

A Feasibility Study of Template-Based Subject-Specific Modelling and Simulation of Upper-Airway Complex

by

Keyi Tang

B.Eng., University of Electronic Science and Technology of China, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Electric & Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

May 2017

© Keyi Tang 2017

Abstract

The upper-airway complex is involved in a number of life-sustaining functions, such as swallowing, speech, breathing and chewing. Disorders associated with these functions can dramatically reduce the life quality of the sufferers. Biomechanical modelling is a useful tool that can bridge the gap between the human knowledge and medical data. When tailored to individual patients, biomechanical models can augment the imaging data, to enable computer-assisted diagnosis and treatment planning.

This thesis introduces a model-registration framework for creating subject-specific models of the upper-airway complex based on 3D medical images. Our framework adapts a state-of-art comprehensive biomechanical model of head and neck, which represents the generic upper-airway anatomy and function. By morphing this functional template to subject-specific data, we create upper-airway models for particular individuals. In order to preserve the functionality of the comprehensive model, we introduce a multi-structure registration technique, which can maintain the spatial relationship between the template components, and preserve the regularity of the underlying mesh structures. The functional information, such as the muscle attachment positions, joint positions and biomechanical properties, is updated to stay relevant to the subject-specific model geometry. We demonstrate the functionality of our subject-specific models in the biomechanical simulations.

Two illustrative case studies are presented. First, we apply our modelling methods to simulating the normal swallowing motion of a particular subject based on the kinematics (of the airway boundary, jaw and hyoid) extracted from dynamic 3D CT images. The results suggest that our model tracks the oropharyngeal motion well, but has limited ability to reproduce the hyolaryngeal movements of normal swallowing. Second, we create two speaker-specific models based on 3D MR images, and perform personalized speech simulations of the utterance /ə-gis/. The models reproduce the speech motion of the tongue and jaw recorded in tagged and cine MRI data with sub-voxel tracking error, predict the muscular coordinating patterns of the speech motion.

This study demonstrates the feasibility of using template-based subject-

Abstract

specific modelling methods to facilitate personalized analysis of upper-airway functions. The proposed model-registration framework provides a foundation for developing a systematic and advanced subject-specific modelling platform.

Lay Summary

After surgical treatment and radiotherapy, oral-cancer patients may have difficulty in certain life-sustaining activities, such as swallowing, breathing and speech. Biomechanical models provide a means to analyze the cause and effect of these upper-airway dysfunctions, and to simulate the surgical changes in bone and muscle structures for prediction of treatment outcomes. These clinically-relevant applications require the models to include as much information as possible from a specific subject. However, creating biomechanical models for coupled upper-airway system relies heavily on expert interaction. The slow process of model creation prevents us from simulating large numbers of individual cases. In order to ease the modelling efforts, this study explores the use of registration methods for model creation: Morph a predefined template model to match with certain subject data. Personalized swallowing and speech simulations are performed to demonstrate the potential of the template-based subject-specific modelling methods for clinically-relevant analysis of upper-airway functions.

Preface

This thesis presented herein was approved by UBC clinical Research Ethics Board, certificate numbers: H16-00016, H16-01546.

Most of the contributions and ideals described in Chapter 3 have been presented previously in the publication [P1]. I was the primary author and main contributor to the design, implementation and testing of the methods developed in the [P1], under supervision of Dr. Sidney Fels. Dr. Negar M. Harandi assisted with the analysis of the results, and gave editorial feedback to the paper. Dr. Yoko Inamoto provided the CT data. Dr. Maureen Stone and Dr. Jonghye Woo provided the MRI data. OPAL project provided the template upper-airway model.

Chapter 4 has not yet been published. I developed the subject-specific upper-airway model, segmented the bone surfaces, proposed and implemented the swallowing simulation methods, analyzed the simulation results. Dr. Yoko Inamoto provided the CT data. Dr. Andrew Ho provided the segmentation of the moving airway boundary. Dr. John Lloyd provided the source code of the forward-dynamics simulation. Dr. Ian Stavness provided the source code of the inverse solver.

Chapter 5 has been partially published in the literature [P2]. I developed the speaker-specific models, proposed and implemented the speech simulation methods, analyzed the simulation results, under supervision of Dr. Sidney Fels. Dr. Negar M. Harandi was the primary author of the publication [P2], and assisted with the analysis of the results. Dr. Maureen Stone and Dr. Jonghye Woo provided the MRI data. Dr. John Lloyd provided the source code of the forward-dynamics simulation. Dr. Ian Stavness provided the source code of the inverse solver.

Journal Manuscripts in Review

[P1] **Tang K**, Harandi NM, Fels S. 2017. Subject-specific modelling of upper-airway complex. *Computer Methods in Biomechanics and Biomedical Engineering (CMBBE)*. Submitted.

Peer-Reviewed Conference Paper in Review

[P2] **Tang K**, Harandi NM, Woo J, Fakhri EG, Stone M, Fels S. 2017. Speaker-specific Biomechanical Model-based Investigation of a Simple Speech Task based on Tagged-MRI. Interspeech. Submitted.

Table of Contents

Abstract	ii
Lay Summary	iv
Preface	v
Table of Contents	vii
List of Tables	x
List of Figures	xi
List of Abbreviations	xiii
Acknowledgements	xiv
Dedication	xv
1 Introduction	1
1.1 Contributions	3
1.2 Outlines	5
2 Background and Previous Work	6
2.1 Data Acquisition and Measurement	7
2.1.1 Cadaver Studies	7
2.1.2 Fluoroscopy and Computed Tomography	7
2.1.3 Magnetic Resonance Imaging	9
2.1.4 Elastography	12
2.1.5 Electromyography	12
2.2 Biomechanical Modelling of Upper-Airway Complex	13
2.2.1 Functional Reference ANatomical Knowledge	14
2.2.2 Combined Multi-Body Finite-Element Simulation	20
2.3 Subject-Specific Modelling	22

Table of Contents

2.4	Biomedical Applications	28
2.4.1	Intralaryngeal Prosthesis Implantation	28
2.4.2	Intensity Modulated Radiation Therapy	29
2.4.3	Hemiglossectomy and Hemimandibulectomy	29
2.4.4	Maxillofacial Surgery	30
2.5	Discussion and Conclusion	30
3	Subject-Specific Modelling of Upper-Airway Complex	32
3.1	Multi-Structure Registration	33
3.1.1	Registration Framework	36
3.1.2	Correspondence Computation	38
3.1.3	As-Similar-As-Possible Deformation	40
3.1.4	Structure-Preserving Free-Form Deformation	43
3.1.5	Experiment	45
3.2	Subject-Specific Modelling	48
3.2.1	Medical Image Data	48
3.2.2	Template Model	49
3.2.3	Modelling Framework	50
3.2.4	FE Quality Control	52
3.2.5	Functionality Transfer	54
3.2.6	Experiments	55
3.3	Discussion and Conclusion	71
4	Data-Driven Swallowing Simulation	75
4.1	Kinematic Data	79
4.2	Inverse Simulation	80
4.3	Results	82
4.4	Discussion and Conclusion	92
5	Data-Driven Speech Simulation	94
5.1	Medical Dataset	96
5.2	Speaker-Specific Modelling	98
5.2.1	Inverse Simulation	101
5.3	Speech Simulation	104
5.3.1	Definition of the Trajectories	104
5.3.2	Results	105
5.4	Discussion and Conclusion	108

Table of Contents

6 Conclusion	110
6.1 Summary	110
6.2 Future Directions	112
6.3 Concluding Remarks	117
Bibliography	118

Appendix

A Muscles and Ligaments in FRANK	136
A.1 Muscles	136
A.2 Ligaments	141

List of Tables

2.1	Summary of FRANK component	15
2.2	A list of deformable (FEM) components in FRANK	19
3.1	Summary of segmentation	56
3.2	Summary of tongue FE mesh	59
3.3	Summary of pharynx FE mesh	60
3.4	Subdivision of FRANK	63
3.5	Registration errors within the subject-specific models	67
3.6	Summary of FE mesh qualities	67
4.1	The peak tracking error during oral transport phase	89
4.2	The peak tracking error during pharyngeal phase	91
5.1	Speaker information	97
5.2	Summary of segmentation	98
5.3	Registration errors within the speaker-specific models	98
5.4	Tracking errors of the template, Speaker A and Speaker B	103
5.5	Tongue tracking error (mm) over 40 control points	107
5.6	Jaw tracking error (mm) over 20 control points.	107
A.1	Jaw, hyoid muscles in FRANK	137
A.2	Face muscles in FRANK	138
A.3	Tongue muscles in FRANK	139
A.4	Soft palate muscles in FRANK	139
A.5	Pharynx muscles in FRANK	140
A.6	Larynx muscles in FRANK	140
A.7	List of all ligaments in FRANK	141

List of Figures

1.1	Mid-sagittal view of the human upper-airway	2
2.1	Digitized fibres from a cadaveric forearm	7
2.2	Midsagittal and axial MSCT images of swallowing	8
2.3	Real Time MRI captured during swallowing	10
2.4	A FE tongue model developed from a DTI data set	11
2.5	Illustration of FRANK	14
2.6	Illustration of all rigid components in FRANK	16
2.7	Illustration of all deformable components in FRANK	17
2.8	Diagrams of two subject-specific modelling categories	22
2.9	Illustration of mesh locking	25
2.10	Virtual intralaryngeal prosthesis implantation	28
3.1	Illustration of the mechanism of multi-level FFD	34
3.2	Proposed multi-structure registration framework	36
3.3	Illustration of multi-structure registration pipeline	37
3.4	Iterative closest point (ICP) correspondence	40
3.5	Notation for triangular mesh.	42
3.6	Illustration of the registration of the synthetic data	46
3.7	Registration error measured over 300 iterations	47
3.8	Mesh quality (η) measured over 300 iterations	47
3.9	Medical images used in this study	49
3.10	Proposed framework for subject-specific modelling and simulation	51
3.11	Illustration of human intervention	52
3.12	Example of element quality distribution of a FE mesh	53
3.13	Illustration of quality-controlling mechanism	53
3.14	Segmentation of the CT and MRI images	57
3.15	Registration results of tongue model	59
3.16	Registration results of pharynx model	60
3.17	Comparison of the resulting element quality by our methods and MMRep	61

List of Figures

3.18	Histograms of the element qualities	62
3.19	Illustration of poor-quality elements	64
3.20	Subject-specific models of upper-airways complex	65
3.21	Histograms of the element qualities of three models	66
3.22	Comparison of the element quality of the resulting models and the template models	68
3.23	View of speech postures of three models	70
3.24	A comparison of muscle activations	72
4.1	Videofluoroscopy of deglutition	77
4.2	Illustration of swallowing modes reported in literatures	78
4.3	Illustration of segmentation and CT images	79
4.4	Illustration of tracking markers and target positions	80
4.5	Methods of determining the target positions	81
4.6	Tracking error of swallowing simulation	82
4.7	Mid-sagittal views of the inverse simulation of swallowing	83
4.8	Oblique views of the inverse simulation of swallowing	84
4.9	Top views of the inverse simulation of swallowing	85
4.10	Activation levels of the extrinsic and intrinsic tongue muscles	86
4.11	Activation levels of the suprahyoid, mastication and palate muscles	87
4.12	Activation levels of the laryngeal and pharyngeal muscles	88
4.13	The model fails to reproduce the target hyoid posture at $t = 0.9s$	90
4.14	The model fails to raise the the corniculate cartilage at $t = 0.9s$	91
5.1	Illustration of two previously reported models for speech sim- ulation	95
5.2	Mid-sagittal slice of cine MR images of the speakers.	97
5.3	Speaker-specific models of upper-airways complex	99
5.4	Comparison of the element quality of the speaker-specific models and the template models	100
5.5	Illustration of the inverse simulation	102
5.6	Illustration of the end postures and the inverse-predicted mus- cle activations	103
5.7	Illustration of results of the speech simulations	106
5.8	Estimated muscle activations	107
5.9	Illustration of erroneous tracking using Tagged MRI	109
6.1	Current subject-specific modelling framework	113
6.2	Target subject-specific modelling framework	114

List of Abbreviations

1D	1-Dimensional
2D	2-Dimensional
3D	3-Dimensional
ANP	Average Nodal Pressure
CBCT	Cone Beam Computed Tomography
CT	Computed Tomography
DoF	Degrees of Freedom
DTI	Diffusion Tensor Imaging
EMG	Electromyography
FB	Fibre Bundle
FE	Finite Element
FEM	Finite Element Model
FOV	Field Of View
FPD	Flat Panel Detector
FRANK	Functional Reference ANatomical Knowledge
HMPS	Hamiltonian Moving Particle Semi-implicit
HP	Harmonic Phase
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MSCT	Multi-Slice Computed Tomography
OSA	Obstructive Sleep Apnea
PCSA	Physiological Cross Section Area
SNR	Signal-to-Noise Ratio
SWE	Shear Wave Elastography
TMJ	Temporomandibular Joint
TR	Repetition Time

Acknowledgements

I would like to express my deepest appreciation to my supervisor, Dr. Sidney Fels, for his support, guidance and mentorship. Without his encouragement and help for my graduate studies as well as personal life, this thesis would not have been possible.

I also received tremendous support from a number of fine researchers throughout my master studies. Thanks must go to Dr. Ian Stavness for bringing me into this research topic that I have no background in. Many thanks to Dr. Negar M. Harandi for sharing her knowledge with me during many discussions and a rewarding collaboration on the publication. Dr. Peter Anderson deserves special thanks for providing assistance with upper-airway modelling. I would like to thank Dr. John Lloyd for his great support with ArtiSynth. Also, I would like to thank Dr. Yoko Inamoto, Dr. Maureen Stone and Dr. Jonghye Woo for generous data sharing.

Thanks also go to colleagues at the Human Communication Technologies (HCT) Laboratory for their support to my research and lab life. Special thanks go to Andrew Ho and Antonio Sánchez for many useful conversations and assistance with the project.

I want to say a heart-felt thank you to my friends and family for their love and support throughout this journey. Special thanks are owed to my families for passing their life-long motivation and inspiration to me. I am also grateful to my girlfriend, Mengyu, for daily accompanying and encouragement.

Dedication

To Lei

Chapter 1

Introduction

The human upper-airway complex is composed of rigid structures including jaw, and hyoid bone, highly deformable muscle-activated tissues such as the tongue, soft palate, pharynx, larynx, an intricate arrangement of many muscles and ligaments, and various constraint situations. It is involved in a wide range of life-sustaining functions, such as breathing, swallowing, and speech production. A number of associated dysfunctions impose adverse effects on the quality of sufferers' life. Dysphagia, or swallowing disorders, is a serious concern for stroke patients as aspiration-related pneumonia leads to 40,000 deaths per year in North America [97]. Obstructive Sleep Apnea (OSA) is a serious disorder involving pauses in breathing during sleep, causing arousal and daytime sleepiness [99], afflicting 12 million people in the United States [136]. Articulation, fluency, and voice disorders afflict more than 7.5 million people in the United States [58].

Due to the complexity of the head and neck anatomy, the inaccessibility of many biomechanical and physiological parameters (e.g. muscular forces and activations) presents significant barriers for traditional data collection methods to provide sufficient insight into normal and abnormal upper-airway functions. The recent advance of computational modelling and simulation techniques greatly facilitates and furthers our understanding of how the body functions. Biological simulations enable the transition from visualizing dynamic movements and measuring physiological parameters, to predicting motions, estimating unobservable variables and exploring motor control mechanisms.

As large inter-subject variation exists in the upper-airway anatomy and physiology, generic computational models – which are suitable for exploring general biomechanical and physiological principles – are not necessarily representative of specific patients. Subject-specific modelling targets at representing geometrical, mechanical and physiological information of particular subjects; hence it enables personalized simulations best suited for clinical usage. Subject-specific modelling and simulation of upper-airway system have

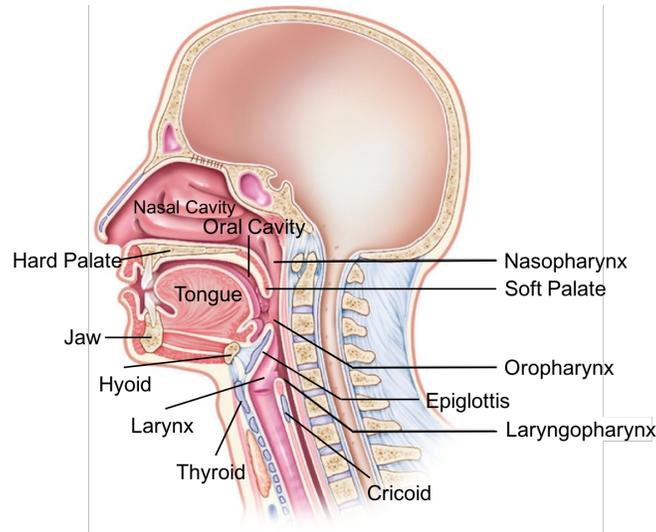


Figure 1.1: Mid-sagittal view of the human upper-airway. © Elsevier (2003), Thibodeau and Patton [144], adapted with permission.

a wide range of viable biomedical applications. For example, subject-specific upper-airway models can be used to predict immeasurable biomechanical quantities that may correlate with dysfunctions, such as OSA, dysphagia and speech pathologies, and therefore enhance diagnosis protocols. Besides, subject-specific modelling and simulation can also aid in planning of treatment and surgical interventions. Personalized biomechanical models can be used to evaluate alternative treatment paths, such as different surgical procedures, or to tailor a particular treatment to a specific patient, such as customizing dosages in radiation therapy. Thus, surgical procedures can be iteratively improved and tailored to a specific patient with little cost and risk. In addition, subject-specific models can also help to guide patient rehabilitation. Given a model of a specific patients reconstruction, simulation of different muscle activation patterns may illuminate new motor strategies to compensate for the altered musculoskeletal structure [136].

Subject-specific modelling of the head and neck has been applied into computer-assisted surgeries, such as the maxillofacial surgery [23, 49, 108, 153] and the jaw reconstruction surgery [156]. However, current subject-specific application is constrained by the limited functional fidelity of biome-

chanical models. Fully reproducing upper-airway functions, such as swallowing and speech production, requires coordination of multiple motor components inside oral cavity, nasopharynx, oropharynx and laryngopharynx. Creating a comprehensive model that couples multiple upper-airway organs relies heavily on expert interaction. The slow process of model creation prevents us from simulating large numbers of individual cases. The present thesis investigates the feasibility of applying registration techniques to facilitating the subject-specific modelling and simulation of upper-airway complex.

1.1 Contributions

This thesis targets at developing biomechanics modelling tools that allow personalized analysis of various life-sustaining functions and dysfunctions associated with the human upper-airway system. We first present a review of data acquisition techniques, the state-of-art upper-airway biomechanical models and subject-specific modelling methods. Then, we identify the challenges facing standard approaches for creating comprehensive upper-airway models for specific subjects. Next, we explore the use of model-registration methods to ease the modelling efforts: Our methods seek to register a pre-defined functional upper-airway template to the subject data. We demonstrate the feasibility of our methods by creating two subject-specific models of upper-airway complex based on two medical data sets. Furthermore, we apply our subject-specific modelling methods to simulating two coupled upper-airway functions: swallowing and speech. Finally, future directions of subject-specific modelling are proposed. The main contributions of the present thesis are summarized as follows.

Developed a model registration framework for subject-specific modelling of upper-airway complex.

- We identified a model-registration strategy for creating comprehensive upper-airway models for particular subjects: Register a functional template to the partial segmentation of the anatomical structures extracted from 3D medical images. In order to preserve the functionality of the registered model, three types of regularity need to be maintained:
 1. Inter-Component Regularity: Maintain the spatial relationship

(including connectivity, topology, relative posture and size) between model components.

2. Intra-Component Regularity: Preserve the regularity of the underlying discretization structures of the template model during registration.
 3. Functional Regularity: Keep the functional information (including coupling attachments between components, muscle attachments and biomechanical properties) similar to the template but relevant to the new model geometry.
- We created a novel multi-structure registration technique, which can preserve both the inter- and intra-component regularity.
 - Based on the multi-structure registration techniques, we developed a registration work flow for creating subject-specific upper-airway models from medical images.
 - We demonstrated the feasibility of our template-based subject-specific modelling methods by creating two comprehensive upper-airway models for particular subjects, and tested their functionality in a set of biomechanical simulations.

Demonstrated the potential of the proposed subject-specific modelling methods for personalized analysis of swallowing biomechanics. We enabled a personalized simulation of realistic normal swallowing using one of the developed models. Our model tracked the oropharyngeal motion well, while having limited ability to reproduce the hyolaryngeal movements of normal swallowing.

Demonstrated the potential of the proposed subject-specific modelling methods for personalized analysis of speech production. Using the proposed model-registration methods, we created two speaker-specific models, and enabled personalized speech simulations of the utterance /ə-gis/. The models reproduced the speech motion of the tongue and jaw based on speaker-specific tagged and cine MRI data, and predicted the corresponding muscular coordinating patterns, which made good agreement with the speech-expert knowledge.

1.2 Outlines

The rest of this thesis is organized as the following. Chapter 2 reviews the state-of-art data acquisition and measurement tools, head-and-neck models, subject-specific modelling methods, and lists a few closely related biomedical applications of the subject-specific upper-airway models. Chapter 3 introduces the subject-specific modelling methods and demonstrates two subject-specific models of upper-airway complex. In Chapter 4, we enable a swallowing simulation using a developed model based on a dynamic CT recording of the corresponding subject. In Chapter 5, we create two speaker-specific models, and enable personalized speech simulations based on the cine and tagged MR recordings of the speakers. Chapter 6 summarizes the thesis contributions, and describes directions for future work. Appendix A summarizes all of the muscles and ligaments included in the template upper-airway complex model (FRANK).

Chapter 2

Background and Previous Work

Subject-specific models would be useful for assessing the biomechanical dysfunction, comparing treatment options when more than one possibilities exists, and predicting clinical outcomes without risking patients. This chapter reviews the tools and techniques used to create and validate subject-specific models of upper-airway system, and then lists a few potential biomedical applications of the subject-specific models.

Current advancement of *in vivo* data acquisition techniques has enabled observation of anatomical structure, recording of physiological data and measurement of tissue properties of the upper-airway complex. Such observational data has widely applied in filed of biomehchanical modelling. However, these techniques are challenged by the low spatial resolution, long acquisition time and limited measurement capability. Section 2.1 provides a review of the data acquisition techniques closely related to biomechanical modelling; their associated limitations and challenges are discussed.

As computational modelling and simulation greatly facilitate analysis of biomechanics and motor control of the head and neck, many biomechanical models of upper-airway sub-structures have been created. Holistic models, which incorporate multiple functional units (organs), further the investigation of complex human behaviours, such as swallowing and speech production. Section 2.2 reviews previously reported biomechanical models and provides a detailed description of the state-of-art functional reference model of upper-airway complex.

Although computational models are suitable for exploring general biomechanical and physiological principles in upper-airway system, they are not necessarily representative of specific patients. Subject-specific modelling enables individualized simulations best-suited for clinical purposes. Section 2.3 reviews the subject-specific modelling methods and identifies the related limitations and issues.

The ultimate goal of subject-specific modelling is applying the computational modelling and simulation techniques into clinical diagnosis and treat-

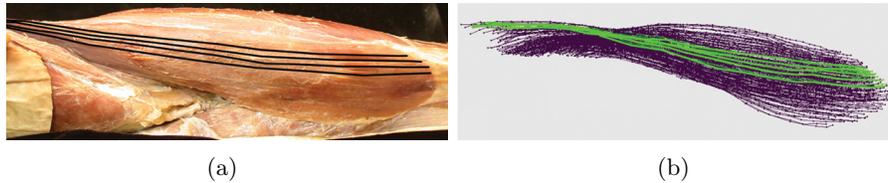


Figure 2.1: Digitized fibres from a cadaveric forearm using a MicroScribeTM 3DX Digitizer. (a) delineated FBs are highlighted in black on an embalmed cadaveric specimen. (b) digitized fibres from specimen. © Taylor & Francis (2014), Li et al. [84].

ment. Section 2.4 reviews a few biomedical applications of biomechanical modelling and simulation closely related to our efforts.

