2). To do this, we need information not only on the surface of the tongue but also within the tongue. Most of the observing techniques reviewed above provide little information within the tongue except CT scanning and MR imaging. Between them, the MRI data show better contrast among soft tissues such as muscles. Also, CT requires a very large dose of ionizing radiation, which makes it impractical for studies with live subjects. However, MR images, the insufficiency in anatomical details (e.g. boundaries between intrinsic muscles,
discussed in Section 6.2.2) remains in our example MRI data acquired from currently available MRI technique and poses difficulties in identifying landmarks.

2.3 Tongue Modeling

The complex articulatory system for speech production comprises the tongue, the lips, the mandible, the larynx and related structures for voice production, the pharynx, and the velum. Many of these organs involve a large number of intrinsic and extrinsic muscles. These controlling muscles generally have independent innervations. Combinations of muscular activities result in a large variety of possible movements. The central component of the articulatory system is the tongue which is the most complex and probably least understood component of the speech apparatus. Since the control signals for this complex system are difficult to obtain, the only viable methodology for understanding the speech production process is to simulate the system and relate the putative control principle to observable articulatory or acoustic variables.

Early studies of tongue modeling originated from the research of articulatory synthesizers (Coker and Fujimura, [23]; Henke, [59]; Mermelstein, [105]). Physiological variables such as tongue body position, tongue-tip position, and lip opening were used as control signals to create continuous variation of the vocal tract configuration in a midsagittal plane from a sequence of target positions that are inherent to discrete phonemic units.
There are several physiologically-oriented models of the tongue using finite element methods (Kiritani et al., [80]; Hashimoto and Suga, [58]; Hirai et al., [60]). Among them, the study by Hirai et al., which is the most comprehensive model to consider interactions among different speech organs, is restricted to the two dimensional midsagittal description. An earlier, ambitious model by Perkell (Perkell, [123]) did not use the finite element method and attempted a physiologically faithful complex nonlinear simulation. The others are 3-d models but describe only static and infinitesimal deformation based on linear elasticity. A dynamic 3-d model for large deformation was recently achieved by Wilhelms-Tricarico ([149]), using a new mathematical method for continuum mechanics and a non-linear, heterotropic tissue description.

2.3.1 2-d Model

Tongue shape description in the articulatory synthesizers mentioned above has been used to account for the coarticulation and vowel reduction effects, using simple low-pass filtering of the parameters (with preselected time constants of the articulatory apparatus) and some ad hoc rules for the mapping from phonetic input data to vocal tract area functions. A critical issue about these models is to find a good parametric representation of the cross-sectional area function of the vocal tract that is in agreement with the acoustic measurements. The "articulatory parameters" in these models are essentially acoustic parameters for specifying the vocal tract area functions rather than articulatory variables considering
motor control factors. In addition, the inertia in dynamic phenomena is not explicitly represented in the model.

Perkell's dissertation ([123]) was the first attempt to construct a physiological model of the human tongue based on the anatomy of intrinsic and extrinsic tongue muscles. In his work the midsagittal plane was modeled as a lumped-parameter lumped-force system. Hirai et al. ([60], [64]) used a finite element method (30 triangular elements) for 2-d tongue deformation as part of their physiological model of speech organs. In their work, articulatory gestures are determined by computing static equilibrium of muscle forces on all of the components pertaining to the tongue, the jaw, the hyoid bone, and the laryngeal cartilages. Recently, a 2-d dynamic finite element tongue model (48 isoparametric elements, 63 nodes) was introduced by Yohan et al. ([152]) to test a control model for speech production. However, their simulation is still limited to linear deformations.

2.3.2 3-d Model

Kiritani et al. ([80]) originated a 3-d computational model of the tongue (86 tetrahedral elements and 33 nodes) that is based on a finite element simulation and treats the tongue as an elastic continuum. Under the assumptions of uniform strain within each element, linear isotropic tissue property, and volume incompressibility of the tongue body, an incremental load method was applied to simulate the static characteristics of the tongue under the constraints of surrounding hard structures. Using a similar method, Kakita et al. ([72]) demonstrated similarities between measured EMG activities in a number of extrinsic and intrinsic lingual muscles
and the computed amount of contraction force for each muscle that is required for static vowel articulation.

Hashimoto and Suga ([58]) used a large number of elements (492 tetrahedral elements and 170 nodes), and approximated midsagittal tongue shapes from X-ray data by least-square methods to find optimal muscle contraction parameters. With the assumptions that there is a linear stress-strain relationship and the strains are small, they can solve a linear matrix equation to obtain an equilibrium configuration of their tongue model after specifying the tensions in the muscles of the model. Hashimoto and Suga also showed that their model could be matched to the measured midsagittal tongue contours, inferring the muscle contraction patterns as the cause of the observed tongue shape. However, the assumption that the stress-strain relation is linear severely limits the applicability of this approach to even static articulatory phenomena.

2.3.3 Dynamic 3-d Model

Wilhelms-Tricarico ([149]) implemented a preliminary 3-d dynamic model of the tongue with 42 cubic elements and eight muscle types for testing his finite-element algorithm, which treats large stress and strain in tissue and takes inertia into account. In order to improve the accuracy of this model for quantitative studies of tongue gestures (see Fujimura, [39], [40]), incorporation of more detailed anatomical knowledge is necessary. Wilhelms-Tricarico arrived at several conclusions about future research strategies by designing and working with the preliminary model:
1) The anatomy of the tongue calls for considerable refinement of his preliminary model. Roughly 150-200 elements may be required to obtain sufficient accuracy in representing the different muscle's cross-sectional areas.

2) For his preliminary model the simplifying assumption was made that within each element of the reference model, the fiber directions of muscles do not vary with respect to their boundaries. Thus, within one finite element in its reference configuration, all fibers of one muscle type are parallel. (During deformation, this generally does not have to be the constraint since the fiber direction varies according to the local deformation in the element.) To obtain a more accurate representation of the muscles, the fiber directions in the elements of the reference model must be specified and not assumed to be uniform for an element. This is necessary because each element in the tongue often has two or several muscle fiber directions; *e.g.* the transversalis and verticalis muscles are often interwoven.

3) The preliminary model has constant passive properties and does not distinguish different tissue types. The real tongue consists of tissue with varying characteristics, such as muscles versus superficial membrane and other passive tissues (tendon, etc.). With more spatial refinement (using more elements), varying tissue properties in different portions of the tongue must be specified. In particular, the mucosa (skin) of the tongue has a distinct elastic property and presumably plays an important role in articulatory gestures.
Part of the goals of this dissertation work is to deal with the first two concerns and thereby make the task of the third feasible.

2.4 Anatomical Mapping Methods

Image registration¹ (or matching) is often necessary when two images obtained by different modalities of medical imaging techniques (e.g. CT and MRI) are integrated to provide complementary information for clinical diagnosis. It requires a mapping function that transforms the coordinates of each point in one image into those of the corresponding point in the other. A mapping function in the context of this chapter is a spatial transformation that maps a set of points from a biological configuration into corresponding points in its warped representation. This section describes general mapping functions developed for digital warping applications in medical imagery and for anatomical atlases. The mapping methods can be affine, perspective, and polynomial transformations. Since affine and perspective transformations can be expressed as first-order polynomials (see Wolberg, [150]), they are thus subsets of polynomial transformations. A more complete review of the mapping functions is available in Schumaker's ([95]) and Barnhill's ([4]) papers. Because of the wide scope of this subject, the discussion below concentrates on the polynomial mapping methods and the thin-plate spline mapping method. In Schumaker's classification, the polynomial mapping methods belong to the global approximation (i.e. approximately fitting the data) category, while the thin-plate

¹Image registration here means finding the correspondence that defines a mapping function between two images.
spline mapping is a global interpolation method (i.e. interpolating the data and keeping references points exactly coincident). In addition, the former uses the least-square technique which is different from the latter’s surface-fitting approach.

2.4.1 Polynomial Mapping Methods

Affine transformations preserve parallel lines and equi-spaced points during a mapping from a source plane (or image) onto a target plane (or image). They are a combination of translation, rotation, and scaling transformations. Affine transformations are mainly used in the field of computer graphics for planar mappings (see Foley et al., [30]), given that coefficients of the transformation matrix are often known in these applications. For example, an affine transformation can map an input triangle into an arbitrary triangle at the output and an input rectangle into an quadrilateral parallelogram. However, they cannot map a rectangle into an arbitrary quadrilateral because only three points are required to specify an affine transformation. To map a rectangle to an arbitrary quadrilateral requires more complex transformation such as perspective and bilinear transformations. Perspective transformations retain parallel lines only when they are parallel to the projection plane whereas bilinear transformations preserve horizontal and vertical lines in the source plane (or image). Both transformations are often used in the fields of remote sensing (Bernstein, [5]) and medical imaging registration (Green, [49]) for image distortion correction. An application of a polynomial mapping
was shown by Wilhelms-Tricarico ([149]). He used trivariate Bézier polynomial mapping method in the design of an initial geometry of his preliminary tongue model.

An example of polynomial transformation was introduced to medical imaging by Singh et al. ([135]). They registered 2-d myocardial perfusion images with first-order polynomials which are usually sufficient for their problems. A similar approach was used by Pickens and Price ([128]) by applying a quadratic polynomial transformation to match intravenous digital subtraction angiographic images (IVDSA). In addition, Maguire et al. ([98]) used Singh's method for registration of images from different medical imaging modalities. A higher order polynomial transformation has been tried by Kenny et al. ([74]) for image registration of lung images in nuclear medicine. However, polynomial transformations do not account well for high-frequency deformations because the behavior of a higher order polynomial is less predictable. Therefore, local deformations (such as tissue shrinking and 3-d images taken from different viewpoints) are often corrected by piecewise polynomial transformations (Goshtasby, [46], [47]) and surface spline interpolations, also known as the thin-plate spline method, (to be discussed in next section).

Although the surface spline interpolation is slow in computation, it provides better accuracy than other polynomial transformations (Franke, [32]; Goshtasby, [48]). Since accurate mapping is our major concern, the thin-plate spline mapping was chosen as the transformation method in our study. A more complete review of various image registration techniques for general and medical imaging systems
can be found in Brown ([16]) and Van den Elsen et al. ([26]). Both papers provide classification of registration techniques and a framework to aid in the selection of a registration method for a specific problem.

2.4.2 Thin-plate Spline Mapping Method

The essence of the thin-plate spline mapping is to fit appropriate surfaces to arbitrarily spaced data points. It is not only one of the most efficient methods for global interpolation of scattered data (see Franke, [32]), but also one of the few exactly understood mathematical methods (see Duchon, [27] and also Meinguet, [103], [104]). The thin-plate spline mapping can be expressed as the solution of the biharmonic equation ($\Delta^2 f = 0$, where $f$ is the deflection function of a plate due to point loads) over an infinite plate with the constraint that the small "deflection" of the plate results in bending only. An analytical solution of the equation and its application for the interpolation of aircraft wing deflections were shown by Harder and Desmarais ([56]).

The thin-plate spline mapping technique, which maps a set of points onto their homologous points in a differently deformed configuration of the tongue, is used in this study. Homologous points are landmarks that define locations having anatomical significance as well as geometric coordinates. They constitute common reference points for the thin-plate spline mapping between any paired data sets, and define the anatomical correspondence or homology between the data sets. This means that the thin-plate spline mapping method maps exactly on homologous points between the data sets. The basic mathematical method
was originally explored by Duchon ([27]) and Meinguet ([104]) in the 1970's for applications in interpolation theory. Grimson ([50]) first applied it to the visual surface reconstruction problem; Terzopoulos ([146]) introduced it to the literature of computer graphics, and Bookstein ([7], [10], [8], [9]) brought it to the attention of the medical imaging community as a suitable general-purpose warping tool. This method has been applied by Evans et al. ([17]) to register patient MR brain images to a reference MR atlas in three dimensions for quantitative neuroanatomical and functional analyses, and by Arad et al. ([113]) to warp 2-d face images of various expressions in a face normalization process. Recently the method was extended by Bookstein and Green ([12]) to include direction mappings.

The thin-plate spline mapping has been used to serve as a tool not only to produce an averaged image (atlas) of a moving or deforming anatomical structure but also to understand specimen-by-specimen variability around the atlas. Bookstein ([11]) showed an example of this application with thirteen landmarks from a set of midsagittal head MR images of nine normal medical students. Furthermore, a 3-d thin-plate spline mapping was used for averaging 3-d curves and surfaces from a set of CT and MR scans to visualize “the skull beneath the skin” in three dimensions (Cutting et al., [24]). The purpose of their study was to find a normal range of skull images. The normalized skull image provides a useful reference for craniofacial surgical treatment plans in many syndromes.
2.5 Linguistic Relevance: Existing Theories/Models of Temporal Organization of Speech

While current phonological theory is committed to a nonsegmental formalization of representations and processes, a segmental view of feature organization largely persists in speech science and technology as well as psycholinguistics. Since speech production theories lack quantitative models of the temporal organization of the inherently multidimensional articulatory process under strong influences of variable prosodic conditions, two new theories of articulatory organization (cf. Browman and Goldstein, [13], [14]; Fujimura, [40], [41]) have emerged to fill this gap. The following presents an overview for these theories/models of temporal organization of speech. For expositional purposes, first, the classical model of segmental organization is outlined. Understanding temporal organization principles constitutes the basis of describing speech phenomena. In order to represent temporal changes of articulatory gestures, we need to be able to map the tongue configuration at one time to that at another. Thus, a descriptive method for the tongue shape change in terms of quantitative mapping provides the basis for any phonetic description of speech phenomena.

2.5.1 Phoneme Concatenation-Coarticulation Model

According to the classical phoneme concatenation-coarticulation model, speech is organized as a string of words, and each word is represented by the concatenation of a string of phonemic segments. The fusion of a string of phonemic segments, consonants and vowels alike, into a connected speech is governed by a
suprasegmental smoothing process called coarticulation (see Lindblom, [94]). Phonetic facts that deviate from this model of speech phenomena, e.g. allophonic variation of the consonant /l/ , are commonly treated with phonological rewrite rules called allophonic rules (see, for example, Halle and Mohanan, [51]; see also Sproat and Fujimura, [140] for an argument against such phonological treatments and for a new approach based on a more powerful phonetic implementation). A review of different models of segmental organization in speech from the classical point of view can be found in papers of Kent ([75]) and Fowler ([31]).

A numerical model of coarticulation proposed early in the 1960’s by Öhman ([115]) treated vowel and consonant gestures independently and superimposed the two types of gestures to account for the coarticulation properties of Swedish dental stops in vowel-consonant-vowel contexts. His approach represented a basic deviation from segmental concatenation models, anticipating the later development of nonlinear phonological theories. Even though a larger unit of segment is adopted later, such as demisyllable (Fujimura, [36]) or syllable (Fujimura and Lovins, [43]), in place of most allophonic rules, a more comprehensive theory for accounting for speech organization and the effects of prosodic conditions of utterances thereon remained to be proposed. Two new frameworks that respond to this need are described below.
2.5.2 Articulatory Phonology

In articulatory phonology, Browman and Goldstein ([14]) hypothesized

"the basic units of phonological contrasts are gestures, which are also abstract characterizations of articulatory events, each with an intrinsic time or duration ... utterances are modeled as organized patterns (constellations) of gestures, in which gestural units may overlap in time."

(p. 155)

They attempt to describe the lexical items in terms of the gestures and their interrelation. They further state

"the patterns of overlapping organization of the gestures and their interrelations can be used to specify important aspects of the phonological structure of particular languages and to account for a variety of different types of variations such as the allophonic variation and fluent speech alternation as well as coarticulation and speech errors."

(p. 155)

A "gestural score", which represents the gestural units and their organizations over time for an utterance, is generated assembling similar specifications of individual lexical items (Browman and Goldstein, ibid.), and this control pattern is used as input to the task-dynamic model (Saltzman and Munhall, [130]). The principle of organization within the lexical representations and in phonological phrases is not explicitly specified so far.
The task-dynamic model describes movement in terms of a task to be performed, using dynamical equation for the implementation of the task. It serves as the controller of an articulatory synthesizer to provide the articulatory movement in an abstract way. A gesture in the gestural score consists of a formation or a release of a constriction in the vocal tract, for example, in lip closure or opening, separately, which is specified in terms of a "tract variable" coordinate system. The dynamics of gestures is modeled by a critically damped linear second order system to characterize the articulatory movement. The transformation of the tract variables into articulator variables is complicated in that several articulators may contribute to a single tract variable (e.g. the upper & lower lips and the jaw are involved in lip protrusion) and that one articulator may be involved in several tract variables (e.g. the lips are involved in defining lip aperture and protrusion tract variables). Detailed information of these transformations is given in Saltzman ([129]) and Saltzman and Munhall ([130]).

Articulatory phonology and the task-dynamic model attempt a unique and quantitative approach to convert the phonological representations to speech signals. It has been used to interpret articulatory data, e.g. from an experiment on the production of the schwa vowel (Browman and Goldstein, [15]). Other investigators adopted this basic framework (see e.g. Kröger, [85], [86]). Browman and Goldstein reported that the best control strategy for speech synthesis was consistent with the hypotheses proposed by articulatory phonology. Perkell ([125]), however, argued
He further argued that the task-dynamic model leaves the actual aerodynamical and biomechanical properties of speech production in the dark. Nevertheless, the task-dynamic model offers an alternative way to investigate experimental data at its current stage and offers interesting and plausible explanations of otherwise an ad hoc set of phenomena such as allophonic variation, segment deletion and insertion, when the prosodic conditions of an utterance are varied.

### 2.5.3 Converter/Distributor Model (C/D Model)

The C/D model proposed by Fujimura ([40], [41]) provides an alternative theory of phonetic implementation, using the syllable rather than phonemic segments, as the "concatenative unit" of speech signals. The C/D model has four computational modules that are organized in a serial order, namely, the converter, the distributor, a parallel set of actuators, and the signal generator.

The process of model implementation is described as follows: (1) The converter converts a prosodic representation of an utterance into a series of syllable/boundary pulses. At its input level, the stress pattern is displayed as a tree structure as described in metrical phonology (see Liberman, [92]), augmented by numerical specifications of utterance-dependent prominence attachments and utterance parameters such as speed and style. The converter computes the magnitudes, and thereby time values, of the pulses. (2) The distributor interprets the feature
specifications for each syllable in terms of selections of elemental gestures, and passes the relevant information, including the magnitude and time value of each syllable or boundary pulse to relevant actuators. In this computation, feature specifications for the utterance are mapped into a set of articulatory gestures, such as apical stop, bilabial nasal, etc. An elemental gesture is a combination of an articulator and an action, and is implemented by the actuator. (3) Each actuator evokes an elemental gesture, which is stored as an impulse response time function to be excited by its pertinent syllable pulse. The output of each actuator is a superpositional sum, as a time function, of individual impulse response functions in each articulatory dimension that represents a component of the physiological activity for the utterance. (4) The signal generator is a physical model that computes the movement of the tongue and other articulatory systems of the vocal tract, such as the lips and the mandible, with the set of control time functions delivered by the actuators as its input signals. The tongue model, for which this dissertation work provides a necessary design method, is the central part of this component of the C/D model.

In summary, the C/D model is a comprehensive theory of phonetics that does not use the traditional concept of segments, but organizes articulatory gestures by a general principle based on feature specifications for syllables. Syllables in this theory are prosodic primitives in the sense that phonetic organization of the phonological substance for any utterance unit is completely specified through determining magnitudes of individual syllables and boundary pulses which form a
linear string. Like articulatory phonology, it deviates from segment-based speech organization theories and it also assembles articulatory gestures to organize speech signals. Unlike articulatory phonology, the C/D model does have a distinction between phonology and phonetics, and recognizes a “segmental” unit (i.e. syllable) for concatenation to phonologically represent lexical entries (as opposed to the gesture score of articulatory phonology which represents lexical items directly in terms of articulatory gestures).
CHAPTER III

Methods

3.1 Introduction

This chapter describes a new method for integrating different anatomical data sets of the human tongue. The integration of these data sets is accomplished by a two- or three-dimensional thin-plate spline mapping method. The method makes use of recently obtained MRI data in combination with a published description of dissected human tongues (Miyawaki, [106]) as well as dissection images of a male cadaver from the Visible Human Project of National Library of Medicine (henceforth, "Visible Man data"). Starting from a description of different data sets that are used in this study, this chapter then describes image analyses of these data and gives an overview of mathematical descriptions of the thin-plate spline mapping, as a means to relate different types of image data to each other. Finally, an example of 2-d thin-plate spline mapping and procedures for a 3-d integration process are presented to conclude this chapter.
3.2 Data Description

A description of Miyawaki’s data is given in Section 2.1. The MRI and Visible Man data are described in this section.

3.2.1 Magnetic Resonance Image (MRI) Data

There are several sets of MRI slices of human heads in vowel articulations at ATR Human Information Processing Research Laboratories in Japan. The data were obtained by Kiyoshi Honda, M.D. and his associates, from several native speakers of Japanese and American English holding stationary articulations of several vowels. The data were acquired by a Shimadzu Magnetic Resonance Imaging machine, SMT-100GUX, installed at the Takanohara Central Hospital. The SMT-100GUX system has a static magnetic field flux density of 1.0 Tesla. The images were recorded, using a spin echo pulse sequence with the relaxation time (TR) of 300 ms for sagittal scanning and 800 ms for transverse and coronal scannings. The excitation time (TE) was 18 ms for all scannings.

The data comprise 7 sagittal slices, 3 mm apart, and 26 transverse and 24 coronal slices, 5 mm apart, for each vowel articulation with no interscan gap. The spatial resolution of the 2-d image is about 1mm × 1mm. Each image slice was in the format of 256 × 256 pixels, covering an area of 25 cm × 25 cm.

From the MRI data, two stacks of transversal sections of the oropharyngeal region were selected for this study. Table 1 shows the specifications of the MRI data.
Table 1: The specifications of the MRI data

<table>
<thead>
<tr>
<th>Axial Imaging Parameters (SMT-100 GUX, 1.0 Static T)</th>
<th>Subject KH (Native Japanese with Tokyo dialect)</th>
<th>Subject DO (Native English from Canada)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR (ms)</td>
<td>800</td>
<td>800</td>
</tr>
<tr>
<td>TE (ms)</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Images</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>Thickness</td>
<td>0.5 cm</td>
<td>0.5 cm</td>
</tr>
<tr>
<td>Interscan skip</td>
<td>0.0 cm</td>
<td>0.0 cm</td>
</tr>
<tr>
<td>Field of View</td>
<td>25 cm</td>
<td>25 cm</td>
</tr>
<tr>
<td>Matrix Size</td>
<td>$256 \times 256$</td>
<td>$256 \times 256$</td>
</tr>
<tr>
<td>Scan Time</td>
<td>4:05 min</td>
<td>4:05 min</td>
</tr>
<tr>
<td>Vowel</td>
<td>Japanese /a/</td>
<td>English /i:/</td>
</tr>
</tbody>
</table>

The data used in this study were obtained from two different speakers. During the MRI recording, the two speakers articulated a stationary vowel, Japanese /a/ in one case, and English /i:/ in the other. From the data, areas of a size of $110 \times 110$ pixels surrounding the tongue were extracted. For a large smoothed display, the samples were interpolated so that a stack of MRI images consisting of $256 \times 256$ pixel images with a pixel size of $0.42$ mm $\times$ $0.42$ mm could be used for further processing. Section 3.3 gives details of this process.