2.1 Data Acquisition and Measurement

2.1.1 Cadaver Studies

Cadaveric study is the most straightforward process for assessing organ geometric data and mechanical properties. Based on cadavers, The Visible Human Project [20] creates complete, anatomically detailed, three-dimensional representations of the normal male and female human bodies. It has supported many projects in the fields of medical image processing and biomechanics. Gerard et al. [44] use a cadaver to measure the mechanical properties of tongue. Li et al. [84] measure the morphologies of muscle fibre bundles (FBs) on a cadaveric forearm, using a MicroScribeTM 3DX Digitizer (0.3mm accuracy; immersion Corporation, San Jose, CA, USA); the cadaveric specimen and the digitized fibres are illustrated in Figure 2.1. However, cadaver studies fail to provide individualized information for living human. Since medical imaging modalities become more and more accurate, standardised and allow *in vivo* measurements, they tend to supplant cadaver studies in the field of biomechanics.

2.1.2 Radiography, Fluoroscopy and Computed Tomography

Radiography is the first modality for *in-vivo* anatomical imaging of internal structure. Based on the measurements of the intensity of X-rays traversing

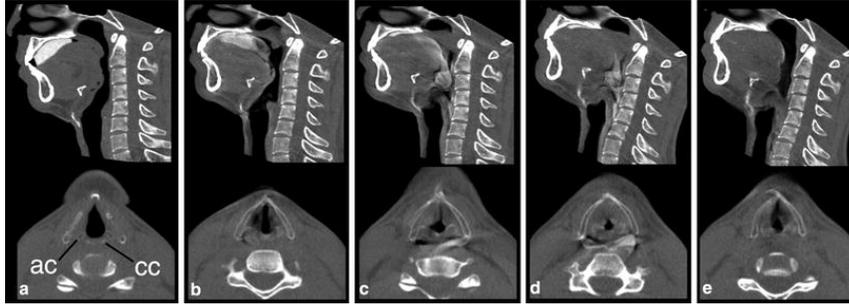


Figure 2.2: Midsagittal and axial MSCT images of swallowing. © Springer (2010), Fujii et al. [37], adapted with permission.

the body, radiography provides a static superimposed 2D representation of all internal structures. Similar to radiography, fluoroscopy uses X-ray to produce real-time 2D images of interior tissues. Stavness et al. [138] use lateral videofluoroscopy to record normal swallowing on three healthy subjects. Since fluoroscopic images do not provide 3D information, this data can only depict motions that are visible in sagittal view.

Computed Tomography (CT) combines multiple X-ray projections to reconstruct 3D images of tissues. CT can provide 3D information, but its acquisition time is much longer than fluoroscopy. To reduce the acquisition time and improve the temporal resolution of dynamic CT scans, Multi-slice CT (MSCT) scanners have recently been equipped with multiple arrays of X-ray detectors that are able to reconstruct a 3D volume from a single rotation. Fujii et al. [37] use a 320-detector-row multislice computed tomography (320-MSCT) scanner for detailed morphological analysis of swallowing (shown in Figure 2.2). A single-phase 3D image covers the area of oral cavity, pharynx, larynx and upper esophagus and is captured in 0.35s. The imaging process is repeated for 29 phases at intervals of 0.1 second, to generate 3D images from oral to the early esophageal stages of swallowing on one volunteer. Inamoto et al. [66] use a similar single phase image protocol to investigate the effects of ages, gender and height on the anatomy of the pharynx and the larynx on 54 healthy volunteers.

Cone Beam CT (CBCT) is another emerging CT imaging technique that uses a cone-beam acquisition geometry and Flat Panel Detector (FPD) to provide relatively low-dose imaging with high isotropic spatial resolution that can be acquired with a single gantry revolution [105]. However, compared with conventional MSCT, CBCT has relatively poor soft-tissue con-

trast. Grauer et al. [53] use CBCT records of 62 nongrowing subjects to evaluate the pharyngeal airway volume and shape. Recently, Glupker et al. [50] use CBCT to measure airway volume changes between open and closed jaw positions for 60 patients with temporomandibular joint disorders.

As X-ray is absorbed by dense tissue and passes through air, CT usually can accurately depict skeleton and airway. High dosage of X-ray exposure is the main drawback of medical CT which hinders its use on healthy volunteers.

2.1.3 Magnetic Resonance Imaging

Hydrogen atomic nuclei can absorb and reemit energy in an external magnetic field. Magnetic Resonance Imaging (MRI) aims at differentiating between body tissues by measuring the released energy of hydrogen atoms when the protons return to the initial state. Tissues react differently depending on the proton density and the duration that the proton resume their initial state (relaxation time). By varying the parameters of the pulse sequence, MRI can produce different contrast between tissues based on their relaxation properties. The MR image shows good soft-tissue contrast and it does not produce harmful ionizing radiation, which makes it well-suited for clinical research involving volunteers.

For years, MRI is limited by its long acquisition time due to inherent trade-off between Signal-to-Noise Ratio (SNR), spatial resolution and temporal resolution. High-resolution MR volumes require a long acquisition time, commonly leading to involuntary movement and introducing motion artifacts. Many methods are proposed to reduce the acquisition time. The simplest modification is to minimize TR by increasing gradient strength. However, this method is limited by engineering cost and human physiology [79]. Echo train imaging is another complementary approach, which acquires more than one phase encode line per TR. However, such method compromises not only contrast but also resolution and in some cases leads to image distortion [79]. Super-resolution imaging techniques are introduced to generate high spatial resolution MR images in relatively short time by combining information from a number of images. Woo et al. [158] report two super-resolution MR volumes of tongue with isotropic spatial resolution of $0.94mm$ and FOV of $240mm \times 240mm \times 240mm$. More recently, Woo et al. [157] use twenty super-resolution MRI volumes to build an high-resolution atlas of vocal tract. Parallel imaging uses multiple receiver coils to augment the time consuming Fourier encoding, which reduces the acquisition time significantly without compromising image contrast. Under-sampling

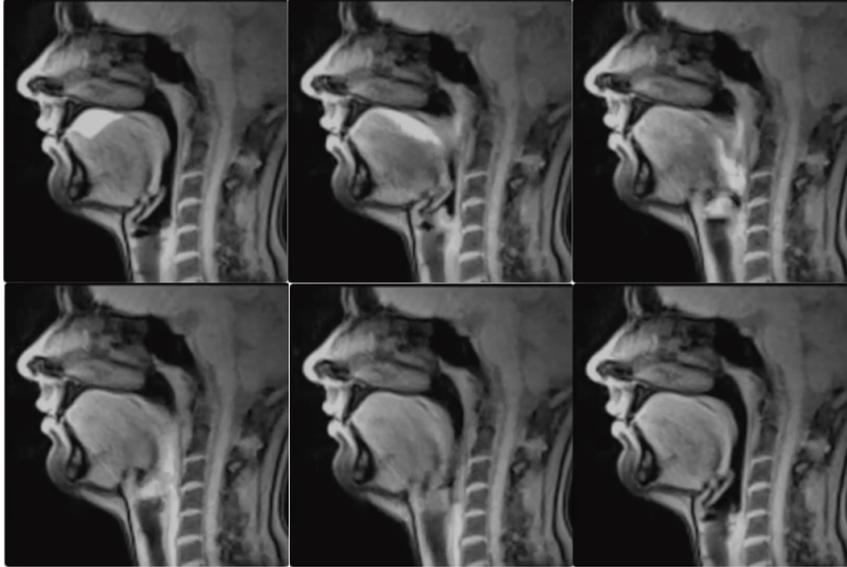


Figure 2.3: Real Time MRI captured during swallowing. © Wikimedia Commons (2011), Uecker et al. [148].

of k-space is another family of acceleration methods. Non-Cartesian sampling trajectories (e.g. radial or spiral) is applied to reduce sampling-aliases. Via combining these acceleration techniques, state-of-art MRI scanners have been successfully applied to capture tissue motions within certain spatial and temporal resolutions. Kim et al. [73] produce dynamic 3D MRI recordings of upper-airway obstruction during natural sleep on eight volunteers. The raw data is captured in real time (2.6 fps for the short scan and 1.7 fps for the long scan). Lingala et al. [86] report a MRI system for study of dynamic vocal tract shaping during speech production. Their MRI recordings achieves spatio-temporal resolutions of $2.4mm \times 2.4mm$ every $12ms$ for slice-slice imaging and every $36ms$ for three-slice imaging.

Although MRI can provide high contrast between soft-tissues, it fails to provide enough contrast to distinguish material points within soft-tissue itself. In order to capture motion information, MR tagging techniques use a special pulse sequence to create temporary features in soft-tissues. Based on the assumption that the Harmonic Phase (HP) value of a fixed material point is time-invariant, the motion of the material points throughout time can be tracked. Xing et al. [160] and Woo et al. [159] use Tagged MRI for analysis of 3D tongue motion during speech.

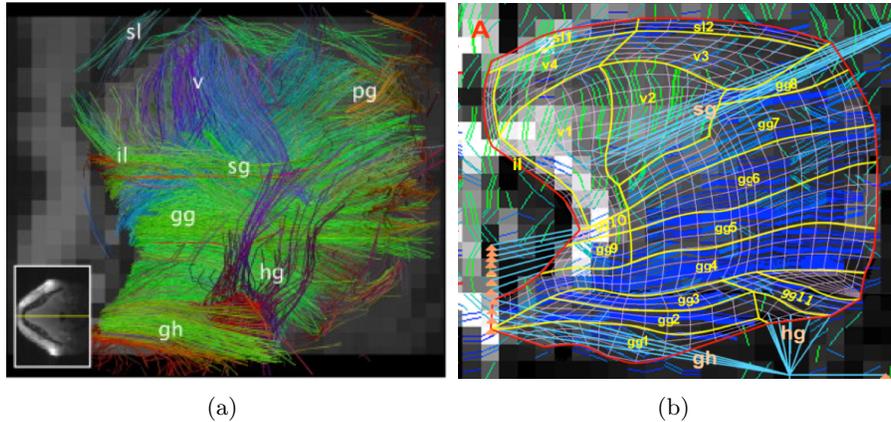


Figure 2.4: Illustration of a FE tongue model developed from images of human lingual myofiber tracts obtained by DTI with tractography. (a) a 3D sagittal view of lingual myoarchitecture based on *in vivo* DTI tractography. © Wiley-Liss (2007), Gaige et al. [39], adapted with permission. (b) the tongue FE mesh overlaid on the DTI image slice. © American Physiological Society (2010), Mijailovich et al. [104].

Fluid flow attenuates MR signal intensity in the direction of the magnetic gradient. Since the self-diffusion of water is restricted by tissue geometry (e.g. fibres direction), by applying different gradient direction, the tissue geometry can be measured with MRI. Diffusion tensor MRI (DTI) aims to measure the restricted diffusion of water in tissues in order to reveal the microscopic details of biological structures. By combining DTI with tractography, the paths of muscle fibres in tissues can be reconstructed. Heemskerk et al. [60] successfully apply 3D DTI to determine overall muscle structure, fibre length, pennation angle and PCSA of mouse. Gaige et al. [39], report the complete 3D myoarchitecture of human tongue and the geometry of intrinsic and extrinsic myofiber populations, which are obtained *in vivo* from DTI tractography. The fibre structures extracted from DTI images are employed by Mijailovich et al. [104] to drive a finite-element (FE) model of lingual deformation during swallowing. DTI provides a non-invasive way to obtain subject-specific fibre directions in upper-airway tissues. However, this imaging technique is still limited by its long acquisition time, insufficient spacial resolution and SNR.

2.1.4 Elastography

Elastography aims to estimate tissue mechanical properties by imaging. Strain elastography measures the mechanical properties of tissues by imaging their deformation under slight compression. Shear Wave Elastography (SWE) achieves the measurement by observing the velocity of mechanically excited shear wave propagation within the tissue of interest. Cheng et al. [26] employ SWE to investigate the viscoelasticity of tongue and soft-palate. Elastography provides a non-invasive way for the measurement of tissue properties. However, as soft tissues exhibit non-linear stress-strain behaviour, responses at larger deformations cannot be inferred from elastography alone [155], which limits its application on biomechanical modelling.

2.1.5 Electromyography

Motor neurons transmit electrical signals that cause muscles to contract. Electromyography (EMG) involves measuring muscle activation indirectly by picking up the electric charge produced by an action potential just as it reaches a muscle [46]. Two current techniques are available for electromyography, i.e. surface EMG and intramuscular EMG. Surface EMG picks up voltage signals on the surface of the skin, which is non-invasive. However, it can only measure the activation of superficial muscles. Moreover, it may pick up signals from different muscles, which will cause confusion and inaccuracy. In contrast, Intramuscular EMG directly inserts the hooked-wire electrodes into the relevant muscle, leading to extremely precise and reliable results [46]. Obviously, hooked-wire electrodes can also cause some discomfort and therefore require more involved ethics approval for research use.

EMG recording has been widely applied to investigate muscle activations during head-and-neck associated activities, such as speech production [11, 41, 98, 147] and swallowing [33, 68, 101, 149]. However, EMG suffers from numerous issues, including cross-talk between adjacent channels, the complexity of anatomy and the discomforting of volunteers. In addition, the relationship between EMG signals and muscle forces is not straightforward. Because these issues remain challenging, nowadays, inverse modelling has become a popular alternative tool for the measurement of muscle activations.

2.2 Biomechanical Modelling of Upper-Airway Complex

Biomechanical modelling of the human upper-airway complex has received a growing interest since it facilitates the analysis of complex human behaviours, such as speech production [17, 42, 43, 57, 96, 116, 117, 131, 154], swallowing [63, 72, 106, 134, 146] and mastication [56]. However, the biomechanical models created in these studies only incorporate a part of upper-airway complex, such as a single tongue model [43, 116] or a single facial model [96, 131]. Although Gerard et al. [42] add surface represented jaw, hyoid, hard and soft palate into the their Finite Element (FE) tongue model to shape oral cavity; these components are modelled as non-dynamic structures (i.e. fixed in space, only contacts with the tongue model are enabled). Based on Gerard's model, Buchaillard et al. [17] then add non-dynamic pharyngeal and laryngeal walls and characterize the hyoid bone as a dynamic rigid-body and connect it with other solid structures using spring-like muscles. This model can provide high fidelity for reproducing the motions happening inside oral cavity, but it has limited functionality for other organs, such as soft palate, temporomandibular joint (TMJ), pharynx and larynx. Kikuchi et al. [72] develop an upper-airway complex model for swallowing simulation using Hamiltonian Moving Particle Semi-implicit (HMPS) method. Their model incorporates the surfaces of tongue, hard palate, soft palate, larynx, pharynx and esophagus. However, the motion of their model is in response to, rather than muscle contraction, manually defined boundary conditions; hence this model lacks the ability to reflect motor-control mechanisms of real human.

In order to support the study of complex human behaviours in different levels of fidelity, a Functional Reference ANatomical Knowledge (FRANK) biomechanical model of the head and neck [7] has been implemented in the ArtiSynth biomechanical simulation toolkit [93]. Multiple anatomical models that have been involved in different studies [7, 17, 24, 45, 107, 110, 137] are tailored to fit together to generate this generic 3D biomechanical model. FRANK is composed of FE models, rigid bodies and functional structures including point-to-point muscles, joints, inter-component attachments and parametrically-controlled skin meshes. On the one hand, bones and cartilages are represented as rigid bodies, which has a relatively low computational complexity and avoids challenges such as the need to construct volumetric meshes. On the other hand, soft tissues are characterized as FE models, which allows simulation of soft tissue deformation. FRANK represents an

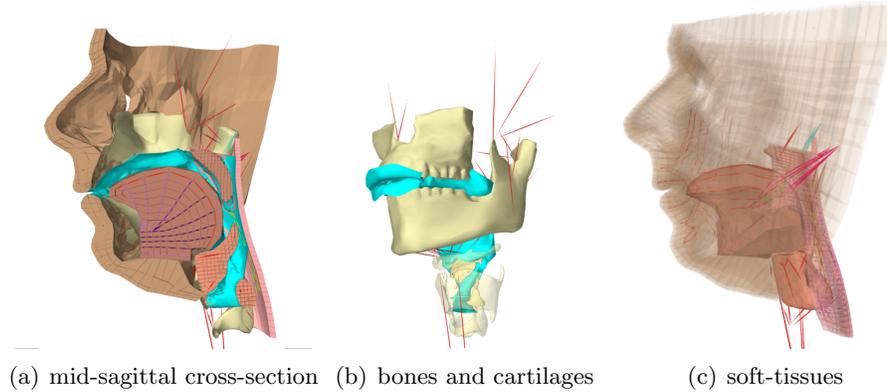


Figure 2.5: FRANK: A template model of the head and neck [7]: (a) Mid-sagittal cross-section. The air-tight airway mesh is shown in cyan. (b) Bones (jaw, maxilla, hyoid) and cartilages (thyroid, cricoid, epiglottis, arytenoid, cuniform) are modelled as rigid-bodies. (c) Soft-tissues (faces, tongue, soft-palate, larynx and pharynx) are modelled as FE deformable-bodies.

average human anatomy and function, which provides a biomechanical template of upper-airway complex. For the purpose of this thesis, we give an overview of this state-of-art upper-airway complex model in following sections.

2.2.1 Functional Reference Anatomical Knowledge (FRANK)

The geometries in FRANK are derived from the work of numerous researchers and multiple data sources; most components differ substantially from the original geometry in order to fit into the FRANK framework using both algorithmic modification and manual adjustments based on related literature and anatomical references [34]. Table 2.1 provides a summary of the model components, including their name, modelling type (either FE models or rigid bodies), mesh type (either hexahedral-dominant mesh or triangular surface mesh) and references to source publications.

Component	Type	Mesh	Reference
Face	FE	Hex	[110]
Tongue	FE	Hex	[17]
Jaw, Hyoid, Maxilla	Rigid	Tri.	[137]
Soft-Palate	FE	Hex	[24, 45]
Pharynx	FE	Hex	[7]
Larynx	FE	Hex	[107]
Larynx Cartilages	Rigid	Hex	[107]

Table 2.1: Summary of components, includes component name, component type (either FE models (FE), rigid body (Rigid)), mesh type (hexahedral-dominant (Hex) or triangular surface mesh (Tri.)) and references to source publications. Larynx Cartilages include thyroid, cricoid, arytenoids, epiglottis, cuneiforms.

Bony Structures

Bony structures are many orders of magnitude stiffer than soft tissue [31], and treating them as rigid bodies can simplify many simulations without significant loss of fidelity. Similarly, FRANK treats cartilage, less stiff than bone but still orders of magnitude stiffer than soft tissues, as rigid. Rigid bodies do not require volumetric mesh but still need high quality surface meshes for correct contact handling. Bones and cartilage have uniform densities, which factors into the inertia calculation of the rigid body. Figure 2.6 illustrates all the bony structures included in FRANK.

The maxilla and upper teeth represent the skull and is fixed in space to serves as an anchor point [7]. The jaw is connected with the maxilla by the TMJ that is modelled as three constraint planes. These planes limit the lateral motion of the jaw and constrain it to follow a pre-defined arc when opening and closing [55]. The hyoid attaches to the base of the tongue, top of the larynx, jaw, and cranium through a series of point-to-point muscles.

Soft Tissues

FRANK represents soft tissues as FE models. Their geometries are illustrated in Figure 2.7.

Tongue The tongue model consists of 946 nodes, 740 hexahedral elements, and 11 pairs of intrinsic and extrinsic muscle bundles (listed in Table A.3) with bilateral symmetry. In the similar approach described in

2.2. Biomechanical Modelling of Upper-Airway Complex

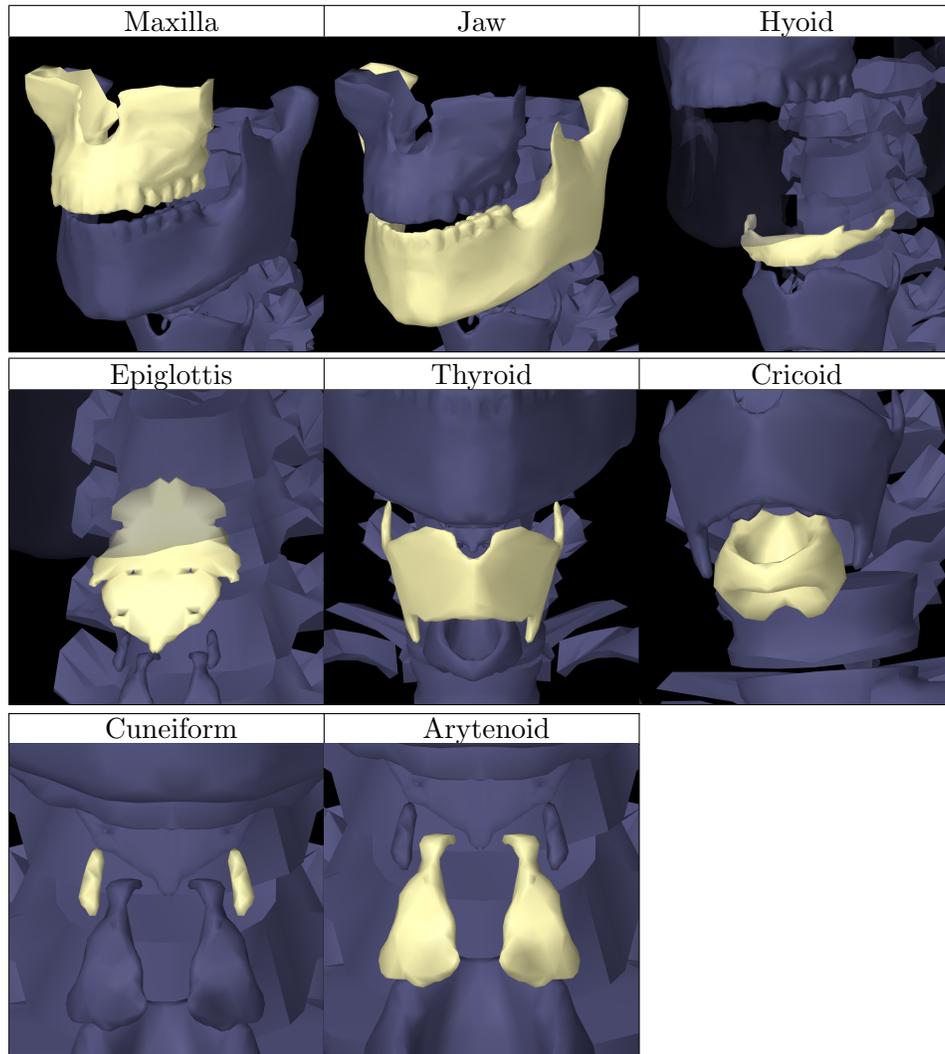


Figure 2.6: Illustration of all rigid components in FRANK.

2.2. Biomechanical Modelling of Upper-Airway Complex

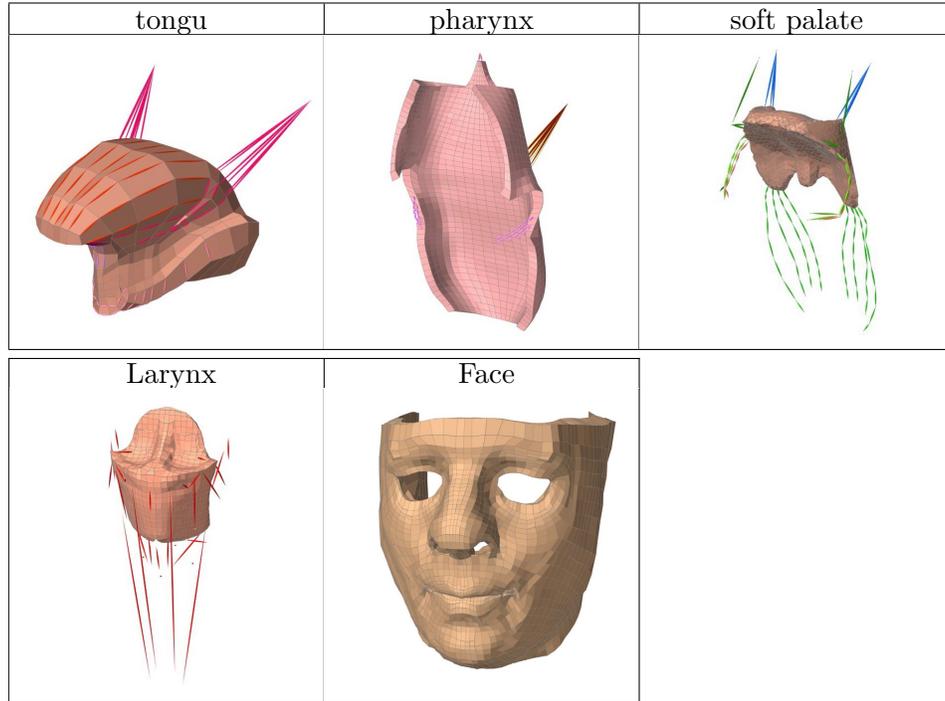


Figure 2.7: Illustration of all deformable (FEM) components in FRANK. Refer to Table 2.2 for density and material properties.

literature [137], the tongue attaches to the hyoid and the jaw via bilateral constraints. Tongue-jaw attachments include the insertion of the genioglossus and geniohyoid onto the mandibular geniotubercle and the insertion of the mylohyoid along the mandibular mylohyoid ridge. Tongue-hyoid attachments include insertions of the geniohyoid, mylohyoid, and hyoglossus muscles with the hyoid.

Pharynx The pharynx model extends from the cranial base position to the lower border of the cricoid cartilage. It attaches to tongue over a manually defined attachment region and forms the posterior wall of oropharynx. It also attaches to thyroid based on proximity. The nodes on the inferior margin of the pharynx are fixed, approximating attachments to the absent organs below it, such as esophagus. The nodes at the pharyngeal raphe are anchored in space using soft bilateral constraints. Seven pharyngeal muscle groups are embedded in the pharynx model, and some of them are subdivided into smaller activation bundles (listed in Table A.5). The pharynx-

thyroid attachments include insertions of the middle constrictors, inferior constrictors, salpingopharyngeus and stylopharyngeus with the thyroid.

Soft Palate The soft palate is attached to the posterior border of the hard palate (the maxilla). The lateral portions of the soft palate geometry conforms with the pharyngeal wall, so that the pharynx can attach to it [7]. Five soft palate muscle groups are incorporated into this model (listed in Table A.4). Only musculus uvulae is strictly intrinsic to the geometry. The other muscles (levator veli palatini, tensor veli palatini, palatoglossus, palatopharyngeus) have extrinsic portions that attach to surrounding components.

Larynx The epiglottis and laryngeal complex complete the lower part of the airway model of FRANK. The deformable larynx model contains a number of cartilage structures: the thyroid, cricoid, epiglottis, left and right cuneiforms, and left and right arytenoids; these are modeled as rigid bodies embedded within the deformable laryngeal tissue, adding substantial stiffness to the component. The pre-epiglottic portion of the larynx attaches to the hyoid based on proximity [7]. The larynx includes point-to-point muscles connected to the rigid bodies (interarytenoid, lateral cricoarytenoid, posterior cricoarytenoid, and thyrohyoid) and sub-divided FEM-internal muscles (vocalis and muscularis portions of the thyroarytenoid, thyroepiglottic, and anterolateral, anteromedial, and posteriolateral portions of the external thyroarytenoid or ventricularis). Table A.6 lists all the muscles associated with the larynx and their properties.

Face The face model contains three layers of elements between the superficial and deep surfaces; the thin outermost layer represents the epidermis and dermis, while the intermediate and deep layers represent the hypodermis [35]. The face attaches to the jaw and the maxilla at a set of manually selected locations [139]. Eleven faical muscle groups are included into the model. The muscles and associated properties are listed in Table A.2.

Material Properties Tongue, larynx and face FE components use a fifth-order Mooney-Rivlin tissue material where the strain energy (W_ϵ) is described as:

$$W_\epsilon = C_{10}(I_1 - 3) + C_{20}(I_1 - 3)^2 + \kappa(\ln J)^2, \quad (2.1)$$

where I_1 is the first invariant of the left Cauchy-Green deformation tensor; C_{10} and C_{20} are the Mooney-Rivlin material parameters, and the term

2.2. Biomechanical Modelling of Upper-Airway Complex

Name	$\rho [kg/m^3]$	Material
tongue	1040	M.R.: $c_{10} = 1037, c_{20} = 486,$ $c_{01}, c_{11}, c_{02} = 0$ Pa, $\kappa = 10 \cdot c_{10}$
soft palate	1040	Lin.: $E = 500$ Pa, $\nu = 0.4995$
pharynx	1040	Lin.: $E = 1500$ Pa, $\nu = 0.49$
larynx	1040	M.R.: $c_{01} = 2500, c_{20} = 1175,$ $c_{10}, c_{11}, c_{02} = 0$ Pa, $\kappa = 10 \cdot c_{01}$
face	1040	M.R.: $c_{10} = 2500, c_{20} = 1175,$ $c_{01}, c_{11}, c_{02} = 0$ Pa, $\kappa = 10 \cdot c_{10}$

Table 2.2: A list of the information of deformable (FEM) components in the FRANK, including density, material properties. A Mooney-Rivlin (M.R.) material is defined by a 5-parameter model and bulk modulus κ , and a linear (Lin.) material is defined by Young's modulus E and Poisson ratio ν .

$\kappa(\ln J)^2$ reinforces the incompressibility. Soft palate and pharynx adopt linear material:

$$\sigma = \mathbf{D}\epsilon, \quad (2.2)$$

where σ is the stress vector; ϵ is the strain vector; \mathbf{D} is the material stiffness matrix:

$$\mathbf{D} = \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} \begin{bmatrix} 1 & \frac{\nu}{1-\nu} & \frac{\nu}{1-\nu} & 0 & 0 & 0 \\ \frac{\nu}{1-\nu} & 1 & \frac{\nu}{1-\nu} & 0 & 0 & 0 \\ \frac{\nu}{1-\nu} & \frac{\nu}{1-\nu} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} \end{bmatrix}. \quad (2.3)$$

E is the Young's Modulus and ν is Poisson's ratio. The densities and material properties of the soft tissues are summarized in Table 2.2.

Muscle Model

The muscles of the upper-airway complex are approximated as point-to-point Hill-type actuators [61]. The contraction force depends on the muscles activation level, a , the muscles length, l , and the speed of muscle shortening, $\partial l / \partial t$. The force equation typically consist of two components, passive and active:

$$f_{muscle}(a, l, \dot{l}) = f_{passive}(l) + a f_{active}(l, \dot{l}). \quad (2.4)$$

The passive force function $f_{passive}(l)$ and active force function $f_{active}(l, \dot{l})$ are defined by muscle-specific parameters such as Physiological Cross-Sectional Area (PCSA), tendon ratio R_t , optimal muscle length L_{opt} , maximal muscle length L_{max} . FRANK adopts the piecewise force functions defined in literature [115]:

$$f_{passive}(l) = \begin{cases} F_{max} & l \geq L_{max} \\ F_{max} \cdot \frac{l - L_{opt}R_t}{L_{max} - L_{opt}} & L_{opt} < l < L_{max} \\ 0 & l \leq L_{opt} \end{cases} \quad (2.5)$$

$$f_{active}(l, \dot{l}) = \begin{cases} \frac{1}{2}F_{max} [1 + \cos(2\pi l_n)] & 0.5 < l_n < 1.5 \\ 0 & \text{otherwise} \end{cases} \quad (2.6)$$

where l_n is the normalized fibre length, $l_n = (l - L_{opt}R_t)/L_{opt}(1 - R_t)$; F_{max} is the maximal muscle force, which is defined as $F_{max} = PCSA \cdot 40N/cm^2$. The point-to-point muscles attach to rigid bodies and fixed points in space; they also pass through FE nodes and exert local forces on them.

Airway

FRANK models the upper-airway mucosa layer as a deformable water-tight mesh (as shown in Figure 2.5), which covers and attaches the bony and soft tissues using a *geometric skinning* method [140]. The upper-airway mucosa layer deforms according to a distance-weighted scheme along with the other airway components. The position of each mucosa layer vertex \mathbf{v}_s is calculated as a weighted sum of contributions from each master component:

$$\mathbf{v}_s = \mathbf{v}_{s_0} + \sum_{i=1}^M w_i f_i(\mathbf{q}_m, \mathbf{q}_{m_0}, \mathbf{v}_{s_0}), \quad (2.7)$$

where \mathbf{v}_{s_0} is the initial position of the mucosa layer vertex; \mathbf{q}_{m_0} is the reset state of the i -th master, w_i is the skinning weight, and f_i is the corresponding blending function. To provide two-way coupling between the skinned mesh and articulators, the forces acting on the mucosa layer are also propagated back to their dynamic masters allowing fluid-solid interaction [7].

2.2.2 Combined Multi-Body Finite-Element Simulation

ArtiSynth achieves full coupling between multibody and FE components by combining the dynamics of all components into a single Lagrangian system

2.2. Biomechanical Modelling of Upper-Airway Complex

with composite positions \mathbf{q} , velocities \mathbf{v} , and forces $\mathbf{f}_{system}(\mathbf{a}, \mathbf{v}, \mathbf{q})$, and a composite mass matrix \mathbf{M} . By applying Newton's second law, we have

$$\mathbf{f}_{system}(\mathbf{a}, \mathbf{v}, \mathbf{q}) = \frac{d\mathbf{M}\mathbf{v}}{dt} = \mathbf{M}\dot{\mathbf{v}} + \dot{\mathbf{M}}\mathbf{v} \quad (2.8)$$

The system forces \mathbf{f} consists of the active forces $\mathbf{f}_{active}(\mathbf{a}, \mathbf{v}, \mathbf{q})$ and passive forces $\mathbf{f}_{passive}(\mathbf{v}, \mathbf{q})$:

$$\begin{aligned} \mathbf{f}_{system}(\mathbf{a}, \mathbf{v}, \mathbf{q}) &= \mathbf{f}_{active}(\mathbf{a}, \mathbf{v}, \mathbf{q}) + \mathbf{f}_{passive}(\mathbf{v}, \mathbf{q}) \\ \mathbf{f}_{active}(\mathbf{a}, \mathbf{v}, \mathbf{q}) &= \Lambda(\mathbf{v}, \mathbf{q})\mathbf{a}, \end{aligned} \quad (2.9)$$

where Λ denotes a nonlinear function that relates the system positions (\mathbf{q}) and the system velocities (\mathbf{v}) to the active muscle forces.

FRANK components attach with each other by several types of bilateral constraints, $\mathbf{B}(\mathbf{q})$. The skull (presented as partial geometry including maxilla and upper-teeth) is anchored in space. The jaw is connected with the maxilla by the TMJ. FE components are attached to other FE components or rigid bodies by nodal attachments. Muscles are approximated as point-to-point Hill-type actuators that may attach to a rigid body or pass through a FE body. The air-tight surface wraps over, and attaches to the FE models and rigid-bodies to create a parametric upper-airway mucosa layer. The contacts between FRANK components are modelled as unilateral constraints, $\mathbf{U}(\mathbf{q})$. These constraints are linearized on velocities, as:

$$\mathbf{B}(\mathbf{q})\mathbf{v} = 0, \quad \mathbf{U}(\mathbf{q})\mathbf{v} \geq 0. \quad (2.10)$$

Forward-dynamics simulation involves solving the above system for the motion in response to forces arising from muscle activations. Inverse dynamics, on the other hand, seeks to estimate the muscle activations \mathbf{a} that produce a given set of target velocities \mathbf{v}^* by solving a quadratic problem:

$$\begin{aligned} \min_{\mathbf{a}} \quad & \frac{w_m}{2} \|\mathbf{v}^* - \mathbf{v}^{i+1}(\mathbf{a})\|^2 + \frac{w_a}{2} \|\mathbf{a}\|^2 + \frac{w_d}{2} \|\dot{\mathbf{a}}\|^2 \\ \text{subject to} \quad & 0 \leq \mathbf{a} \leq 1 \end{aligned}, \quad (2.11)$$

where \mathbf{v}^* is the target velocity trajectory of tracking points; \mathbf{v}^{i+1} is the velocity vector of the tracking points in next time step, which is a linear function of the muscle activations \mathbf{a} . The second term in Equation 2.11 is a l^2 -regularization term. The third term is a damping term; $\dot{\mathbf{a}}$ denotes the time-derivative of the activations \mathbf{a} . The weights w_m , w_a and w_d are used to trade off between the cost terms. As a result, at each timestep of inverse

2.3. Subject-Specific Modelling

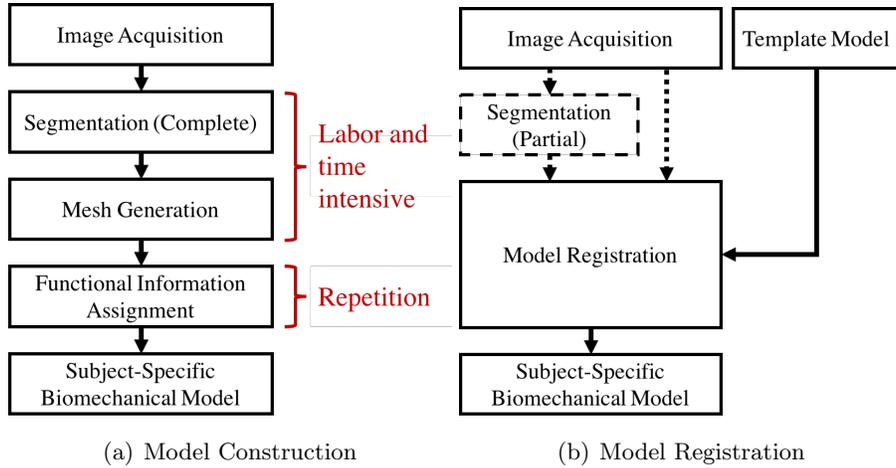


Figure 2.8: Diagrams of two subject-specific modelling categories. (a) direct model construction from subject’s medical images. (b) register a predefined template model to subject-specific medical images or segmentation.

simulation, Equation 2.11 is solved to provide muscle activations to advance the forward dynamics system defined by Equation 2.8 and 2.10. Due to the redundancy of biomechanical systems, one target motion can be produced with multiple sets of activations combinations.

2.3 Subject-Specific Modelling

Computer-aided diagnosis and treatment relies on biomechanical models to predict musculoskeletal behaviour (e.g. bone kinematics, tissue deformation, tissue degeneration, tissue reconstruction, etc.) from morphology, kinematic constraints, mechanical constraints or neuromuscular impulses [48]. The large inter-subject variability of anatomy and physiology requires the transition of biomechanical research from generic understanding of biomechanical phenomena to subject-specific studies addressing biomechanics of a particular individual. Subject-specific modelling is typically done by creating a three-dimensional computation reconstruction of the anatomy of the tissue or mathematical model of the organ of interest in the individual subject, based on imaging scans or other individualized parameters [40].

Currently, subject-specific modelling of human anatomy can be organized into two major categories, as illustrated in Figure 2.8. The first cate-

gory is direct model construction from image data. This approach usually requires image segmentation, surface extraction (i.e. segmentation), meshing and functional information assignment (i.e. boundary condition, muscle definition, coupling attachments, joints, biomechanical properties, etc.). For example, Nithiarasu et al. [118] use segmentation-and-meshing method to build a subject-specific upper-airway-surface FE model (using tetrahedron mesh) based on a 3D CT scan. However, this approach suffers from several limitations. Firstly, it requires complete geometry for the target organ, which may be unavailable due to inconsistent image quality. Secondly, automated FE meshing, especially for Hexahedral or Hexahedral-dominant meshing, remains to be a challenging problem. The manual mesh generation requires significant time and operator effort to complete even a single mesh. Thirdly, direct construction of biomechanical model needs redesigning functional feature, such as muscle definition or coupling attachments, for each subject model. Because of these issues, this approach can be extremely time-consuming and labour-intensive.

The second category is registration. Instead of direct construction from medical data, registration aims at finding a deformation that morphs a template model to a subject dataset. This approach takes advantage of the prior knowledge about the average organ shape, the morphological variability and the feasible organ functionality in the population, thus reducing the ambiguity introduced by medical data and easing the re-designing efforts. Registration has been extensive studies in the field of imaging processing and computer graphics. With respect to the purpose of this thesis, we give a brief overview of registration based subject-specific modelling method.

Articulated Modelling Articulated models use a skeleton as a basis for modelling human motion. These models assume that bones undergo large rigid transformations and local nonrigid surface deformation occurs near joints. Template based articulated subject-specific modelling has been extensively studied by computer graphics communities for realistic biomechanics-based animation. These models are represented as either polygonal meshes or point-cloud. Lewis et al. [82] introduce an articulated model that the displacements of its vertices are generated by a weighted set of (usually linear) influences from neighboring joints. Allen et al. [4] propose a template articulated model that is represented as a posable subdivision surface. Angelov et al. [8] introduce a statistical human shape template model that spans variation in both subject shapes and poses; based on a limited set of available markers specifying the target subjects, the statistical template

model evolves along the degrees of freedom (pose and shape deformation subspace) to obtain the best fit. Corazza et al. [29] expand the statistical human shape space described in [8] by incorporating a space of embedded kinematic models (anatomically feasible joint centers locations). The subject-specific model generated with this method is capable of accurate motion tracking with Markerless Motion Capture (MMC) systems. In order to create subject-specific model from medical image segmentation, Gilles et al. [47] incorporate feasible joint limits (angular limits and translation limits) into their registration framework, to avoid unrealistic joint transformation that may be generated due to noise and local minimum. Their framework treats both local and global deformations in successive regularization steps: Smooth elastic deformations are represented by an displacement field between the reference and current configuration of the template; global and discontinuous displacements are estimated through a projection onto a statistical shape model.

FE Modelling FE modelling, in the past three decades, has provided considerable understanding to the area of musculoskeletal biomechanics. Numerical models based on the FE method become very popular because it is able to address the complex geometries, the anisotropic material properties and the specific boundary conditions associated with living tissues [18]. The accuracy and efficiency of FE simulations (i.e. the solution to the partial differential equations) is highly predisposed to the quality of the finite element mesh [122]. Since the 4-noded tetrahedral mesh can be generated automatically if the information about organ geometry is available as closed surfaces, linear tetrahedral mesh is the most popular choice for subject-specific modelling. However, 4-noded tetrahedral mesh suffers from artificial stiffening, well-known as volumetric locking, when applied in modelling of incompressible (or nearly incompressible) continua [65], such as tongue and brain; the mesh-locking effect is illustrated in Figure 2.9. One approach to address this issue is to improve the tetrahedral elements to prevent locking. For example, Average Nodal Pressure (ANP) prevents volumetric locking by defining nodal volumes and evaluating average nodal pressures in terms of these volumes [14, 71]. ANP provides much better results for nearly incompressible materials than the standard tetrahedral element with only small increase in the computational cost. Another approach is to adopt hexahedra or hexahedra-dominant mesh. Hexahedra mesh not only avoids volumetric locking but also reduces the computational complexity of FE analysis, making it preferable for modelling incompressible tissues.

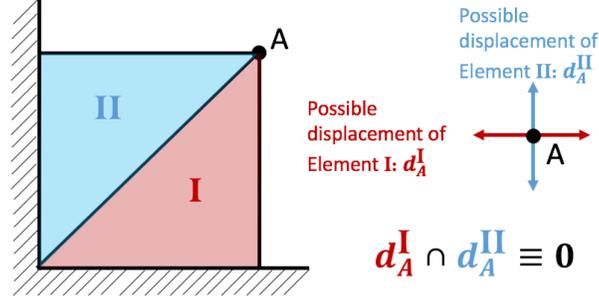


Figure 2.9: Illustration of the mesh-locking effect by incompressible material in a 2D case. Node A is the only free node. To maintain the area of the triangle element I, node A can only move in the horizontal direction; to maintain the area of triangle element II, node A can only move in the vertical direction. Therefore, maintaining the areas of the two triangle elements leads to zero displacements.

As aforementioned in this section, FE model (mesh) can be generated directly from organ geometry (surface) for each subject, known as meshing. This method is challenged by two facts. Firstly, automatic meshing techniques cannot identify distinct mesh features. For example, the 3D mesh of the face model in [23] consists of two distinct layers of elements: the outer layer represents the dermis tissues, while the inner layer models the hypodermis tissues. Secondly, automatic meshing for producing high-quality hexahedral FE mesh for complicated organ shapes remains to be an ill-posed problem. The semi-automated meshing, however, still needs an excessive amount of manual effort to achieve satisfactory results [18, 155].

Several studies have applied registration method to generate subject-specific FE models. Couteau et al. [30] propose a Mesh-Mathing (M-M) algorithm, which firstly computed the volumetric function \mathbf{T} (octree-splines) that transforms external nodes of the template FE mesh onto the target surface, and then apply \mathbf{T} to all internal nodes, leading to a new 3D mesh. Based on Couteau’s MM method, Bucki et al. [18] add a constraint on space distortion, which ensures the non-folding property at every point in space. Grosland et al. [54] morph a template hexahedral mesh (with good quality) to a target surface using FE method. Multiple levels of mesh refinement is utilized in their registration framework, i.e. the template mesh initialized with a low resolution and end with a fine resolution. Wang et al. [152] introduce a statistical template based approach to automate subject-specific FE

2.3. Subject-Specific Modelling

modelling. They first construct a statistical atlas from a shape population, including the statistical shape model and the FE model of the mean shape; then, based on correspondence established between the template and a given subject shape, the template shape evolves along the degrees of freedom (shape variation of the population) to obtain the best fit; finally, using Free-Form Deformation (FFD), they morph the internal nodes of the FE model of mean shape to conform with the new geometry. Campbell et al. [22] apply a similar approach to automate the subject-specific FE modelling of lumbar spine using a statistical shape model. Although the registration methods, mentioned above, start from good-quality template meshes, they may produce excessive spatial distortion when reposition the internal FE nodes and result in poor-quality elements. Several mesh untangling and quality improvement techniques have been proposed by previous studies [36, 76, 77]. However, they have no constraints on reallocating surface nodes, which may cause loss of the registration accuracy.

To generate good-quality subject-specific FE meshes, Luboz et al. [95] add an extra mesh-repair step to correct invalid elements after Mesh-Matching [30]. They repair the registered mesh using iterative approach: In each iteration, for the irregular elements, their nodes that have negative Jacobian value $|\mathbf{J}|$ are displaced in the gradient direction of $|\mathbf{J}|$. In order to maintain the registration accuracy, the maximal node displacements (distance between the initial and final position) are constrained. Following this Mesh-Match-and-Repair (MMRep) ideal, Bucki et al. [18] achieve mesh repair by a two-fold process: 1. recover the regularity of inverted elements ($|\mathbf{J}| < 0$); 2. improve the qualities of the elements to an acceptable level. In the second process, they use Jacobian Ratio (JR) as the measure of the overall distortion of an element. JR is defined as $|\mathbf{J}_n^e|/|\mathbf{J}_{max}^e|$, where $|\mathbf{J}_n^e|$ is the Jacobian value at node n in element e , and $|\mathbf{J}_{max}^e| = \max_{n \in e} |\mathbf{J}_n^e|$. They set $1/30$ as the minimum JR value and $5mm$ as the maximal nodal displacement in the repair step. However, Buchi's subject-specific FE modeling methods suffer from two limitations. First, MMRep only maintains a minimal mesh quality. When large deformation involves in biomechanical simulation, the resulting FE model may cause instability. Second, MMRep uses nodal test to decide the element regularity, i.e. an element is valid if and only if the Jacobian values at every nodes of it are positive. However, this conjecture is false for the hexedral element [75], which may cause the mesh repair to fail when MM produce excessive spatial distortion.

This state-of-art FE-mesh registration technique MMRep has been applied to many subject-specific modelling cases. For example, Chabanas

et al. [23] generate patient-specific FE models of the face by registering a generic face model to skin and skull surfaces segmented from CT scans. Their models are used to predict facial soft-tissue deformations resulting from bone repositioning in maxillofacial surgery. Harandi et al. [57] apply the MMRep method to generate a subject-specific model of oropharynx by registering FRANK components (tongue, jaw and hyoid) to a volumetric cine-MRI dataset captured during speech. MMRep provides a fast subject-specific FE mesh generation method. However, none of them focuses on registration of hybrid models that couple multiple rigid bodies and FE components.

Multi-structure Modelling A number of musculoskeletal models with multiple anatomical structures, such as bones, skins, muscles and tendons, have been proposed to enable realistic biomechanical simulations [7, 10, 80, 117, 137, 142]. In order to create fictional characters with precise internal anatomy for realistic animation, Ali-Hamadi et al. [3] propose an anatomy transfer method, which transfers a template anatomical model to an arbitrary target character defined by its boundary representation (skin). They first compute the registration of the template and the target skin. Then the template bones are transferred to the target character. Finally, the bone layer, along with the target skin eroded using the fat thickness information, are then used to define a volume where we map the internal anatomy of the template model to the target character using harmonic (Laplacian) deformation. This anatomy-transfer method can quickly generate anatomical models for a wide range of target characters. However, it has two unacceptable drawbacks in terms of subject-specific biomechanical modelling. Firstly, in most cases, the correspondence information extracted from the subject medical data is not the boundary of the unknown anatomical structures; hence, both interpolation and extrapolation are needed in order to transfer the anatomy. Besides, this method fails to preserve mesh quality of template models; it is very likely to generate invalid elements in FE models when large deformation is needed. Currently, creating subject-specific models with multiple anatomical structures attracts much less attention than subject-specific modelling of skeletons and isolated organs. There remains a need for an efficient method that can map multi-structure biomechanical models to given subject data and maintain their functionality.