3.2.2 Visible Man Data

The Visible Human Project for the male cadaver was conducted by Spitzer and Whitlock (1991) at the University of Colorado. The Visible Man data sets consist of transverse MRI, CT, and cryosection (dissection of deep frozen cadaver) images. Each transverse anatomical image is a $2048 \times 1216$ pixel image with 24 bits of
color for each pixel. There are 1871 transverse images for cryosection mode at 1 mm intervals. They are directly comparable with CT axial (transverse) images which have a resolution of $512 \times 512$ pixels with a 12-bit grey scale for each pixel. Axial MRI images of the head and neck were obtained at 4 mm intervals. The MRI image has a $256 \times 256$ resolution with a 12-bit grey scale for each pixel.

The Visible Man data used for this dissertation work are from "The Visual Man CD-ROM", a CD-ROM version of cryosection images derived from the original data by Data Express Inc., Sunrise, Florida. The CD-ROM data have a resolution of $1024 \times 608$ pixels with 8 bits of color for each pixel. The pixel spacing in the "XY" (transverse) plane is doubled to 0.66 mm, and the spacing in the "Z" (vertical) direction is the same as the original data. The CD-ROM data were down-loaded and cropped out (using a public domain program X-view) to produce a $150 \times 150$ pixels oral cavity portion of the image slice number 160 through 240 (81 slices). These cropped data were saved as graphics interchange format (GIF\textsuperscript{1}) files in our Sun-4 computer. Examples of the visible man data on three cutting planes are shown in Figure 1. Sagittal and coronal sections were synthesized from the compiled stack of transverse sections.

3.3 Image Analysis

In order to make use of MRI data effectively, the general anatomy of the oral cavity and vocal tract must be correctly understood. We consulted with the anatomical

\textsuperscript{1}GIF and Graphics Interchange Format are registered trademarks of Compuserve Inc.
Figure 1: Display examples of the Visible Man data on three cutting planes. The top-left panel is a display on the sagittal (S) plane (viewed from the right side; the bottom-left panel is a display on the transverse (T) plane (viewed from below, the upper half is the right side; the top-right panel is a display on the coronal (C) plane (viewed from the front). Cross lines indicate the cutting locations for the other two displays and the number shows the location in pixel value.
drawings of cross sections of the human tongue (Miyawaki, *ibid.*) and anatomy atlas (McMinn, [102]). Miyawaki's study contains drawings of three different excised tongue halves, microtomed in slices (2.5-3 mm in thickness) of three roughly perpendicular planes. In each drawing, the projection of tongue muscle fiber directions were marked, giving information for estimating the fiber directions for most locations in the tongue in three dimensions. The unusually detailed drawings are particularly valuable for our study as a guide in this anatomical interpretation, since the MRI data do not reveal details of muscle tissue directions.

To facilitate a better visualization of the data on the computer screen, the MR images were digitally enlarged by a linear factor of 2.2 to a compatible size with the same image resolution (72 digits per inch (dpi)) as Miyawaki's sketches. This process was implemented with the Adobe Photoshop software on a Macintosh personal computer. Figure 2(a) shows an example of transverse slices after the processes of grabbing and enlargement from the Japanese subject data.

Using the same software and an image scanner, Epson 6000, Miyawaki's tongue sketches were scanned and stored into computer in a raw data image format with a resolution of 72 digits per inch. The original tongue sketches show only either the left or right half of a tongue. In order to achieve a better visualization for subsequent work, a mirror image was created to produce laterally symmetric pictures for the coronal (front view) as well as the transverse (horizontal) set of tongue sketches (assuming the usually observed approximate lateral symmetry of tongues). From the sets of symmetrized sketches, the transverse set of tongue
Figure 2: (a) Enlarged oral portion of the MRI transverse section at a level of the third vertebra from the Japanese vowel /a/. (b) Full tongue sketch of transverse section at a comparable level of Figure 2(a).

sketches was mapped into the transverse set of MRI slices of selected vowels. The resulting full size tongue sketches are exemplified in Figure 2(b).

To reduce the amount of image data, polygon data were generated manually by mouse input with visual feedback, (see Figure 3). As far as possible, the polygon data circumscribe regions that can be associated with anatomical structures. In many cases it was not easy to retain the tissue-air boundary between the surface of the tongue and the hard and soft palates. The teeth are usually not visible but their position can be estimated because they appear as "weak holes" in the MRI.

The result of the image analysis is a 3-d representation of the tongue that makes it possible to display the muscle fiber directions in 2-d projections at any location in the tongue, when such fiber information is obtained from dissection data such as Miyawaki’s data (see below).
Figure 3: (a) Bold polygon represents the tongue contour of MRI data in Figure 2(a). (b) Closed polygons (in colors) circumscribe regions of different muscles and line segments for muscle fiber directions in Figure 2(b).

3.4 Mathematical Methods of Anatomical Description

The following sections present the steps to integrate the MRI data and anatomical drawings in two dimensions and the method to construct a structural model of the human tongue with the thin-plate spline mapping method. We shall start with an overview of the algebra of the thin-plate spline mapping.

3.4.1 The Thin-plate Spline Mapping Technique

The 2-d thin-plate spline mapping not only maps two sets of corresponding points \([(x_i, y_i), (X_i, Y_i)], \ i = 1, \ldots n \) exactly onto each other, it is also a unique mapping in the sense that it minimizes the bending energy defined as the integral shown in this section. In this dissertation, the set of points with coordinates \(x_i, y_i\) refer to the \(i - th\) landmark of Miyawaki’s data and the set with coordinates \(X_i, Y_i\) to that of the MRI data. We want to determine two smooth thin-plate (surface)
splines \( f_x(x, y) (= X) \) and \( f_y(x, y) (= Y) \) that specify the two coordinate values of the mapped point in the \((X, Y)\) 2-d space. The thin-plate spline describes an infinite plate that is deformed under some point load (force) at points \((x_i, y_i)\). Its analytical equations (Equations 3.1 and 3.4) are given in Harder and Desmarais ([56]) as separate \( X - \) and \( Y - \) projections of a 2-d vector \( f(P) \) (see Equation 3.8).

\[
X = f_x(x, y) = a_1 + a_xx + a_yy + \sum_{i=1}^{n} F_i r_i^2 \ln r_i^2 \quad (i = 1, \ldots n), \tag{3.1}
\]

for the \( X \)-component of the mapping function \( f(x, y) \), where \( n \) is the number of point loads, \( r_i^2 = (x - x_i)^2 + (y - y_i)^2 \). Parameters \( a_1, a_x, a_y, \) and \( F_i, i = 1, \ldots n \) of Equation 3.1 are determined from

\[
\sum_{i=1}^{n} F_i = \sum_{i=1}^{n} x_i F_i = \sum_{i=1}^{n} y_i F_i = 0 \tag{3.2}
\]

and

\[
f_x(x_j, y_j) = a_1 + a_xx_j + a_yy_j + \sum_{i=1}^{n} F_i r_{ij}^2 \ln r_{ij}^2 \quad (j = 1, \ldots n) \tag{3.3}
\]

where \( r_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 \). The coordinate values of any pair of \( n \) load points are represented by \((x_i, y_i)\) and \((x_j, y_j)\).

In a similar way, the \( Y \)-component of mapping function is determined as

\[
Y = f_y(x, y) = b_1 + b_xx + b_yy + \sum_{i=1}^{n} G_i r_i^2 \ln r_i^2. \tag{3.4}
\]

The algebraic treatment of the thin-plate mapping functions that were used in this dissertation was originally given in Bookstein ([7]). It is given below for a complete mathematical description of this method.
Let \( U \) be the function \( U(r) = r^2 \log r \) (in three dimensions, \( U(r) = |r| \)), and let \( P_i = (x_i, y_i), \ i = 1, \ldots, n, \) be \( n \) points in the ordinary Euclidean plane. \( U \) is a fundamental solution of the biharmonic equation, \( \Delta^2 U = 0, \) the differential equation relating bending deflections to loads of a plate in the following way. \( \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \) represents the Laplacian operator. Denoting \( U(r_{ij}) \) by \( U_{ij}, \) where \( r_{ij} = |P_i - P_j|, \) let us define the matrices \( K, Q, \) and \( L \) as follows:

\[
K = \begin{pmatrix}
0 & U_{12} & \cdots & U_{1n} \\
U_{21} & 0 & \cdots & U_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
U_{n1} & U_{n2} & \cdots & 0
\end{pmatrix}, \quad (3.5)
\]

\[
Q = \begin{pmatrix}
1 & x_1 & y_1 \\
1 & x_2 & y_2 \\
\vdots & \vdots & \vdots \\
1 & x_n & y_n
\end{pmatrix}, \quad (3.6)
\]

\[
L = \begin{pmatrix}
K & Q \\
Q^T & O
\end{pmatrix}, \quad (n + 3) \times (n + 3), \quad (3.7)
\]

where \( O \) is a \( 3 \times 3 \) matrix of zeros.

The thin-plate splines \( f(P) = [f_x(P), f_y(P)] \) having a 2-d \( h_i \) \( (X_i, Y_i) \) at points \( P_i = (x_i, y_i), i = 1, \ldots, n, \) is the function

\[
f(P) = a_1 + a_x x + a_y y + \sum_{i=1}^n w_i U(|P - P_i|), \quad (3.8)
\]

where the coefficients \( w_1, \ldots, a_x, \) and \( a_y \) are obtained as:

\[
W = (w_1 \ldots w_n \ a_1 \ a_x \ a_y)^T = L^{-1} Y, \quad (n + 3) \times 2, \quad (3.9)
\]

with

\[
Y = (h_1 \ h_2 \ldots h_n \ 0 \ 0 \ 0)^T, \quad (n + 3) \times 2. \quad (3.10)
\]
Note that the \( w \)'s multiply copies of the basis function \( U(r) = r^2 \log r \), evaluated for each disjunct pair of points: \( \{P_i, P_j\} \), with \( i \neq j, \forall i, j \). The coefficients \( a_1, a_x, a_y \) (and \( a_z \) for the three-dimensional case) represent the behavior of the function at infinity.

Then the function \( f(P) \) holds the following properties:

1) \( f(P) = h_i, i = 1, \ldots n, and \)

2) The function \( f \) minimizes the positive definite integral

\[
I_f = \int_{\mathbb{R}^2} \left( |\frac{\partial^2 f}{\partial x^2}|^2 + 2|\frac{\partial^2 f}{\partial x \partial y}|^2 + |\frac{\partial^2 f}{\partial y^2}|^2 \right) dx dy
\]

over the entire surface. This is called the bending energy. In three dimensions, \( I_f \) includes partials with respect to \( z \), an additional term \( a_z z \) appears in \( f(P) \), and \( U(r) = |r| \). The coefficients \( w_i \) and \( a_x, a_y, a_z \) are determined from a given set of landmark pairs.

3) The value of \( I_f \) is proportional to

\[
W^T K W = W^T Y = Y^T L_n^{-1} Y,
\]

where \( L_n^{-1} \) is the \( n \times n \) upper left submatrix of \( L^{-1} \). This integral is zero only when all the components of \( W \) are zero: in that case, the computed splines, \( f(P) = a_1 + a_x x + a_y y \), are flat surfaces.
3.4.2 An Example of 2-d Thin-plate Spline Mapping

An example of the 2-d mapping was implemented to demonstrate its applicability for integrating anatomical drawings with shape information obtained from MRI scans (see Figure 4).

In Figure 4, panels A and B are examples of a typical sketch of a Miyawaki tongue drawing and an outline of tongue obtained from MRI image, respectively. Panel C shows an example of hypothetical landmarks (27 points) for each of panels A and B and their correspondence. With a thin-plate spline mapping function determined from these 27 landmark pairs, a mapped result from panel A onto panel B is shown in panel D as an integrated image. Panel F demonstrates how this mapping function deformed a regular grid as seen in E.

To obtain the set of displays in Figure 4, the following steps are necessary:

(a) Miyawaki's tongue sketches were scanned and stored in computer image data format. Mouse drawn polygons were used to circumscribe and classify regions within the drawings, using simple Matlab procedures (see Appendix A). The polygons were labeled for classification (e.g. with a color code for each muscle type), as shown in Figure 3(b).

(b) Based on the MRI image display, tongue surfaces and hard palate shapes were outlined and the resulting polygon data were stored (see Figure 3(a)).

(c) In the polygon data from steps (a) and (b), several clearly corresponding anatomical landmarks were labeled and stored. These landmarks, to be used as homologous points, as shown in Figure 4(c) (cross marks for anatomical drawing
Figure 4: Demonstration of the mapping with 2-d thin-plate splines. (Handdrawn polygons were used in this example to show the concept.) (a) A typical sketch of a tongue section as in Miyawaki's tongue drawings. (b) Outline of tongue as obtained from MRI. (c) Specification of corresponding landmarks (here 27 landmark pairs for A and B). (d) Panel A mapped onto panel B with thin-plate spline to make all landmarks coincide. (e) - (f) Demonstration of the mapping by using a grid.
and circles for MRI data, with connecting line segments showing correspondence for each pair).

(d) The mapping function (Equation 3.8) maps the homologous point pairs exactly on top of each other and maps other points between the landmarks by interpolation, fulfilling the minimum bending energy constraint (Equation 3.11).

(e) The resulting mapping was checked for reasonability. For example, structures inside the tongue should be mapped inside, etc. If needed, the homologous points have to be repositioned to make the mapping more accurate.

The tasks (a-e) involved data preparation and software programming. Figure 5 shows results of mapping with different number of landmark points between MRI (thin solid contour) and Miyawaki's data (bold solid polygons). In Figure 5, landmark points are labeled as circles (in red color) for the tongue sketch and cross marks (in green color) for the hand-drawn contour of MRI data. Solid line segments (in red color) that connect circles and cross marks indicate the correspondence between the two images in terms of landmarks. Two more landmark pairs that were added in Figure 5(b) obviously improved mapping in the area of interest over Figure 5(a). In general, there is no limit in principle on the number of landmark pairs, but the appropriately smooth mapping is usually obtained with relatively few mapping landmarks requiring reasonable computing time.

The current results of 2-d mapping are based on a subjective selection of compatible slices of MRI data and tongue sketches. The number of MRI slices is different from the number of tongue drawing sketches. Each Miyawaki drawing is
a sample of a particular tongue at a certain spatial sampling interval (2.5-3.0 mm), while each MRI slice represents a different slice of the tongue at a different spatial sampling interval (5 mm for transverse direction). The preliminary results show the feasibility of the method for the 2-d cases assuming meaningful correspondence between a selected slice in one data set to another selected slice in the other data set. A more reliable slice-to-slice correspondence can be obtained only through preparing a 3-d image using a sufficiently densely sampled stack of slices in each data set. Such a method of mapping is demonstrated in this dissertation in the next section using the Visible Man data in place of MRI.
3.4.3 3-d Matching of Anatomical Drawings and Visible Man Data

The computational method for integrating different sets of 3-d anatomical data of the human tongue was performed basically in two stages of transformation. The first stage of transformation is the alignment of each data set in three dimensions. The second stage of transformation takes the aligned set of anatomical drawings and matches it to the Visible Man data with the 3-d thin-plate spline mapping technique. The alignment means that there is a section-to-section coherence for each data set. It consists of a procedure of translation, rotation, and scaling (see Section 5.2).

Anatomical drawings taken from Miyawaki's dissection work provide the intrinsic structure for the tongue. Visible Man data rather than MRI data were used for the main body of this dissertation work to provide the outer shape of the tongue, the palate, and other parts of the articulatory apparatus\(^2\) that are integrated with muscle information. For the mapping between the Visible Man and Miyawaki's data, transverse tongue sketches of the Miyawaki data were scanned and stored into computer for the second time with a resolution of 200 digits per inch, using the Adobe Photoshop software and an HP image scanner. These data were saved in the graphics interchange format in our Sun-4 computer. Examples of this second set of the scanned-in Miyawaki data and its extracted polygon file are

\(^2\)The MRI data which were made available to us were not sufficient in anatomical details for our purpose. The MRI data for our purpose must have sufficient resolution in intrinsic muscles so that landmarks within the tongue can be defined reliably. We are preparing a new experiment of MRI data acquisition, which will be used in subsequent studies.
Figure 6: The left panel: the new scanned-in transverse slice number 9 of Miyawaki’s data. The right panel: the extracted polygon file of the left panel (V: Verticalis; T: Transversus; SL: Superior Longitudinalis; GG: Genioglossus muscle). Shown in Figure 6. In the right panel of Figure 6, filled polygons (color patches) represent regions of different muscles (V: Verticalis; T: Transversus; SL: Superior Longitudinalis; GG: Genioglossus). Bold circles (red) in the right panel of Figure 6 are the control points that were used for the alignment mapping. The fine line segments inside the filled polygons in the right panel stand for the fiber direction for the pertinent muscle.

The Miyawaki and Visible Man data need to be integrated as a single 3-d image to obtain an approximately quantified structure of the human tongue, so it may serve as a basis for biomechanical simulation studies. The integration procedure was split into the following tasks:

(a) The areas of the various muscles and the projected fiber directions of the muscles were manually marked. Miyawaki’s tongue sketches consist of three sets of
hand-drawn tongue specimen slices cut in three different main directions roughly perpendicular to each other. By translating, rotating and nonuniformly deforming the 2-d figures, they were approximately brought to a match for a 3-d display, according to the profile figure in Miyawaki's article. This part of the work resulted in a geometric model of the intrinsic 3-d structure of a human tongue in some state of deformation (of the cadaver samples).

(b) The Visible Man data were aligned already when the data were down-loaded to the computer. The data representation is organized in a voxel format of 150 × 150 × 81.

(c) A strategy was devised for defining a set of landmark point pairs between two different anatomical data sets of the human tongue. The selected landmark pairs constituted common reference points for thin-plate spline mapping between paired data sets and defined the anatomical correspondence or homology between the data sets.

(d) The integration of these two data sets of the human tongue is based on this correspondence (step (c)), and it is achieved by the 3-d thin-plate spline mapping. The thin-plate spline mapping is also used to evaluate the effectiveness of interpolation using the landmark set that was chosen to compare the 3-d tongue images. This was done by comparing the original image data before the mapping with the image obtained by another similar mapping of the deformed shape after mapping back to the original shape.
The next two chapters (Chapter IV, Chapter V) describe the selection of the landmark point pairs between the anatomical drawings and the Visible Man data, the implementation of the 3-d thin-plate spline mapping, and the evaluation of accuracy.
CHAPTER IV

Landmarks for 3-d Tongues

4.1 Introduction

The thin-plate spline mapping, like other point-mapping methods (see Section 2.4), needs to determine a set of corresponding points (known as control points or homologous points) in both image data, in order to define a mapping function that maps these corresponding points onto each other between the images. In this dissertation work, the homologous points are landmarks that define locations having anatomical significance and identifiable geometric coordinates. Finding the homologous points in the paired data sets for tongue images is a difficult task, but it plays a crucial role in the mapping process. How good the mapping depends on how good the landmark points are. In this chapter, we will describe what the features of the landmark points are in general and how they are chosen from both data sets.

4.2 Types of Landmarks

It is desirable that all landmarks are robust and easily recognizable with respect to anatomical structures or their readily computable and physically meaningful
representative points (such as the centers of gravity). The first step to select landmarks from the paired data sets is to find common features which can be identified in both data sets. Typical features used for the selection of landmarks are extrema on a contour, intersection among different anatomical structures, maxima of curvatures, and centroids of closed-boundary regions, etc. These features were further classified by Bookstein ([10]) into three principal types of landmarks that are frequently used in morphometrics of biological organs. They are:

- **Type 1: Juxtapositions of tissues**
  
  A landmark of this type is found in an intersection among different anatomical structures and can be defined as a centroid of a small closed-boundary region, such as branching points of bronchial or arterial systems and the eyes or nuclei of the brain.

- **Type 2: Maxima of curvatures**
  
  Landmarks of the second type may be characterized as maxima of curvatures such as "tips of extrusions" and "valleys of invaginations". For example, tips of the bony processes and centers of the bony fossa where muscular attachment is centered belong to this category.

- **Type 3: Extremal points**
  
  This category consists of extrema on a contour. Included are endpoints of maximum diameters on anatomical structures.

Examples of these three types of landmarks are shown in Figure 7.
Figure 7: The three types of landmarks. Type 1: Juxtapositions of tissues. Type 2: Maxima of curvature. Type 3: Extremal points given a selected coordinate system. (adaptation from Figure 3.3.1 in Bookstein, [10])

The identification of these landmarks can be either manually detected or automatically determined by a matching process using cross-correlation or other feature detection techniques (Goshtasby, [48]). Since it is easier for automatic techniques to compute geometric properties of outlines or curves than to detect automatically defined landmark points, efforts made by Bookstein’s group tend to integrate the thin-plate spline mapping with other automatic algorithms for detecting types 2 and 3 landmarks whenever processing of huge amount of image data is necessary.

In addition to the positions of landmarks, Bookstein ([10]) argued that other properties can be included in the landmark information for 2-d and 3-d data. Information of “tangent direction” and “outline curvature” at each landmark (point) for 2-d data and of “normal direction” and “surface curvature” at every landmark point for 3-d data are examples of additional information to the mapping. Related theoretical extension in this direction and application for averaging 3-d curves
and surfaces of human skull have been explored in his recent publication with his colleagues (Bookstein and Green, [12], [6]; Cutting et al., [24]).

4.3 Landmark Identification for Tongue Image Data

Based on the preceding classification of landmarks, this section introduces a strategy of landmark identification for tongue image data. A list of landmarks and examples of landmark configurations in our data sets (Miyawaki's data vs. MRI data and Miyawaki's data vs. Visible Man data) are given to show how these guidelines were followed to choose landmarks in this dissertation.

4.3.1 Strategy for Landmark Selection

With a highly deformable nature and individual-to-individual variability of the geometry of the tongue, the selection of robust landmarks between different data sets of the tongue is rather difficult to objectivize and quantify. Especially for the currently available data, finding proper landmark points within the tongue body is not an easy task. Although many tissue boundaries inside the tongue can be clearly defined as landmark points on Miyawaki's tongue sketches, the same tissue boundaries are often not well visualized in MRI data. On the other hand, many surrounding oral cavity bony structures such as mandible and teeth can be defined as landmarks in our available MRI data, but Miyawaki's data were based on a tongue specimen that was cut out from its surrounding structures. For this reason, in order to develop a general strategy for future MRI (and other) studies, we used the Visible Man data as a substitute for the MRI data. We propose a
strategy for selecting landmarks for the tongue among different data sets in the form of several guidelines listed below:

- **In order to have a smooth mapping, the selection of landmarks should be spread as widely as possible over an entire target image or object.** This implies that the number of landmarks on the surface of the tongue should be made comparable to the number within the tongue.

- **For type 1 landmarks, locations at which two or three tongue muscles unite are useful.** Examples are intersection of the styloglossus (SG) and the hyoglossus (HG) muscles posterolaterally inside the tongue, and the anterior apex of the lateral septum (LS) where the inferior longitudinalis (IL), the styloglossus, and the hyoglossus muscles meet anterolaterally within the tongue.

- **Landmarks that are located on the anterior- and posterior-most points of transverse slices at top-, mid-, and bottom-cutting level are chosen as surface landmarks of the tongue.** This guideline fits to the definition of the type 3 landmark. Caution should be exercised to avoid sensitivity to incomparable slice levels. This issue can be addressed by the manipulation of the vertical position of the homologous points looking at the global match (see later section).

- **Locations of the main lingual arteries and nerves in the areas of fibrous septa inside the tongue are also candidate landmarks for the tongue if they can be**
identified in the data sets. Caution should be exercised regarding individual-to-individual anatomical variation.