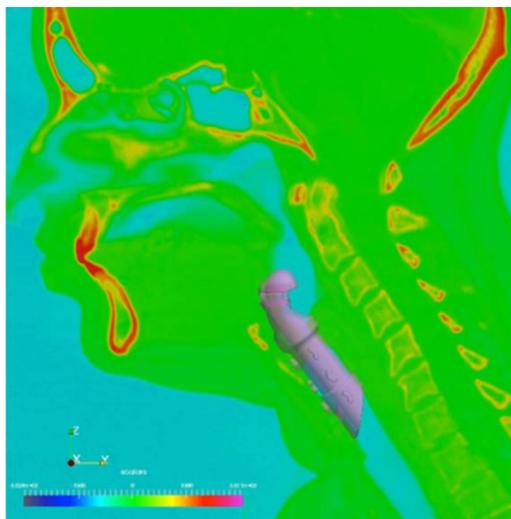


Figure 2.10: Virtual implantation of intralaryngeal prosthesis on a 3D reconstruction from a CT scan of a patient. © Wiley (2016), Raguin et al. [123], adapted with permission.

2.4 Biomedical Applications

Subject specific modelling can be put into the practice to assist in managing a wide range of different medical situations, such as analyzing dysfunctional cases, comparing treatment options when more than one possibility exists, and postoperative prediction. Here we review several potential biomedical applications that are most closely related to our efforts.

2.4.1 Intralaryngeal Prosthesis Implantation

Intralaryngeal prosthesis proves to be the optimal solution for all those patients who have dysphagia due to an impairment of the pharyngeal stage: Due to a deficiency in the brain or nervous system, the epiglottis is not able to prevent food and liquids from penetrating into the pulmonary tract. Raguin et al. [123] apply virtual implantation on 3D reconstruction from a CT scan sequence of the patient in pre-implantation step to determine the size of the prosthesis (Figure 2.10). We expect subject-specific models of upper-airway complex to assist with pre-implantation planning and customizing prosthesis by predicting the postoperative biomechanics of the associate behaviours such as swallowing and breathing.

2.4.2 Intensity Modulated Radiation Therapy

Intensity-modulated radiation therapy (IMRT) is an advanced mode of high-precision radiotherapy that uses computer-controlled linear accelerators to deliver precise radiation doses to a malignant tumor or specific areas within the tumor. Due to high radiation doses, patients with head and neck cancer have risk of suffering from swallowing disorder (or dysphagia) following the IMRT treatment. Swallowing-sparing intensity-modulated radiation therapy (sw-IMRT) tends to shape doses of radiation to avoid salivary glands and oropharyngeal structures thought to be essential to swallowing [113]. Song et al. [133] apply a mathematical model of trabecular bone to identify patients with high risk for IMRT treatment-related bone fracture. We expect to apply subject-specific models of upper-airway complex to assess swallowing muscle function of specific subjects, and for understanding of the relationship between the swallowing dysfunction risk factor and the radiation dose delivered by an IMRT plan, which can provide valuable information for IMRT treatment planning.

2.4.3 Hemiglossectomy and Hemimandibulectomy

Treatment of oral cancer commonly involves surgical resection of cancerous tissue, in addition to radiation therapy and chemotherapy. Depending on the size and location of the lesion, tissue resection can involve the portions of the mandible (hemimandibulectomy), tongue (hemiglossectomy), floor of mouth, and associated muscles [136].

Biomechanical tongue models have been used to simulate the effect of hemiglossectomy [16, 38]. These studies modified the structure of a tongue model to mimic tongue resection and reconstruction with free-flap soft-tissue grafts. The reported simulations deal primarily with the effect of stiffening a sub-region of the tongue, representing a lesion or reconstruction, on tongue movements.

Vascularized osteocutaneous, osteomyocutaneous and alloplastic grafts are commonly used to restore mandibular continuity after hemimandibulectomy [1, 59, 102, 119, 127]. Whether or not the jaw is reconstructed, Hemimandibulectomy can significantly alters jaw biomechanics and deficiencies in mastication, speech and other orofacial functions are often observed [56]. Numerical simulation methods have been applied to investigate the influence of jaw reconstruction on mandible movement and bite force [56, 136]; Stavness [136] use inverse simulation to find muscle activation strategy that can compensate for functional deficits after hemimandibulectomy.

Subject-specific modelling can provide anatomical and physiological information for a specific patient with oral cancer. It enables physicians to take account of the inter-subject variability of oropharyngeal biomechanics. We expect subject-specific models of upper-airway complex to predict the surgical influence on human behaviours, such as speech, swallowing and mastication, and to assist with biomedical treatment and rehabilitation planning.

2.4.4 Maxillofacial Surgery

Orthognathic surgery is addressed for patients suffering from maxillofacial dysmorphism of the lower part of the face, i.e. from disequilibrium between the mandible, the upper jaw and the face [23]. Subject specific modelling method has been applied to predict aesthetic outcomes of maxillofaical surgeries [23, 49, 108, 153]. The reported studies modify skull structure to mimic a surgical procedure, and then use passive FE models of the facial tissue to predict the resulting impact on the face surface. This approach can be expanded to incorporate active deformation (deformation in respond to muscle contraction), thereby allowing prediction of different post-operative facial expressions.

2.5 Discussion and Conclusion

Cadaveric study is the most straightforward process for assessing organ geometric data and mechanical properties. In recent years, a wide range of *in vivo* medical imaging techniques have been applied to visualize the complex anatomical structures of bones, soft tissues and muscles of the head and neck. As a natural extension of observational analysis, biomechanical modelling combines the geometric and mechanical information within a mathematical representation thereby enabling functional analysis. Many biomechanical models have been proposed in the literature for the jaw, tongue, face, and other upper-airway sub-components. By coupling these sub-components, a state-of-art holistic model of upper-airway complex, i.e. FRANK (Section 2.2.1), has been generated, representing the average human anatomy and function. Subject-specific modelling enables the transition from generic understanding of biomechanical phenomena to addressing biomechanics of a particular individual. Creating a biomechanical model relies heavily on expert interaction. The slow process of model creation prevents us from simulating large numbers of individual cases. To ease the efforts of subject-specific modelling, registration methods tends to morph

2.5. Discussion and Conclusion

a predefined template model to align with certain subject data. However, few published methods focus on registration of holistic models of the coupled upper-airway system, such as FRANK. A patient-specific model of the upper-airway system can enable motor-control simulations of complex human behaviours, such as swallowing and speech production. Thus, numerous biomedical applications will become feasible, including analyzing dysfunction and planning treatment for patients.

This chapter has presented an overview of biomechanical modelling of the upper-airway system, subject-specific modelling methods and the related biomedical applications. We have identified areas of research that require further investigation. The state-of-art biomechanical model of upper-airway complex are generic and irrelevant to medical data and measurement of any particular subjects. Currently available subject-specific modelling methods do not apply to hybrid and modularized models. The improvement of the functional resolution of the subject-specific model enables the simulations of complex human behaviours and can be applied to the potential treatment. In the following chapters we describe our contributions to these open research problems.

Chapter 3

Subject-Specific Modelling of Upper-Airway Complex

Subject-specific modelling enables individualized simulations best suited for clinical purposes. Subject-specific models may be directly reconstructed from medical image data. However, this approach proves to be time-consuming and labour-intensive in case of complex models. Registration methods aim at reducing the ambiguity introduced by medical data and easing the re-designing efforts: a deformation field is sought that morphs a pre-designed template model to a target dataset. In previous chapter, we reviewed a functional template of upper-airway complex: FRANK. In this chapter, we introduce a registration method for transferring this holistic model from the generic space into a specific subject space.

Model registration for constructing subject-specific upper-airway models has two main challenges. First, complete medical data for all components involved in FRANK is usually hard to obtain for a single subject; hence the model geometry partially is unknown in most cases. To address this issue, model registration needs to minimize the morphological deviation from the average anatomy (i.e. the template) when enforcing the correspondence constraints.

The second challenge of registration methods is preserving the functionality of the resulting model in terms of two aspects: 1. the numerical stability and accuracy; 2. the correct motor control behaviours. As the mesh quality will significantly influence the numerical accuracy and stability of FE models, maintaining mesh quality is one of the most important requirements for the registration. Besides, in order to obtain correct motor control behaviours for a hybrid and modularized model, such as FRANK, the coupling constraints, such as the connectivity between the FE components, the location of joints, the positions of muscle attachments need to be consistent with the new model geometry.

We summarize the requirements mentioned above as three types of regularity: 1. **inter-component regularity**: maintaining the spatial relationship, including connectivity, topology, relative posture and size, between

the model components; 2. **intra-component regularity**: preserving the shape regularity of the underlying discretization structures of every sub-components; 3. **functional regularity**: keeping the functional information (including coupling attachments between components, muscle attachments and biomechanical properties) similar to the template but relevant to the new model geometry.

In following sections, we describe a subject-specific upper-airway complex modelling method that aims at morphing the holistic template to the subject data and maintaining the three types of regularity. This method is based on a novel multi-structure registration technique, which is able to maintain both the inter- and intra-component regularity, and to accommodate different geometric discretization. We first introduce the multi-structure registration technique in Section 3.1. Then, the detailed description of our proposed subject-specific modelling framework follows in Section 3.2. This framework is evaluated by registering FRANK onto two volumetric medical image dataset of the human head and neck. As a result, we obtain two subject-specific models of upper-airway complex and demonstrate their functionality in biomechanical simulation of simple speech motions. Finally, the discussion and conclusion are given in Section 3.3.

3.1 Multi-Structure Registration

Efforts to register a complete holistic model, such as FRANK, into the subject domain mostly suffer from excessive inter- and intra-component distortion, specially when there is a large morphological difference or sparse correspondences between the template and the subject data.

Free-form deformation (FFD) [52, 129] is frequently used for non-rigid registration [2, 18, 125, 126, 128, 163]. Meshes or images are embedded into a deformable virtual grid, and morphed alongside, while the grid vertices move. Through regularizing the grid deformation, FFD minimizes the distance between the correspondence pairs, while smoothly interpolating in the regions where no explicit correspondence exists. Tensor product B-spline volume is the most popular FFD model for registration due to its low computational complexity. To model a large range of deformations, multi-resolution B-spline FFDs are commonly used in previous studies [2, 18, 125, 126, 128, 163]; FFD grids are progressively refined to sequentially reduce the registration error. This approach maintains an adjustable registration accuracy, and is often sufficient to avoid spatial folding, as illustrated in Figure 3.1. However, multi-resolution B-spline FFD does not

3.1. Multi-Structure Registration

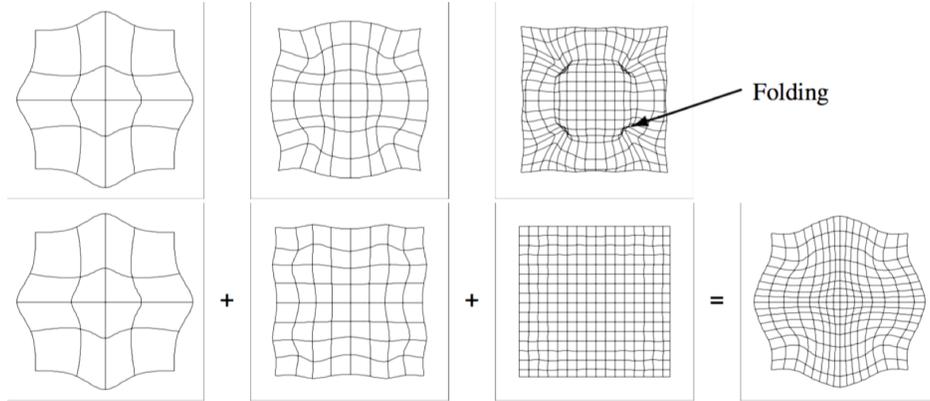


Figure 3.1: Top: Low FFD resolutions limit the registration accuracy; high FFD resolutions may develop spatial folding (see arrow). Bottom: Multi-resolution FFD can maintain an adjustable registration accuracy, without developing any spatial folding. © Springer (2001), Schnabel et al. [128], adapted with permission.

guarantee a homeomorphism (i.e. continuous, invertible transformations). To address this issue, Edwards et al. [32] model the non-folding property as a soft constraint. Choi and Lee [27], however, derive sufficient conditions for FFD injectivity which are represented in terms of control point displacements. Rueckert et al. [126] blend these injectivity conditions into their composite-FFD framework in order to achieve diffeomorphic image registration. Homeomorphic FFD maintains the spatial relationship between embedded components. However, the overall regularity of the grid deformation cannot guarantee the regularity within each embedded component itself (i.e. FFD is not shape-aware). When large deformation is required for registration, quality of the embedded meshes is very likely to be undermined. Besides, the deformation in the interpolated region can be arbitrary, i.e. the deformed morphologies may vary for different initial settings of the virtual grid when the correspondence is sparse.

Local Transformation Model (LTM) is another popular deformation model for non-rigid registration [5, 6, 64, 81, 83, 85, 112, 143, 161]. The LTM mapping function is defined as $F(\mathbf{p}_i) = \mathbf{T}_i \mathbf{p}_i + \mathbf{t}_i$, where \mathbf{p}_i is the i -th vertex position; \mathbf{T}_i and \mathbf{t}_i are the local linear transformation and the translation respectively. Smuth et al. [143] and Huang et al. [64] define the local transformations \mathbf{T}_i as the rotations, $\mathbf{T}_i = \mathbf{R}_i$, leading to As-Rigid-As-Possible

(ARAP) deformations. ARAP deformations are length-preserving; hence they can effectively avoid mesh distortion in registration. However, these rigid local transformation models are usually over-constrained for registration; they are incapable of handling the correspondences with different sizes or those which undergo large local stretching [162]. To increase the flexibility of the deformations, several As-Similar-As-Possible (ASAP) mapping methods are proposed in previous studies [81, 112, 161] for registration. The local transformations \mathbf{T}_i are defined as a combination of scaling and rotation, $\mathbf{T}_i = s_i \mathbf{R}_i$. ASAP can effectively avoid mesh distortion and is flexible enough to allow both local and global scaling. As-Affine-As-Possible (AAAP) deformations allow more freedom to capture fine details in registration [5, 6, 83, 85]. Their local transformations are defined as affine transformations, $\mathbf{T}_i = \mathbf{A}_i$. Since AAAP deformations have large degrees of freedom, they are very likely to make the registration under-constrained, particularly when the correspondence is sparse. Therefore, additional constraints are often applied to regularize the registration system. For example, Allen et al. [5] and Amberg et al. [6] add smooth constraints into their registration frameworks to ensure adjacent affine transformations are similar. LTMs are useful for avoiding distortion within a single component. However they cannot preserve the spatial relationship between different components.

Noticing that FFD and LTM can compensate the shortcomings of each other, we choose to combine a B-Spline FFD and an ASAP mapping into a single transformation model: Structure-Preserving Free-Form Deformation (SPFFD). SPFFD permits two exciting deformation properties: 1. the embedded shapes are mapped to a target configuration by a homeomorphism, i.e. the mapping function is continuous and topology-preserving; 2. for each shape, the mapping function is a similarity transformation (nearly), which avoids excessive distortion while being flexible enough to allow local and global scaling. Moreover, the proposed SPFFD can accommodate different geometric discretization (e.g. surface mesh, volumetric mesh). We first demonstrate the overall framework of the proposed multi-structure registration in Section 3.1.1. Then we introduce the generalized ASAP deformation in Section 3.1.3. The full description for SPFFD is given in Section 3.1.4. Finally, we demonstrate the performance of our proposed registration framework using a synthetic data set; the experiments and results are demonstrated in Section 3.1.5.

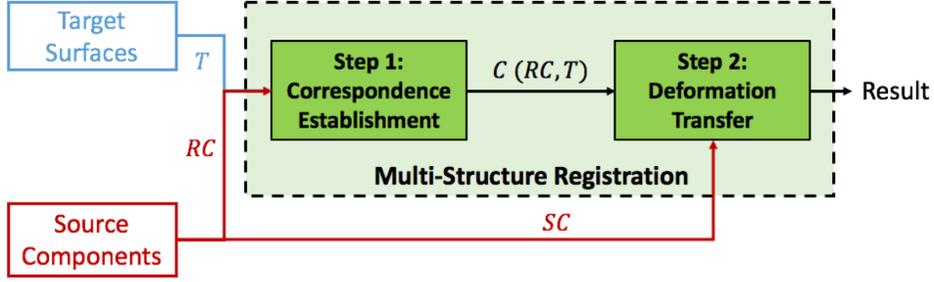


Figure 3.2: Proposed multi-structure registration framework.

3.1.1 Registration Framework

Our multi-structure registration pipeline is illustrated using a simple 2D didactic case in Figure 3.3. Given a set of Source Components (SC) (rectangles in Figure 3.3(a)), we assume some of them have point-to-point correspondences with certain target surfaces (T) (ellipse in Figure 3.3(a)); we refer to these components as Reference Components (RC) (red rectangle in Figure 3.3(a)). As shown in Figure 3.2, we register the source components to the target surfaces in two sequential steps:

Step 1. Correspondences Establishment: The first step is to establish the correspondences between the reference components and the target surfaces, we refer to them as $\mathcal{C}(RC, T)$ (Figure 3.3(c)). We find $\mathcal{C}(RC, T)$ by registering the surface of RC to the target surfaces using a modified extrinsic Iterative Closet Point (ICP) method, which is described in Section 3.1.2. By minimizing the correspondence energy E_{corr} , we maximize the closest-point shape similarity between the reference components and the target surfaces.

To allow enough flexibility for registration of two shapes with different morphology, and to preserve the fine details of the template, we employ an As-Similar-As-Possible (ASAP) mapping. The ASAP mapping is obtained by minimizing a deformation energy, $E_{ASAP} = \sum_n E_{local}(n)$, where E_{local} is local energy that measures the difference between deformation gradients and the corresponding similarity transformations. Details of our ASAP mapping are introduced in Section 3.1.3.

Thus we established the correspondence $\mathcal{C}(RC, T)$ by minimizing the following energy:

$$E_{CE} = E_{corr} + w_{ASAP}E_{ASAP}. \quad (3.1)$$

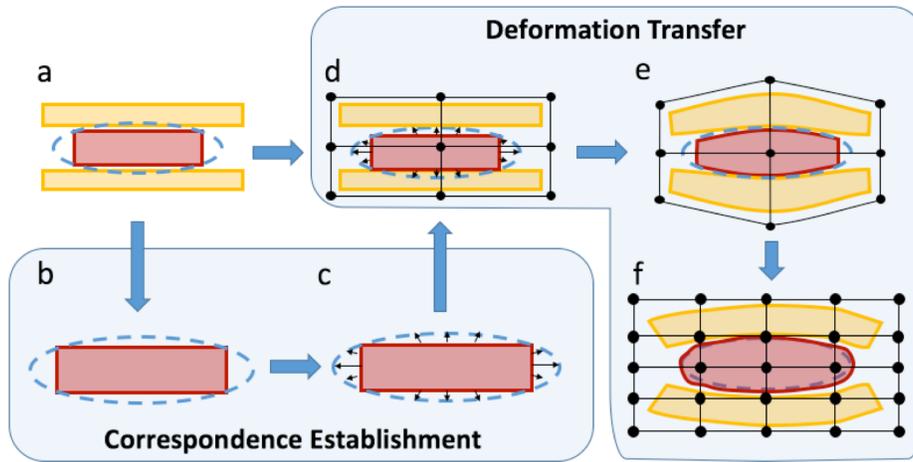


Figure 3.3: Illustration of multi-structure registration pipeline. (a) the initial setting: all the rectangles are source components; the red rectangle represents the reference component, and other two rectangles are expected to deform along with the red. The ellipse by dash line represents the target surface. (b)-(c) establish the correspondences between the red rectangle and the ellipse by computing the registration between them. (d)-(f) all the rectangles are embedded into a SPFFD grid; by moving the SPFFD control points (black dots), the distances between the correspondences are minimized while making minimal morphological change on the rectangles. In order to convexify the objective function and prevent overfitting, a coarse-to-fine strategy is applied: SPFFD starts from a coarse deformation grid, and then gradually increases the grid resolution.

3.1. Multi-Structure Registration

We initialize the registration with high ASAP weights $w_{ASAP} = 1.0$. If the relative correspondence energy did not change significantly between the iteration j and $j + 1$ (i.e. $|E_{corr}^j - E_{corr}^{j+1}|/E_{corr}^j < 0.1$), we additionally relax the regularization weights $w_{ASAP}^{j+1} = w_{ASAP}^j - 0.01$, until $w_{ASAP} < 0.02$. The adaptation of weights initially favors global rigid alignment and subsequently lowers the stiffness of the surface of RC to allow increasing deformation as the optimization progresses. Hand tuning is not necessary in the correspondences-establishment step.

Step 2. Deformation Transfer: In the second step, we apply the proposed Structure-Preserving Free-Form Deformation (SPFFD) to generate a homeomorphic, shape-aware mapping function, which transfers the deformation of the reference components to all the source components through an energy minimization (from Figure 3.3(d) to Figure 3.3(f)). The known correspondences, $\mathcal{C}(RC, T)$, act as constraints for our deformation transfer process. Since such constraints are over-deterministic, we enforce them softly by minimizing a correspondence energy $E_{corr'}$, which is formulated in Section 3.1.2. SPFFD maintains the inter- and intra-component shape regularity by minimizing a smooth energy E_{smooth} and a shape-preserving energy E_{shape} respectively. The details of our proposed SPFFD are given in Section 3.1.4.

Sum the individual energy terms from above to form the full objective function of the optimization:

$$E_{DF} = E_{corr'} + w_{smooth}E_{smooth} + w_{shape}E_{shape}. \quad (3.2)$$

We keep the two regularization weights (the smooth weight w_{smooth} and shape-preserving weight w_{shape}) at the same level, i.e. $w_{smooth} = w_{shape} = w_{reg}$. The regularization weight w_{reg} controls the flexibility of the source components, and can be tuned to allow appropriate levels of deformation for different applications.

3.1.2 Correspondence Computation

For discrete-shape registration, correspondences are defined for each vertex as the displacement that maximizes a certain similarity measure [47]. Extrinsic similarity measures are based on the current configuration of the surfaces in the Euclidean space. Closest point criterion is the most popular extrinsic similarity measure for discrete-shape registration for its simplicity [15]. Iterative Closest Point (ICP) algorithm [12] seeks to register

3.1. Multi-Structure Registration

two shapes by finding the best transformation that minimizes the extrinsic distance between them. Since the error function of the closest point registration framework is non-linear and non-smooth, ICP algorithm is very likely to converge at certain local minima, particularly for non-rigid registration. Intrinsic methods, on the other hand, use distance metrics embedded in lower dimensions as the shape similarity measures [15, 64, 74, 87, 100]. These intrinsic properties of shapes are quasi-invariant under object pose and deformation [15]. However, intrinsic similarities are sensitive to topological noise, which makes them inappropriate for partial registration. Probabilistic correspondences are another class of shape-similarity measurements [28, 51, 69, 90, 91, 109, 145]. Instead of assuming the one-to-one (binary) correspondence, one-to-many relaxations are applied to allow for fuzzy correspondences [69]. Probabilistic similarity measures are less sensitive to the missing correspondence and outliers. In addition, from optimization point view, probabilistic registration frameworks can replace a non-linear, non-smooth error function by a convex smooth approximation [69]. These properties make them valuable when the template and target shapes are complex, noisy or partially overlapping. However, probabilistic similarity measures often lead to considerably larger computational cost.

For the simplicity of implementation and the relatively low computational cost, we establish the correspondences between the surfaces of RC and the target surfaces using a pairwise ICP method. At each iteration, the template surfaces are projected to the corresponding target surfaces and vice-versa, until a satisfactory error distance is obtained (Figure 3.4(a)). In order to avoid local minima, we assign a larger standard deviation of distances in the tangent direction compared to the normal direction, as shown in Figure 3.4(b). The correspondence energy is then defined as:

$$E_{corr} = \sum_c (\mathbf{v}_c - \mathbf{t}_c)^T \mathbf{Cov}_c^{-1} (\mathbf{v}_c - \mathbf{t}_c), \quad (3.3)$$

where c is the index for correspondence pairs; \mathbf{v}_c is the vertices on the surface of RC, and \mathbf{t}_c is their target positions; \mathbf{Cov}_c is the 3×3 covariance matrix for the c -th correspondence, which is calculated as:

$$\mathbf{Cov}_c = \mathbf{R}_c \cdot \begin{bmatrix} \epsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{R}_c^T, \quad (3.4)$$

where ϵ is a small constant representing covariance along the surface normal. \mathbf{R}_c is a 3×3 orthogonal matrix; the first column of \mathbf{R}_c is the normal vector

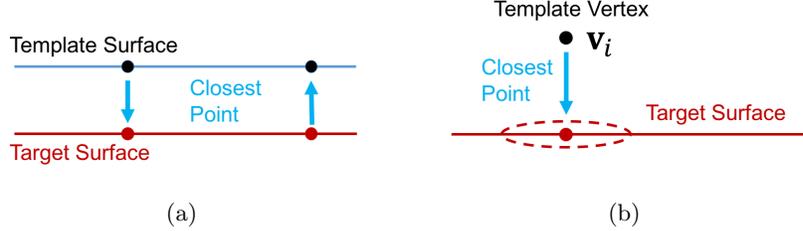


Figure 3.4: Iterative closest point (ICP) correspondence. (a) The template surface is projected to the target surface and vice-versa. (b) In order to avoid local minimum, a larger standard deviation of distances is assigned in the tangent direction compared to the normal direction.

of the target surfaces at position \mathbf{t}_c ; and the other two columns contain the two basis vectors of the tangent plane.

In the deformation-transfer step, the correspondence energy $E_{corr'}$ is similar to Equation 3.3, as:

$$E_{corr'} = \sum_{c \in \mathcal{C}(RC, T)} (\mathbf{v}_c - \mathbf{t}_c)^T \mathbf{Cov}_c^{-1} (\mathbf{v}_c - \mathbf{t}_c). \quad (3.5)$$

Different from Equation 3.3, the target positions \mathbf{t}_c are fixed in this equation.

3.1.3 As-Similar-As-Possible Deformation

Yamazaki et al. [161] proposed an ASAP mapping method for triangle meshes that proves to have good convergence rate and relatively low computational cost. We generalize Yamazaki’s method to accommodate different geometric discretization. Suppose U_n is a certain local unit on the 3D mesh M . The global deformation energy E_{ASAP} of mesh M is the weighted sum of local energies $E_{ASAP}(n)$ measured at local units U_n , denoted as:

$$E_{ASAP} = \sum_n w_n E_{ASAP}(n), \quad (3.6)$$

where w_n are weights for local energies.

Suppose \mathbf{v}_i is the position of the i -th vertex on the mesh M ; \mathbf{v}'_i is its deformed position. Each local unit U_n associates with a set of neighbour vertices $\mathcal{N}(n)$. Our ASAP mapping restricts the local linear transformation

3.1. Multi-Structure Registration

for U_n to a similarity transformation. We define the local energy for U_n as:

$$\begin{aligned} E_{ASAP}(n) &= \left(\frac{1}{s_n} \tilde{\mathbf{R}}_n^T \mathbf{x}'_n - \mathbf{x}_n \right)^T \mathbf{K}_n \left(\frac{1}{s_n} \tilde{\mathbf{R}}_n^T \mathbf{x}'_n - \mathbf{x}_n \right) \\ &= \mathbf{x}_n^T \mathbf{K}_n \mathbf{x}_n - \frac{2}{s_n} \mathbf{x}_n'^T \tilde{\mathbf{R}}_n \mathbf{K}_n \mathbf{x}_n + \frac{1}{s_n^2} \mathbf{x}_n'^T \tilde{\mathbf{R}}_n \mathbf{K}_n \tilde{\mathbf{R}}_n^T \mathbf{x}'_n \end{aligned} \quad (3.7)$$

where \mathbf{x}_n is the concatenated vectors of vertices $\mathbf{v}_i \in \mathcal{N}(n)$; \mathbf{K}_n is a local stiffness matrix associated with unit U_n ; \mathbf{R}_n and s_n are the local rotation matrix and the scaling factor at unit U_n ; a tilde represents taking the Kronecker product, $\tilde{\mathbf{R}}_n = \mathbf{I}_{3 \times 3} \otimes \mathbf{R}_n$, where $\mathbf{I}_{3 \times 3}$ is an identity matrix.

Following Sorkine and Alexa [135], we minimize E_{ASAP} by iterative alternating between local and global phases:

- **Local Phase:** Fix the vertices, and compute best fit local rotations and scaling factors:

We determine the local rotations by:

$$\mathbf{R}_n = \underset{\mathbf{R}_n}{\operatorname{argmin}} E_{ASAP}(n); \quad (3.8)$$

If the local stiffness \mathbf{K}_n is rotational-invariant (i.e. $\mathbf{K}_n = \tilde{\mathbf{R}}_n \mathbf{K}_n \tilde{\mathbf{R}}_n^T$), Equation 3.8 can be expanded as:

$$\mathbf{R}_n = \underset{\mathbf{R}_n}{\operatorname{argmax}} \mathbf{x}_n'^T \tilde{\mathbf{R}}_n \mathbf{K}_n \mathbf{x}_n; \quad (3.9)$$

In contrast, if \mathbf{K}_n in Equation 3.7 is not rotation-invariant, i.e. $\mathbf{K}_n \neq \tilde{\mathbf{R}}_n \mathbf{K}_n \tilde{\mathbf{R}}_n^T$, we assume that the incremental local rotation $\delta \mathbf{R}_n$ is negligible in each iteration so that $[\tilde{\mathbf{R}}_n]_{t+1} \mathbf{K}_n [\tilde{\mathbf{R}}_n]_{t+1}^T \approx [\tilde{\mathbf{R}}_n]_t \mathbf{K}_n [\tilde{\mathbf{R}}_n]_t^T$, where $[\mathbf{R}_n]_t$ is the local rotation at the t -th iteration. Then we can determine the optimal $\delta \mathbf{R}_n$ by:

$$\delta \mathbf{R}_n = \underset{\delta \mathbf{R}_n}{\operatorname{argmax}} \mathbf{x}_n'^T \delta \tilde{\mathbf{R}}_n [\tilde{\mathbf{R}}_n]_t \mathbf{K}_n \mathbf{x}_n; \quad (3.10)$$

We optimize Equation 3.9 and 3.10 using the Singular Value Decomposition (SVD) method proposed by [9].

To update the local scaling factors, we use the updated \mathbf{R}_n values to enforce that the partial derivatives of Equation 3.7 w.r.t s_n vanish; this leads to:

$$\begin{aligned} s_n &= \underset{s_n}{\operatorname{argmin}} E_{ASAP}(n) \\ &= \frac{\mathbf{x}_n'^T \tilde{\mathbf{R}}_n \mathbf{K}_n \tilde{\mathbf{R}}_n^T \mathbf{x}'_n}{\mathbf{x}_n'^T \tilde{\mathbf{R}}_n \mathbf{K}_n \mathbf{x}_n} \end{aligned} \quad (3.11)$$

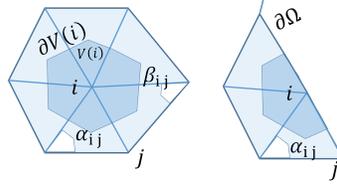


Figure 3.5: Notation for triangular mesh.

- **Global Phase:** Fix the local rotations and scaling factors, and update vertices by minimizing the deformation energy E_{ASAP} in Equation 3.6.

Due to the fact that our model is closely related to volumetric mesh and triangle mesh, without losing generality, we formulate the ASAP mapping for these two discretizations as follow.

Triangle Mesh Given a triangle mesh S_t , we interpret the unit U_i as the one-ring neighbours denoted by $\mathcal{N}(i)$, which is a set of vertices connected to the i -th vertex, as shown in Figure 3.5. Analogue to [135], we define the local energy as:

$$E_{ASAP}(i) = \sum_{j \in \mathcal{N}(i)} w_{ij}^2 \left\| \frac{1}{s_i} \mathbf{R}_i^T (\mathbf{v}'_i - \mathbf{v}'_j) - (\mathbf{v}_i - \mathbf{v}_j) \right\|^2, \quad (3.12)$$

which can be compactly written in the form of Equation 3.7. Noted that $\mathbf{K}_i = \tilde{\mathbf{R}}_i \mathbf{K}_i \tilde{\mathbf{R}}_i^T$, the local rotation \mathbf{R}_i is determined using Equation 3.8. To avoid discretization bias, we adopt the cotangent weighting formula [103, 121]:

$$w_{ij} = \begin{cases} \frac{1}{2}(\cot \alpha_{ij} + \beta_{ij}) & \{ij\} \text{ is an interior edge;} \\ \frac{1}{2}(\cot \alpha_{ij}) & \{ij\} \text{ is a boundary edge;} \\ 0 & \text{otherwise,} \end{cases} \quad (3.13)$$

where α_{ij}, β_{ij} are the angle opposite of the mesh edge (i, j) . To compute the global deformation energy, we set the weights in Equation 3.6 as $w_i = \frac{1}{A_i}$, where A_i is the Voronoi area of cell at center vertex \mathbf{v}_i [103].

Volumetric Mesh Given a volumetric mesh S_v , using Finite Element(FE) method, we interpret the unit U_e as the e -th element; the nodes of U_e are

denoted by $\mathcal{N}(e)$. We update the global phase by minimizing the mechanical strain energy E_ϵ as:

$$E_\epsilon = \int_{\Omega} \sigma : \epsilon \, dV, \quad (3.14)$$

where σ and ϵ are the stress and the strain tensor respectively. The local stiffness can be denoted as:

$$\mathbf{K}_e = \int_{\Omega_e} \mathbf{B}_e^T \mathbf{D}_e \mathbf{B}_e \, dV, \quad (3.15)$$

where \mathbf{D}_e is a constitutive matrix and \mathbf{B}_e is a differential operator that maps the nodal-displacement vector to the strain vector. This local stiffness matrix is not rotational-invariant; therefore, we determine the local rotation using Equation 3.10.

3.1.4 Structure-Preserving Free-Form Deformation

We embed the template shape into a B-spline FFD virtual grid as in [126]. We then deform the shape by manipulating a lattice of control points in a continuous deformation field. The deformed position \mathbf{v}' of an arbitrary embedded point \mathbf{v} is determined by:

$$\mathbf{v}' = \sum_{i=0}^n \sum_{j=0}^m \sum_{k=0}^l N_{i,p}(u) N_{j,q}(v) N_{k,r}(w) \mathbf{p}_{ijk}, \quad (3.16)$$

where \mathbf{p}_{ijk} is the control point labelled with the lattice index i, j, k . And n, m, l are the total number of control points in each direction; $N_{x,y}()$ represents the x -th basis function (of degree p) of the 1D B-spline. The parametric coordinates u, v, w of an arbitrary point are found using the Newton-Raphson method, which searches within the bounds determined by the knots span where \mathbf{v} belongs. In order to reduce computation cost we choose linear basis functions. In its compact form, Equation 3.16 can be written as:

$$\mathbf{v}' = \mathbf{B}\mathbf{p}, \quad (3.17)$$

where \mathbf{B} is a sparse matrix (local support property [120]) which contains all the basis functions; \mathbf{p} is the concatenated vector of the control points.

To avoid undesired distortion, we regularize the FFD grid deformation using its first-order derivatives, which define a global smoothness constraint, as:

$$E_{smooth} = \int_{\Omega} \left(\|\nabla d_x\|^2 + \|\nabla d_y\|^2 + \|\nabla d_z\|^2 \right) \, dV, \quad (3.18)$$

3.1. Multi-Structure Registration

where ∇ is the gradient operator; d_x, d_y, d_z are the displacements in the three canonical orthogonal directions. Using the variational principle, E_{smooth} can be minimized by changing the positions of the FFD control points.

The global smoothness constraint in Equation 3.18 is not shape-aware; hence it cannot maintain the regularity of the individual embedding mesh structures. We introduce shape-aware local rigidity by coupling our ASAP energy (Equation 3.6) with the FFD (Equation 3.17):

$$\begin{aligned}
 E_{shape}(M) &= \sum_n \left(\frac{1}{s_n} \tilde{\mathbf{R}}_n^T \hat{\mathbf{B}}_n \mathbf{p} - \mathbf{x}_n \right)^T \mathbf{K}_n \left(\frac{1}{s_n} \tilde{\mathbf{R}}_n^T \hat{\mathbf{B}}_n \mathbf{p} - \mathbf{x}_n \right) \\
 &= \mathbf{p}^T \left[\sum_n \frac{1}{s_n^2} \hat{\mathbf{B}}_n^T \hat{\mathbf{K}}_n \hat{\mathbf{B}}_n \right] \mathbf{p} \\
 &\quad - 2\mathbf{p}^T \left[\sum_n \frac{1}{s_n} \hat{\mathbf{B}}_n^T \hat{\mathbf{K}}_n \left(\tilde{\mathbf{R}}_n \mathbf{x}_n \right) \right] + \sum_n \mathbf{x}_n^T \mathbf{K}_n \mathbf{x}_n,
 \end{aligned} \tag{3.19}$$

where $\hat{\mathbf{B}}_n$ is the concatenated matrix of the basis function \mathbf{B} (in Equation 3.17), so that $\mathbf{x}'_n = \hat{\mathbf{B}}_n \mathbf{p}$; $\hat{\mathbf{K}}_n$ is the warped stiffness matrix: $\hat{\mathbf{K}}_n = \tilde{\mathbf{R}}_n \mathbf{K}_n \tilde{\mathbf{R}}_n^T$. The last term is a constant. Equation 3.19 gives the deformation energy of the embedded objects, which is controlled by the positions of the FFD control points. Therefore, The shape-preserving energy $E_{shape}(M)$ contributes a local stiffness for mesh structure (or component) M , which enables the FFD volume to be shape-aware. The overall shape-preserving energy E_{shape} in Equation 3.2 is formulated as the weighted sum of the energies of individual mesh components, i.e. $E_{shape} = \sum_M w_M E_{shape}(M)$. For the reference components that have complete target geometry we set their weights w_M as 1.0; For the components that do not have target information, we set their weights w_M as 3.0 to enforce larger morphological constraints.

Similar to the composite FFD proposed by [126], a coarse-to-fine strategy is applied to avoid over-fitting: the FFD starts from a coarse grid, and gradually increases its resolution until a satisfactory error distance is obtained. In each level, when no significant energy decreases is obtained by moving the control points, the FFD grid is re-meshed with additional DoF (as shown in Figure 3.3(e) and (f)). The pose of the FFD grid in world coordinates is initialized to align with the three principle directions of the embedded objects. The initial state of the global smooth energy (Equation 3.18) is restored when the FFD grid goes to the next level, while accumulating the shape energy (Equation 3.19) of each embedded object. This strategy allows large deformations while maintaining the regularity of embedded shapes. In order to avoid spatial folding, we set the maximal displacement

of each control point as one third of the FFD grid spacing [27]; this leads to a homeomorphism: bijective, continuous, non-folding mapping.

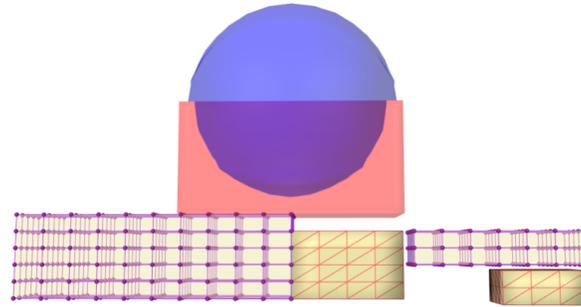
3.1.5 Experiment

To demonstrate the performance of our registration method, we arranged a simple hybrid registration problem that involves both volumetric and surface meshes with sparse correspondences. The initial setting is illustrated in Figure 3.6(a). The red box surface is (the only) reference component in this experiment, and the blue sphere is its target surface. All the components placed at the bottom, including two volumetric meshes and two surface meshes, are expected to deform along with the red box.

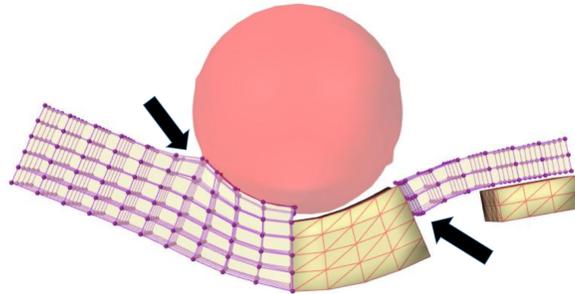
As described in Section 3.1.1, we first established the correspondence between the red box and the blue sphere. Then, our proposed structure-preserving FFD (based on linear B-spline volumes) was used to deform all the source components enforcing the correspondence constraints. The final result is illustrated in Figure 3.6(b). We compared our method with the diffeomorphic FFD (based on cubic B-spline volumes) proposed by [126]. The resulting configuration by their method is shown in the Figure 3.6(c). As it can be seen in the figure, both methods preserved the spatial relationship between source components; our method maintained the shape regularity of the meshes, while the diffeomorphic FFD caused excessive mesh distortion.

Figure 3.7 and 3.8 show the registration error and mesh quality over the course of 300 iterations. The registration error is calculated as the surface distance between the reference component (red box) and its target surface (blue sphere). The mean values of the registration error is similar between the two methods. Our method shows larger maximum error, since it compromises some correspondence constraints to preserve shape regularity.

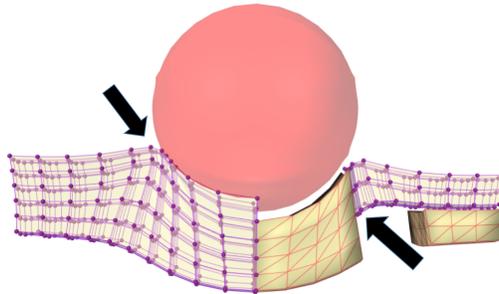
We calculate the mesh quality (for both volumetric meshes and surface meshes) using a shape-measure metric: mean ratio (η) [70, 89]. Shape-measure metrics are invariant under translation, rotation, reflection and uniform scaling of the element; they attain the maximum (1.0) when the element is in ideal configuration, and minimum (0.0) when the element is degenerated [76]. For both the volumetric and surface meshes, our method shows considerably better mesh quality. The grid resolution increases with the number of iterations which, in turn, provides higher degrees of freedom to both methods. Our method is able to recover the mesh quality after sufficient iterations (i.e. around 50 for η_{mean} , and 175 for η_{min}). However, since the diffeomorphic FFD is not shape-aware, it fails to recover the mesh quality.



(a) Initial setting



(b) Proposed method



(c) Diffeomorphic

Figure 3.6: Illustration of the registration performance on the synthetic data: (a) the initial setting; the red box and the blue sphere are the reference and target surfaces. Four other meshes – two volumetric FE meshes and two triangular surface meshes (yellow) – are expected to deform along with the red box. (b) result of our proposed method; both the inter- and intra-component regularities are preserved. (c) result of the diffeomorphic registration [126]; the method preserves the spatial relationship between the source components, but generates excessive mesh distortion, e.g. the elements located at the pointed region are excessively warped, which does not happen in the result by our method.

3.1. Multi-Structure Registration

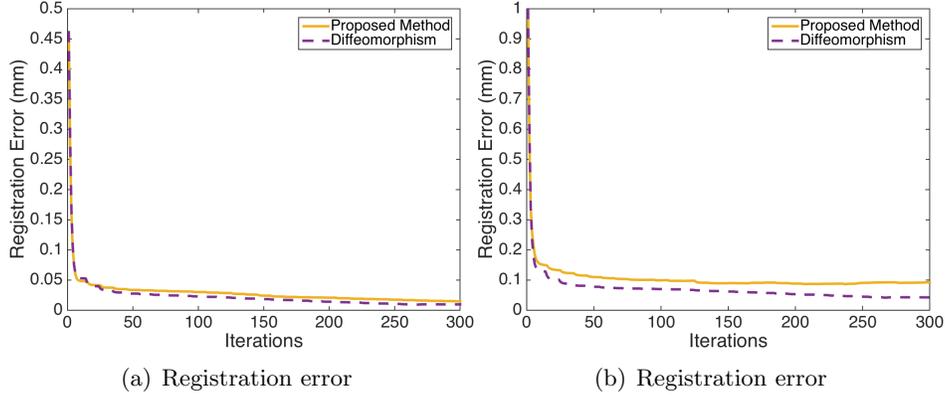


Figure 3.7: Registration error (surface distance) measured over the courses of 300 iterations. Our results (in solid lines) are compared with the diffeomorphic registration by Rueckert et al. [126] (in dashed lines): (a) the average registration error (Error_{mean}); and (b) maximum registration error (Error_{max}).

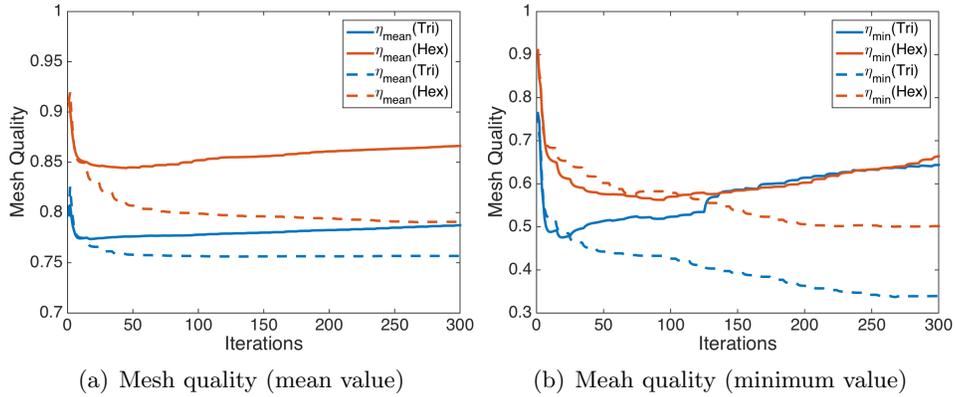


Figure 3.8: Mesh quality (η) measured over the courses of 300 iterations. Our results (in solid lines) are compared with the diffeomorphic registration by Rueckert et al. [126] (in dashed lines): (a) the average mesh quality (η_{mean}); and (b) the minimum mesh quality (η_{min}).

3.2 Subject-Specific Modelling

Figure 3.10 shows the framework for subject-specific modelling of upper-airway complex as we move from the medical images, and the hybrid template model, to creating and simulating our subject-specific models. Our medical data has been briefly introduced in Chapter 2; more details are given in Section 3.2.1. We revisit the comprehensive template model (FRANK) and briefly outline its major features in Section 3.2.2. As shown in the figure, we morph the template into the subject-space in three steps: Correspondence establishment, anatomy and functionality transfer. The details of this framework are described in Section 3.2.3

3.2.1 Medical Image Data

We build subject-specific models based on two volumetric medical images, shown in Figure 3.9.

CT Dataset : Volumetric images of a male subject during a swallow were captured in a single shot using a 320-detector-row CT scanner (Toshiba Medical Systems, Otawara, Japan) [67]. The spatial resolution of the images is $0.5mm \times 0.5mm \times 0.5mm$, and the temporal resolution is 10 Hz. This dataset has 21 time frames in total. The subject was seated on the chair in a semi-reclining position at an angle of about 45° . We morphed the FRANK template into the last frame of the dynamic CT scans, where the subject and the FRANK are in the similar posture, as shown in Figures 3.9(a)–(c).

MR Imaging Dataset : An atlas image of the human head-and-neck was constructed using 3D MR images captured on 20 normal subjects in the supine position [157]. All MRI scanning was performed on a Siemens 3.0T Tim Trio system (Siemens Healthcare, Inc., Malvern, PA, USA). A T2-weighted Turbo Spin Echo sequence with an echo train length of 12 and TE/TR of $62ms/2500ms$ was used. The dimensions of the 3D MR images are $255 \times 255 \times z$ (where z ranges from 10 to 24), with $0.94mm \times 0.94mm$ in-plane resolution and 3mm slice thickness. The resulting atlas is an isotropic volume with spatial resolution of $0.94mm$. Detailed description of the involved image processing techniques can be found in [157]. The MR atlas image is shown in Figures 3.9(d)–(f).