- When bony structures (such as the hyoid bone and the mandible) are available in the data sets, the process (the protrusion of the bone) or fossa (the invagination of the bone) to which tongue muscles are attached can be considered as possible landmarks. This fits to the definition of the type 2 landmark.

- When more landmark points are needed for an accurate mapping, functional definitions of the tongue, such as the tongue tip (the portion nearest to the lower incisors), the blade (the part closest to the upper alveolar ridge), in the phonetic rest position (if applicable, e.g. in MRI images comparing different subjects) may be useful.

We followed these guidelines (except the last) and used anatomy atlases as references (Abd-el-Malek, [28]; Gray's anatomy, [147]; Lillie and Bauer, [93]) to choose landmarks from the Visible Man and Miyawaki data. A set of eight landmark pairs between these two data sets was identified and is listed in Table 2. In Table 2, the name, the sequence number, and the type of the landmarks are given together with a short anatomical description for each landmark.

The first landmark that was selected to set up the correspondence between two data sets is at the anterior-most location of the tongue. It was called the Anterior Tip point. The selection of landmark #2 was suggested by Negulesco (personal communication) due to its robustness with respect to individual-to-individual
variation, and we called it the SG-HG point. Landmarks #3 and #4 were chosen to cover the correspondences between data sets in the anterior and posterior regions inside the tongue blade. They were named the Anterior Apex of lateral septum (LS) and the Posterior Apex of paramedian septum (PS).

The first half of the eight landmark pairs were spread mostly in the lower half of the tongue. To follow the guideline for wide-spread landmarks, the second half of the landmarks were sought in the upper half of the tongue. The structure of the upper half of the tongue consists mostly of the intrinsic muscles such as the vertical and transverse muscles. They are often interdigitated with each other; therefore, only points between the boundary of the superior longitudinalis (SL) and muscle block of the interwoven verticalis and transversus within the midsagittal plane were considered candidates of the second half of the landmarks, even though this involved some subjective decision about the third coordinate (height) of the corresponding landmark pairs between the two data sets. These points were chosen as the first three landmarks of the second half of landmark pairs and were called lateral, anterior, and posterior SL-VT point, respectively. The last landmark was chosen to cover the top region of the tongue. It was located at the most superoposterior point (apex or nadir) of the superior longitudinalis muscle on the midsagittal plane within the coronal (top) region of the tongue. This last landmark was called the Top End SL point.

1“Coronal” is a term used in phonetics.
Table 2: Landmarks for 3-d tongues, as shown in Figure 16

<table>
<thead>
<tr>
<th>Number</th>
<th>Name</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ant. Tip point</td>
<td>3</td>
<td>Ant. most points on the tongue</td>
</tr>
<tr>
<td>2</td>
<td>SG-HG point</td>
<td>1</td>
<td>Intersecting point of HG and SG</td>
</tr>
<tr>
<td>3</td>
<td>Ant. Apex of LS</td>
<td>1</td>
<td>location where IL, HG, and SG unite</td>
</tr>
<tr>
<td>4</td>
<td>Pos. Apex of PS</td>
<td>1</td>
<td>location where IL, GG, T &amp; V unite</td>
</tr>
<tr>
<td>5</td>
<td>Lat. SL-VT point</td>
<td>1 &amp; 3</td>
<td>location where SL, V &amp; T meet laterally</td>
</tr>
<tr>
<td>6</td>
<td>Ant. SL-VT point</td>
<td>1</td>
<td>location where SL, V &amp; T meet anteriorly</td>
</tr>
<tr>
<td>7</td>
<td>Pos. SL-VT point</td>
<td>1</td>
<td>location where SL, V &amp; T meet posteriorly</td>
</tr>
<tr>
<td>8</td>
<td>Top End SL point</td>
<td>2</td>
<td>location of SL's most superoposterior point</td>
</tr>
</tbody>
</table>

In the next section, the identification of the landmark pairs in the Visible Man and Miyawaki data is discussed, and the results of these landmark identifications are shown.

4.3.2 Locating Landmarks

The identification of landmark #1 was easier than that of the others. The section image that contains these two landmarks was first identified in a midsagittal view of two data sets. Then, the coordinates of the landmarks from two data sets were determined. Figure 8 shows the results of this procedure. Note that it was chosen off the midsagittal plane.
Figure 8: The left panel: landmark #1 is identified by a filled (red) circle at mid right portion of the tongue (V: Verticalis; T: Transversus; SL: Superior Longitudinalis; GG: Genioglossus; HG: Hyoglossus; SG: Styloglossus muscle). The downward protrusion of the lower (left) half of the tongue was caused by a missing tooth. The right panel: the transverse slice shown is number 9. The filled circle at the upper left extreme is the corresponding landmark #1 in Miyawaki's data.

Choosing the landmark #2 within the tongue body for the Visible Man data requires additional steps. First of all, the contours of the hyoglossus and the styloglossus muscles in the 2-d projection were manually extracted and displayed in 3-d, starting from their attached bony structures (the hyoid bone for the hyoglossus muscle and the styloid process for the styloglossus muscle) to the points where they disappear in the data sets. The next step was to fit the anterior-most and posterior-most points of the hyoglossus muscle in each slice and the superior-most and inferior-most points of the styloglossus muscle to third-order polynomial curves. Finally, the intersecting or the closest point between these two muscles was determined from the inferior boundary curve of the styloglossus muscle and
the posterior boundary curve of the hyoglossus muscle. Figure 9 shows how landmark #2 was identified. In Figure 9, the vertical stripes are the boundaries of the styloglossus muscle that were circumscribed from the slices of coronal projection and the horizontal stripes are the boundaries of the hyoglossus muscle that were circumscribed from the slices of transverse projection. The bold polynomial curves are the projected surface boundaries of the styloglossus (green) and the hyoglossus (blue) muscles. The filled (red) circle is the landmark #2, the SG-HG point. The lowest horizontal stripe is the origin of the hyoglossus muscle that is attached to the greater horn of the hyoid bone and the left-most vertical stripe is the origin of the styloglossus muscle that is attached to the styloid process.

The selection of landmark #2 for the Miyawaki data was determined by visual inspection of the transverse slice number 15 by consulting the sagittal and the coronal sections of the data. Landmark #2 for the Miyawaki data is shown in Figure 10.

The location of landmark #3 is at the anterior apex of the lateral septum where the styloglossus, the hyoglossus, and the inferior longitudinalis muscles unite. In Miyawaki's data, this point is located on the transverse slice number 12. The selection of landmark #3 in the Visible Man data was identified interactively on a display of three cutting planes similar to Figure 1. They are shown in Figure 11.

In a way similar to the selection for landmark #3, landmark #4 was located on the transverse slice number 11 of Miyawaki's data and the Visible Man data, respectively, at the location where four muscles (the genioglossus, the inferior
Figure 9: The vertical stripes are the projected boundaries of the styloglossus muscle and the horizontal stripes are the projected boundaries of the hyoglossus muscle. The bold polynomial curves are the fitted boundaries of the styloglossus (green) and the hyoglossus (blue) muscles. The filled (red) circle is the landmark #2, the SG-HG point. The anterior of the tongue is on the right-hand side and the coordinate system is in pixel number.
Figure 10: The transverse slice shown here is number 15. The filled circle is the corresponding landmark #2 in Miyawaki's data.

Figure 11: The left panel: landmark #3 is shown in filled (red) circle (right of IL) (V: Verticalis; T: Transversus; IL: Inferior Longitudinalis; GG: Genioglossus; HG: Hyoglossus; SG: Styloglossus muscle). The right panel: the transverse slice shown here is number 12. The filled circle is the corresponding landmark #3 in Miyawaki's data.
Figure 12: The left panel: landmark #4 is shown in filled (red) circle (to the left of IL) (V: Verticalis; T: Transversus; IL: Inferior Longitudinalis; GG: Genioglossus; HG: Hyoglossus; SG: Styloglossus muscle). The right panel: the transverse slice shown here is number 11. The filled circle is the corresponding landmark #4 in Miyawaki's data.

The selection of the second half of landmark pairs is similar to the preceding ones. The anterior, posterior, and the lateral SL-VT points were located on the transverse slice numbers 5 and 7. The top end SL point was on the transverse slice number 3. The corresponding landmarks were also identified in the Visible Man data. The identification for all the landmarks in the Visible Man data was implemented with a set of Matlab procedures (see Appendix B). The results of these landmark selections are shown in Figures 13, 14, and 15.

A 3-d display of the results for all landmark pairs chosen from both the Visible Man and Miyawaki data is shown in Figure 16 as an overall view of their
Figure 13: The left panel: landmark #5 is shown in filled (red) circle (above V & T) (V: Verticalis; T: Transversus; SL: Superior Longitudinalis; GG: Genioglossus; HG: Hyoglossus; SG: Styloglossus muscle). The protrusion of the lower (left) half of the tongue was caused by a missing tooth. The right panel: the transverse slice shown here is number 7. The filled circle is the corresponding landmark #5 in Miyawaki’s data.

Figure 14: The left panel: landmarks #6 and #7 are shown in filled (red) circles (to the right of SL & right of V & T) (V: Verticalis; T: Transversus; SL: Superior Longitudinalis muscle). The right panel: the transverse slice shown here is number 5. The filled circles are the corresponding landmarks #6 and #7 in Miyawaki’s data.
Figure 15: The left panel: landmark #8 is shown in filled (red) circle (at SL) (V: Verticalis; T: Transversus; SL: Superior Longitudinalis; GG: Genioglossus muscle). The right panel: the transverse slice shown here is number 3. The filled circle on the midsagittal (upper) edge is the corresponding landmark #8 in Miyawaki's data.

distribution in the tongue. In Figure 16, landmark numbers were also given for all landmarks to show their correspondence. These landmark pairs were then used to map Miyawaki's data onto the Visible Man data via the 3-d thin-plate spline mapping. The robustness and sufficiency of the selection of landmark pairs will be discussed below. Further description on the implementation of the 3-d thin-plate spline mapping and the evaluation of the mapping accuracy will be discussed also in the next chapter.

4.4 Summary

In this chapter, we gave a description of three principal landmark types and discussed our guidelines for identifying landmarks. Based on these guidelines, eight landmark pairs were located in the Visible Man data and the Miyawaki data. A
Figure 16: The left panel: a 3-d display of all landmarks open (red) circles identified from the Visible Man data. The right panel: the corresponding landmarks are displayed in cross (green) marks for Miyawaki’s data. The landmark numbers are given in both panel to show their correspondences.

...sequence of figures showed the landmarks identified in pertinent 2-d images. A 3-d display of the landmarks summarized the result for each of the data sets. In the next chapter, a procedure to integrate the Visible Man data and the Miyawaki data with the 3-d thin-plate spline method based on the landmark pairs will be described.
CHAPTER V

Integration of Different Data Sets of the Human Tongue

5.1 Introduction

The computational method for integrating different sets of 3-d anatomical data of the human tongue is basically composed of two stages of transformation. In this study, the Miyawaki hand-drawn dissection data were integrated with the Visible Man data. The first stage of transformation is the alignment of 2-d slice images in each data set to construct a datum in three dimensions. The second stage of transformation maps the aligned anatomical drawings into the Visible Man data with the 3-d thin-plate spline mapping technique. This chapter describes the procedure for implementing the integration of human tongue data sets and readjusting the precise locations of the landmarks.

5.2 Alignment Mapping

Alignment of Miyawaki's data is a construction of a 3-d stack based on a series of transverse 2-d tongue drawings. Miyawaki's data are more or less qualitative, focusing on identification of muscles and their fiber directions within each 2-d slice.
Figure 17: The outline of a tongue was viewed from the side. The parallel line shows where the transverse section was made. The number on the right side is the slice number (adaptation from Figure 5-i in Miyawaki, [106]).

We need to obtain an anatomically consistent reconstruction of the tongue as a 3-d structure in order to unwarp the moderate global deformation that may have been caused by the slicing of the original 3-d object and the slice-by-slice drawing of the sections. We used affine transformation to make the slices consistent with each other. It includes rotation, translation, and scaling. We consulted the profile figure in Miyawaki's article (see Figure 17) to align the transverse 2-d tongue drawings. Figure 17 is Figure 5-i in Miyawaki's article and it was scanned into a computer with a resolution of 200 digits (pixels) per inch. Its coordinates are in pixel number. The intersection points between the profile outline of the tongue and the parallel lines in the left (front) and right (rear) sides of the figure are the control points used for the alignment mapping. In Figure 17, the ordinate represents the z-coordinate of the control point.
Figure 18: The definition of the control points based on hypothesis (a). They are shown in filled (red) circles identified by numbers 1 and 2. Midsag. Pl. stands for the midsagittal plane.

We considered two possible interpretations of Figure 17 with respect to the location of each control point.

- Hypothesis (a): If we assume the profile is a lateral view of the tongue before dissection, then we need to define the control points to be the x-direction extrema of the outline in each section projected onto the midsagittal plane. Figure 18 shows the definition of the control points (red) in the transverse slice number 12 according to this hypothesis. The contour is the skin surface of the tongue.

- Hypothesis (b): If we assume that the profile figure is a view of the tongue in the midsagittal cut plane, then the control points have to label the beginning and end of the line segment delimited by the profile contour in each section. The control points from the same slice number 12 defined according to this
Figure 19: The definition of the control points based on hypothesis (b). (cf. Figure 18).

hypothesis are shown in Figure 19. The contour is the skin surface of the tongue.

Our impression is that alignment of the midsagittal cuts (hypothesis (b)) rather than the midsagittal projections results in less inconsistency between the profile figure and the lengths of the drawn slices (discussed in Sec 6.2.1). Therefore, our alignment was performed according to hypothesis (b), which means that Figure 17 is assumed to be a display of the midsagittal end-points. In other words, the y-coordinate values of the tongue after the mapping were assumed to be zero for all control points. The mapping is to align each line segment that connects two control points \(((x_{1i},y_{1i}),(x_{2i},y_{2i}))\), \(i = 1, \ldots 18\) on the midsagittal plane of each transverse slice number \(i\) with each line segment (i.e. connecting \(((u_{1i},v_{1i}),(u_{2i},v_{2i}))\), \(i = 1, \ldots 18\), where \(v_1 = v_2 = 0\). This 2-d (x-y plane) mapping comprises two 2-d translations, one uniform scaling, and one rotation.
The mapping begins with a 2-d translation by \((-x_1, -y_1)\) for each transverse slice and brings the control point \((x_1, y_1)\) to the origin of its coordinate system. It is followed by a rotation by an angle \(\alpha = \arcsin(H)\), for which the general equation

\[
H = \frac{(x_2-x_1)(v_2-v_1)-(y_2-y_1)(u_2-u_1)}{\sqrt{(u_2-u_1)^2+(v_2-v_1)^2}}
\]

about the origin. The rotation is followed by a uniform scaling (as determined by the translations of two end points, assuming the same scaling factor in both x-direction \((s_x)\) and y-direction \((s_y)\) is the same,

\[
s_x = s_y = \gamma = \frac{\sqrt{(u_2-u_1)^2+(v_2-v_1)^2}}{\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}}
\]

and concludes with the second 2-d translation by \((u_1, v_1)\) to move the control point from the origin to \((u_1, v_1)\) \((v_1 = 0)\) for each point on the left side of Figure 17 where the parallel lines intersect with the profile of the tongue. Figure 20 illustrates this alignment transformation without scaling. The resulting alignment transformation matrix is represented with homogeneous coordinates as

\[
T = \begin{pmatrix}
    s_x \cos(\alpha) & -(s_x \sin(\alpha)) & X_t \\
    s_y \sin(\alpha) & s_y \cos(\alpha) & Y_t \\
    0 & 0 & 1
\end{pmatrix},
\]

where \(X_t = u_1 - s_x \ x_1 \ \cos(\alpha) + s_x \ y_1 \ \sin(\alpha)\), and \(Y_t = v_1 - s_y \ y_1 \ \cos(\alpha) - s_y \ x_1 \ \sin(\alpha)\).

The alignment mapping of the z-coordinate was simply a translation transformation since it is assumed to be a constant for each transverse slice. It is given as

\[
z_1 = z_1 + z_t,
\]

where \(z_t = 0.5 \ [(w_1 + w_2) - (z_1 + z_2)]\).

A list of transformation parameters, the angle \((\alpha)\) of rotation and the scaling factor \((\gamma)\) for each transverse slice, is shown in Table 3. Other parameters used for
Alignment Transformation

Figure 20: Graphic illustration of the alignment transformation

the alignment mapping are also listed in Tables 4, 5, and 6. In Tables 4 and 5, N is the slice number of the transverse sections.

The final result of the alignment mapping for the polygon files of the transverse slices in the Miyawaki data is shown in Figure 21. The tongue tip in this figure points to the left side. The filled (red) circles indicate the location of each control point (two for each transverse slice) after the alignment mapping. Parameters for the alignment mapping listed in the above tables were also used to transform the landmark pairs during the process of 3-d thin-plate spline mapping.

The Visible Man data were already aligned when they were downloaded to the computer. The data representation was in a voxel format of 150 × 150 × 81. A color index was assigned to each voxel where its coordinates were specified by the
Table 3: Parameters of the alignment mapping for each transverse slice

<table>
<thead>
<tr>
<th>Slice Number</th>
<th>Angle (α, in degree)</th>
<th>Scaling (γ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2800</td>
<td>0.7815</td>
</tr>
<tr>
<td>2</td>
<td>-0.4492</td>
<td>0.7692</td>
</tr>
<tr>
<td>3</td>
<td>-1.2271</td>
<td>0.7148</td>
</tr>
<tr>
<td>4</td>
<td>0.3045</td>
<td>0.7848</td>
</tr>
<tr>
<td>5</td>
<td>0.0956</td>
<td>0.7629</td>
</tr>
<tr>
<td>6</td>
<td>1.0998</td>
<td>0.7377</td>
</tr>
<tr>
<td>7</td>
<td>0.2657</td>
<td>0.7113</td>
</tr>
<tr>
<td>8</td>
<td>0.2796</td>
<td>0.6838</td>
</tr>
<tr>
<td>9</td>
<td>1.1224</td>
<td>0.7094</td>
</tr>
<tr>
<td>10</td>
<td>0.1977</td>
<td>0.7079</td>
</tr>
<tr>
<td>11</td>
<td>0.9086</td>
<td>0.7822</td>
</tr>
<tr>
<td>12</td>
<td>2.7010</td>
<td>0.9431</td>
</tr>
<tr>
<td>13</td>
<td>0.6290</td>
<td>0.8039</td>
</tr>
<tr>
<td>14</td>
<td>2.0370</td>
<td>0.7690</td>
</tr>
<tr>
<td>15</td>
<td>0.6811</td>
<td>0.6982</td>
</tr>
<tr>
<td>16</td>
<td>-3.3528</td>
<td>0.5745</td>
</tr>
<tr>
<td>17</td>
<td>-1.7822</td>
<td>0.5283</td>
</tr>
<tr>
<td>18</td>
<td>-0.8421</td>
<td>0.4583</td>
</tr>
</tbody>
</table>
Table 4: Coordinates (in pixel) of the control points from Figure 17

<table>
<thead>
<tr>
<th>N</th>
<th>v1</th>
<th>u1</th>
<th>v1</th>
<th>u2</th>
<th>v2</th>
<th>w2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>248.9724</td>
<td>0</td>
<td>33.9646</td>
<td>402.2588</td>
<td>0</td>
<td>32.9803</td>
</tr>
<tr>
<td>2</td>
<td>182.2186</td>
<td>0</td>
<td>57.5866</td>
<td>485.0829</td>
<td>0</td>
<td>55.6181</td>
</tr>
<tr>
<td>3</td>
<td>141.4246</td>
<td>0</td>
<td>81.2087</td>
<td>532.0578</td>
<td>0</td>
<td>77.2717</td>
</tr>
<tr>
<td>4</td>
<td>110.5201</td>
<td>0</td>
<td>106.7992</td>
<td>569.1432</td>
<td>0</td>
<td>101.8780</td>
</tr>
<tr>
<td>5</td>
<td>87.0327</td>
<td>0</td>
<td>131.4055</td>
<td>586.4497</td>
<td>0</td>
<td>124.5157</td>
</tr>
<tr>
<td>6</td>
<td>62.3090</td>
<td>0</td>
<td>156.9961</td>
<td>601.2839</td>
<td>0</td>
<td>149.1220</td>
</tr>
<tr>
<td>7</td>
<td>54.8920</td>
<td>0</td>
<td>180.6181</td>
<td>613.6457</td>
<td>0</td>
<td>174.7126</td>
</tr>
<tr>
<td>8</td>
<td>52.4196</td>
<td>0</td>
<td>208.1772</td>
<td>626.0075</td>
<td>0</td>
<td>202.2717</td>
</tr>
<tr>
<td>9</td>
<td>52.4196</td>
<td>0</td>
<td>233.7677</td>
<td>633.4246</td>
<td>0</td>
<td>226.8780</td>
</tr>
<tr>
<td>10</td>
<td>53.6558</td>
<td>0</td>
<td>256.4055</td>
<td>638.3693</td>
<td>0</td>
<td>253.4528</td>
</tr>
<tr>
<td>11</td>
<td>70.9623</td>
<td>0</td>
<td>276.0906</td>
<td>638.3693</td>
<td>0</td>
<td>274.1220</td>
</tr>
<tr>
<td>12</td>
<td>155.0226</td>
<td>0</td>
<td>300.6969</td>
<td>635.8970</td>
<td>0</td>
<td>296.7598</td>
</tr>
<tr>
<td>13</td>
<td>199.5251</td>
<td>0</td>
<td>324.3189</td>
<td>632.1884</td>
<td>0</td>
<td>322.3504</td>
</tr>
<tr>
<td>14</td>
<td>246.5000</td>
<td>0</td>
<td>346.9567</td>
<td>626.0075</td>
<td>0</td>
<td>344.9882</td>
</tr>
<tr>
<td>15</td>
<td>325.6156</td>
<td>0</td>
<td>368.6102</td>
<td>616.1181</td>
<td>0</td>
<td>368.6102</td>
</tr>
<tr>
<td>16</td>
<td>383.7161</td>
<td>0</td>
<td>388.2953</td>
<td>604.9925</td>
<td>0</td>
<td>387.3110</td>
</tr>
<tr>
<td>17</td>
<td>408.4397</td>
<td>0</td>
<td>410.9331</td>
<td>593.8668</td>
<td>0</td>
<td>408.9646</td>
</tr>
<tr>
<td>18</td>
<td>434.3995</td>
<td>0</td>
<td>434.4936</td>
<td>580.2688</td>
<td>0</td>
<td>432.4636</td>
</tr>
</tbody>
</table>
Table 5: Coordinates of the control points for each transverse slice of Miyawaki’s data

<table>
<thead>
<tr>
<th>N</th>
<th>x1 (in pixel)</th>
<th>y1 (in pixel)</th>
<th>z1 (in mm)</th>
<th>x2 (in pixel)</th>
<th>y2 (in pixel)</th>
<th>z2 (in mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>39.1726</td>
<td>222.2009</td>
<td>51.0</td>
<td>235.3218</td>
<td>221.2423</td>
<td>51.0</td>
</tr>
<tr>
<td>2</td>
<td>46.6538</td>
<td>277.5521</td>
<td>48.0</td>
<td>440.3585</td>
<td>280.6388</td>
<td>48.0</td>
</tr>
<tr>
<td>3</td>
<td>10.6405</td>
<td>262.5610</td>
<td>45.0</td>
<td>557.0367</td>
<td>274.2653</td>
<td>45.0</td>
</tr>
<tr>
<td>4</td>
<td>15.4637</td>
<td>321.7224</td>
<td>42.0</td>
<td>599.8387</td>
<td>318.6166</td>
<td>42.0</td>
</tr>
<tr>
<td>5</td>
<td>5.1815</td>
<td>335.2247</td>
<td>39.0</td>
<td>659.8185</td>
<td>334.1319</td>
<td>39.0</td>
</tr>
<tr>
<td>6</td>
<td>31.8177</td>
<td>378.3462</td>
<td>36.0</td>
<td>762.3419</td>
<td>364.3221</td>
<td>36.0</td>
</tr>
<tr>
<td>7</td>
<td>10.1306</td>
<td>372.4719</td>
<td>33.0</td>
<td>795.6952</td>
<td>368.8292</td>
<td>33.0</td>
</tr>
<tr>
<td>8</td>
<td>17.6875</td>
<td>413.0832</td>
<td>30.0</td>
<td>856.5484</td>
<td>408.9900</td>
<td>30.0</td>
</tr>
<tr>
<td>9</td>
<td>29.0993</td>
<td>418.3284</td>
<td>27.0</td>
<td>847.8968</td>
<td>402.2858</td>
<td>27.0</td>
</tr>
<tr>
<td>10</td>
<td>24.1795</td>
<td>417.2129</td>
<td>24.0</td>
<td>850.1170</td>
<td>414.3627</td>
<td>24.0</td>
</tr>
<tr>
<td>11</td>
<td>63.9052</td>
<td>425.2684</td>
<td>21.0</td>
<td>789.2581</td>
<td>413.7653</td>
<td>21.0</td>
</tr>
<tr>
<td>12</td>
<td>194.1827</td>
<td>462.1907</td>
<td>18.0</td>
<td>703.5224</td>
<td>438.1621</td>
<td>18.0</td>
</tr>
<tr>
<td>13</td>
<td>48.4988</td>
<td>476.5396</td>
<td>15.0</td>
<td>586.7004</td>
<td>470.6310</td>
<td>15.0</td>
</tr>
<tr>
<td>14</td>
<td>29.7534</td>
<td>546.8949</td>
<td>12.0</td>
<td>522.9692</td>
<td>529.3528</td>
<td>12.0</td>
</tr>
<tr>
<td>15</td>
<td>44.8956</td>
<td>505.7776</td>
<td>9.0</td>
<td>460.9540</td>
<td>500.8313</td>
<td>9.0</td>
</tr>
<tr>
<td>16</td>
<td>82.8581</td>
<td>444.1692</td>
<td>6.0</td>
<td>467.3742</td>
<td>466.6961</td>
<td>6.0</td>
</tr>
<tr>
<td>17</td>
<td>48.9323</td>
<td>470.5358</td>
<td>3.0</td>
<td>399.7185</td>
<td>481.4509</td>
<td>3.0</td>
</tr>
<tr>
<td>18</td>
<td>72.7500</td>
<td>345.8850</td>
<td>0.0</td>
<td>390.9960</td>
<td>350.5629</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Table 6: Translation of the z-coordinate for each transverse slice

<table>
<thead>
<tr>
<th>Slice Number</th>
<th>Zt (in pixel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-17.5276</td>
</tr>
<tr>
<td>2</td>
<td>8.6024</td>
</tr>
<tr>
<td>3</td>
<td>34.2402</td>
</tr>
<tr>
<td>4</td>
<td>62.3386</td>
</tr>
<tr>
<td>5</td>
<td>88.9606</td>
</tr>
<tr>
<td>6</td>
<td>117.0591</td>
</tr>
<tr>
<td>7</td>
<td>144.6653</td>
</tr>
<tr>
<td>8</td>
<td>175.2244</td>
</tr>
<tr>
<td>9</td>
<td>203.3228</td>
</tr>
<tr>
<td>10</td>
<td>230.9291</td>
</tr>
<tr>
<td>11</td>
<td>254.1063</td>
</tr>
<tr>
<td>12</td>
<td>280.7283</td>
</tr>
<tr>
<td>13</td>
<td>308.3347</td>
</tr>
<tr>
<td>14</td>
<td>333.9724</td>
</tr>
<tr>
<td>15</td>
<td>359.6102</td>
</tr>
<tr>
<td>16</td>
<td>381.8031</td>
</tr>
<tr>
<td>17</td>
<td>406.9488</td>
</tr>
<tr>
<td>18</td>
<td>433.4786</td>
</tr>
</tbody>
</table>
Figure 21: The result of the alignment mapping for the polygon files of the transverse slices of Miyawaki's data (lateral view). The filled (red) circles indicate the locations of the control points. The numbers indicate the landmark numbers for Miyawaki’s data.

pixel value in each (x, y, or z) direction. An example of a display in 3 different projections (sagittal, coronal, and transverse) is shown in Figure 1.

5.3 3-d Thin-plate Spline Mapping

Once the data were aligned and the landmark pairs were chosen according to the guidelines proposed in Section 4.3.1, the tongue data sets are ready for 3-d thin-plate spline mapping. The 3-d thin-plate spline mapping was conducted with the following procedure:

- The first step to implement the 3-d thin-plate spline mapping was to transform the coordinates of the selected landmarks of the Miyawaki data according to the parameters of the alignment mapping. The aligned landmarks and
their corresponding landmarks from the Visible Man data were then used in the next step of the mapping.

- The aligned polygon files of the Miyawaki data as a 3-d object, as shown in Figure 21, were warped with the 3-d thin-plate splines to match the Visible Man 3-d object based on the selected landmark pairs. During the mapping, the 3-d coordinates of the mesh points were transformed from the coordinate system of the Miyawaki data into the coordinate system of the Visible Man data.

- The result of the mapping was a display of a 3-d stack of polygon files, each of which was deformed three dimensionally into a warped plate fitted within the Visible Man data.

- An iterative process comprising the steps above followed after readjustments of landmark positions when the mapping was not satisfied.

Following the procedure above, the result of the mapping based on the initially chosen landmark pairs from Chapter IV is shown in Figure 22. In Figure 22, the filled (red) circles indicate the locations of the control points, i.e. the endpoints of the midsagittal lines for slices (the tongue tip points to the right). If the selection of the landmark pairs was appropriate, the 3-d shape of this figure should look like the 3-d tongue shape of the Visible Man data. The question remains whether the landmark pairs were initially chosen correctly or not. In fact, as discussed later, we need to seek optimal positions of the landmark points by evaluating the overall
Figure 22: The result of the 3-d thin-plate spline mapping for the polygon files of the transverse slices of Miyawaki's data (lateral view). The filled (red) circles indicate the locations of the control points and closed points with numbers indicate the landmark numbers for the Visible Man data.

3-d image as the function of readjustment in locating each landmark point. The sensitivity of the overall matching of corresponding images to slightly incomparable slice levels and the small variation of the x-y coordinates in the landmark pairs is the central issue for the evaluation of the mapping accuracy.

5.4 Evaluation of the Accuracy of Mapping

As mentioned in Section 4.3.1, caution should be exercised to avoid matching images at incomparable slice levels during the selection of the landmark pairs. A mismatch in vertical position of the landmark pairs caused by the incomparable slice levels where the landmarks are located often visibly deteriorates the result
of the mapping. Mapping errors of this kind were anticipated given that the slice correspondence between the two sets of data and the precise vertical position within a slice in Miyawaki's data are not necessarily clear, and therefore the best vertical coordinate values of landmarks needed to be inferred. We evaluated the sensitivity to incomparable slice level and the resulting inaccuracy of mapping is described below:

- A slice-wise correspondence between the transverse slices of the two data sets was established between the Miyawaki data and the Visible Man data. Table 7 gives a list of such corresponding slice pairs (the numbers are slice numbers). For example, slice number 7 in the Miyawaki data approximately corresponds in vertical level to slice number 182 in the Visible Man data. These corresponding slice numbers were determined by comparing the shapes of anatomical structures identified on the slices of both data sets.

- A mesh with 3-d coordinates was generated to cover the region of an image from each transverse slice of the Miyawaki data. The mesh was first transformed with the alignment mapping. The mesh was then warped with the 3-d thin-plate spline based on the landmark pairs located in the Miyawaki image and the corresponding Visible Man image. During the mapping, the coordinates of the mesh were transformed from the 3-d coordinate system of the Miyawaki data into the 3-d coordinate system of the Visible Man data.
• The warped mesh resulting from this mapping was used to sample the Visible Man data block. The result of this sampling is a Miyawaki surface in 3-d space with color indices of the Visible Man data on the coordinates of the warped Miyawaki mesh.

• The corresponding slice pairs on which the landmark pairs were not identified were chosen to test the validity of the mapping.

• If the landmark pairs were correctly chosen, the reverse mapping of this Visible Man image pattern onto the transverse flat plane should look like the image of the original corresponding transverse slice of the Miyawaki data. More often, an iteration of the alignment and the thin-plate spline mappings back and forth involving manipulation of the landmark positions was necessary before a "global matching" of the warped Miyawaki data and the Visible Man data was achieved. Global matching is judged to have been achieved when the reverse mapping of each warped slice matches the image in the original corresponding transverse slice of the Miyawaki data\textsuperscript{1}. Landmark manipulation was performed by interactively readjusting the coordinates of the landmark position of the Miyawaki data with visual feedback from the computer display for the overall image matching.

\textsuperscript{1}This may sound gross and subjective. The fact is, however, that the overall image changes obviously when a slight change in landmark position is introduced. Thus, the overall visual impression is a sensitive measure for evaluating the accuracy of landmark location. At the same time, image evaluation requires experience and some knowledge of the internal muscle structures. Therefore, it is difficult at present to automate this mapping process.
Table 7: Slice-wise correspondence between the Miyawaki data and the Visible Man data

<table>
<thead>
<tr>
<th>Miyawaki's data</th>
<th>Visible Man data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>164</td>
</tr>
<tr>
<td>2</td>
<td>167</td>
</tr>
<tr>
<td>3</td>
<td>170</td>
</tr>
<tr>
<td>4</td>
<td>173</td>
</tr>
<tr>
<td>5</td>
<td>176</td>
</tr>
<tr>
<td>6</td>
<td>179</td>
</tr>
<tr>
<td>7</td>
<td>182</td>
</tr>
<tr>
<td>8</td>
<td>185</td>
</tr>
<tr>
<td>9</td>
<td>188</td>
</tr>
<tr>
<td>10</td>
<td>191</td>
</tr>
<tr>
<td>11</td>
<td>194</td>
</tr>
<tr>
<td>12</td>
<td>197</td>
</tr>
<tr>
<td>13</td>
<td>200</td>
</tr>
<tr>
<td>14</td>
<td>203</td>
</tr>
<tr>
<td>15</td>
<td>206</td>
</tr>
<tr>
<td>16</td>
<td>209</td>
</tr>
<tr>
<td>17</td>
<td>212</td>
</tr>
<tr>
<td>18</td>
<td>215</td>
</tr>
</tbody>
</table>

A graphic superposition of the original transverse slice image of the Miyawaki data on the resulting projection image of the Visible Man data facilitated this evaluation process. It was implemented by adding an extra color index to the original indices used by the Visible Man data and displaying both images in the same graphic window with the new set of color indices. The result was a pseudo-transparent image of the original transverse slice of the Miyawaki data superimposed on top of the resulting projection image of the Visible Man data. The fiber direction of the Miyawaki data was thus displayed with the new add-on color index.
Following the above evaluation procedure, a mapping result based on the initially chosen landmark pairs from Chapter IV is shown in Figure 23. Transverse slice number 4 was used to evaluate the validity of the 3-d thin-plate spline mapping because there are no landmarks selected from this slice. The left panel of Figure 23 shows the mapped result of the Miyawaki transverse slice image, (cf. the left panel of Figure 23 and the upper half of the tongue in the right panel of Figure 24). The right panel of Figure 23 shows the mapped result of the Miyawaki transverse slice image superimposed on the left panel of Figure 23 to assist evaluation of the mapping result. In the right panel of Figure 23, a good match overall of this superimposed image indicates that the landmarks in the upper half of the tongue are chosen appropriately. Note that, in order to make a valid judgment about the image match, we need to apply anatomical knowledge and experience with color displays of the Visible Man data, as seen in the interpretation of the protruded area to the right extrema, which is relatively light in color due to the type of tissue.

A mismatch in the orientation of the hyoglossus (HG) muscle was found when slice number 17 was chosen to test the mapping. In Figure 25, the left panel shows an enlarged image after mapping of the original (Miyawaki) transverse slice image filled with color indices of the Visible Man data; the right panel displays the mapped result of the Miyawaki transverse slice image superimposed on the left panel image to assist evaluation of the mapping result. Shown in the left panel
Figure 23: The left panel: the mapped result of transverse slice number 4 of Miyawaki's data. The right panel: the original contour of the same Miyawaki slice superimposed on the left panel image after 3-d thin-plate spline mapping (V: Verticalis; T: Transversus; SL: Superior Longitudinalis muscle).

Figure 24: The left panel: the scanned-in transverse slice number 4 of Miyawaki's data. The right panel: the corresponding slice in the Visible Man data to the left panel (V: Verticalis; T: Transversus; SL: Superior Longitudinalis muscle).
Figure 25: The left panel: the mapped result of transverse slice number 17 of Miyawaki's data. The right panel: the contour of transverse slice number 17 of Miyawaki's data superimposed on the left panel image (V: Verticalis; T: Transversus; GG: Genioglossus; HG: Hyoglossus; MY: Mylohyoid muscle).

of Figure 26 is the original (Miyawaki's) transverse slice image; the corresponding slice in the Visible Man data is shown in the right panel of Figure 26.

The right panel of Figure 25 indicates a mismatch with respect to the HG region in the lower half of the tongue. This mismatch was caused by an inappropriate selection of landmark positions in the lower half of the tongue. For a comparison with Table 7, a summary of positions from initially selected landmark pairs that resulted in this figure is given in Table 8. In Table 8, the vertical positions of the landmark pairs are shown by the slice numbers on which they are located. The first column shows the sequence number of each landmark. The transverse slice numbers of Miyawaki's data and the Visible Man data are shown in the second column and third column, respectively.
Figure 26: The left panel: the scanned-in transverse slice number 17 of Miyawaki's data. The right panel: the corresponding slice in the Visible Man data to the left panel (V: Verticalis; T: Transversus; GG: Genioglossus; HG: Hyoglossus; MY: Mylohyoid muscle).

Table 8: Summary of initially evaluated correspondence in vertical position (slice number) for each landmark pair (identified by landmark number to the left) that were used for mapping shown in Figure 25

<table>
<thead>
<tr>
<th>Number</th>
<th>Miyawaki's Data</th>
<th>Visible Man Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>t9</td>
<td>187</td>
</tr>
<tr>
<td>2</td>
<td>t15</td>
<td>201</td>
</tr>
<tr>
<td>3</td>
<td>t12</td>
<td>196</td>
</tr>
<tr>
<td>4</td>
<td>t11</td>
<td>195</td>
</tr>
<tr>
<td>5</td>
<td>t7</td>
<td>182</td>
</tr>
<tr>
<td>6</td>
<td>t5</td>
<td>176</td>
</tr>
<tr>
<td>7</td>
<td>t5</td>
<td>176</td>
</tr>
<tr>
<td>8</td>
<td>t3</td>
<td>171</td>
</tr>
</tbody>
</table>
In Tables 8, we find that landmark #2 was in the t15 slice in Miyawaki’s data and in slice #201 in the Visible Man data. In Table 7, which shows the slice numbers used in readjusted landmark locations, we find that the Visible Man slice number that corresponds to Miyawaki’s t15 is slice #206 in the Visible Man data. The correspondence shown in Table 8 (i.e. #201) was erroneously low. This initial error in finding correspondence between Miyawaki’s slice and the Visible Man slice could explain why the effect of level sensitivity is greater in transverse slice number 17, as we discussed with Figure 25, than other higher slices as in Figure 23. Since the lower half of the tongue is closer to landmark #2, the effect by the poor selection of landmark #2 is greater than the upper half of the tongue. To adjust the position of landmark #2 was the first step to improve the global matching. We believe that the position of landmark #2 in the Visible Man data (shown in Figure 9) was selected more accurately than that of the Miyawaki data because of the 3-d graphic capability of the former. We needed to adjust the location of landmark #2 from Miyawaki’s data in three different directions (i.e. x, y, and z direction) to improve the global matching. A z-direction readjustment implies that the correct landmark point in Miyawaki's data could be outside any of the transverse slices, or somewhere between two slices. We need to readjust the position of the selected landmark point (i.e. landmark #2) until an optimal overall image matching has been obtained. During the process of position manipulation, we found that changing the position of landmark #2 in the z direction
Figure 27: The transverse slice shown here is number 16. The filled circle is the new landmark #2 in Miyawaki’s data (T: Transversus; GG: Genioglossus; HG: Hyoglossus; SG: Styloglossus muscle).

(Vertical movement) was more effective than readjustment in the x and y directions for optimizing the match.

After several iterated processes of the evaluation procedure by moving the landmark #2 in three different directions, the new location of landmark #2 was found to be in transverse slice number 16 of the Miyawaki data. The corrected location of landmark #2 is one slice lower (i.e. 3 mm lower) in the vertical direction (z direction) from the initially estimated position (as in Table 8) with small changes in the x (front-back) and y (medial-lateral) directions. The new location of landmark #2 is shown in Figure 27. It is indicated by an arrow and specified as “SG-HG Pt.”
Figure 28: The left panel: another Visible Man slice corresponding to transverse slice number 4 of Miyawaki's data. The right panel: A similar Visible Man slice corresponding to transverse slice number 17 of Miyawaki's data (V: Verticalis; T: Transversus; GG: Genioglossus; HG: Hyoglossus; MY: Mylohyoid; SL: superior longitudinalis muscle).

The mapping results using the new landmark #2 for Miyawaki's t4 and t17 are shown in Figure 28. In Figure 28, the left panel displays no deterioration on the mapped result of transverse slice number 4 (compared with the right panel of Figure 23) using the new landmark location. The right panel shows a good improvement in the orientation of the hyoglossus muscle. The orientation of the hyoglossus muscle in the mapped image better agrees with that in the superimposed original (Miyawaki's) transverse slice image, and the mylohyoid muscle of the mapped Visible Man image (as indicated by MY) is correctly outside the skin (solid contour) of the original transverse slice image of the tongue (cf. Figure 25).

We show transverse slice numbers 9 and 12 as examples to further demonstrate that a global matching is achieved with the newly selected landmark #2.
Figure 29: The mapped result of transverse slice number 9 of Miyawaki's data (V: Verticalis; T: Transversus; GG: Genioglossus muscle).

three dimensionally. Figure 29 shows the Visible Man slice corresponding to transverse slice number 9 of Miyawaki's data superimposed onto the original image of Miyawaki's data. The bright area in the lower left side of the figure is the teeth that are outside of the tongue (indicated by a solid contour). The result of the mapping in Figure 29 can be compared with the upper half of the tongue in the left panel of Figure 8.

Figure 30 displays the warped (as in Figure 23) transverse slice number 12 of Miyawaki's image superimposed onto the transformed Visible Man image. Labels of muscles in the figure indicate the identified muscle sub-structures for both images. The upper left of this figure shows the area of the sublingual gland tissue. This figure shows a good match for transverse slice number 12. The mapped result
Figure 30: The mapped result of transverse slice number 12 of Miyawaki's data (V: Verticalis; T: Transversus; GG: Genioglossus; IL: Inferior Longitudinalis; MY: Mylohyoid; HG: Hyoglossus; SG: Styloglossus muscle).

of Figure 30 can also be compared to the upper half of the tongue in the left panel of Figure 11 (rotated by 180 degrees).

We summarize the new landmark pairs after the evaluation procedure as a list of their $x$, $y$, $z$ coordinates in Table 9. In Table 9, $N$ is the landmark number. $P$ indicates the landmark points of Miyawaki's data and $H$ represents the landmarks of the Visible Man data.

5.5 Summary

In this chapter, we demonstrated the integration of the Miyawaki and Visible Man data, using the 3-d thin-plate spline mapping method. The integration process was
Table 9: Revised coordinates of the landmark pairs for the Visible Man data and the Miyawaki data

<table>
<thead>
<tr>
<th>N</th>
<th>$P_x$(in pixel)</th>
<th>$P_y$(in pixel)</th>
<th>$P_z$(in mm)</th>
<th>$H_x$(in pixel)</th>
<th>$H_y$(in pixel)</th>
<th>$H_z$(in pixel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17</td>
<td>52</td>
<td>27</td>
<td>73</td>
<td>142</td>
<td>53</td>
</tr>
<tr>
<td>2</td>
<td>335</td>
<td>341</td>
<td>6</td>
<td>34</td>
<td>52</td>
<td>39</td>
</tr>
<tr>
<td>3</td>
<td>157</td>
<td>213</td>
<td>18</td>
<td>57</td>
<td>91</td>
<td>43</td>
</tr>
<tr>
<td>4</td>
<td>476</td>
<td>130</td>
<td>21</td>
<td>60</td>
<td>66</td>
<td>45</td>
</tr>
<tr>
<td>5</td>
<td>498</td>
<td>315</td>
<td>33</td>
<td>44</td>
<td>59</td>
<td>58</td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>13</td>
<td>39</td>
<td>79</td>
<td>119</td>
<td>64</td>
</tr>
<tr>
<td>7</td>
<td>447</td>
<td>14</td>
<td>39</td>
<td>75</td>
<td>71</td>
<td>64</td>
</tr>
<tr>
<td>8</td>
<td>474</td>
<td>62</td>
<td>45</td>
<td>74</td>
<td>65</td>
<td>69</td>
</tr>
</tbody>
</table>

based on the selected landmark pairs. We proposed a sequence of steps to evaluate the validity of the selection of the landmark pairs and applied the procedures to correct the originally poor location of landmark #2. A 3-d global matching was achieved with the newly selected landmark pairs. Superposition of the Miyawaki transverse slice images with their corresponding mapped results in the Visible Man data were shown to demonstrate the sufficiency of the selected landmark pairs (8 points) in the integration of different anatomical data sets of the human tongue.

In the next, and final, chapter, we will discuss limitations of our approach and example data, review the contributions of this dissertation work, and give pointers to future work.
CHAPTER VI

Discussion

6.1 Introduction

This dissertation is concerned with a general computational method for evaluating quantitative relations among the external shapes and internal muscular structures of the tongue observed by various methods in speech production. We have introduced a strategy in the form of several guidelines for defining a set of landmarks in reference to a general 3-d tongue structure representation. The computational method for integrating different data sets of the human tongue was based on the principle to determine pairs of corresponding sets of landmarks and this process was achieved by the 3-d thin-plate spline mapping. We have shown preliminary results of integrating the Miyawaki data and the Visible Man data, and proposed a sequence of steps to evaluate the validity of the selection of the corresponding landmark pairs. In this chapter, we will discuss the limitations presented by the currently available data and how these limitations have affected our computational procedure. Finally, the contribution of this dissertation work and pointers to future work will be discussed.
6.2 Limitation of the Current Data

In the course of this dissertation work, a great portion of time was spent on processing and analyzing the data which we used. Limitations of the currently available data posed difficulties in landmark identification and alignment mapping. In this section, limitations of the data are described and discussed.