3.2. Subject-Specific Modelling

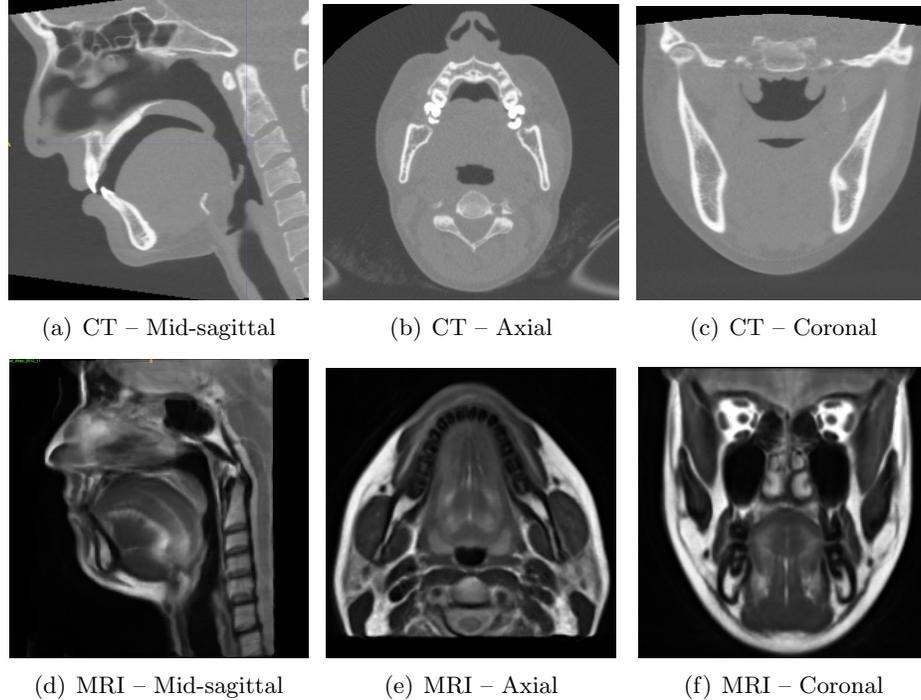


Figure 3.9: Medical images used in this study: CT (a–c) and MRI (d–f).

3.2.2 Template Model

In order to simulate complex motions coordinating multiple upper-airway organs, we adopt the Functional Reference ANatomical Knowledge (FRANK) as the template model for subject-specific modelling of upper-airway complex. As shown in Figure 2.5, FRANK couples 5 FE models (tongue, soft palate, pharynx, larynx and face), 3 rigid-body-represented bones (jaw, hyoid and maxilla), 7 rigid-body-represented laryngeal cartilages and a number of spring-like structures (muscles and ligaments) in a modularized approach. FRANK components attach with each other by several types of bilateral constraints. The maxilla is anchored in space. The jaw is connected with the maxilla by the temporomandibular joint that is modelled as three constraint planes; these planes limit the lateral motion of the jaw and constrain it to follow a pre-defined arc when opening and closing. FE components are attached to other FE components or rigid bodies using nodal attachments. Muscles are approximated as point-to-point Hill-type actuators that may at-

tach to a rigid body or pass through a FE body. The air-tight surface wraps over, and attaches to the FE models and rigid-bodies to create a parametric upper-airway mucosa layer. The details of FRANK have been introduced in Section 2.2.1

3.2.3 Modelling Framework

Subject-specific modelling *by registration* strives to find a deformation field which morphs a biomechanical template model to a target dataset, while preserving the morphology and functionality of the model. Here, the template model is hybrid, i.e. it consists of components with different geometric discretization; and the target is represented in the form of surfaces, point clouds, or landmarks extracted from medical images.

It is often hard to obtain complete correspondence information between the medical images and the template. We refer to all the rigid bodies and FE models of the template model as the Source Components (SC). Some of these components (such as the jaw and hyoid in CT registration) have explicit correspondences to the image (i.e. target positions); these components are referred to as the Reference Components (RC). List of RC and SC components is included in Table 3.2.6.

This method is a three-step sequential procedure, as shown in Figure 3.10. The first two steps are similar to the multi-structure registration procedure proposed in Section 3.1. The last step is to update the auxiliary components, such as various bilateral constraints and physical properties, and keep them relevant to the new geometry.

Step 1. Correspondences Establishment: The first step is to establish the correspondences between the template model and the subject data, we refer to them as $\mathcal{C}(RC, Seg)$. We find $\mathcal{C}(RC, Seg)$ using the correspondence establishment method introduced in Section 3.1.1.

Step 2. Anatomy Transfer: In the second step, using the deformation-transfer method introduced in Section 3.1.1, we generate a smooth, homeomorphic, shape-aware mapping, which transfers the geometry of the template organs from generic space to subject space based on the correspondence constraints $\mathcal{C}(RC, Seg)$. Since the mapping function is topology preserving, we subdivide the hybrid model into subgroups if topology changes are needed during registration; in each subgroup, the mapping function preserves the spatial relationship between the organs and maintains their regularity.

3.2. Subject-Specific Modelling

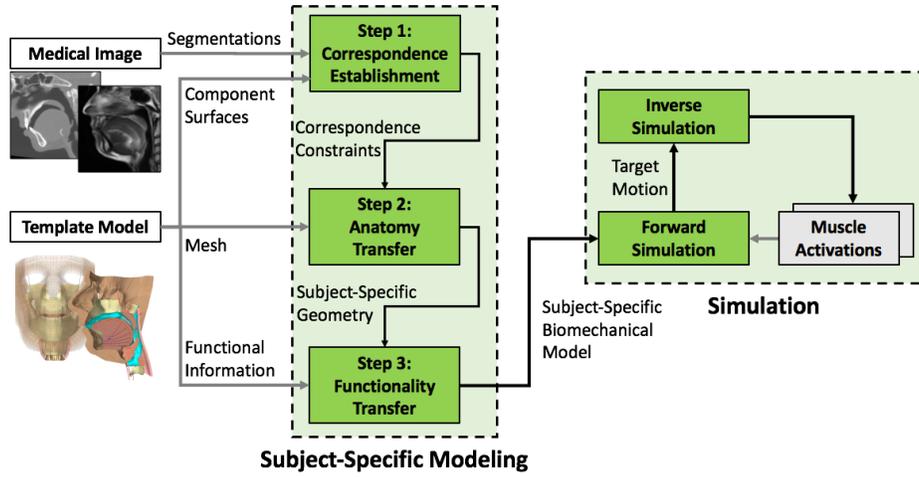


Figure 3.10: Proposed framework for subject-specific modelling and simulation of the upper-airway complex.

To maintain the functionality of the model, we add an element-quality-controlling energy into the anatomy transfer process to prevent the poor-quality elements (of the template FE meshes) from further degrading, as described in Section 3.2.4.

Our anatomy transfer method is able to maintain the spatial relationship between adjacent components of the template model. However, when there exists a large morphological difference between the template and subject, the resulting geometry can mismatch with the subject data (see Figure 3.11(a)). Disparity in morphology can influence the functional behaviours of the subject-specific model by inducing false muscle line directions or joint locations, which may affect the validity of the clinically relevant simulation. To maintain the registration accuracy, we allow landmarks (picked manually) to be added in order to guide and correct the deformation. Figure 3.11(b) illustrates the registration result where the nose and the lower part of the pharynx are corrected by adding landmarks that guide the anatomy transfer process.

Step 3. Functionality Transfer: In the final step, we use an approach to transfer the functional information of the template model, such as the morphology of muscles or the joint locations, to the registered subject-specific meshes. Our functionality transfer methods, as described in Section 3.2.5,

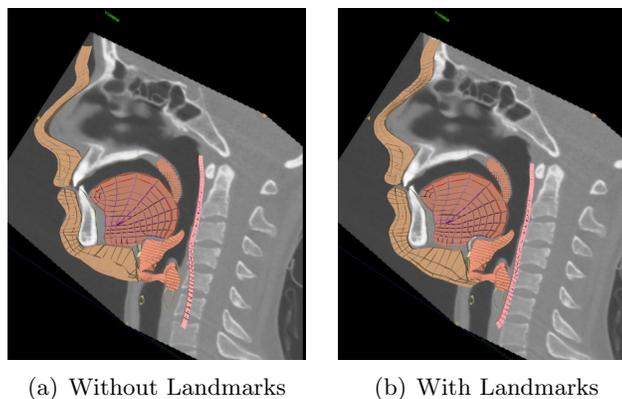


Figure 3.11: The registration result where the nose and the lower part of the pharynx are corrected by adding landmarks that guide the anatomy transfer process.

are automated, and do not require additional manual effort.

3.2.4 FE Quality Control

Our shape-aware FFD can maintain the overall mesh quality during anatomy transfer. However, when poor quality sub-domains exist in the template FE meshes, their quality can degrade even more so after registration. This situation may result in invalid elements, as shown in Figure 3.2.4. Several mesh untangling and quality improvement techniques have been previously proposed [36, 76, 77]; but they impose no constraints on reallocation of the nodes, which may cause loss of the registration accuracy. [19] proposed a Jacobian based relaxation procedure to recover element regularity and improve element quality after registration. However, this Mesh-Match-and-Repair (MMRep) algorithm only maintains a minimum mesh quality which may not be sufficient to maintain FE stability under large deformation during simulations. Moreover, the nodal test based on Jacobian metric, used by MMRep may fail to detect inverted elements [75] and thus cause the repair to fail.

3.2. Subject-Specific Modelling

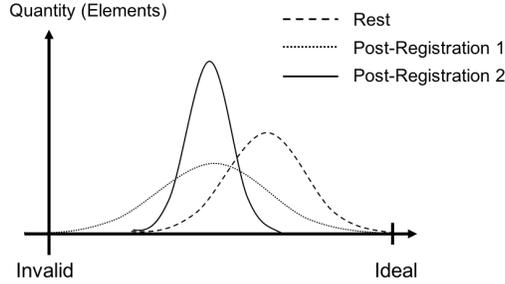


Figure 3.12: Example of element quality distribution of a FE mesh before and after registration. The dashed line represents the mesh quality distribution in rest configuration. After registration, the mesh quality level is decreased by the resulting deformation. A wide distribution produces invalid elements (round dot dash line); a narrow distribution (solid line), maintains the quality of every elements above an acceptable level.

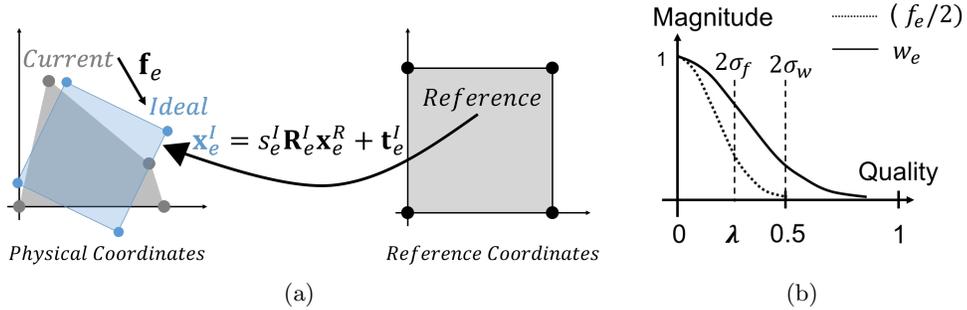


Figure 3.13: Illustration of quality-controlling mechanism. (a). for each element, we first find its ideal configuration by transforming the reference configuration to make the best fit with the current configuration; then a force \mathbf{f}_e is applied to drive the element deforming to its ideal configuration. (b). The magnitudes of the forces \mathbf{f}_e and weights w_e are Gaussian-distribution functions of the element quality. The standard deviation of the weight functions σ_w is set to 0.25. The standard deviation of the force function σ_f is set to $\lambda/2$, where λ is a value so that ten percent of elements have quality values smaller than it; we limit the maximum value of λ as 0.5.

In order to avoid local irregularity, we aim at preserving and improving

3.2. Subject-Specific Modelling

the element qualities in every step of our registration. A force-based element-quality constraint is added into the anatomy transfer process, which maps the actual shapes of the poor quality elements to their ideal configurations. As shown in Figure 3.13, a body-force \mathbf{f}_e is applied to the e -th element. We define the mesh-quality-controlling energy $E_{quality}$ as the summation of the mechanical potential energy (based on the principle of virtual work) [78] of each element:

$$E_{quality} = \sum_e w_e \int_{\Omega_e} \sigma_e : \epsilon_e - \mathbf{f}_e \cdot \mathbf{u}_e \, dV \quad (3.20)$$

where $\int_{\Omega_e} \sigma_e : \epsilon_e \, dV$ is the mechanical strain energy of a single element; \mathbf{u}_e is the displacement field, and \mathbf{f}_e is the body-force per unit volume, which is defined as:

$$\begin{aligned} \mathbf{f}_e &= \frac{f_e}{s_e^I V_R} (\mathbf{x}_e^I - \mathbf{x}_e) \\ \mathbf{x}_e^I &= s_e^I \mathbf{R}_e^I \mathbf{x}_e^R + \mathbf{t}_e^I, \end{aligned} \quad (3.21)$$

where V_R is the volume of the reference element; \mathbf{x}_e^R is a point in the reference element frame, and s_e^I , \mathbf{R}_e^I , \mathbf{t}_e^I are scaling factor, rotation, translation respectively, which transform the reference element to its target configuration (i.e. ideal element). The target configuration of the e -th element is determined by minimizing $\int_{\Omega_e} \|\mathbf{f}_e\|^2 \, dV$. The magnitudes of the weights w_e and forces f_e are defined as Gaussian-distribution functions (Figure 3.13(b)) of the element-quality values; the poor quality elements are assigned large weights and strength to prevent them from further degrading, thus leading to a uniform quality distribution across the FE mesh (solid line in Figure 3.2.4). We use the mean-ratio value [70, 88] as the mesh-quality metric.

3.2.5 Functionality Transfer

As aforementioned in this chapter, in order to preserve the functionality of the registered model, the functional information, such as the position of the muscle attachments, the locations of joints and the physical properties, should stay relevant to the new model geometry.

Muscle attachments: In biomechanical models, muscles and ligaments are either embedded into FE components (soft-tissues) or attached to the rigid bodies (bones, cartilages). In the first case, after the FE mesh is morphed into the subject space, we transform the embedded muscles using the FE interpolation functions (i.e. shape functions). In the second case,

we apply a surface-oriented FFD [132] to transform the muscle attachment points. The transformation preserves the barycentric coordinates in the surrounding polymesh faces; thus, the locations of the muscle attachments are determined by the local geometry of their associated rigid bodies.

Component attachments: The coupling attachments – which are used to connect adjacent components – need to be reconstructed to maintain the functionality of the subject-specific model. Based on the types of parametrization of organs within the hybrid model, we summarize the coupling attachments into three categories, as follow:

- **Rigid body – Rigid body:** The pose of the joints between the rigid-body pairs are transformed (rotated and translated) to fit with the deformed geometry; rotations and translations are defined based on the rigid transformations of the vertices that surround the joints.
- **FE model – FE model:** The proposed anatomy-transfer method preserves the embedding topology (i.e. the boundary connections between the FE models are maintained); therefore the FE-FE attachments are transferred to the deformed geometry without any additional re-designing efforts.
- **FE model – Rigid body:** The locations of the node-attachments are updated to the deformed positions of the corresponding nodes directly.

Physical properties: We translate the center of mass of the rigid-bodies using the centroid displacements of their associated meshes. We assume that the densities of organs are consistent between different subjects; therefore, the mass of each component is re-calculated according to their volumes. Based on the assumption that the template model and the subject data are in the same posture, we proportionally scale the morphological parameters (i.e. rest length and the optimal fascicle length) of the muscles and the ligaments according to their length change.

3.2.6 Experiments

We evaluate the proposed functionality-preserving non-rigid registration method on two 3D medical Images (CT and MRI) of the human head and neck. The template model of upper-airway complex (FRANK) has been introduced in Section 2.2.1. We extracted the segmentation for RC from the CT and MRI images. The segmentation information is summarized in Table 3.2.6. The segmentation results are shown in Figure 3.14.

3.2. Subject-Specific Modelling

Organs (RC)	Dataset	Method	Geometry Result
Bones	CT	Auto	Complete
Tongue	CT	Manual	Complete
Soft-palate	CT	Manual	Partial
Epiglottis	CT	Manual	Partial
Airway	CT	Auto	-
Bones	MRI	Manual	Partial
Tongue	MRI	Manual	Complete
Airway	MRI	Auto	-

Table 3.1: Summary of segmentation. Bones include jaw, maxilla and hyoid.

In section 3.2.6, we compare the performance of our method with a state-of-art FE mesh registration technique, Mesh-Match-and-Repair[18], in two registration scenarios with complete and partial correspondences. In Section 3.2.6, we present our subject-specific models of upper-airway complex.

Registration of the FE models

Mesh-quality of FE models is essential for biomechanical simulation; bad mesh-quality will cause inaccuracy and instability [21]. We validate the proposed FE mesh registration method (SPFFD) by comparing with MMRep algorithm in two different cases:

- Complete organ geometry can be retrieved from the medical data: the template tongue mesh is registered to the tongue surfaces segmented from the CT and MRI dataset.
- Partial or adjacent organ geometry is available: the airway surface in the template model was used as the reference component and is morphed onto the airway segmentation (from the CT and MRI dataset). As the source component, the template pharynx FE mesh is deformed along with the template airway surface.

We used the same correspondence-search algorithm (single-direction ICP) in the evaluation process. We rigidly aligned the FE model and the target surface as the initial setting for both our proposed algorithm and MMRep. Then, the two registration methods were applied to generate subject-specific FE meshes. We set the stop criteria as 0.45mm mean surface distance, which is smaller than the voxel-size in both images.

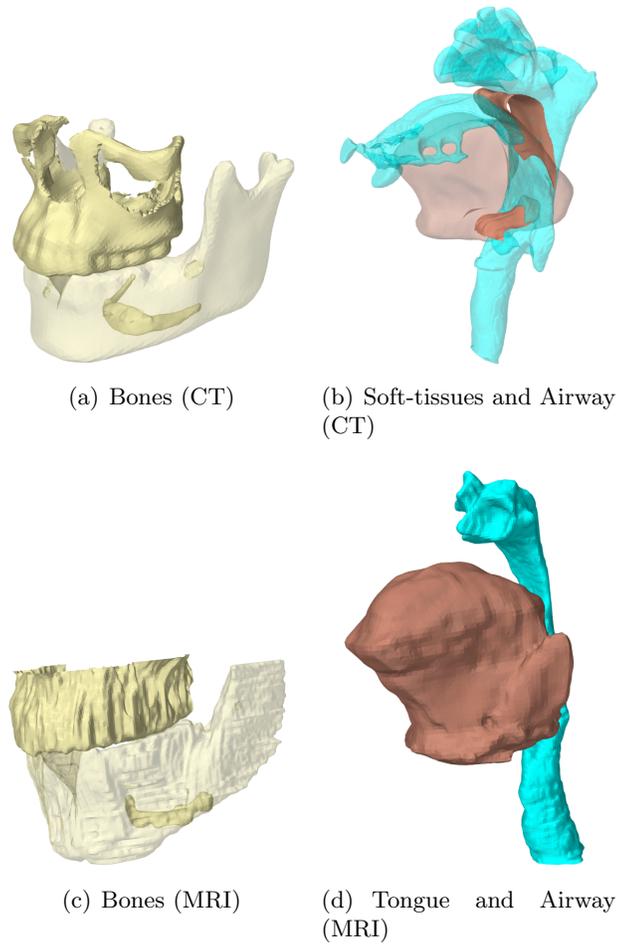


Figure 3.14: Segmentation of the CT and MRI images. Some segmentation surfaces are set to be transparent to illustrate the covered components. (a). Segmentation of bones from the CT image, including complete geometry of jaw, maxilla and hyoid. (b). Segmentation of soft-tissues and airway boundary from the CT image, including complete geometry of tongue and soft-palate, partial geometry of epiglottis. (c). Segmentation of bones from the MRI image, including partial geometry of jaw, maxilla and hyoid. (d). Segmentation of complete geometry of tongue and airway from the MRI image.

The registration results are shown in Figures 3.15 and 3.16. In Figure 3.2.6, we illustrate the element-quality histograms of the deformed FE meshes. Our method produced more uniform element qualities (narrower distribution) over the resulting mesh. Table 3.2 and 3.3 summarizes the performance of the two registration procedures for tongue and pharynx respectively. When the two methods achieved similar registration accuracy (mean value of surface distance), our method has smaller standard deviation. Notably, for both image datasets, the average and minimum mean-ratio values of the resulting models from our method are considerably larger than the MMRep results (Figure 3.17). Compared to the template tongue and pharynx model, the average and the minimum mean-ratio values (η_{mean}) have increased with our method. Although the MMRep method holds the Jacobian-Ratio values (JR) of the resulting meshes above a threshold (0.03), it fails to repair the elements with small mean-ratio value ($\eta < 0.1$). We compare the resulting pharynx geometry with the template model (as shown in Figure 3.16(a) – 3.16(f)); when the available correspondence is sparse, our results have smaller morphological change from the template than the MMRep results.

3.2. Subject-Specific Modelling

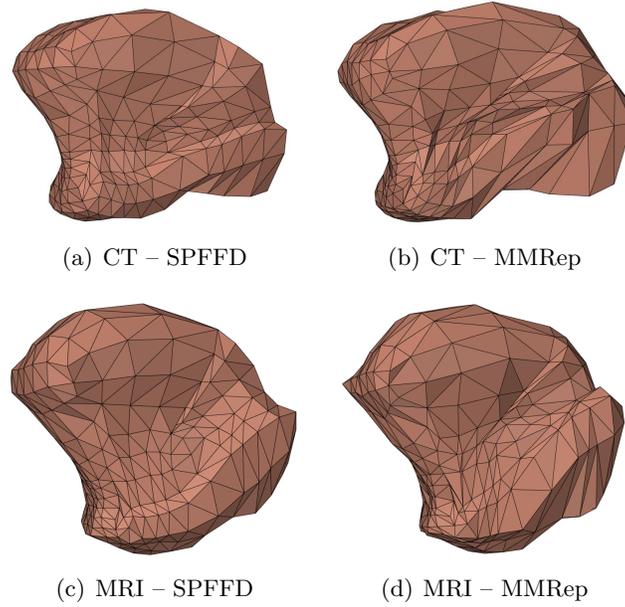


Figure 3.15: Registration results of tongue model.

FE Mesh	$D_{mean} \pm \sigma(mm)$	$\eta_{mean}(\%)$	$\eta_{min}(\%)$	$JR_{min}(\%)$
Template	-	48.72	0.00	0.54
CT - SPFFD	0.33 ± 0.35	58.60	26.96	6.25
CT - MMRep	0.44 ± 0.67	44.85	0.00	3.32
MRI - SPFFD	0.36 ± 0.36	57.93	19.58	2.73
MRI - MMRep	0.39 ± 0.65	46.82	0.00	3.35

Table 3.2: Summary of tongue FE mesh. D_{mean} and σ are the mean surface distance and the standard deviation. η_{mean} and η_{min} are the mean and minimum mean-ratio value [70]. JR_{min} is the minimum Jacobian ratio value.

3.2. Subject-Specific Modelling

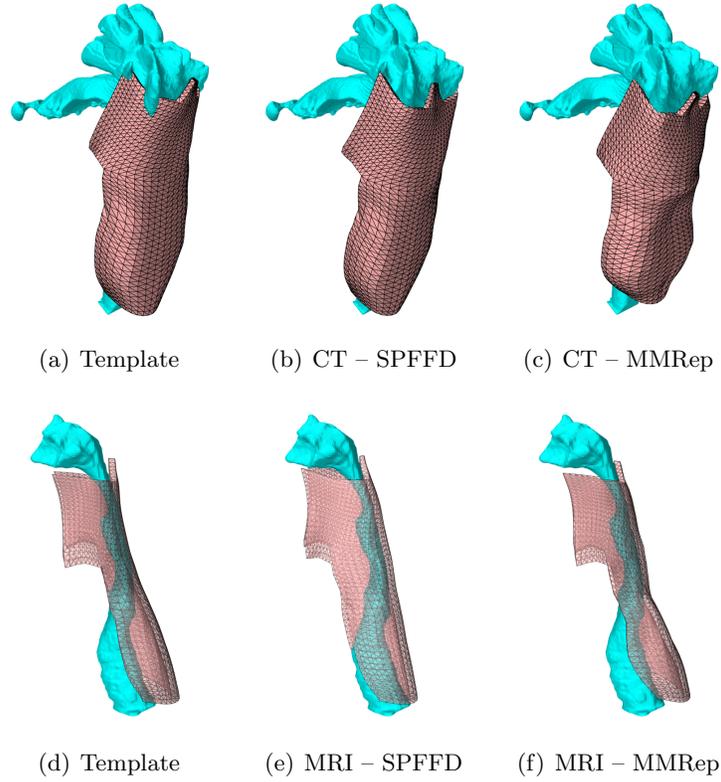
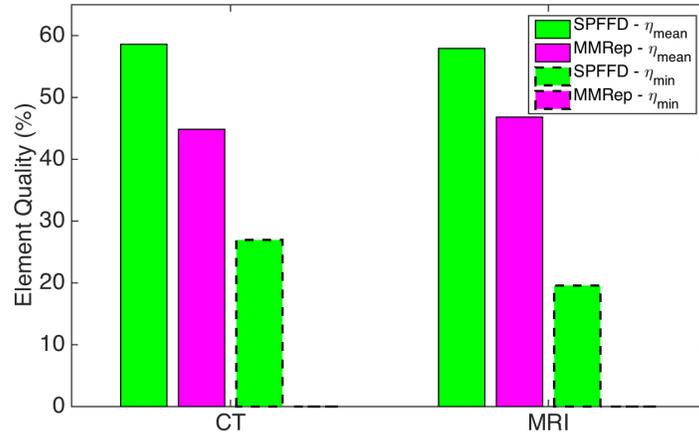


Figure 3.16: Registration results of pharynx model. The template and the registered pharynx aligned with the subject airway meshes are shown. When available correspondence is sparse, morphological deviation (from the template geometry) of our results is smaller than deviation of MMRep results.