6.2.1 Miyawaki's Data

The Miyawaki data were chosen for this dissertation work due to the unusually detailed drawings of tongue muscles (see Figure 6). The detailed identification of fiber directions is not only crucial to the study of biomechanical 3-d tongue modeling but also important for the identification of the landmarks within the tongue. However, the Miyawaki data were based on a tongue specimen cut out from its surrounding structures\(^1\). For this reason, the information relative to its surrounding structures was not available when we tried to compare the dissection images with other data sets acquired from different observing techniques such as the MR imaging. Missing information of this sort complicated the alignment process and consequent reconstruction into a 3-d stack. An examination of the alignment mapping as discussed below provides an illustration of such complications. In addition, the effective value of the third coordinate (z-coordinate) of each point

---

\(^1\)Note that if we did use the fixed points on bones such as the maxilla and the mandible, along with landmarks within the tongue, we would be able to describe the articulatory gestures entirely in terms of the distorted grids (in 3-d space) of the mapping between articulatory gesture images and reference (e.g. resting or neutral articulation) state. This would require a mapping for a very large and nonlinear distortion, but it could be handled effectively by the 3-d thin-plate spline method.
in the slice image needed to be indirectly inferred in conformity with the profile figure (Figure 17) and the slice thickness information to optimize the result.

Our investigation of the alignment problem mainly concerned how the control points of the mapping were to be chosen and how the scaling of the mapping was to be performed. We tried an alignment with only scaling in the x direction (front-back), so that the y coordinates (left-right) were not affected, except for translation. The idea behind this was that the hyoglossus muscles in the original drawings seemed to follow a consistent pattern. For example, if we measured the distance (in cm) from the midsagittal plane to the closest point of the contour of the hyoglossus muscle in the figures T:10 through T:18 (pages 47-49 in the original article), we get Table 10. These measurements may involve errors up to 2 mm in our image scale, which means less than 10% when compared to our measured distances. This would be accurate enough for our purpose in identifying the converging shape of the hyoglossus muscle superoinferior toward the hyoid bone.

We know that the hyoglossus muscles must converge towards the bottom of the tongue to attach to the greater horn of the hyoid bone on each side. It also comes closer to the midsagittal plane towards the bottom of the tongue. This can be verified from the Visible Man data looking at the coronal sections. The monotonically decreasing distance evaluated independently within each transverse section between the midsagittal plane and the closest point of the contour of the hyoglossus muscle, as shown in Table 10, is thus reasonable (except slice numbers
Table 10: The distance from the midsagittal plane to the closest point of the contour of the hyoglossus muscle in transverse slices of the Miyawaki data

<table>
<thead>
<tr>
<th>Slice Number</th>
<th>Distance (in cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>3.5</td>
</tr>
<tr>
<td>11</td>
<td>3.0</td>
</tr>
<tr>
<td>12</td>
<td>2.6</td>
</tr>
<tr>
<td>13</td>
<td>2.7</td>
</tr>
<tr>
<td>14</td>
<td>2.3</td>
</tr>
<tr>
<td>15</td>
<td>2.4</td>
</tr>
<tr>
<td>16</td>
<td>2.2</td>
</tr>
<tr>
<td>17</td>
<td>2.2</td>
</tr>
<tr>
<td>18</td>
<td>2.4</td>
</tr>
</tbody>
</table>

14 and 18). Also, the angles at which the hyoglossus muscles can be seen within the transverse section seem reasonable.

However, when we used this x-only scaling, the angles of the hyoglossus muscle became very odd. Figure 31 compares this result with the result of x-y scaling as described in Section 5.2. It shows a top view (looking from top) of the hyoglossus muscles which were extracted from the transverse slices and represented as filled polygons (patches surrounded by thick black contours). The upper right direction indicates lateromedial and superoinferior orientation. The left panel was the result of the alignment mapping under 2-d (x-y) uniform scaling and the right panel was for scaling in the x direction only. The left panel shows a pattern of the hyoglossus muscle converging to the hyoid bone toward the direction of the midsagittal plane. The y axis from bottom to top indicates a direction from lateral to midsagittal line. Obviously there is something wrong in the drawings. In studying this issue,
Figure 31: The left panel: top view of the hyoglossus muscle under a uniform scaling. The right panel: top view of the hyoglossus muscle under a scaling in x direction only. The slice # 10-18 (as shown) goes from top to bottom along the vertical dimension.

We also compared the measurements for the length of each transverse slice under hypotheses (a) and (b), discussed in the preceding chapter, and an additional measurement of a 200% enlarged xerox copy of the profile figure (figure 5-i in Miyawaki's article, [106]).

In Table 11, N is the slice number of the transverse sections, L1 is the length of the projection of the section onto the midsagittal plane, L2 is the length of the midsagittal cut, and L3 is a measurement that we obtained from a 200% enlarged xerox copy of the profile figure, measuring the extrema-to-extrema distance in the x-direction of each slice. All measurements are in cm. If we now normalize L3 in Table 11, so that the transverse slice 10 (t10) has a length of 10.5 cm in L3, we get Table 12. For example, the midsagittal cut of transverse slice number 18, which has a midsagittal line length of 4 cm in the original drawing T:18, should
Table 11: Lengths of the transverse slices of the Miyawaki data

<table>
<thead>
<tr>
<th>N</th>
<th>L1 (in cm)</th>
<th>L2 (in cm)</th>
<th>L3 (in cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>10.5</td>
<td>10.5</td>
<td>14.3</td>
</tr>
<tr>
<td>11</td>
<td>10.0</td>
<td>9.3</td>
<td>14.2</td>
</tr>
<tr>
<td>12</td>
<td>8.5</td>
<td>6.5</td>
<td>12.2</td>
</tr>
<tr>
<td>13</td>
<td>7.0</td>
<td>6.5</td>
<td>11.0</td>
</tr>
<tr>
<td>14</td>
<td>6.5</td>
<td>6.3</td>
<td>9.7</td>
</tr>
<tr>
<td>15</td>
<td>6.5</td>
<td>5.5</td>
<td>7.4</td>
</tr>
<tr>
<td>16</td>
<td>5.5</td>
<td>4.8</td>
<td>5.4</td>
</tr>
<tr>
<td>17</td>
<td>5.5</td>
<td>4.7</td>
<td>4.7</td>
</tr>
<tr>
<td>18</td>
<td>4.5</td>
<td>4.0</td>
<td>3.7</td>
</tr>
</tbody>
</table>

be 2.7 cm if it had the length shown in the profile figure (relative to t10). This is the reason for us to adopt hypothesis (b) rather than (a).

Table 3 in Chapter V was also computed based on hypothesis (b) for the slice alignment and the uniform (x-y) scaling method.

We see that for the last four slices (numbers 15-18) in Table 3, the scaling parameter computed is comparatively small, and therefore the effect of scaling is large. If this scaling is applied only in the x-direction along the length of the tongue, the angles between the hyoglossus direction and the midsagittal plane (connect the anterior-most point and posterior-most point on the hyoglossus contour by a straight line and consider the angle between this line and the midsagittal line) must be most affected in the last four slices (i.e. lowest). This is not what we find in the relation between the line lengths in the profile figure and the x-direction line.
Table 12: Normalized Lengths (L3) of the transverse slices of the Miyawaki data

<table>
<thead>
<tr>
<th>N</th>
<th>L1 (in cm)</th>
<th>L2 (in cm)</th>
<th>L3 (in cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>10.5</td>
<td>10.5</td>
<td>10.50</td>
</tr>
<tr>
<td>11</td>
<td>10.0</td>
<td>9.3</td>
<td>10.43</td>
</tr>
<tr>
<td>12</td>
<td>8.5</td>
<td>6.5</td>
<td>8.96</td>
</tr>
<tr>
<td>13</td>
<td>7.0</td>
<td>6.5</td>
<td>8.08</td>
</tr>
<tr>
<td>14</td>
<td>6.5</td>
<td>6.3</td>
<td>7.12</td>
</tr>
<tr>
<td>15</td>
<td>6.5</td>
<td>5.5</td>
<td>5.43</td>
</tr>
<tr>
<td>16</td>
<td>5.5</td>
<td>4.8</td>
<td>3.97</td>
</tr>
<tr>
<td>17</td>
<td>5.5</td>
<td>4.7</td>
<td>3.45</td>
</tr>
<tr>
<td>18</td>
<td>4.5</td>
<td>4.0</td>
<td>2.72</td>
</tr>
</tbody>
</table>

lengths in the midsagittal plane of Miyawaki’s data. The angles seem reasonable in the drawings when we use the uniform scaling in the x and y direction.

Even with these optimal choices of alignment and scaling methods, however, there still seems to remain some inconsistency in dimension measures in the Miyawaki data. If we assume that the drawings are roughly to scale (with equal scale), then this profile should have a much thicker (longer) base. In particular, the length of slice t18, which is 1.9 cm in Miyawaki’s profile figure, should be 7.4cm/10.5cm*4.0cm = 2.85 cm according to our interpretation of the transverse drawing (7.4 cm is the length of t10 in profile figure and 10.5 cm is its length in the transverse drawing T:10, 4.0 cm is the length of figure T:18 of the transverse drawing). This inconsistency may reflect some deformation of the specimen encountered during the dissection study, which aimed at as accurate as possible muscle fiber identification as opposed to geometric measurements.
6.2.2 MRI Data

Contrary to the Miyawaki data, the available MRI data provide other surrounding oral cavity structures and useful living tongue shapes (human subjects) of sustained phonetic articulations (vowels). Its insufficiency in anatomical details (e.g. boundaries between muscles) within the tongue posed difficulties in identifying landmarks. This difficulty was primarily caused by the course slice sampling in the vertical position in the MRI data. We have tried to look into other orientations of the same MRI data (i.e. coronal and sagittal data sets of the same subjects). However, they are no better than the transverse data that we used, except that some of coronal slices reveal good distinctions in the areas of septa tissues, which we could use according to our current guidelines to identify the landmarks. The sagittal slices, due to their limited numbers (7 for each subject in total), were not considered for the mapping even though they show good boundaries among different muscles. Aided with the experience gained by the use of the Visible Man data integrated with the Miyawaki data, it may now be possible to make use of available MRI data more effectively. Future work with newly acquired MRI data seems even more promising.

6.2.3 Visible Man Data

The Visible Man data were used as a substitute to the MRI data in developing a general strategy for landmark identification. The Visible Man data have better anatomical details in terms of their distinct boundaries among different muscles
partly due to the smaller slice thickness (1 mm). However, the missing teeth contribute to a rather large deformation of the tongue (see Figure 13). As a result, the midsagittal plane of the tongue is laterally considerably asymmetric. We used the right half of the tongue in the Visible Man data for the mapping. In general, however, the Visible Man data provide very useful anatomical information for our purpose as well as for other studies of the human anatomy.

6.3 Implications of this Work

This dissertation work was focused on two areas. First, to find reliable landmarks for the tongue, we introduced a set of guidelines to identify the homologous points using example data. Second, based on the selected landmark pairs, we developed computational tools for integrating the example data of the tongue as a means to supplement tongue shape data with tissue boundary and fiber direction information of the internal organization of the tongue. In this section, we will discuss possible applications of this computational method in other research.

6.3.1 Application for Morphological Studies of Soft Tissues

The thin-plate spline mapping, which we used in our computational method, serves as a tool for morphological studies in bony structure (such as the human skull) not only to produce an averaged image (Cutting et al., [24]) but also to understand specimen-by-specimen variability (Bookstein, [11]) in the atlases. In addition to

---

2 Most atlas books of the human anatomy do not provide good reference (coordinates) for 3-d reconstruction of these 2-d drawings.
the application to the human tongue, our computational method for integrating images from different sources with different types of information is also applicable to describing other biomedical structures, particularly those comprising highly deformable tissues such as cardiac muscles. The landmark pairs can be defined to measure the structure and deformation of organs such as the human heart. Such information could be useful to identify the difference between normal and pathological hearts. In addition, guidelines for selecting landmarks introduced in this work and the selected landmarks for our tongue specimen could provide anatomists and pathologists a tool for finding differences between groups of data or integrating partial sets of information from different resources (such as MRI and CT scans) to form a complete and useful data set.

Our computational method will also be useful for measuring the lengths of major muscles in the tongue, for example, from MRI data of many individuals (Kiyoshi Honda, personal communication). It is hoped also that this approach could improve the current situation of methodological difficulty in speech production research, especially in EMG studies of articulatory organs, by assisting needle insertion and surface electrode signal interpretation, for example.

### 6.3.2 Applications to Phonetic Studies

As mentioned earlier, while the current work does not contribute to phonetic knowledge with data for various articulations, it was motivated for this purpose, and has provided a substantive methodological preparation for future research in this direction in the form of computational tools and example data. We are interested
immediately in several issues. One pertains to normalization of tongue shapes and surrounding structures in speech production for each specified phonological unit or phonetic gesture, such as a vowel, by individuals speaking the same language. Anatomical normalization is a key issue for studies of phonetic differences among individuals as well as to derive a "normal" pattern to be compared with pathological patterns. Another is comparing articulatory gestures produced by the same subject for different phonological units (e.g., vowels). This has been the focus of studies in segmental phonetics, but no quantitative explanation of observed articulatory variation has been obtained, largely because of the difficulty of comparing different subjects without anatomical normalization. Separating geometric differences of individuals' anatomies from differences in phonetic habits (phonetic disposition, see Fujimura, [39]) has been extremely difficult.

Phonetic and most other muscular studies furthermore require understanding non-stationary dynamics. While dynamic information cannot be obtained for accurate 3-d images of the tongue with the present MRI or CT technology, a computer simulation as proposed in the C/D model (detail in Section 2.5.3) is an ultimate solution when sufficient anatomical data and understanding of speech motor control processes are provided (see Section 6.4 for further discussion). The method explored here hopefully forms a significant step toward quantitatively accurate dynamic simulations in articulatory research. In particular, the C/D model equipped with such biological data will be tested against partial but quantitative data in
different aspects of real-life speech phenomena, as obtained by electromyography, Cineradiography, microbeam, etc.

6.3.3 Application to Biomechanical System Modeling

This dissertation work has developed a new computational method for integrating different 3-d anatomical data of the human tongue for use in a 3-d tongue model. The computational method can integrate different types of data sets including a variety of phonetic gestures (vowels), etc. with detailed fiber direction information into a 3-d tongue structural model on which a finite element mesh is constructed. Simulation of this finite element model will provide a theoretical basis for characterizing subject-to-subject variation in mechanical movement patterns of the organ via the quantitative parametric description of the subjects' anatomy. The same approach can be applied to model other moving organs, such as the cardiac muscular system. Simulation of the finite element mesh generated by this approach can be used to understand how normal and pathological hearts respond to any putative testing condition, and how a partial prothesis would function for an individual case, before surgery.

6.4 Future Work

The original intention of this dissertation work was to use MRI data as a source of tongue shape information for comparison of different vowel gestures by the same subject and the same vowel by different subjects. It is considered as part of the development of a computational model of speech utterances in realistic
conversational situations. Due to the limitations of currently available MRI data (see Section 6.2.2), this work used the Visible Man data instead to develop a general strategy for future MRI (or other) research. We have succeeded in integrating Miyawaki's data, which is extremely deformed in shape, with the Visible Man data, and introduced a strategy to identify a set of landmarks as well as a procedure to evaluate its accuracy (see Section 5.4). This procedure was used to manipulate the landmark pairs until a global matching of the Miyawaki data and the Visible Man data was achieved. Global matching was considered when the transverse projection of each 3-d surface was consistent with the image in the original corresponding transverse slice of the Miyawaki data.

The immediate future work in this direction of research is to include handling of phonetic gestures (e.g. vowels). There is no doubt that we need MRI data designed for this particular purpose. Many researchers are trying to do the job now in different parts of the world, but to date no data are completely adequate for our quantitative comparison of different gestures or different subjects. We need to experiment with different machine parameters and image display conditions to optimize the image clarity for the tongue. In addition, a complete image set of all orientations (i.e. the coronal, the transverse, and the sagittal orientation), as a good source of information is needed for landmark identification and 3-d reconstruction. In particular, a complete set of sagittal slices is needed to assist the identification of landmarks within the tongue. Given that new MRI techniques
are available, a smaller slice thickness (e.g. 1 mm, if possible), rather than 3 mm or 5 mm, with no gap would be desirable for our purpose.

Nevertheless, even before the desirable data for our purpose is made available, now being equipped with our experience with cadaver sample images, we may be able to succeed in mapping 3-dimensionally a selected MRI datum among existing data onto dissection data with muscle fibers identified, in particular, the integrated Visible Man-Miyawaki data as we have obtained. This would provide not only a demonstration of computational means for comparing different subjects and different vowel gestures as discussed above, but also substantive data as the basis for integrating MRI data in forms of muscular structure.

The next step toward extending the current work, is to expand the computational method to obtain a 3-d reconstructed volume from Miyawaki's data and map it as a whole onto the MRI volume data, rather than map individual slices of Miyawaki's data onto the MRI data. This would directly lead to the creation of a finite element model for a 3-d tongue modeling. In fact, this is our next immediate goal. The adaptation of a finite element mesh to the warped anatomical volume created by the expanded method just described involves mapping from an abstract spatial structure, which we call the "topological model", to the master model of the particular individual. The topological model is composed of geometrically simple blocks (hexahedra, tetrahedra, and prisms) which are further subdivided by flat planes. The master model is then anatomically normalized to fit the individual speaker's geometry and used as a reference model in the simulation of the tongue
deformation in speech articulation. The tongue model, together with other articulatory models (e.g. lips, jaw), comprise the main component of the signal generator in the C/D model, which can be used to compute the speech signals at different levels such as muscular contractions (to be compared with the EMG data), tongue and other articulatory organs’ sampled flesh-point data (X-ray microbeam pellets or magnetic coils), patterns of the tongue surface contact with the palate (dynamic palatography), the vocal fold vibration pattern (electroglottography, etc.), and the resulting acoustic signal (microphone signal).

Quantitative and exact identification of the physiological characteristics of voice quality control in different articulations, for instance, can be achieved by using the nonlinear 3-d mapping technique. We may be able to obtain MRI data from trained subjects producing a vowel in different voice qualities. Most probably, however, we will need to combine MRI image information with information obtained from microbeam measurement, since MRI takes a long time to record one gesture, and perhaps can never be used to capture time varying realistic speech utterances. If we do succeed in this plan of comprehensive modeling of speech production processes, this will be a major contribution to our understanding of the nature of speech signals, hitherto almost totally unexplored but crucial for real progress in speech technology in the future.
BIBLIOGRAPHY


APPENDIX A

Program Description and Listing for 2-d Thin-plate Spline Mapping

A.1 Overview

There are two Matlab scripts that were written to implement 2-d thin-plate spline mapping. There are located in the directory /home3/wucm/matlab/tpmp2d in Hattori.

`drawcolorsc.m` is an interactive program that can input the images of Miyawaki's tongue sketches, MRI data, and the Visible Man data and identify regions of tissues in the images with color polygons. The polygons are labeled for classification (i.e., with a color code for muscle type). The areas of the various muscles and the projected fiber directions of the muscles were drawn on the images of Miyawaki's tongue sketches (see section 3.3). The mouse drawn polygons, either a closed or open polygon (line segments and points) for each sketch, were then saved to files for 2-d thin-plate spline mapping. Based on the MRI image display, tongue contours were manually outlined and the resulting polygon data were saved to files. Examples of the resulting polygon
data on both tongue sketches and MRI data were demonstrated in Figures 3(a) and 3(b).

`tpmp2dsc.m` is an interactive program that can read polygon files created by the program `drawcolorsc.m`, defines homologous points, and implements the mapping with 2-d thin-plate spline method. The program can read and save polygon files from MRI data and Miyawaki's data. The homologous point pairs were obtained interactively using the mouse to select points on a display of both polygon files in the same window. Examples of the resulting 2-d thin-plate mapping was shown in Figure 5(a).

A.2 Program Description

This section describes functions that are contained in programs of `drawcolorsc.m` and `tpmp2dsc.m`.

A.2.1 Program Description for `drawcolorsc.m`

The graphic display of the program `drawcolorsc.m` is shown in Figure 32. It has two menus and five control boxes. The `Muscles` menu contain several commands specified by muscular names (e.g. Genioglossus, Hyoglossus etc.). Each command has two different functions (i.e. Make polygon and Edit polygon). The `File` menu have six commands listed as below

Remove Polygon
Save Polygons to File
Load Polygon File
Figure 32: A display produced by the program `drawcolorsc.m`.

Control boxes are located at the bottom and in the upper right hand corner of Figure 32. A list of their functions with initial labels is shown below.

- **Image** = off: Click on this box to toggle displayed image on and off.
- 
- **Polygon** = off: Click to toggle displayed polygons on and off.
- **Muscle** indicates currently chosen muscle's name.

In Figure 32, a polygon data file (T9.dat) and its image file (T9-72dpiF) for Figure 3(b) is shown as an example. In the following, we show how to draw and edit polygons step by step.
• **Draw polygon**

To draw polygon over a muscle (*e.g.* Hyoglossus) with

1. Input the desired image file name in the control box (*e.g.* T9-72dpiF).
2. In the **File** menu, choose the command **Load New Image**, and then choose the command **Load Gray image**.
3. Use the **Muscles** menu to select the command **Hyoglossus** and then choose the command **Make polygon**.
4. Click the left button of the mouse to start drawing a polygon around the pertinent region of the hyoglossus muscle and click the right button to end the drawing.
5. Repeat steps 1-4 for other muscles.
6. Input the desired polygon file name in the control box (*e.g.* T9.dat).
7. Use **File** menu to select the command **Save Polygons to File** to save polygon data to file.

• **Edit polygon**

To edit polygon over a muscle:

1. Repeat steps 1-2 in **Draw polygon**.
2. Input the desired polygon file name in the control box (*e.g.* T9.dat).
3. Use **File** menu to select the command **Load Polygon File**.
4. Click the control box to toggle the image off.
5. Use the Muscles menu to select the command Hyoglossus and then choose the command Edit polygon.

6. Click the left button of the mouse to select the desired point, then move mouse to a new location and click to complete editing the polygon; click the middle button to add a new point to the polygon and click the right button to end the drawing.

7. Click the control box to toggle the image on again.

8. Use the File menu to select the command Save Polygons to File to save polygon data to file.

A.2.2 Program Description for tpmp2dsc.m

The graphic display of the program tpmp2dsc.m is shown in Figure 33. It has two menus and two control boxes. The Mapping menu contains two commands, Find Corres and ThinplateMap. The File menu has four commands:

Remove Polygon
Save Polygons to File
Load Polygon File
Quit.

Control boxes are located at the bottom of Figure 33. A list of their functions with initial labels is shown below.

Polygon=off toggle displayed polygons on and off.
*.dat input polygon file name.
Figure 33: A display produced by the program \texttt{tpmp2dsc.m}.

In Figure 33, a result of 2-d thin-plate spline mapping for Figure 5(a) is shown as an example. In the following, we show how to implement the 2-d thin-plate spline mapping step by step.

- **2-d Thin-plate Spline Mapping**

To implement a 2-d thin-plate spline mapping from tongue sketch polygons onto MRI tongue contour:

1. Input MRI polygon file name in the control box (e.g. \texttt{hat17.dat}).
2. Use the \texttt{File} menu to select the command \texttt{Load Polygon File}.
3. Repeat steps 1-2 for Miyawaki tongue sketch polygon file.
4. Use the \texttt{File} menu to select the command \texttt{Find Corres}.
5. Click the left button of the mouse to select a homologous point in the
tongue sketch polygons and click the left button again to select the corresponding point in the MRI tongue contour; use the right button to end this step.

6. Use the File menu to select the command ThinplateMap to execute the 2-d thin-plate spline mapping.

7. Use the File menu to select the command Save Polygons to File to save the mapped result with the file name in the control box.

### A.3 Program Listing

This section lists source codes of the programs drawcolorsc.m and tpmp2dsc.m.

#### A.3.1 Program Listing for drawcolorsc.m

```matlab
% drawcolorsc.m is an interactive
% program that can input the images of Miyawaki's
% tongue sketches, MRI data, and the Visible Man data
% and identify regions of tissues in the images with
% color polygons. Chao-Min Wu 03/06/96
% drawf=figure('numbertitle','off','name','Drawpolygon Figure',... 
  'colormap',gray(240));
axis equal;
fid = fopen('S1-72dpi','r');
[a, count] = fread(fid,[256,256],'uchar');
close(fid);
imageh=image(a');
musclenames=str2mat(...
  'Genioglossus',... 
  'Genioglossus-post',...
  'Styloglossus',...
  'Hyoglossus',...
  'Transversus',...
`
getfilenameout=[...
'polygonfile=get(filenameeditor,''string'');',...
'disp([''File Named: '' polygonfile]);'];
togglepolygon=
uicontrol('style','checkbox',...
'position',[0.48 0.01 0.20 0.05],'+units',+normalized',...
'backgroundcolor',[0.8 0.8 0.8],'string','Polygon=off',...
'callback',...
['for k=1:nummuscles,','...
  for l=1:nummuscles,','...
    set(lin(k,l),''visible'',''off'');','...
  end;',...
end;',...
end;',...
'if get(togglepolygon,''value'')==0,','...
'for k=1:nummuscles,','...
  for l=1:nummuscles,','...
    set(lin(k,l),''visible'',''on'');','...
  end;',...
end;',...
end;',
getimagefilecallback=[...
'imagefile = get(imagefileeditor,''string'');',...
'disp([''Load image file named: '' imagefile]);',...
delete(imageh);',...
'fid = fopen(imagefile,''r'');',...
'[a, count] = fread(fid,[256,256],''uchar'');',...
'fclose(fid);',...
colormap(gray(240));',...
a=((a/256).^0.5);',...
'imageh=image(a);',...
eval(createobject);'];
getgifcallback=[...
'giffile = get(imagefileeditor,''string'');',...
disp([''Load gif file named: '' giffile]);',...
delete(imageh);',...
'[a,map] = gifread(giffile);',...
colormap(map);',...
'imageh=image(a);',...
'eval(createobject);']
rot90image=[...
'a=rot90(a,3);',...
'imageh=image(a'');',...
'eval(createobject);']
curvmusclemenu = uimenu('Label','Muscles','backgroundcolor',...[0.7 0.7 0.7]);
drawpolygon=[...
disp('''Make Polygon''');',...
'xx = zeros(1,1000);',...
'yy = zeros(1,1000);',...
'but = 0;',...
'for i=1:nummuscles,'
  x = get(lin(curvenumber,i),''xdata'');',...
  if x(1) == 0,'
    last = i;break;',...
  end;',...
'end;',...
'kk =0 ;',...
'while but ~= 3,'
  kk = kk+1;...
  [xx(kk),yy(kk),but] = ginput(1);',...
    if but == 1,'
        if kk == 1,'
            lin(curvenumber,last) = line(xx(1),yy(1),',
                ''color'' ',''r'',''erasuremode'',''none'');',...
        else,'
            set(lin(curvenumber,last),''xdata'',xx(1:kk),',
                ''ydata'',yy(1:kk));',...
        end;',...
    elseif but == 2,'
        disp(''Not Used'');',...
    elseif but == 3,'
        xx(kk) = xx(1);',...
        yy(kk) = yy(1);',...
        set(lin(curvenumber,last),''xdata'',xx(1:kk),',
            ''ydata'',yy(1:kk));',...
    end;',...
end;

set(lin(curvenumber,:), 'color', musclecolors(curvenumber));

curmmusshow = uicontrol('style', 'pushbutton',
'position', [0.75 0.96 0.25 0.05], 'units', 'normalized',
'backgroundcolor', 'g', 'string', 'Muscle');

curveMenu = [...
disp(curvenumber);

set(curmmusshow,'string',musclenames(curvenumber,:));

curcuf=num2str(curvenumber);

eval(drawpolygon);

editpolygon= [...
disp('Edit Polygon');

set(curmmusshow,'string',musclenames(curvenumber,:));

xe = zeros(nummuscles,50);

ye = zeros(nummuscles,50);

for i=1:nn

    x = get(lin(curvenumber,i),'xdata');

    len = length(x);

    xe(i,1:len) = get(lin(curvenumber,i),'xdata');

    ye(i,1:len) = get(lin(curvenumber,i),'ydata');

    if xe(i,1) == 0,

        des = i-1;

    break;

    end;

end;

mb = 1;dist=3;

while mb ~= 3,
    [xm,ym,mb]= ginput(1);
    if mb == 3, return;end;
    kclose = 0;
    nline = 0;
    lon = 1;
    for i=1:des,
        len = length(xe(i,:));
        lenl = len;
        for k=1:lenl,
for k = 1:nummuscles,
curn = int2str(k);
muscle(k) = uimenu(curvemusclemenu,'Label',musclenames(k,:),...  
'backgroundcolor',[0.7 0.7 0.7], 'foregroundcolor',...  
musclecolors(k));
uimenu(muscle(k),'Label','Make Polygon',...  
'callback',['curvenumber=' curn '; eval(curvemenu);']);
uimenu(muscle(k),'Label','Edit Polygon',...  
'callback',['curvenumber=' curn '; eval(edtpolygon);']);
end;
curvedefinemenu = uimenu('Label','File','backgroundcolor',...  
[0.8 0.8 0.8]);
savepolygonfile=[...  
'disp(''Save Polygon to'');','...  
'eval(getfilenameout);','...  
'fido=fopen(polygonfile,''w'');','...  
'for k=1:nummuscles,'...
  '  for l=1:nummuscles,'...
    'x=get(lin(k,l),''xdata'');','...  
    'y=get(lin(k,l),''ydata'');','...  
    'le=length(x);','...  
    'fprintf (fido, ''%.s %d %d\n'', musclenames(k,:),k,le);','...  
    'for i=1:le, fprintf(fido,''%.5f %.2f\n'',x(i),y(i));','...  
      '  end;',
      '  end;',
    'end;',
'end;'];
loadpolygonfile=[...  
'disp(''Load Polygon from'');','...  
'eval(getfilenameout);','...  
'eval(createobject);','...  
'fidi=fopen(polygonfile,''r'');','...  
'if fidi <0 ',
  'disp([''File '' polygonfile '' could not be opened'']);',
else ',
  'for k=1:nummuscles,'
    '  for l=1:nummuscles,'
      'goof = fscanf(fidi,''%.s'',1);',
...
'le = fscanf(fidi,'''d''',1);',...
'le = fscanf(fidi,'''d''',1);',...
'xi=zeros(1,le); yi=zeros(1,le);',...
'xymat=fscanf(fidi,'''f %f''',[2,le]);',...
'set(lin(k,1),''xdata'',xymat(1,:),''ydata'',xymat(2,:),',...
' ''color'',musclecolors(k),''visible'',''on'');',...
' end;',...
'end;',...
'fclose(fidi);',...
'end;'];
removepolygon=[...
'eval(deleteobject);',...
'eval(createobject);'];
shutdown = [...
'close;'];
uimenu(curvedefinemenu,'Label','Remove Polygons',...
'backgroundcolor',[0.8 0.8 0.8],'callback',...
'eval(removepolygon)','accelerator','d');
uimenu(curvedefinemenu,'Label','Save Polygons To File',...
'backgroundcolor',[0.8 0.8 0.8],'callback',...
'eval(savepolygonfile)','accelerator','s');
uimenu(curvedefinemenu,'Label','Load Polygon File',...
'backgroundcolor',[0.8 0.8 0.8],'callback',...
'eval(loadpolygonfile)','accelerator','p');
loadimageh=uimenu(curvedefinemenu,'Label','Load New Image',...
'backgroundcolor',[0.8 0.8 0.8]);
uimenu(loadimageh,'Label','Load Gray Image',...
'callback',eval(getimagefilecallback));
uimenu(loadimageh,'Label','Load Gif Image',...
'callback',eval(getgifcallback));
uimenu(curvedefinemenu,'Label','CCW Rot90 Image',...
'backgroundcolor',[0.8 0.8 0.8],'callback',...
'eval(rot90image)','accelerator','r');
uimenu(curvedefinemenu,'Label','Quit','backgroundcolor',[0.8 0.8 0.8],
'callback',eval(shutdown)','accelerator','q');
A.3.2 Program Listing for tpmp2dsc.m

% tpmp2dsc.m is an interactive program that can read polygon
% files created by the program drawcoloresc.m, defines
% homologous points and implements the mapping with
% 2-d thin-plate spline method. Chao-Min Wu 03/06/96
%
tpmp2d=figure('numbertitle','off','name','2-D Thin-Plate Mapping');
set(gcf,'units','pixel');
axis equal;
musclenames=str2mat(...
'Genioglossus',...
'Genioglossus-post',...
'Styloglossus',...
'Hyoglossus',...
'Transversus',...
'Verticalis',...
'Longitudinalis-inf.',...
'Longitudinalis-sup.',...
'Digastric-Ant.',...
'Outline',...
'Bone-Cartilage');
musclecolors='bbcgmryyckk';
ummuscles = length(musclenames(:,1));
createobject=[...
'for k =1:nummuscles,'
'  for l =1:nummuscles,'
'    lin(k,l) = line([0 1],[0 0],'visible','off');','
'  end;',...
'end;'];
eval(createobject);
deleteobject=[...
'for k =1:nummuscles,'
'  for l =1:nummuscles,'
'    delete(lin(k,l));','
'  end;',...
'end;'];
filenameeditor = ...
uicontrol('style','edit',...
'position',[0.62 0.01 0.30 0.05], 'units', 'normalized', ...
'backgroundcolor','y','string','*.dat');
getfilenameout=[...
'polygonfile=get(filenameeditor, 'string');','...
'disp([''File Named: '' polygonfile']);'];
togglepolygon=
uitable('style','checkbox',...
'position',[0.40 0.01 0.19 0.05], 'units', 'normalized', ...
'backgroundcolor',[0.8 0.8 0.8], 'string','Polygon=off', ...
'callback', ...
['for k=1:nummuscles,','...
' for l=1:nummuscles,','...
' set(lin(k,l),''visible'',''off'');','...
' end;', 'end;','...
'if get(togglepolygon, 'value')==-0,','...
'for k=1:nummuscles,','...
' for l=1:nummuscles,','...
' set(lin(k,l),''visible'',''on'');','...
' end;', 'end;end;']);
mappingmenu = uimenu('Label','Mapping','backgroundcolor',...
[0.7 0.7 0.7]);
correspondence=[...
'P=[];','....
'H=[];','....
'bt = 1;','....
'plin=[];','....
'hlin=[];','....
'phlin=[];','....
'count=0;','....
'while bt < 2,','....
' count = count+1;','....
' [x,y,bt] = ginput(1);','....
' P(count,1) = x; P(count,2) = y;','....
' plin(count)=line(x,y,'color','','m','linestyle','','o');','....
' [x,y,bt] = ginput(1);','....
' H(count,1) = x; H(count,2) = y;','....
' hlin(count)=line(x,y,'color','g','linestyle','x');',...
' phlin(count)=line([P(count,1), H(count,1)],',...
' [P(count,2), H(count,2)],'color','r','linestyle','...',
' ...');',...
'end;'];
thinplatemap=[...
'K = length(P(:,1));',...
'LL = zeros(K+3,K+3);',...
'for k=1:K',...
' for l = k+1:K, ...
' x = P(k,1)-P(l,1);',...
' y = P(k,2)-P(l,2);',...
' r2 = x^2+y^2;',...
' LL(k,l) = 0.5*r2*log(r2);',...
' end;',...
' end;'];

'LL(1:K,K+1) = ones(K,1);',...
'LL(1:K,K+2:K+3) = P;',...
'LL = LL+LL'';',...
'Y = [H;[0 0];[0 0];[0 0]];',...
'W = inv(LL)*Y;',...
'for k=1:nummuscles',...
' for l=1:nummuscles, '
' x=get(lin(k,l),'xdata');',...
' y=get(lin(k,l), 'ydata');',...
' M = length(x);',...
' r2 = (P(:,1)*ones([1,M]) - ones([K,1]) * x).^2 + ',...
' (P(:,2)*ones([1,M]) - ones([K,1]) * y).^2 ;',...
' r2= 0.5*r2.*log(r2);',...
' Y = zeros(size([x' y']));',...
' Y(1:M,1)= sum(r2 .* (W(1:K,1)*ones([1,M]))'' + ',...
' W(K+1,1) *ones([M,1])+W(K+2,1)*x''+W(K+3,1)*y'';',...
' Y(1:M,2)= sum(r2 .* (W(1:K,2)*ones([1,M]))'' + ',...
' W(K+1,2) *ones([M,1])+W(K+2,2)*x''+W(K+3,2)*y'';',...
' set(lin(k,l),'xdata',Y(:,1),'ydata',Y(:,2),'...',
' 'color',musclecolors(k),'visible','on');',...
' end;',...
'end;'];

'K = length(P(:,1));',...
'LL = zeros(K+3,K+3);',...
'for k=1:K',...
' for l = k+1:K, ...
' x = P(k,1)-P(l,1);',...
' y = P(k,2)-P(l,2);',...
' r2 = x^2+y^2;',...
' LL(k,l) = 0.5*r2*log(r2);',...
' end;',...
' end;'];
thinplatemap=[...
'K = length(P(:,1));',...
'LL = zeros(K+3,K+3);',...
'for k=1:K',...
' for l = k+1:K, ...
' x = P(k,1)-P(l,1);',...
' y = P(k,2)-P(l,2);',...
' r2 = x^2+y^2;',...
' LL(k,l) = 0.5*r2*log(r2);',...
' end;',...
' end;'];
thinplatemap=[...
uimenu(mappingmenu,'Label','Find Corres.','backgroundcolor',... 
[0.8 0.8 0.8], 'callback', 'eval(correspondence)'); 

uimenu(mappingmenu,'Label','ThinplateMap','backgroundcolor',... 
[0.8 0.8 0.8], 'callback', 'eval(thinplatemap)');

curvedefinemenu = uimenu('Label', 'File', 'backgroundcolor',... 
[0.8 0.8 0.8]);

savepolygonfile=[...
'disp(''Save Polygon to'');',...
'eval(getfilenameout);',...
'fido=fopen(polygonfile,''w'');',...
'for k=1:nummuscles,',...
'  for l=1:nummuscles,',...
'  x=get(lin(k,l),''xdata'');',...
'  y=get(lin(k,l),''ydata'');',...
'  le=length(x);',...
'  fprintf(fido,''/%s %d %d\n'',musclenames(k,:),k,le);',...
'  for i=1:le, fprintf(fido,''%.5f %.2f\n'',x(i),y(i));end;',...
'  end;',...
'end;',...
'fclose(fido);',...
'end;']);

loadpolygonfile=[...
'disp(''Load Polygon from'');',...
'eval(getfilenameout);',...
'eval(createobject);',...
'fidi=fopen(polygonfile, ''r'');',...
'if fidi <0',...
'disp([''File '' polygonfile '' could not be opened'']);',...
'else',...
'for k=1:nummuscles,',...
'  for l=1:nummuscles,',...
'    goof = fscanf(fidi, ''%s'',[1]);',...
'    le = fscanf(fidi, ''%d'',[1]);',...
'    le = fscanf(fidi, ''%d'',[2,1]);',...
'    xymat=fscanf(fidi, ''%f %f'',[2,le]);',...
'    set(lin(k,l),''xdata'',xymat(:,1),''ydata'',xymat(:,2),',...
'    ''color'',musclecolors(k),''visible'',''on'');',...
'  end;',...
fclose(fidi);
end;

removepolygon=[...
cla;

eval(createobject);
end;

shutdown = [...
close;

uimenu(curvedefinemenu,'Label','Remove Polygons',... 
'backgroundcolor',[0.8 0.8 0.8],'callback',... 
'eval(removepolygon)','accelerator','d');

uimenu(curvedefinemenu,'Label','Save Polygons To File',... 
'backgroundcolor',[0.8 0.8 0.8],'callback',... 
'eval(savepolygonfile)','accelerator','s');

uimenu(curvedefinemenu,'Label','Load Polygon File',... 
'backgroundcolor',[0.8 0.8 0.8],'callback',... 
'eval(loadpolygonfile)','accelerator','p');

uimenu(curvedefinemenu,'Label','Quit','backgroundcolor',
[0.8 0.8 0.8],'callback','eval(shutdown)','accelerator','q');
APPENDIX B

Program Description and Listing for 3-d Thin-plate Spline Mapping

B.1 Overview

There are several Matlab scripts and functions that were written to implement 3-d thin-plate spline mapping. These are located in the directory

/home3/wucm/matlab/tpmp3d/.

Newtracetongue.m is another interactive program that can input the GIF images of Miyawaki’s tongue sketches and identify regions of tissues in the images with color polygons and patches. The areas of the muscles and their projected fiber directions have been drawn on the images of Miyawaki’s tongue sketches. There are several different objects generated by this program, e.g. lines that represent outlines or skin surface (object type 1) and patches that represent regions of muscles or other tissue (object type 2). The object type is specified either by menu selection (for types 1 and 2), or by means of mouse button. The mouse drawn polygon objects for each sketch were then saved into files for later use in the alignment transformation. An example of the resulting polygon data file is shown in Figure 6(b).
New2Colorcut.m is a program that can interactively display sections of the Visible Man data in three different orientations (i.e. sagittal, coronal, and transverse). This program also can trace regions of different types of tissue with polygons over a section of the Visible Man data and save polygon files for later use (used in locating landmark #2).

profilealigngra.m is used to produce parameters (Table 3) and result (Figure 21) for alignment transformation.

loadsketch.m is the program to display an image of the Miyawaki data and specify the $x, y$ coordinates of the selected landmark point on the image.

loadman.m is the program to display an image of the Visible Man data and specify the $x, y$ coordinates of the selected landmark point on the image.

alignPgra.m is used to transform the coordinates of landmark points (selected from Miyawaki's data by loadsketch.m) based on the parameters of the alignment transformation.

profilealigntpmp.m is used to warp the aligned polygon files of the Miyawaki data, as shown in Figure 21, with the 3-d thin-plate splines based on the selected landmark pairs. The mapped result is shown in Figure 22.

tpmpaligngridgra.m is the program that evaluates the sensitivity to incomparable slice level and the resulting accuracy of mapping as described in Section 5.4.
B.2 Program Description

B.2.1 Program Description for Newtracetongue.m

The graphic display of the program Newtracetongue.m is shown in Figure 34. It has three menus and six control boxes. The Tissue type menu contains several commands specified by muscular names (e.g. Genioglossus, Hyoglossus etc.). Each command can produce different object types by pressing different mouse buttons (i.e. the left button for patch object, the middle button for direction object, and the right button for point object). The Images menu is used for displaying transverse image of Miyawaki's data. The Visibility in Fig.3 menu is useful only when the program profilealigngrap.m is implemented to create the third window containing the aligned polygon files of the Miyawaki data, as shown in
Figure 21. The **Visibility** in **Fig. 3** menu is used for switching the visibility of polygon object for individual tissue on and off.

Control boxes are located at the bottom of Figure 34. Clicking the left button of the mouse inside a particular control box execute its corresponding function listed below.

- **Reset Stack** resets all variable values to zero.
- **DELETE OBJECT** deletes last entered object.
- **Quit** ends the program and closes all the displayed windows.
- **SAVE** saves the drawn polygon objects to a file.
- **LOAD** inputs the corresponding polygon data file on the 2nd window.
- **Pt.1** indicates currently chosen point.

In Figure 34, an example of Miyawaki's GIF image (t9.gif) is shown. Its corresponding polygon file (t9.graph) is shown in the right panel of Figure 6(b). In the following, we show a step by step description on how to produce a polygon file for a Miyawaki image.

### • Make polygon object file

To make a polygon object file for an image *(e.g. t9.gif)*:

1. Use the **Images** menu to select the command **t9.gif**.
2. Use the **Tissue type** menu to select the command **Genioglossus**.
3. For outlining the region the left button of the mouse is used. Polygon points are gathered as long as the mouse button is pressed down and moved. If the muscle region is defined, the user can use the middle button of the mouse and draw dashes along the fibers. These will be registered as 2-point...
Figure 35: A display produced by the program `New2Colorcut.m`.

lines between the beginning point and the end point of the dashes. Marking points (for alignment etc.) are defined by depressing the right button of the mouse and releasing it. These objects are displayed in the second window.

4. Repeat steps 1-3 for other tissues.

5. Click the left button of the mouse in the SAVE control box to save the polygon object file (i.e. `t9.graph`).

B.2.2 Program Description for New2Colorcut.m

The graphic display of the program `New2Colorcut.m` is shown in Figure 35. It has two menus and six control boxes. The Tissue type menu contains several commands specified by muscular names (e.g. Genioglossus, Hyoglossus etc.). Each command can produce patch object on the second window by clicking the middle
button around the selected tissue and using the right button of the mouse to end this function. The **File** menu has two different commands. The **Save Patch to File** command is used for saving patch objects from memory stack to a file. The **Define Midsag Points** command is used for defining midsagittal points either on the surface of or inside the tongue in the midsagittal plane by clicking the right button of the mouse. Clicking the left button of the mouse is used for selecting a cutting location in a display on any (sagittal, coronal, or transverse) plane. Cross lines indicate the cutting locations for the other two displays and the number shows the cutting location in pixel value.

Control boxes are located at lower right side of Figure 35. Clicking the left button of the mouse inside a particular control box executes its corresponding function (except the control box *.dat) listed below.

- ***.dat** enters polygon file name.
- **del obj** deletes last entered object.
- **Quit** ends the program and closes all the displayed windows.
- **Store** saves the drawn patch object to memory stack.
- **Reset** resets all variable values to zero.
- **Carved-out Display** shows a 3-d carved-out display.

In the following, we show a step by step description on how to produce a polygon file for an image in the Visible Man data.
• Make polygon object file

To make a polygon object file for an image:

1. Click the left button of the mouse to select the desired image.
2. Click the left button of the mouse in the Reset control box to reset all the variables.
3. Use the Tissue type menu to select a command, e.g. Hyoglossus.
4. Click the left button of the mouse to start drawing polygon around the pertinent region, e.g. hyoglossus muscle, and click the right button to end the drawing. The patch object is displayed in the second window.
5. Click the left button of the mouse in the Store control box to save the drawn object to memory stack.
6. Repeat steps 2-5 for other tissues.
7. Use the File menu to select the Define Midsag Points command to define midsagittal points, if needed.
7. Enter the desired file name in the editable *.dat control box.
8. Use File menu to select the Save patch to File command to save patch object to the entered file name.

Other programs described in Section B.1 will be discussed in Appendix C since they all represent their own single function. However, their source codes will be included in the next section.
B.3 Program Listing

This section lists source codes of the programs described in Section B.1.

B.3.1 Program Listing for Newtrace tongue.m

% Newtrace tonguen. A program for tracing the Miyawaki drawings.
% Needs subroutine:
% Either gifread function in image processing toolbox, or the
% subroutine readsgif.m without an image processing toolbox
% license. If you have a license set the variable
% haveimageslicense to 1 down in the script.
% Further subroutines needed: safefigstack.m, add2menu.m, and
% loadfigstack.m
% We have now several different objects:
%
%1: Lines that represent outlines or skin surface
%2: Patches that represent regions of muscles or other tissue
%3: Fiber Directions (shown as points with a line)
%4,5,6: Points (reference points for alignment transformation)
% The object type is specified either by menu selection
% (for type 1 and 2), or by means of mouse button. If button 2
% is pressed, muscle fiber directions are drawn and if button 3
% is pressed, reference points are defined. Typically, the user
% selects a tissue type (such as Genioglossus) and defines a
% region for the muscle in figure 1. For outlining the region
% the button 1 of the mouse is used. Polygon points are gathered
% as long as the mouse button is pressed down and moved. If the
% muscle region is defined, the user can use button 2 and draw
% dashes along the fibers (tangentially). These will be registered
% as 2-point lines between the beginning point and the end point
% of the dashes. Marking points (for alignment etc) are defined
% by depressing the mousebutton number 3 and releasing it.
%
% When the save button is pushed, the objects are saved. It is
% recommended to reset the stack before a new image is processed.
% The data corresponding to one image are saved in one file.
%
% All userdata of objects need to contain
% a) number of userdata
% b) size of the image where the coordinates where obtained
% c) slice number (e.g. 6 for slice number 6)
% d) plane type (1: Sagittal, 2: Coronal, 3: Transverse)
% e) tissue type number (e.g. 1 for Genioglossus muscle)
% f) curve type number for a line (1), a patch (2),
% fiber direction (3) or a reference point (4,5,6).

% put here your own colorlist:
colorlist = [...
0.5979 0.0983 0.0514;
0.6672 0.2904 0.1029;
0.7152 0.4642 0.1361;
0.6787 0.6854 0.1854;
0.5310 0.7121 0.2121;
0.3789 0.7357 0.2357;
0.3866 0.7703 0.4290;
0.3967 0.8024 0.5847;
0.4099 0.8404 0.7910;
0.4194 0.7258 0.8043;
% 0.4379 0.4916 0.8293;
% 0.5094 0.4351 0.8412;
% 0.6588 0.4557 0.8526;
% 0.8081 0.4643 0.8846;
% 0.9784 0.4784 0.8545;
% 0.9866 0.4866 0.7403;
0.6962 0.8056 0.8056;
0.9132 0.9444 0.9444;
0.8637 0.7968 0.6424;
0.0 0.0 1.0;
0.9688 0.9800 0.7000;
0.9688 0.8000 0.7000;
0.0 0.0 1.0;
0.0 0.8 0.0;
0.9 0.3 0.3];
closecurve=1;nfig =1;nshowfig=2;maxnpts=1000;
figure(1); colormap([1 1 1; 0 0 0]);
figure(2); figure(1); mindis=0.0001; colorstr = 'ymcrgbk';
ncol=1; stacklen=500;
planetype = 3; % 1: Sagittal, 2: Coronal, 3: Transverse;
zval = 0;
xtransl = 0; ytransl = 0; % translation
dpi = 200.0; inch=2.56; % dots per inch
set(1, 'units', 'normalized');
clearlastline = ['if stackptr>0, ',
' figure(1); if stackfig1(stackptr)~=0, ',
' delete(stackfig1(stackptr));end;',
' figure(2); delete(stackfig2(stackptr));',
' stackptr=stackptr-1; ',
' figure(1);',
' end;',
']；

preset = [...
' poso = get(gca,'position');[sizex,sizex] = size(X);'];
% assume X contains image
gifpath = '/usr/users32/wucm/matlab/gif/';
saveall = [...
'save drawings ndat pointer objtype',...
'tissuetype objsize xdata ydata zdata ;'];
loadall = 'load drawings;';
excludenonimage= ['kind=get(gca,''children'');',
' for k = 1:length(kind),',
' if strcmp(get(kind(k),''type''),''image'')==0,',
' delete(kind(k)); end;',
' end;'];
resetStack = [...
' stackptr=0; stackfigl=zeros(stacklen,1);',
' stackfig2=zeros(stacklen,1);figure(2);',
' eval(excludenonimage);figure(1); eval(excludenonimage);'];
eval(resetStack);
savestack = 'safefigstack(stackfig2,stackptr,fil);';
loadstack = [...
' figure(2);',
' [stackfig2, stackptr]=loadfigstack(fil,stacklen,colorlist);',
' stackfig1=zeros(stacklen,1);'];
shutdown = [...
 'close all;';...
];
ui = uicontrol('position',[0.01 0.01 0.2 0.04],'units',...
 'normalized', 'style','pushbutton', 'string','Reset Stack',...
 'callback','eval(resetStack);');
ui = uicontrol('position',[0.21 0.01 0.2 0.04],'units',...
 'normalized', 'style','pushbutton', 'string','DELETE OBJ',...
 'callback','eval(clearlastline);');
ui = uicontrol('position',[0.41 0.01 0.18 0.04],'units',...
 'backgroundcolor','r ','style','pushbutton', 'string','Quit',
 'callback','eval(shutdown);');
ui = uicontrol('position',[0.59 0.01 0.2 0.04],'units',...
 'backgroundcolor','g ','style','pushbutton', 'string','SAVE',
 'callback','eval(savestack);');
ui = uicontrol('position',[0.79 0.01 0.2 0.04],'units',...
 'backgroundcolor','g ','style','pushbutton', 'string','LOAD',
 'callback','eval(loadstack);');
toggleno = [...
 'alignpointnr = alignpointnr+1;';...
 'if alignpointnr>maxalignpoints, alignpointnr=1; end;','...
 'sprit = sprintf("Pt. %i",alignpointnr);','...
 'set(alignsui,'string'.sprit);','...
];
% ui to control alignment point number which is added
% to the userdata of the defined points.
alignsui = uicontrol('position',[0.85 0.05 0.08 0.03],...
 'units', 'normalized', 'backgroundcolor', 'y', 'style',...
 'pushbutton', 'string', 'Pt. 1', 'callback', 'eval(toggleno);');
alignpointnr=1;maxalignpoints =2;matlist = [];
pulld = uimenu('label','Tissue type');
matlist=add2menu(matlist,pulld,'Genioglossus', [colorlist(1,:),2]);
matlist=add2menu(matlist,pulld,'Geniohyoid', [colorlist(2,:),2]);
matlist=add2menu(matlist,pulld,'Hyoglossus', [colorlist(3,:),2]);
matlist=add2menu(matlist,pulld,'Mylohyoid', [colorlist(4,:),2]);
matlist=add2menu(matlist,pulld,'Digastric', [colorlist(5,:),2]);
matlist=add2menu(matlist,pulld,'Sup. Longitudinalis',...
 [colorlist(6,:),2]);
matlist\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Inf. Longitudinalis'}),...
\text{colorlist}(7,:)\text{,}2)\};
\text{matlist}\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Verticalis}', \text{colorlist}(8,:)\text{,}2)\};
\text{matlist}\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Transversalis}',\text{colorlist}(9,:)\text{,}2)\};
\text{matlist}\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Styloglossus'}),...
\text{colorlist}(10,:)\text{,}2)\};
\text{matlist}\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Bone}', \text{colorlist}(11,:)\text{,}2)\};
\text{matlist}\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Tooth}', \text{colorlist}(12,:)\text{,}2)\};
\text{matlist}\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Skin Cut}', \text{colorlist}(13,:)\text{,}1)\};
\text{matlist}\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Outline}', \text{colorlist}(14,:)\text{,}1)\};
\text{matlist}\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Gland}', \text{colorlist}(15,:)\text{,}2)\};
\text{matlist}\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Fat}', \text{colorlist}(16,:)\text{,}2)\};
\text{matlist}\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Align up}', \text{colorlist}(17,:)\text{,}4)\};
\text{matlist}\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Align dwn}', \text{colorlist}(18,:)\text{,}5)\};
\text{matlist}\text{=}\text{add2menu}(\text{matlist},\text{pulld},'\text{Align Cut}', \text{colorlist}(19,:)\text{,}6)\};
h1\text{=}\text{str2mat}(\ldots
'\text{Genioglossus}', \ldots
'\text{Genichyoid}', \ldots