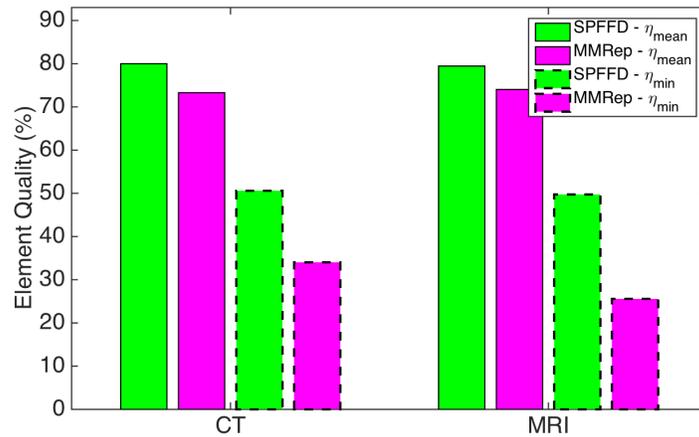
FE Mesh	$\eta_{mean}(\%)$	$\eta_{min}(\%)$	$JR_{min}(\%)$
Template	79.05	35.71	0.34
CT – SPFFD	80.00	50.59	38.62
CT – MMRep	73.29	34.03	30.11
MRI – SPFFD	79.45	49.74	33.80
MRI – MMRep	74.03	25.57	16.14

Table 3.3: Summary of pharynx FE mesh. η_{mean} and η_{min} are the mean and minimum mean-ratio value [70]. JR_{min} is the minimum Jacobian ratio value.

3.2. Subject-Specific Modelling



(a) Tongue Models



(b) Pharynx Models

Figure 3.17: Comparison of the mesh quality by our methods and MMRep. The bars with solid outlines and dash lines illustrate the mean mean-ratio values (η_{mean}) and minimum mean-ratio values (η_{min}) of the resulting FE models respectively. Figure (a) compares the element qualities of the tongue models by two methods; the mean-ratio values (η_{mean} and η_{min}) have been listed in Table 3.2. Note that the minimum mean-ratio values by our method are 0.0. Figure (b) compares the element qualities of the pharynx models by two methods; the mean-ratio values (η_{mean} and η_{min}) have been listed in Table 3.3.

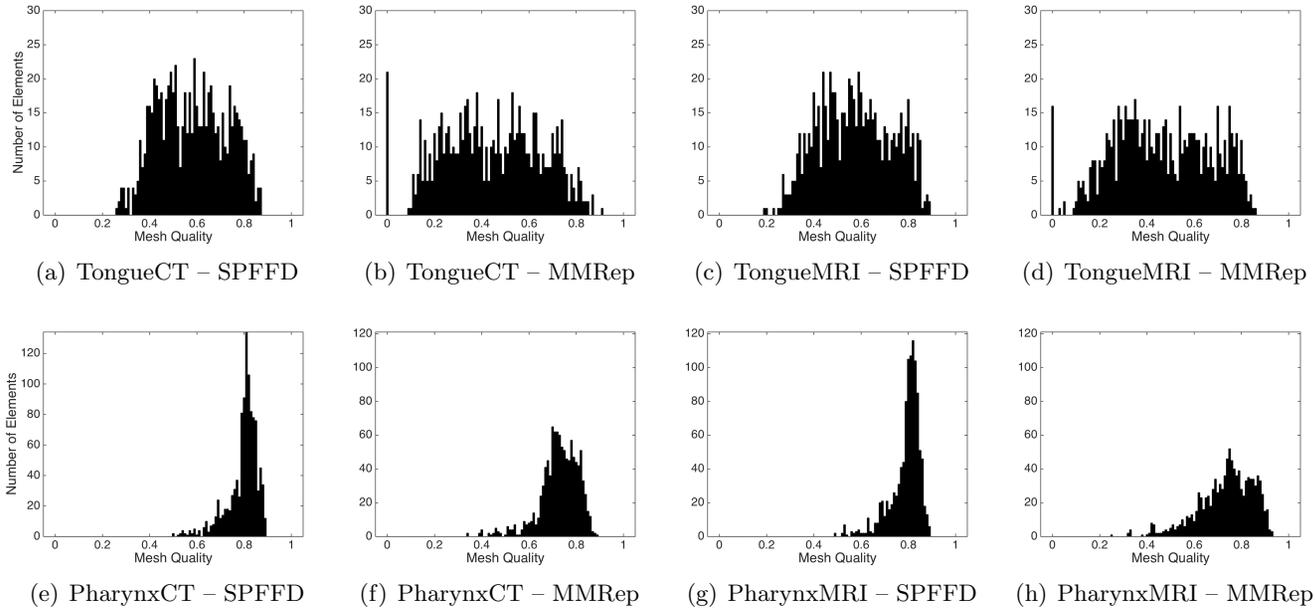


Figure 3.18: Histograms of the element qualities (mean ratio) over tongue and pharynx FE meshes. TongueCT and TongueMRI are the results from CT and MRI respectively; PharynxCT and PharynxMRI are the results from CT and MRI respectively.

3.2. Subject-Specific Modelling

Groups	RC	SC
CT1	tongue, soft-palate, airway	tongue, soft-palate, pharynx
CT2	airway, epiglottis, pharynx	larynx, laryngeal cartilages
CT3	jaw, maxilla, hyoid, tongue, thyroid	jaw, maxilla, hyoid, face
MRI1	tongue, airway	tongue, soft-palate, pharynx
MRI2	tongue, airway, pharynx	larynx, laryngeal cartilage
MRI3	jaw, maxilla, hyoid, tongue, thyroid	jaw, maxilla, hyoid, face

Table 3.4: Subdivision of FRANK.

Registration of FRANK

For each dataset, we divided the FRANK model into three sub-groups and performed the registration to the segmented surfaces in each subgroup (from Figure 3.14) separately. The RC and SC of each group are illustrated in Table 3.2.6. Because the face and larynx models do not connect to other FE models, this sub-division scheme does not affect the functionality transfer process described in Section 3.2.5.

The registration results, Model 1 (from CT dataset) and Model 2 (from MRI dataset) are illustrated in Figure 3.20. Registration errors are shown in Table 3.2.6. For each RC, our method achieved submillimeter registration accuracy (mean surface distance). Larger registration errors were localized around absent or ill-defined features, such as the temporal process in the CT image (which is absent in the template model), shown in Figure 3.14(a)), the vocal fold which is invisible in the MRI image and the lateral boundary between the tongue and pharynx. In these areas, the closest point correspondences were inconsistent with the structure-preserving constraints (defined by Equation 3.19); as a result, the local rigidity prevented the points in these areas from converging to bad positions, thus leaving the larger registration errors.

The histograms of the mesh quality of the template and the resulting FE models are shown in Figure 3.21. The mesh-quality information is also summarized in Table 3.2.6. Our method showed similar performance on the two datasets. After registration, the average element qualities of the resulting FE models were similar to the values of the template(Figure 3.22(a)). The numbers of elements whose mean-ratio values were smaller than 0.1 were reduced(Figure 3.22(b)). For tongue and pharynx, the minimum element qualities were noticeably increased. However, our methods are unable to improve the poor-quality elements located at the anatomical sub-regions that have high-curvature or fine geometric details. For example, Figure 3.19 illustrates

3.2. Subject-Specific Modelling

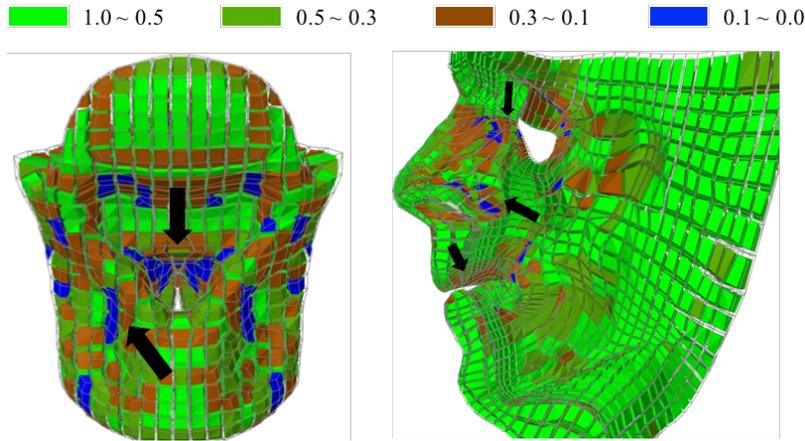


Figure 3.19: Illustration of poor-quality elements in larynx and face. The top row shows the mean-ratio values and their corresponding colors. Arrows point out the poor-quality sub-regions.

poor-quality elements in template larynx model. Since our method maintained and improved the mesh qualities compared the template model, the resulting subject-specific models demonstrated comparable stability during biomechanical simulation.

Since the CT and MRI dataset were captured in different postures (as described in Section 3.2.1), the rest configurations of the tongue are noticeably different in Model 1 and Model 2. During the scanning process, the subject in the CT data was lying backwards at an angle of 45° . In contrast, the MRI dataset was captured in the supine position; hence the tongue deformed towards the hard and soft palate due to the effect of gravity. As our functionality-transfer methods assume the subjects in the medical data are in stress-free configurations, the external-loading influence on the biological structures are neglected.

For the purpose of examining the functionality of the two subject-specific models, we performed forward and inverse simulations on three simple speech postures and compared the results with similar simulations on the template model.

3.2. Subject-Specific Modelling

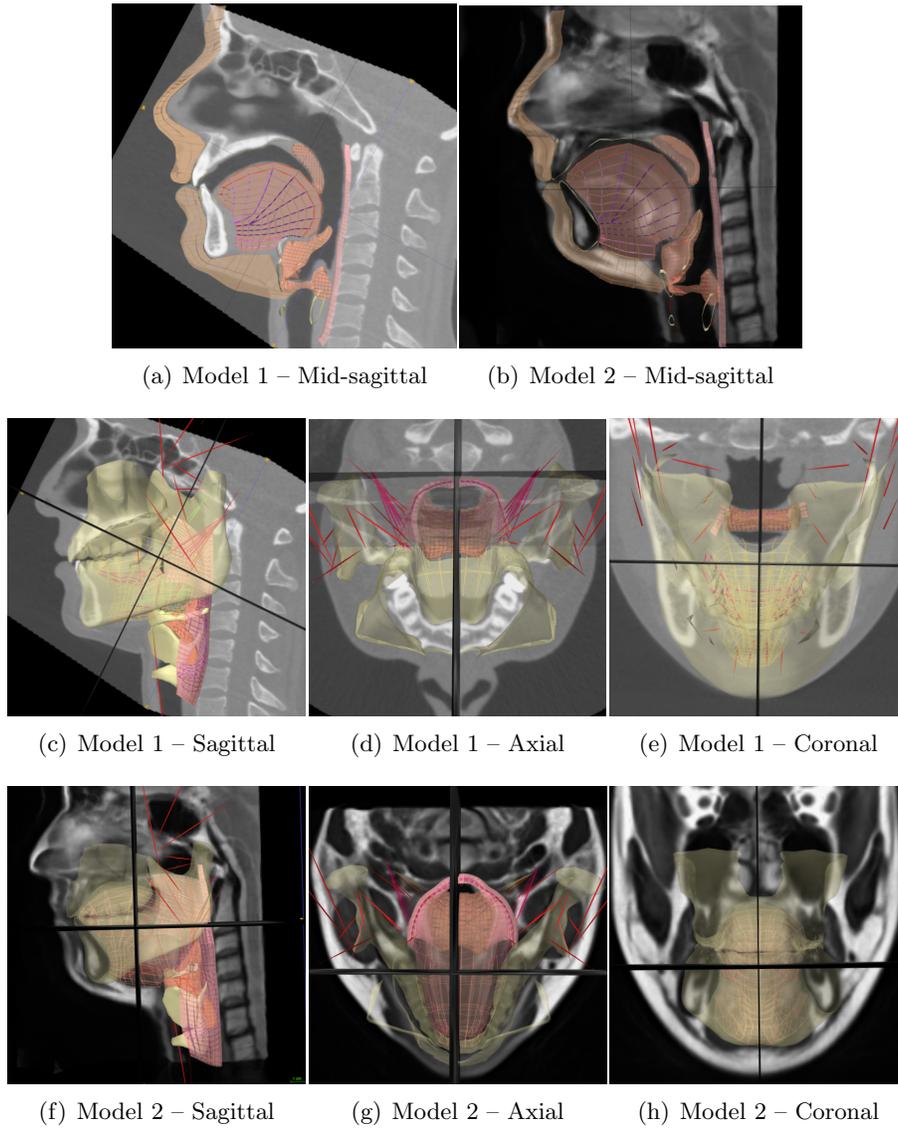


Figure 3.20: Subject-specific models of upper-airways complex

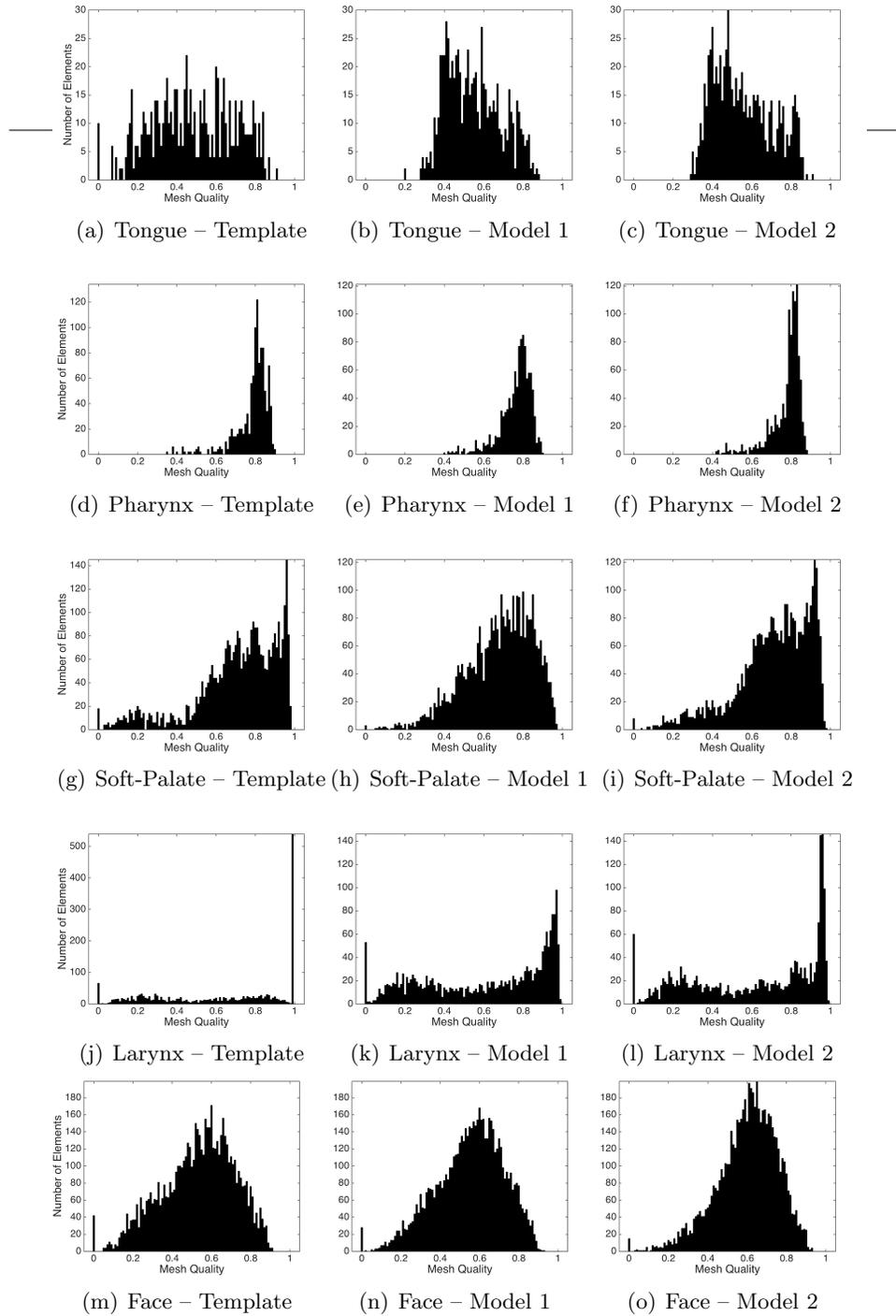


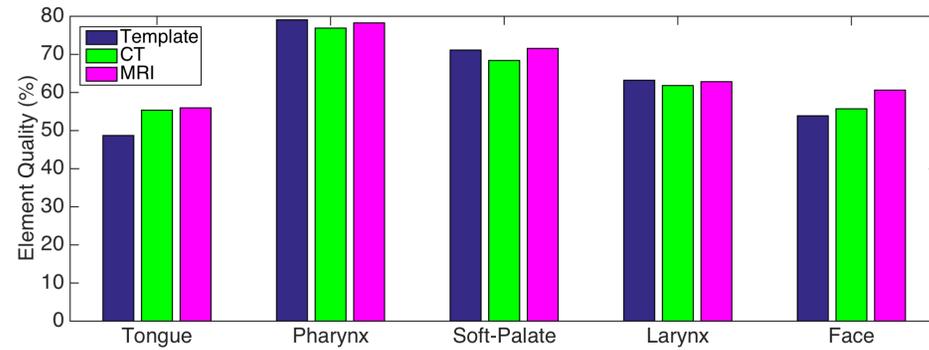
Figure 3.21: Histograms of the element qualities (mean ratio) over FE meshes within the three models of upper-airway complex. Figures (a)(d)(g)(j)(m) show the mesh qualities of FRANK model. Model 1 and Model 2 are the registration results of the CT and MRI dataset respectively.

Images	Metric	Jaw	Maxilla	Hyoid	Tongue	Soft-palate	Epiglottis	Airway
CT	$D_{mean}(mm)$	0.30	0.41	0.16	0.74	0.55	0.67	0.86
	$D_{max}(mm)$	2.14	6.29	0.76	4.67	3.59	2.74	3.81
	$\sigma(mm)$	0.27	0.47	0.13	0.68	0.56	0.49	0.79
MRI	$D_{mean}(mm)$	0.39	0.44	0.29	0.62	-	-	0.93
	$D_{max}(mm)$	2.33	2.38	1.40	3.65	-	-	5.02
	$\sigma(mm)$	0.34	0.35	0.25	0.62	-	-	0.89

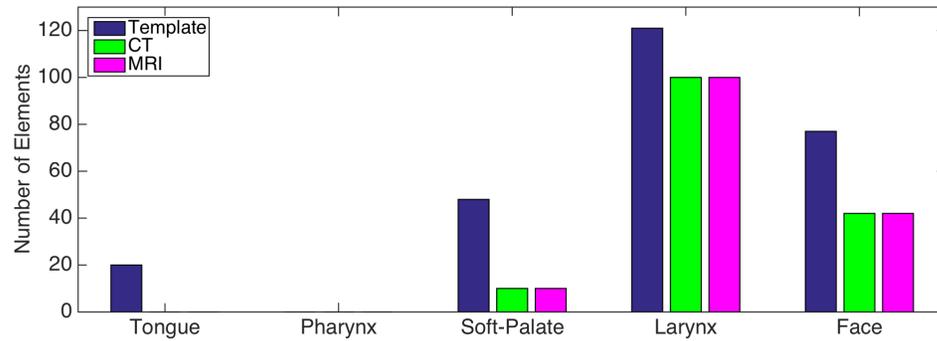
Table 3.5: Registration errors within the two subject-specific models of upper-airway complex. D_{mean} , D_{max} and σ are the mean surface distance, the maximum surface distances and the standard deviation.

	Metric	Tongue	Pharynx	Soft-palate	Larynx	Face
Template	$\eta_{mean}(\%)$	48.72	79.05	71.14	63.22	53.88
	$\eta_{min}(\%)$	0	35.71	0	0	0
	$\#(\eta < 0.1)$	20	0	48	121	77
Model 1 (from CT)	$\eta_{mean}(\%)$	55.35	76.90	68.40	61.83	55.72
	$\eta_{min}(\%)$	20.10	40.07	0	0	0
	$\#(\eta < 0.1)$	0	0	10	100	42
Model 2 (from MRI)	$\eta_{mean}(\%)$	55.96	78.26	71.56	62.84	60.63
	$\eta_{min}(\%)$	29.29	42.30	0	0	0
	$\#(\eta < 0.1)$	0	0	10	100	42

Table 3.6: Summary of FE mesh qualities. η_{mean} and η_{max} are the mean and maximum mean-ratio value [70]. $\#(\eta < 0.1)$ is the number of elements whose mean-ratios value are smaller than 0.1.



(a)



(b)

Figure 3.22: Comparison of the element quality of the resulting models and the template models. Figure (a) illustrates the average element qualities (η_{mean}) of the each FE components, which have been listed in Table 3.2.6. Figure (b) illustrates the number of elements whose mean-ratio values are smaller than 0.1 ($\#(\eta < 0.1)$), which has been listed in Table 3.2.6.

Forward Simulation To demonstrate the functionality of the models, we performed forward simulation using 3 sets of muscle activations for speech production – corresponding to the vowels /**a**/, /**i**/ and /**u**/ – which were defined in [7]. Below we list the muscles involved followed by the percent activation in parenthesis:

- /**a**/ Formed using tongue and jaw muscles. Tongue: genioglossus middle (0.08), anterior (0.12); hyoglossus (0.15); verticalis (0.05). Jaw opening: anterior and posterior digastric, sternohyoid (all at 0.1).
- /**i**/ Formed with tongue, face, and jaw muscles. Tongue: genioglossus posterior(0.5), middle (0.02), anterior (0.02); superior longitudinal (0.05); inferior longitudinal (0.05); mylohyoid(0.10). Face: risorius (0.05) and zygomaticus (0.05). Jaw closing: temporalis (anterior, middle, posterior), masseter, median pterygoid (all at 0.01).
- /**u**/ Formed using tongue and face muscles. Tongue: styloglossus (0.1). Face: orbicularis oris middle ring (0.35).

We fed these activations into the three models, i.e. the template model (FRANK), Model 1 (from CT) and Model 2 (from MRI). Each model was allowed to settle under gravity for 0.1s, before the muscle activations were linearly increased to their maximum level at 0.3s. The models settled again without further activation changes until 0.4s. The end postures are illustrated in Figure 3.23. Three models are able to maintain their stability and to produce reasonable motions in response to the muscle activations in the forward-dynamics simulations.

Inverse Simulation To demonstrate the inverse simulation capabilities of the resulting models, we recreated the posture /**a**/ by supplying the forward-simulation motion into the inverse solver. The inverse simulation sought to calculate activations used to attain the posture. We picked 35 tracking points on the tongue back surface and 11 tracking points on the jaw. The tracking weight (w_m), regularization weight (w_a) and damping weight (w_d) were set to 1.0, 0.001 and 0.0001 respectively.