'\text{Hyoglossus}', \ldots
'\text{Mylohyoid}', \ldots
'\text{Digastric}', \ldots
'\text{Sup. Longitudinalis}', \ldots
'\text{Inf. Longitudinalis}', \ldots
'\text{Verticalis}', \ldots
'\text{Transversalis}', \ldots
'\text{Styloglossus}');
h2 = \text{str2mat}(\ldots
'\text{Bone}', \ldots
'\text{Tooth}', \ldots
'\text{Skin Cut}', \ldots
'\text{Outline}', \ldots
'\text{Gland}', \ldots
'\text{Fat}', \ldots
'\text{Align up}', \ldots
'\text{Align dwn}', \ldots
'\text{Align Cut}');
\text{Tissuename} = \text{str2mat}(h1,h2);
\text{kind} = \text{get}(\text{pulld},'\text{children}');
for k=1:length(kind),
    cb = get(kind(k),'callback');
    cb = [cb, ' curvetype=typedata(4);'];
    set(kind(k),'callback',cb);
end;
%The following command are used for monochrome monitor.
%safematlist = matlist;
%if length(matlist) ~= 0 ,
%    matlist = zeros(size(matlist));
%end;
imamenu = uimenu('label',  'Images'); 'backgroundcolor','g') ;
% The below was changed to reading .sgif files which is a format % invented on the spot to avoid calling gifread for which I don't % have a licence for the images toolbox. % This can be easily changed back if the license becomes available. %

haveimageslicense = 1; %0 without an image processing toolbox
if haveimageslicense == 1,
    showpict = [...
'figure(1); [X,cmap] = gifread(fil); image(X);',...
'title(fil); eval(preset);'];
for k=1:18,
    str = sprintf ('t%i.gif',k);
    fil = [gifpath,str];
    zstr=sprintf ('%i', 18*3-3*k);
    pl = sprintf ('planeno=%i;',k);
    uimenu(imamenu,'label',str,'callback',['fil=',fil,''],
      'xtransl=0;ytransl=0;zval=',zstr,';',pl,' eval(showpict);'])
end;
else % No license:
    showpict = [...
'figure(1); X = readsgif(fil);image(X);title(fil);',...
'eval(preset);'];
for k=1:18,
    str = sprintf ('t%i.sgif',k);
    fil = [gifpath,str];
    zstr=sprintf ('%i',18*3-3*k);
    pl = sprintf ('planeno=%i;',k);
uimenu(imamenu,'label',str,'callback',['fil=''','fil,''''';'],
'xtransl=0;ytransl=0;zval='''zstr,''';',pl,' eval(showpict);']);
end;
end;

% Install an additional menu for controlling visibility
% in figure 3.
figure(1);
visitoggle = uimenu('label','Visibility in Fig.3',...
'backgroundcolor','w');

for k=1:length(Tissuename(:,1)),
  selcommand = sprintf('what = %i;',k);
  uimenu(visitoggle,'label',Tissuename(k,:),...
    'backgroundcolor',colorlist(k,:),'callback',...
    [selcommand, 'togglevisible(3,what);figure(1);']);
end;

uimenu(visitoggle,'label','Toggle all','backgroundcolor','w',...
[1 1 1], 'callback','togglevisible(3,0); figure(1);');
uimenu(visitoggle,'label','All visible','backgroundcolor','w',...
[1 1 1], 'callback','togglevisible(3,-1); figure(1);');
uimenu(visitoggle,'label','None visible','backgroundcolor',...
[1 1 1], 'callback','togglevisible(3,-2); figure(1);');

addpoint=[...
'val=get(nfig,''CurrentPoint'');',...'
'if npts>0,',...'
  'dis = (val(1)-xx(npts))^2+(val(2)-yy(npts))^2;',...'
  'if dis>mindis, ',...'
    'npts=npts+1; if npts>maxnpts, npts=maxnpts; end;',...'
    'xx(npts) = val(1); yy(npts)=val(2);',...'
  'end;',...'
'else ',...'
  'npts=1;xx(npts) = val(1); yy(npts)=val(2);',...'
'bb = get(nfig,''selectiontype'');',...'
'if strcmp(bb,''normal''), bb = 1;',...'
'elseif strcmp(bb,''extend''), bb = 2;',...'
'elseif strcmp(bb,''alt''), bb = 3;end;',...'
'if curvetype==2, closecurve=1; else closecurve=0; end;',...'
'if bb==2, curvetype=3; closecurve=0; end;',...
startrun = [...
'npts = 0;',...
'xx=zeros(maxnpts,1);',...
'yy=zeros(maxnpts,1);',...
'eval(addpoint);',...
'set(nfig,''WindowButtonMotionFcn'',''eval(addpoint);'');'];
finalrun=[...
'set(nfig,''WindowButtonMotionFcn'','' ','); ',...
'if closecurve==1, npts=npts+1; xx(npts)=xx(1);',...
'yy(npts)=yy(1);end; zcm=ones(1,npts)*zval;',...
xpix = (xx(1:npts)-poso(1))*sizex/poso(3)+1;',...
'ypix = (yy(1:npts)-poso(2))*sizey/poso(4);',...
xcm = xpix*inch/dpi+xtransl;',...
ycm = ypix*inch/dpi+ytransl;',...
'showcurve=line(xpix,sizey-ypix,''color'',patchcolor);',...
'userdat = [sizex,sizey,planeno,planetype,nitem,curvetype];',...
'figure(2);',...
'if curvetype==1, ',...
'lastgraphobj =line(xpix,ypix,zcm,''color'',patchcolor,'' ',
''linewidth'',3,''userdata'',userdat);'',
'elseif curvetype==2, ',...
'lastgraphobj=patch(xpix,ypix,zcm,patchcolor,''userdata'',userdat)',
'elseif curvetype==3, ',...
'lastgraphobj=line([xpix(l);xpix(npts)],[ypix(l);ypix(npts)],[zcm(l);zcm(npts)],'',
''color'',''w'',''userdata'',userdat);'',
'elseif curvetype>=4, ',...
'lastgraphobj=line(xpix(1),ypix(l),zcm(l),''linestyle'',
''o'',''markersize'',5,''linewidth'',3, '',
''color'',patchcolor,''userdata'',userdat),
[userdata,alignpointnr]);'',
'tx = sprintf(''%i'','alignpointnr);'',
'text(xpix(1)+10,ypix(l)+10,zcm(l),tx,''color'',patchcolor);'',
'eval(toggleno);'',
'end;'...
B.3.2 Program Listing for New2Colorcut.m

% New Display interactively sections of the Visible Man Data.  
% Colorversion  
% Author: Reiner Wilhelms-Tricarico, M.I.T. RLE. 1995  
% This program assumes that the images have been loaded  
% into memory.  
%  
% Required Predefined Variables:  
% Cindex: Array of size (im_m*im_n,im_no) containing color indices.  
% map: global color map  
% im_no: number of images  
% im_m: vertical format of images (y length)  
% im_n: horizontal format of images (x length)  
%  
% Data file "newcolorindex.mat" contains variables mentioned  
% above.  
% Functions for tracing polygon over Visible Man Data with  
% different color were added by Chao-Min Wu 10/13/95  
% Needs subroutine: addtomenum, cubedesign.m, and getbpoints.m
figure(1); clf;
Cindextrans = Cindex';  
% memory hog but 5 times faster  
% to create coronal sections.  
set(1,'numbertitle','off','name','Visible Man Sections');
colormap(map);
global nitem lastgraphobj patchcolor closecurve;
global midsagcurvepoints midsagplanepoints;
closecurve=1; sagim = 0; transim =0; coroim =0;
nosag=75; notrans=10; nocoro=20;  
% initial display
sagcross=0; transcross=0; corocross=0;
clearlastline = [...
've if lastgraphobj=0,'
'  figure(2);'
'  delete(lastgraphobj); lastgraphobj=0,'
'  figure(1);'
've end,'
];
addtodata=[
've if ndat==0, xyzpoint = 1,'
've else xyzpoint = pointer(ndat)+objsize(ndat)+1; end,'
've ndat = ndat+1,'
've if lastgraphobj ~= 0,'
've  objects = [objects; lastgraphobj];'
've  figure(2);'
've  x = get(lastgraphobj, ''xdata'');'
've  y = get(lastgraphobj, ''ydata'');'
've  z = get(lastgraphobj, ''zdata'');'
've  len = length(x);'
've  tissue= [tissue; nitem];'
've  objsize = [objsize;len];'
've  objtype = [objtype; closecurve];'
've  xdata = [xdata; x(:)];'
've  ydata = [ydata; y(:)];'
've  zdata = [zdata; z(:)];'
've  pointer = [pointer; xyzpoint];'
've end,'
];
reset = [''objects=[]; objtype=[]; raidsagcurvepoints=[];'','
'  ndat=0; xdata=[]; ydata=[]; zdata=[],'
'  midsagplanepoints=[],'
'  pointer=[]; objsize=[]; tissue=[]''];
fileperf = uimenu('label','File');
%uimenu(fileperf,'label','Save all','callback','eval(saveall)');
%uimenu(fileperf,'label','Load all','callback','eval(loadall)');
uimenu(fileperf,'Label','Save Patch To File','callback','eval(savepatchfile)');
midsag=uimenu(fileperf,'Label','Define Midsag. Points');
uimenu(midsag,'Label','Midsagplane Points','callback','nitem=-1');
uimenu(midsag,'Label','Midsagcurve Points','callback','nitem=-2;');
ui = uicontrol('position',[0.58 0.2 0.2 0.04],'units',...
'normalized','style','pushbutton','string','Reset',...
'callback','eval(reset);');
ui = uicontrol('position',[0.78 0.2 0.2 0.04],'units',...
'normalized','style','pushbutton','string','Store',...
'callback','eval(addtodata);');
ui = uicontrol('position',[0.58 0.5 0.2 0.04],'units',...
'normalized','style','pushbutton','string','del obj',...
'callback','eval(clearlastline);');
ui = uicontrol('position',[0.78 0.5 0.2 0.04],'units',...
'normalized','style','pushbutton','string','Quit',...
'callback',['close all']);
ui = uicontrol('position',[0.58 0.16 0.27 0.04], 'units',...
'normalized','style','pushbutton','string','Carved-out Display','callback','cubedesign;');
filenameeditor = ...
ucicontrol('style','edit',...
'position',[0.58 0.42 0.27 0.05],'units','normalized',...
'string','*.dat');
getfilenameout='...
'patchfile=get(filenameeditor,"string");','...
'disp(['File Named: ' patchfile']);','
];
savepatchfile=[...
'disp('Save patch to');','
'eval(getfilenameout);','
'figure(2);','
'kind = get(gca,"children");','
'fid = fopen(patchfile,"w");','
'fprintf(fid,"i\n",length(matlist(:,1)));','
'fprintf(fid,"d d d d \n",matlist');','
'fprintf(fid,"i\n",length(kind));','
'for k=1:length(kind),','
'  xcm = get(kind(k),"xdata");','
'  ycm = get(kind(k),"ydata");','
'  zcm = get(kind(k),"zdata");','
'  udat = get(kind(k),"userdata");','
'for k=1:length(kind),...
if length(udat)==0, udat = 0; end;
fprintf(fid,'\%i \%i \%i\n',length(xcm),udat);
fprintf(fid,'\%d \%d \%d\n',[xcm',ycm',zcm']);
end;
fclose(fid);
disp(’done’);

matlist=[]; pulld = uimenu(’label’,’Tissue type’);
matlist=addtomenu(matlist,pulld,’Bone’,0.6962, 0.8056, 0.8056);
matlist=addtomenu(matlist,pulld,’Tooth’,0.9132, 0.9444, 0.9444);
matlist=addtomenu(matlist,pulld,’Skin’,0.8637, 0.7968, 0.6424);
matlist=addtomenu(matlist,pulld,’Gland’,0.9688, 0.9800, 0.7000);
matlist=addtomenu(matlist,pulld,’Fat’,0.9688, 0.800, 0.7000);
matlist=addtomenu(matlist,pulld,’Genioglossus’,0.7265, 0.4714,...
0.4714);
matlist=addtomenu(matlist,pulld,’Geniohyoid’,0.9841,0.0159,1.0);
matlist=addtomenu(matlist,pulld,’Hyoglossus’,0.8333, 0, 0);
matlist=addtomenu(matlist,pulld,’Mylohyoid’,0.4603, 0.5397,1.0);
matlist=addtomenu(matlist,pulld,’Digastric’,0.9325, 0.5828,...
0.3712);
matlist=addtomenu(matlist,pulld,’Sup. Longitudinalis’,0.2,...
0.8, 0.5);
matlist=addtomenu(matlist,pulld,’Inf. Longitudinalis’,0.2,...
0.8, 0.6);
matlist=addtomenu(matlist,pulld,’Transv. and Verticalis’,1.0,...
0.62, 0.4);
matlist=addtomenu(matlist,pulld,’Verticalis’,0.5, 1.0, 0.83);
matlist=addtomenu(matlist,pulld,’Transversus’,0.5, 0.5, 1.0);
axsag = axes(’position’,[0.1 0.64 0.4 0.3],’units’,’normalized’);
set(get(axsag,’ylabel’),’string’,’T’);
axtrans = axes(’position’,[0.1 0.08 0.4 0.5],’units’,’normalized’);
set(get(axtrans,’xlabel’),’string’,’C’);
set(get(axtrans,’ylabel’),’string’,’S’);
axcoro = axes(’position’,[0.58 0.64 0.4 0.3],’units’,’normalized’);
set(get(axcoro,’xlabel’),’string’,’S’);
axtext = axes(’position’,[0.5 0.3 0.4 0.1],’units’,’normalized’);
axes(axtext); axis(’off’);
loctext = text('position',[0.1,0.1,0.4],'color','yellow',...  'units','normalized','fontsize',20,'string','-');
windfc = ['currax = gca; ',...
  'val = get(gca,''currentpoint'');',...
  'xv = round(val(2,1))+1; yv= round(val(2,2))+1;'];
locationtext=['sst=sprintf ('  'S= %d, T= %d, C= %d',',',...
  'nosag,notrans,nocoro');','set(loctext,''string'',sst);'];

% initial:
  eval(locationtext);
sagmod= ['nocoro=xv; notrans=yv;',...
  'eval(transview); eval(coroview);','axes(axsag);','
  'if sagcross~=0, delete(sagcross); sagcross=0; end;',...
  'sagcross=line(''xdata'',[nocoro nocoro NaN 0 im_m],',',...
    '"ydata'' ,[0 im_no NaN notrans notrans],',',...
    '"color'' ,'',yellow''});','
  'set(sagcross,''buttondownfcn'',''if getbpoints(axsag)==1',',...
  'axes(axsag);eval(windfc);eval(sagmod);end;'');','
  'eval(locationtext);'];
transmod= ['nosag=yv; nocoro=xv;',...
  'eval(sagview); eval(coroview);','axes(axtrans);','
  'if transcross~=0, delete(transcross); transcross=0; end;',...
  'transcross=line(''xdata'',[nocoro nocoro NaN 0 im_m],',',...
    '"ydata'' ,[0 im_n NaN nosag nosag],',',...
    '"color'' ,'',yellow''});','
  'set(transcross,''buttondownfcn'',''if getbpoints(axtrans)==1',',...
  'axes(axtrans);eval(windfc);eval(transmod);end;'');','
  'eval(locationtext);'];
coromod= ['nosag=xv; notrans=yv;',...
  'eval(sagview); eval(transview);','axes(axcoro);','
  'if corocross~=0, delete(corocross); corocross=0; end;',...
  'corocross=line(''xdata'',[nosag nosag NaN 0 im_n],',',...
    '"ydata'' ,[0 im_no NaN notrans notrans],',',...
    '"color'' ,'',yellow''});','
  'set(corocross,''buttondownfcn'',''if getbpoints(axcoro)==1',',...
  'axes(axcoro);eval(windfc);eval(coromod);end;'');','...
eval(locationtext);]

sagview = [...
'Imsat = Cindex((nosag-1)*im_m+1:nosag*im_m,:);', ...
'axes(axsag); if sagim==0, delete(sagim); sagim=0; end;', ..., 
'if sagcross==0, delete(sagcross); sagcross=0; end;', ...
'sagim=image(Imsat); set(axsag,''userdata'',[nosag,1,im_n,im_no,im_m,nitem]); axis(''equal'');', ...
'set(sagim,''buttondownfcn'',''if getbpoints(axsag)==1,'', ...
'axes(axsag);eval(windfc);eval(sagmod);end;'');', ...
'sagcross=line(''xdata'',[nocoro nocoro NaN 0 im_m],', ...
(''''ydata'',[0 im_no NaN notrans notrans],', ...
(''''color'',''red'');', ...
'set(sagcross,''buttondownfcn'',''if getbpoints(axsag)==1,'', ...
'axes(axsag);eval(windfc);eval(sagmod);end;'');', ...
'set(get(axsag,''ylabel''),''string'',''T'');', ...
]
);
transview = [...
'Imtrans = zeros(im_m,im_n);', ...
'Imtrans(:) = Cindex(:,notrans);Imtrans=Imtrans;', ..., 
'axes(axtrans);if transim==0, delete(transim); transim=0; end;', ..., 
'if transcross==0, delete(transcross); transcross=0; end;', ...
'transim=image(Imtrans); set(axtrans,''userdata'',[notrans,2,im_n,im_m,im_no,nitem]);axis(''equal'');', ...
'transcross=line(''xdata'',[nocoro nocoro NaN 0 im_m],', ...
(''''ydata'',[0 im_n NaN nosag nosag],', ...
(''''color'',''red'');', ...
'set(transcross,''buttondownfcn'',''if getbpoints(axtrans)==1,'', ...
'axes(axtrans);eval(windfc);eval(transmod);end;'');', ...
'set(transim,''buttondownfcn'',''if getbpoints(axtrans)==1,'', ...
'axes(axtrans);eval(windfc);eval(transmod);end;'');', ...
'set(get(axtrans,''xlabel''),''string'',''C'');', ...
'set(get(axtrans,’’ylabel’’),''string'',''S'');', ...
]
);
coroview = [...
'ix = [0:im_n-1]*im_m+nocoro;', ...
'Imcoro = Cindextrans(:,ix);', ...
'axes(axcoro);if coroim==0, delete(coroim); coroim=0; end;', ..., 
'if corocross==0, delete(corocross); corocross=0; end;', ...
}
'coroim=image(Imcoro); set(axcoro,''userdata'','');...
[nocoro,3,im_m,im_no,im_n,nitem]); axis(''equal'');,'
'corocross=line(''xdata'',[nosag nosag NaN 0 im_n],''...'
ydata'',[0 im_no NaN notrans notrans],''...'
'color'',''red'');','...
'set(corocross,''buttondownfcn'',''if getbpoints(axcoro)==1,''
'axes(axcoro);eval(windfc);eval(coromod);end;'');'',
'set(coroim,''buttondownfcn'',''if getbpoints(axcoro)==1,''
'axes(axcoro);eval(windfc);eval(coromod);end;'');'',
'set(get(axcoro,''xlabel''),''string'',''S'');'',
]
';
eval(sagview); eval(transview); eval(coroview);
disp('K lick in any image to change view');

B.3.3 Program Listing for other programs

function profilealigngra(colorlist)
% Use not only as supplement to Newtrace tongue but also
% use independently to align 2-d polygons files (*.graph)
% with the profile figure in Miyawaki's article.
% Call functions: markersinfiggra.m;
% linetransform.m;
% ltransformfiggra.m. Chao-Min Wu 03/06/96
%
% x1        y1        x2        y2
% ------------------------
% line along front     Line along back
% X,Y coordinates of the control points in the scanned-in
% profile figure
figpts = ...  
[248.9724  33.9646  402.2588  32.9803; 32.9803; bottom of section 1
182.2186  57.5866  485.0829  55.6181;
141.4246  81.2087  532.0578  77.2717;
110.5201 106.7992  569.1432 101.8780;
 87.0327 131.4055  586.4497 124.5157;
 62.3090 156.9961  601.2839 149.1220;
 54.8920 180.6181  613.6457 174.7126;
The following funny operation is necessary to take the fact into account, that we always look at the top view of the slabs of 3mm thickness. So the x coordinates have to be shifted to the next lower plane.

\[
[x, i] = \text{min}(\text{figpts}(; , 1));
\]

\[
[x, j] = \text{max}(\text{figpts}(; , 3));
\]

\[
y_1 = \text{mean}(\text{diff}(\text{figpts}(; , 2)));
\]

\[
y_2 = \text{mean}(\text{diff}(\text{figpts}(; , 4)));
\]

\[
\text{figpts} = [\text{figpts}; [0 \text{figpts}(17,2)+y_1, 0, \text{figpts}(17,4)+y_2]];
\]

\[
\text{figpts}(i+1:18,1) = \text{figpts}(i:17,1);
\]

\[
\text{figpts}(j+1:18,3) = \text{figpts}(j:17,3);
\]

\[
\text{figure}(1);
\]

for \(k = 1:18,\)

\[
\text{sec} = k;
\]

\[
\text{name} = \text{sprintf}('t'; i', \text{graph}', \text{sec});
\]

\[
[\text{xx}, \text{yy}, \text{zz}] = \text{markersinfiggra(name)};
\]

\[
\text{zt} = 0.5*(\text{figpts}(k,2)+\text{figpts}(k,4)) - 0.5*(\text{zz}(1)+\text{zz}(2));
\]

\[
\text{lineA} = [\text{xx}(1), \text{yy}(1), \text{xx}(2), \text{yy}(2)];
\]

\[
\text{lineB} = [\text{figpts}(k,1), 0, \text{figpts}(k,3), 0];
\]

\[
\text{pars} = \text{lins} \text{transform(lineA, lineB)}; \quad \% \text{now 6 params.}
\]

\[
\text{lins} \text{transformfiggra(name, colorlist, [pars, 500-zt])};
\]

end;

return;

function \(x = \text{loadsketch}(k)\)

\% \text{loadsketch.m} is the program to display an image of the Miyawaki
\% data and specify the x, y coordinates of the selected
% landmark point on the image. Chao-Min Wu 03/06/96
fname = sprintf('t%d.gif',k)
[x,map] = gifread(fname);
colormap(map);
image(x);
set(gca,'fontsize',16)
xlabel('X (in pixel)');ylabel('Y (in pixel)');
title(['Slice No. T.',int2str(k)]);
return;

function x=loadman(k)
% loadsketch.m is the program to display an image of the Visible
% Man data and specify the x, y coordinates of the selected
% landmark point on the image. Chao-Min Wu 03/06/96
fname = sprintf('vm%d.gif',k)
[x,map] = gifread(fname);
colormap(map);
image(x');
axis equal;
set(gca,'fontsize',16)
xlabel('C (in pixel)');ylabel('S (in pixel)');
title(['Slice No.',int2str(k)]);
brighten(0.2);
cmap=colormap;
return;

function palign=alignPgra(P)
% Function palign=align(P) was used to align landmarks P
% for Miyawaki's data. Chao-Min Wu 03/06/96
% when the
% \[ zt = 0.5*(\text{figpts(sec,2)}+\text{figpts(sec,4)}) - 0.5*(\text{zz(1)}+\text{zz(2)}) \]
% is used, then the \[ pz = pz + (500-zt) \]; otherwise,
% \[ pz = pz + zt \], where \[ zt = pz-0.5*(\text{zz(1)}+\text{zz(2)}) \].
%
% \[ % x1 y1 x2 y2 
% \]
% line along front Line along back
%
palign=[];
figpts = ...

\[
\begin{bmatrix}
248.9724 & 33.9646 & 402.2588 & 32.9803; \% \text{bottom of section 1}
182.2186 & 57.5866 & 485.0829 & 55.6181;
141.4246 & 81.2087 & 532.0578 & 77.2717;
110.5201 & 106.7992 & 569.1432 & 101.8780;
87.0327 & 131.4055 & 633.4246 & 226.8780;
62.3090 & 156.9961 & 638.3693 & 253.4528;
48.920 & 180.6181 & 635.8970 & 274.1220;
155.0226 & 276.0906 & 632.1884 & 296.7598;
199.5251 & 300.6969 & 616.1181 & 306.3504;
246.5000 & 324.3189 & 604.9925 & 322.3504;
325.6156 & 346.9567 & 593.8668 & 344.9882;
383.7161 & 368.6102 & 580.2688 & 368.6102;
434.3995 & 388.2953 & 586.4497 & 387.3110;
410.9331 & 580.2688 & 408.9646; \% \text{bottom of section 18}
\end{bmatrix}
\]

\[x,i] = \text{min(figpts(:,1))};
[x,j] = \text{max(figpts(:,3))};
y1 = \text{mean(diff(figpts(:,2)))};
y2 = \text{mean(diff(figpts(:,4)))};
figpts = [figpts; [0 figpts(17,2)+y1, 0, figpts(17,4)+y2]];}
figpts(i+1:18,1) = figpts(i:17,1);
figpts(j+1:18,3) = figpts(j:17,3);
alignf=figure('numbertitle','off','name',...
'Aligned Landmark P Figure');
axis equal;set(gca,'fontsize',16);xlabel('X (in pixel)');
ylabel('Y (in pixel)');zlabel('Z (in pixel)');

\%
\%
\%

\%
% for *.graph alignment.

paras=...

\[0.2800 \ 0.7815;
-0.4492 \ 0.7692;
-1.2271 \ 0.7148;
0.3045 \ 0.7848;\]
for k=1:length(P),
    sec=(51-P(k,3))/3+1;
    angle=paras(sec,1).*pi./180;
    scal=paras(sec,2);
    sina=sin(angle);cosa=cos(angle);
    px=P(k,1);py=P(k,2);pz=P(k,3);
    name = sprintf('t%i.graph',sec);
    [xx,yy,zz] = markersinfiggra(name);
    zt = 0.5*(figpts(sec,2)+figpts(sec,4)) - 0.5*(zz(l)+zz(2));
    xl = xx(l);y1 = yy(1);u1 = figpts(sec,1);v1 = 0;
    sx = scal;sy = scal; % uniform scaling:
    t11 = sx*cosa; % sx Cos[a];
    t12 = -sx*sina; % -(sx Sin[a])
    t13 = u1 - sx*x1*cosa + sx*y1*sina;
    % u1 - sx x1 Cos[a] + sx y1 Sin[a]
    t21 = sy*sina; % sy Sin[a]
    t22 = sy*cosa; % sy Cos[a]
    t23 = v1 - sy*y1*cosa - sy*x1*sina;
    % v1 - sy y1 Cos[a] - sy x1 Sin[a]
    px = px*t11 + py*t12 + t13;
    py = px*t21 + py*t22 + t23;
    pz = pz + (500-zt);
    line('xdata',px,'ydata',py,'zdata',pz,'linestyle','x',...
        'color','g','linewidth',3,'markersize',8);
s = sprintf('%i',k);
text(px+10,py+10,pz+10,s,'color','b','fontweight','...
' 'bold','fontsize',14);palign=[palign;px py pz];
end;
return;

function profilealignpmp(P,H,colorlist)
% Used independently to align 2-d polygons files
% (*.graph) with the profile figure in Miyawaki's article
% and create a a thin-plate spline mapping result of this
% aligned 3-d stack.
% Call functions: markersinfiggra.m;
% linetransform.m;
% litpmpfig.m. Chao-Min Wu 03/06/96
%
% x1    y1       x2       y2
% -------------       -------------
% line along front   Line along back
% X,Y coordinates of the control points in the
% scanned-in profile figure
figpts = ...
[248.9724 33.9646 402.2588 32.9803;
  182.2186 57.5866 485.0829 55.6181;
  141.4246 81.2087 532.0578 77.2717;
  110.5201 106.7992 569.1432 101.8780;
  87.0327 131.4055 586.4497 124.5157;
  62.3090 156.9961 601.2839 149.1220;
  54.8920 180.6181 613.6457 174.7126;
  52.4196 208.1772 626.0075 202.2717;
  53.6558 233.7677 633.4246 226.8780;
  70.9623 256.4055 638.3693 253.4528;
  155.0226 276.0906 635.8970 274.1220;
  199.5251 300.6969 626.0075 296.7598;
  246.5000 324.3189 626.0075 322.3504;
  325.6156 346.9567 616.1181 344.9882;
  383.7161 368.6102 604.9925 368.6102;
  408.4397 388.2953 593.8668 387.3110;
  434.3995 410.9331 580.2688 408.9646];
\[ [x, i] = \text{min}(\text{figpts}(:, 1)); \]
\[ [x, j] = \text{max}(\text{figpts}( :, 3)); \]
\[ y_1 = \text{mean}(\text{diff}(\text{figpts}( :, 2))); \]
\[ y_2 = \text{mean}(\text{diff}(\text{figpts}( :, 4))); \]
\[ \text{figpts} = [\text{figpts}; [0 \text{figpts}(17, 2) + y_1, 0, \text{figpts}(17, 4) + y_2]]; \]
\[ \text{figpts}(i+1:18, 1) = \text{figpts}(i:17, 1); \]
\[ \text{figpts}(j+1:18, 3) = \text{figpts}(j:17, 3); \]
\[ \text{figure}(3); \]
\[ \text{for } k=1:18, \]
\[ \sec = k; \]
\[ \text{name} = \text{sprintf('t\%i.graph', sec);} \]
\[ \text{poly} = \text{figpts with sizey-y} \]
\[ \text{figpts}(k, 2) + \text{figpts}(k, 4)) - 0.5*(zz(1)+zz(2)); \]
\[ \text{lineA} = [\text{xx}(1), yy(1), xx(2), yy(2)]; \]
\[ \text{lineB} = \text{figpts}(k, 1), 0, \text{figpts}(k, 3), 0]; \]
\[ \text{pars} = \text{linetransform(lineA, lineB); } \]
\[ \text{litpmpfig(name, P, H, colorlist, [pars, 500-zt]);} \]
\[ \text{end}; \]
\[ \text{return; } \]

\text{function [pic, xs, ys, zs]=tpmpaligngridgra(P, H, sec, impars, ...} \]
\[ \text{Cindex, cmap, myimage)} \]
\[ \text{see todolast.tex for detail; Chao-Min Wu 03/06/96} \]
\[ \text{P, H} \quad \text{Homologous point pairs} \]
\[ \text{sec} \quad \text{number of section} \]
\[ \text{impars} \quad \text{block size parameters [im_no, im_m, im_n, whitepixel]} \]
\[ \text{Cindex} \quad \text{Image block of Visible man data} \]
\[ \text{Nx, Ny} \quad \text{Grid size [Ny, Nx] = size(myimage)} \]
\[ \text{Call functions: markersinfiggra.m; hullofstackgra.m;} \]
\[ \text{linetransform.m; thin3dtrans.m;} \]
\[ \text{transformgrid.m; sampleblock.m.} \]
\[ \text{im_no = impars(1);} \]
\[ \text{im_m = impars(2);} \]
\[ \text{im_n = impars(3);} \]
\[ \text{whitepixel = impars(4);} \]
\[ \text{defined by whitepixel=length([map；1 1 1]);} \]
\[ \text{111} \]
figpts = ...
[248.9724 33.9646 402.2588 32.9803;
 182.2186 57.5866 485.0829 55.6181;
 141.4246 81.2087 532.0578 77.2717;
 110.5201 106.7992 569.1432 101.8780;
  87.0327 131.4055 586.4497 124.5157;
  62.3090 156.9961 601.2839 253.4528;
  52.4196 208.1772 626.0075 274.1220;
  53.6558 233.7677 633.4246 226.8780;
  70.9623 256.4055 638.3693 253.4528;
 155.0226 276.0906 635.8970 274.1220;
 199.5251 300.6969 632.1884 296.7598;
 246.5000 324.3189 626.0075 322.3504;
 235.6156 346.9567 616.1181 344.9882;
 383.7161 368.6102 604.9925 368.6102;
 408.4397 388.2953 593.8668 344.9882;
 434.3995 410.9331 580.2688 408.9646];

[x,i] = min(figpts(:,1));
[x,j] = max(figpts(:,3));
yl = mean(diff(figpts(:,2)));
y2 = mean(diff(figpts(:,4)));
figpts = [figpts; [0 figpts(17,2)+yl, 0, figpts(17,4)+y2]];
figpts(i+1:18,1) = figpts(i:17,1);
figpts(j+1:18,3) = figpts(j:17,3);
name = sprintf('t%.i.graph*sec');
[xx,yy,zz] = markersinfiggra(name);
zt = 0.5*(figpts(sec,2)+figpts(sec,4)) - 0.5*(zz(1)+zz(2));
[xm,ym,zm] = hullofstackgra(name);
[Ny,Nx] = size(myimage)
[xg,yg] = meshgrid(1:1:Nx,1:1:Ny);
xg = (xg-1) * (xm(2)-xm(1))/(Nx-1) + xm(1);
yg = (yg-1) * (ym(2)-ym(1))/(Ny-1) + ym(1);
zg = ones(size(xg))*(zm(1)+zm(2))/2;
lineA = [xx(1),yy(1),xx(2),yy(2)];
lineB = [figpts(sec,1), 0, figpts(sec,3), 0];
pars = linetransformdineA,lineB);  % now 6 params.
[xg,yg,zg]= ltransformgrid(xg,yg,zg,[pars,500-zt]);
datarange(xg)
datarange(yg)
datarange(zg)%display data range
[xs,ys,zs]=thin3dtrans(P,H,xg,yg,zg);
datarange(xs)
datarange(ys)
datarange(zs)
pic = sampleblock(xs,ys,zs,im_no,im_m,im_n,Cindex,whitepixel);
gridf=figure('numbertitle','off','name',' Mapping Figure',...
'colormap',cmap);
pic(find(myimage==2)) = pic(find(myimage==2))*0+255;
c cmap = [cmap; 0.8, 0.8, 0.8];
colormap(ccmap);
image(pic);axis('equal');
set(gca,'fontsize',16);
xlabel('X (in pixel)');ylabel('Y (in pixel)');
title(['Mapped Slice T',int2str(sec)]);
% view(80,10);
return;
APPENDIX C

Technical Details for Program Implementation

C.1 Computer Environment

C.1.1 Getting Started

This section gives a description of computer environment that was used. We used MATLAB Version 4.2c with MATLAB Toolbox Version 4.2a and Image Processing Toolbox Version 1.0b on Sun-4/SPARC system. It was installed on the Sun-4 server hattori at Department of Speech and Hearing (Dr. Fujimura's research laboratory). It is in the directory /home1/usr/local/bin/Matlab. MATLAB needs X Window System Server (X11R4 or above) to run the software. For more information on system requirements for the UNIX implementation of MATLAB software, please refer to the UNIX installation guide that comes with the software.

To prepare a UNIX account for invoking MATLAB, one creates a matlab directory under his home directory by typing mkdir matlab at the UNIX prompt. Then go to the matlab directory by typing cd matlab at the UNIX prompt. To invoke MATLAB, type matlab at the UNIX prompt. The MATLAB banner appears with the following display:
Commands to get started: intro, demo, help help
Commands for more information: help, whatsnew, info, subscribe

>>

The MATLAB prompt >> is waiting for user to enter command. We will describe how to set up the path to access programs and data used in this dissertation.

**C.1.2 Setting MATLABPATH**

Programs and data used in this dissertation are located in the directories listed below.

/home3/wucm/matlab/tpmp2d keeps programs for 2-d thin-plate spline mapping.

/home3/wucm/matlab/tpmp3d keeps programs for 3-d thin-plate spline mapping.

/home3/wucm/mri keeps all MRI data.

/home3/wucm/gif keeps all GIF images of the Miyawaki data and their polygon files.

/home3/wucm/drawing keeps all Miyawaki tongue sketches for 2-d thin-plate spline mapping.

MATLAB has a search path that is specified by the environment variable MATLABPATH. To access programs and data used in this dissertation, users must include these directories as lines added to their own .cshrc file:
setenv MATLABPATH /home3/wucm/matlab/tpmp2d:
/home3/wucm/matlab/tpmp3d:/home3/wucm/mri:
/home3/wucm/gif:/home3/wucm/drawing

Next time you login the UNIX system and invoke the MATLAB, these directories will be included as search paths for MATLAB.

C.2 Program Implementation

C.2.1 Programs for 2-d Thin-plate Spline Mapping

To run the programs drawcolorsc.m and tpmp2dsc.m in Appendix A, one has to invoke MATLAB. With the MATLAB prompt, type the command whitebg(0) to set the background color to white color for all the windows created by MATLAB. The command whitebg(0) is implemented only at the beginning of MATLAB program. Then, type

```
>>drawcolorsc;
```

or

```
>>tpmp2dsc;
```

as needed. Once the windows for the programs appear, one can follow the instruction described in Appendix A to draw polygon or implement 2-d thin-plate spline mapping.

C.2.2 Programs for 3-d Thin-plate Spline Mapping

Steps to run the program Newtracetongue.m are similar to those for the programs we have just described except that we enter Newtracetongue after the
MATLAB prompt. Once the windows for the program appear, follow the instruction described in Appendix B to create and save polygon files for each Miyawaki tongue sketch for later use in the alignment transformation. To run the program New2Colorcut.m, the file newcolorindex.mat has to be loaded before New2Colorcut.m can be executed. The file newcolorindex.mat contains variables needed for New2Colorcut.m. Steps to run New2Colorcut.m with MATLAB prompt are:

```matlab
>> whitebg(0);
>> load newcolorindex;
>> New2Colorcut;
```

Once the window for the program appear, one can follow the instruction described in Appendix B to draw polygon object over the Visible Man data. Programs Newtracetongue.m and New2Colorcut.m are also used to assist locating landmarks.

Programs loadsketch.m and loadman.m are the application programs used to display images of the Miyawaki data or the Visible Man data, and determine the x, y coordinates of the selected landmark point on the images, respectively. For example, the right panel of Figure 15 was obtained by typing

```matlab
>> loadsketch(3);
>> ginput(1);
```

and using the left button of the mouse to determine the x, y coordinate of the selected landmark point on the image. Similar steps were used to obtain the left panel of Figure 15 except replacing loadsketch(3) by loadman(171). After
landmark pairs were identified, the landmark points of the Miyawaki data have to be transformed to the aligned position, as shown in the left panel of Figure 16. It is obtained by typing

```matlab
>> Pa = alignPgra(P);
```

at the MATLAB prompt. The variable P contains original x, y, z coordinates of the landmark points and the variable Pa has x, y, z coordinates of the landmark points in the aligned position.

Programs `profilealigngra.m` and `profilealigntpmp.m` were implemented with the following steps to obtain Figures 21 and 22, respectively.

```matlab
>> load tvmlamkpair;
>> load colorlist;
>> profilealigngra(colorlist);
>> profilealigntpmp(Pa, Ha, colorlist);
```

The file `tvmlamkpair.mat` contains variables (P, Pa, Ha, impars, cmap) for the implementation of the programs described above. The file `colorlist.mat` contains one variable, colorlist, which has various color specification for different tissue, and the variable Ha has x, y, z coordinates of the landmark points in the Visible Man data.

The program `tpmpaligngridgra.m` is the main program for evaluation. The implementation of evaluation for transverse slice number 4 of Miyawaki's data, as shown in the left panel of Figure 28, was demonstrated as an example by the following steps.
If the result of this implementation did not show a good match, an iteration of adjusting landmark position, the alignment and thin-plate spline mapping discussed in this section is necessary before a “global matching” of the Miyawaki data and the Visible Man data is achieved.