The resulting postures (from the inverse simulations) were similar to what is shown in Figure 3.23. The tracking errors – average distances between the tracking points and their corresponding target positions – of the template, CT and MRI models were $0.24 \pm 0.09mm$, $0.21 \pm 0.09mm$ and $0.05 \pm 0.03mm$ respectively. The predefined forward simulation activations are compared with the inverse-predicted activations in Figure 3.24. The

3.2. Subject-Specific Modelling

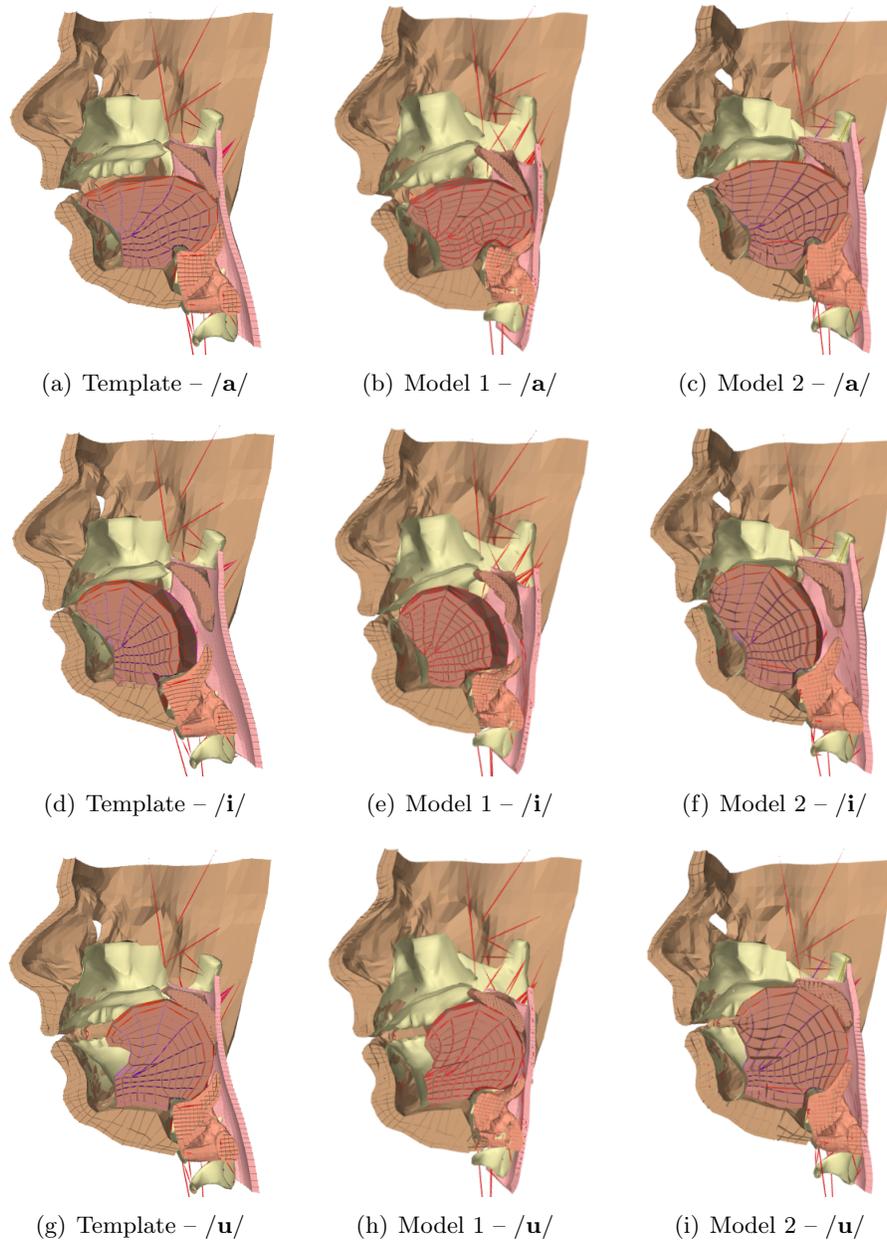


Figure 3.23: View of end postures of three models, Template (FRANK), Model 1 (from CT dataset) and Model 2 (from MRI dataset), in vowels (/a/, /i/ and /u/) productions.

muscles – whose activation levels were greater than 1% – were considered as active muscles; other muscles are not shown in Figure 3.24. The three models yielded similar predictions of muscle activations for vowel /**a**/ production. The active muscles picked by the inverse solver were the same as the predefined combinations.

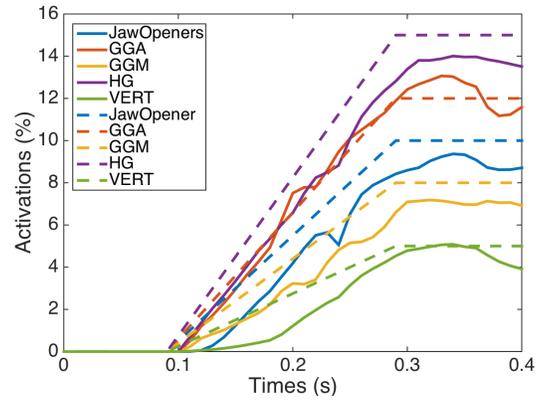
3.3 Discussion and Conclusion

For the first time, motor-driven subject-specific models of upper-airway complex were created. Based on correspondences established between a template model and subject data, we transfer a functional template from the generic space into a specific subject space. In order to preserve the functionality of the template, we seek to maintain three types of regularity:

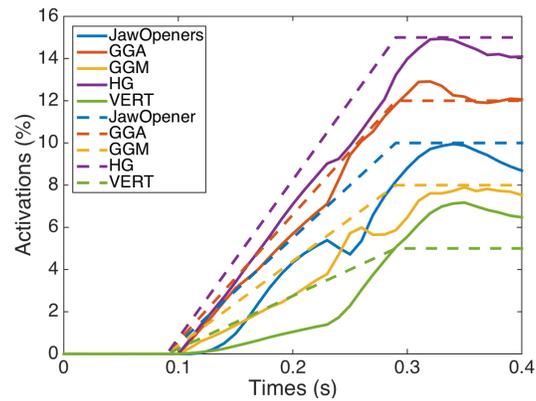
1. **Inter-Component Regularity:** Maintain the spatial relationship (including connectivity, topology, relative posture and size) between model components. We apply volumetric B-Splines to interpolate between sparse correspondences so that the model can deform along with the sub-anatomical-structures whose geometries are available.
2. **Intra-Component Regularity:** Preserve the underlying discretization structures of the template model during registration. For each component, we enforce their local transformations to be similarity-invariant.
3. **Functional Regularity:** Keep the functional information (including coupling attachments between components, muscle attachments and biomechanical properties) similar to the template but relevant to the new model geometry.

In order to maintain the inter- and intra-component regularity, we first propose a multi-structure registration technique (in Section 3.1), which tends to superimpose a collection of template meshes (with different types of geometric discretizations) onto certain target dataset (e.g. surfaces or point cloud), while minimizing their morphological deviation. We performed an experiment on a synthetic data set and compare the registration performance with a traditional diffeomorphic registration method [126]. Both methods are capable of maintaining the spatial relationship between components. However, our method demonstrates a better performance in terms of preserving the mesh quality and shape regularity of each individual component.

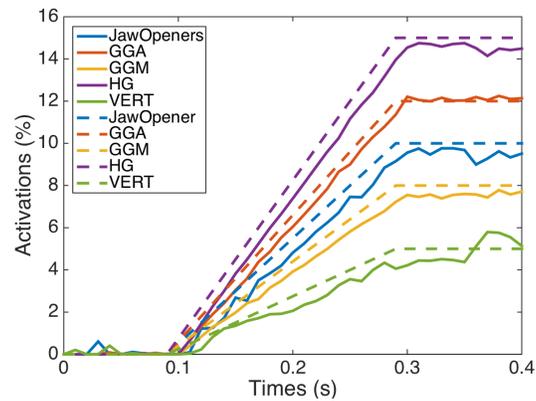
3.3. Discussion and Conclusion



(a) Template (FRANK)



(b) Model 1 (from CT dataset)



(c) Model 2 (from MRI dataset)

Figure 3.24: A comparison of predefined muscle activations (dashed lines) and the inverse-predicted activations (solid lines).

3.3. Discussion and Conclusion

Based on the multi-structure registration technique, we then propose a registration framework for subject-specific modelling of upper-airway complex 3.2. We compared the performance of our methods with a state-of-art FE-model-registration method (MMRep) [18] (in Section 3.2.6). Our experiments indicate that our methods outperform MMRep with respect to preserving the FE mesh quality and addressing sparse correspondence on these particular datasets.

We created two subject-specific models of upper-airway complex based on the CT and MRI dataset (in Section 3.2.6). For every source components (components that have correspondence), our method achieved sub-millimeter surface-distance-represented optimization accuracy. To demonstrate the functionality of the resulting models, we performed activation-driven simulations to model speech production on three vowels /**a**/, /**i**/ and /**u**/. Our models can maintain their stability during the simulation and properly respond to the muscle activations and external loadings (such as gravity). The kinematics obtained from the forward simulations of /**a**/ were used to drive the inverse simulations. Compared with the template model, our models showed comparable tracking performance and yielded similar muscle-activation predictions.

Our subject-specific registration methods can benefit from several improvements. **Firstly**, our free-form-deformation starts from a uniformly spaced grid at every resolution level. When the registration errors are unevenly distributed in space, this approach can introduce unnecessary computational cost during registration. This problem can be avoided by replacing FFD with a shape-customized embedded deformation. **Secondly**, we achieve mesh-quality control during registration, which avoids the loss of registration accuracy that happens in the mesh repair step after registration. However, for partial registration, the force-based energy may lead to unrealistic deformation at positions where geometry is not fully determined by correspondence constraints. Future study should explore better mesh quality preserving mechanisms. **Thirdly**, due to the complexity of the organ geometry, our template FE model contains poor-quality elements. Redesigning efforts, to improve the initial mesh quality, will as well enhance the numerical stability of the registered models. **Fourth**, our methods can benefit from replacing the deterministic template with a statistic atlas, which not only can represent the average information of human anatomy but also can describe the inter-subject variations. Replacing the deterministic template with a statistic atlas can help with reducing the morphological uncertainty and increasing our confidence on the registration results when the subject organ geometry is partially unknown from the medical data. Wang et al. [152]

propose a statistical atlas based subject-specific FE modelling framework. Future work should explore a more generalized framework to accommodate hybrid and modularized biomechanical models, such as FRANK. **Fifth**, we expect to move our registration method into the image space in order to reduce or eliminate segmentation efforts.

Currently, we assume that our target organ geometries extracted from the medical images with similar posture to our template model are in a stress-free configuration. Such assumption does not hold in most *in vivo* conditions, due to the external forces (such as gravity) and tetanic contractions. A good approximation of the unloading configuration can help with minimizing the bias of the modelled biomechanics. Vavourakis et al. [151] proposed an inverse analysis method to derive the pressure-free configuration of human aortas, and the gravity-free shape of the female breast. However, their studies only focus on isolated FE models. Further study is needed to determine the unloading reference configuration of hybrid models (such as FRANK).

Finally, our methods do not personalize the elastic properties of soft-tissues. Tissue properties may influence the muscular coordinating patterns of a specific physiological task and are needed for reliable subject-specific modelling. Measuring the tissue properties *in vivo* is still an ill-posed problem. Elastography provides a non-invasive way for such measurement. For example, Cheng et al. [26] applied magnetic resonance elastography to measure the viscoelastic properties of tongue and soft palate. However, as soft tissues exhibit non-linear stress-strain behaviour, tissue properties at larger deformations cannot be inferred from elastography alone [155]. Future work should further the investigation of *in vivo* tissue-property measurement.

Our study shows the potential of the proposed registration techniques for subject-specific modelling of coupled upper-airway structures. However, careful validation is required for applying subject-specific models to clinical usage. From the anatomy view, the consistency between the resulting model geometry and subject anatomy is needed to be quantified. From the functionality view, the fidelity of the biomechanics and the motor-control behaviours of the subject-specific models is also needed to be measured.

Chapter 4

Data-Driven Swallowing Simulation: Towards Postoperative Prediction of Swallowing Biomechanics

Our primary motivation for developing the subject-specific modelling tools is in their application to biomedicine. Computer-assisted diagnosis and treatment of coupled upper-airway functions and dysfunctions requires comprehensive subject-specific models that coordinate multiple motor systems. In the previous chapter we have described a subject-specific modelling method for generating holistic models of upper-airway complex for specific individuals. In this chapter, we apply one of our developed models to simulating the normal swallowing motion, and assess the feasibility of postoperative prediction of swallowing biomechanics.

Swallowing is a rapid and complex sequence of movements which requires precise coordination of more than 30 muscles located within the oral cavity, pharynx and larynx [130]. In normal conditions, swallowing is divided into four primary phases:

- **Oral Preparatory Phase:** food is broken down (via mastication) and a cohesive bolus is formed [94].
- **Oral Transport Phase:** food (bolus) is propelled posteriorly through the oral cavity and into the oropharynx [94].
- **Pharyngeal Phase:** food (bolus) transits through the oropharynx and Upper Esophageal Sphincter (UES). Epiglottis is automatically closed to prevent food and liquid to penetrate into the upper airways and lungs [130].
- **Oesophageal Phase:** food (bolus) is pushed from the oesophagus to the stomach.

Swallowing disorder is manifested in the older patients in the case of stroke, neurodegenerative disease, and dementia [124]. It can also be caused by treatment of oropharyngeal cancer, such as laryngectomy. Swallowing disorders are associated with a higher incidence of patient suffering from aspiration [123], due to an impairment of the pharyngeal phase: because of a deficiency in the brain or nervous system the epiglottis is not able to prevent food and liquids from penetrating into the pulmonary tract; the accumulation of organic matter could lead to an inflammation of the lungs and airways [123]. Patients who will not respond or are not eligible for swallowing rehabilitation, the standard treatment is a tracheostomy. In addition, some patients may also have to be fed via gastrostomy. An alternative option to the invasive surgical methods for aspiration prevention is intralaryngeal prosthesis implantation. For example, a case of a patient with swallowing disorder (caused by partial laryngectomy) receiving an intralaryngeal prosthesis is reported in [123]. The deglutition videofluoroscopy before and after intralaryngeal prosthesis implantation is shown in Figure 4.1. In order to determine the size of a such prosthesis in advance, virtual implantation is created on a 3D reconstruction from a CT scan of the patient [123], as shown in Figure 2.10. However, no prediction and assessment of the postoperative bolus-transit biomechanics has been accomplished before surgery.

In recent years, computational models, as an alternative investigative tool, have been applied to study human swallowing process. Mizunuma et al. [106] and Sonomura et al. [134] apply FE models of human upper airway (as shown in Figure 4.2(a)) to investigate the swallowing of jelly and liquid bolus. Based on manually defined movements of the tongue and pharyngeal wall, they simulate the motions of the rheological models (jelly and fluid bolus) happening during the swallowing. Following Mizunuma et al. and Sonomura et al., Kikuchi et al. [72] use a HMPS model (described in Section 2.2) to reproduce the swallowing movements recorded in by videofluoroscopy (VF). However, the upper-airway models used in these studies do not include bony components (such as hyoid) and muscle driven structures. Although their models can assist with the visualization of upper-airway movements, they are unable to reveal the motor control mechanisms of swallowing. In order to investigate the influence of muscle aging on hyoid motion during swallowing, Tsou et al. [146] build a muscle-driven laryngopharynx model based on cadaver data, as shown in Figure 4.2(b). By applying the inverse-dynamics simulator to tracking the motion data extracted from VFs, Tsou's model successfully reproduces the hyoid motion of swallowing within a certain range. More recently, Ho et al. [63] use the oropharyngeal part of FRANK (Section 2.2.1) to simulate swallowing move-

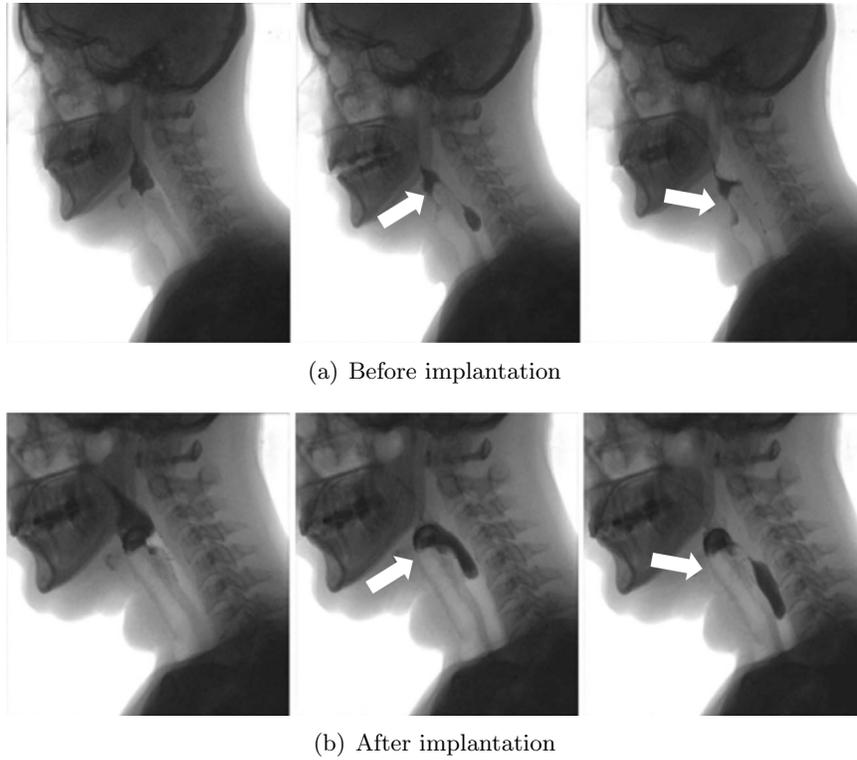


Figure 4.1: Three successive sequences from videofluoroscopy of deglutition before intralaryngeal prosthesis implantation (a) and after implantation (b). © Wiley (2016), Raguin et al. [123], adapted with permission. Arrows illustrate the food bolus penetrating into the trachea of the patient before and after implantation.

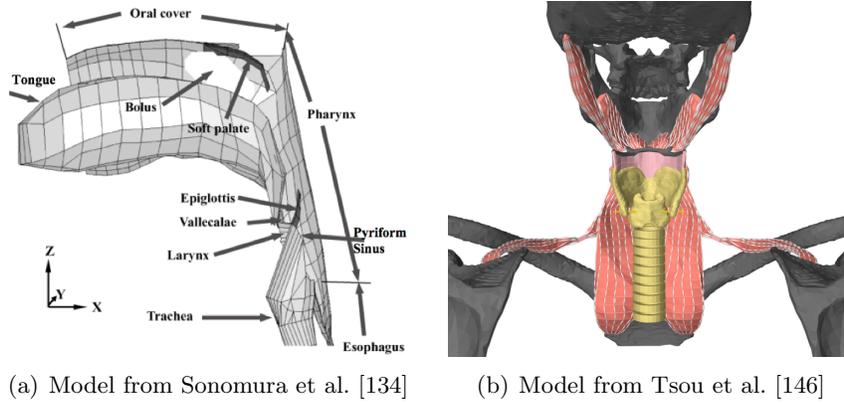


Figure 4.2: Illustration of two biomechanical models for swallowing simulation reported in previous studies. (a) FE model of upper-airway complex from Sonomura et al [134]. This model does not include bony components (such as hyoid) and muscle driven structures; the organ movements are driven by manually defined boundary conditions and external forces. © Wiley (2011), Sonomura et al [134], adapted with permission. (b) Coupled biomechanics model of laryngopharynx from Tsou et al. [146]. Muscles are modelled as Hill-type actuators embedded in FE meshes; bony structures and cartilages are modelled as rigid bodies. The geometry of this laryngopharynx model is extracted from cadaveric data.

ment; then, the moving organ surfaces are extracted as the boundary condition to simulate food bolus motion using Smooth Particle Hydrodynamics (SPH) method. Ho’s methods provide a promising way to simulate high-fidelity human swallowing. However, their study is not based on realistic anatomy and motion trajectories of a specific subject.

In this chapter, we apply a subject-specific model of upper-airway complex to simulate the real swallowing behaviour recorded in medical data. Our simulation is based on dynamic 3D CT scans, which capture the organ motion in oral cavity, pharynx, larynx and upper esophagus during normal swallowing [37]. We first describe the details of the kinematic data used to drive the swallowing simulation in Section 4.1. Then, the details of the inverse simulation are given in Section 4.2. The simulation results are presented in Section 4.3. The discussion and conclusion follow in Section 4.4.

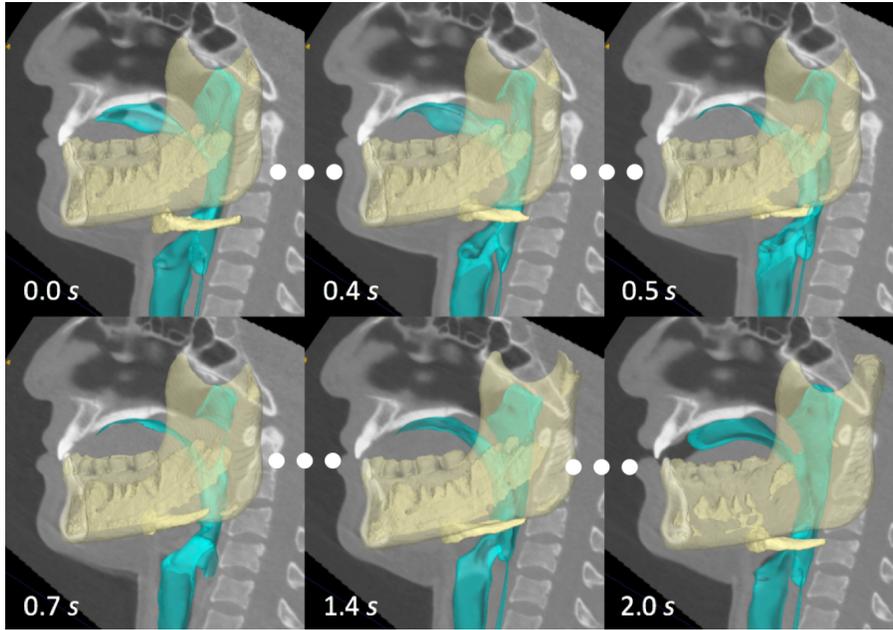


Figure 4.3: Segmentation of airway (cyan), jaw and hyoid bones (yellow) superimposed with the mid-sagittal slices of the sequential CT images (at 6 frames).

4.1 Kinematic Data

The dynamic CT dataset used in this study has been introduced in Section 3.2.1. It has 21 time frames in total with the temporal resolution of 10 *fps*. This sequential CT images record the swallowing motion from oral transit to the early esophageal phases. Based on the last frame, we construct a subject-specific upper-airway model, as shown in Figure 3.20.

We aim at reproducing the swallowing motion from oral transport phase to pharyngeal phase that is described in the CT image sequence. However, the CT images do not provide enough contrast ratio inside the soft-tissues to depict the motion of internal tissue points. Therefore, we define the tracking markers as the points distributed on component surfaces; the target positions of these surface markers are located on the organ boundaries in the sequential CT images. Ho et al. [62] manually segment the moving airway boundary from the 21 sequential CT images that we use in this study; the airway boundary depicts critical movements of the upper-airway soft-tissues.

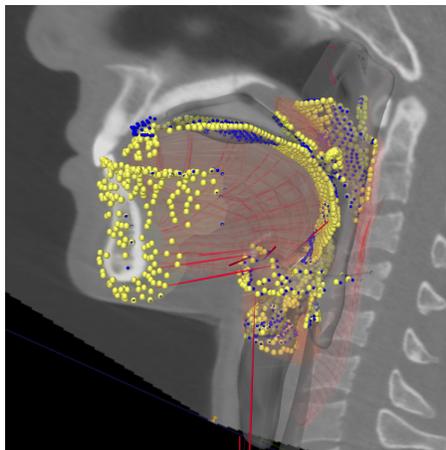


Figure 4.4: Illustration of tracking markers (yellow) and their target positions (blue).

In order to generate target motion trajectory for rigid bodies, including jaw and hyoid, we extract their surfaces from the CT volumes using the automatic segmentation tool in ITK-SNAP toolkit [164]. The segmentation (airway and bones) and CT images are illustrated in Figure 4.3.

4.2 Inverse Simulation

We perform swallowing simulations based on the segmented bone surfaces and Ho's airway segmentation. During the simulation, each frame ($0.1s$) of the dynamic CT was divided into fifty numerical time steps. We pick tracking markers in the model (representing the motion of the model) by uniformly sampling on the segmented airway and bones (jaw and hyoid) surfaces (yellow points in Figure 4.4). Since rigid bodies have much smaller DoF than FE components, the density of tracking points picked on the jaw and the hyoid geometry is set as one sixth of the density on the airway. We used ICP method to search their correspondences on the segmentation, and then linearly interpolate between the starting posture and the correspondences of each frame to generate the target motions (blue points in Figure 4.4) for every numerical time steps. As illustrated in Figure 4.5, for rigid components, we first find the best-fit rigid transformations that map their meshes to the corresponding segmentation; and then apply the rigid transformations to their associated tracking markers to generate their tar-

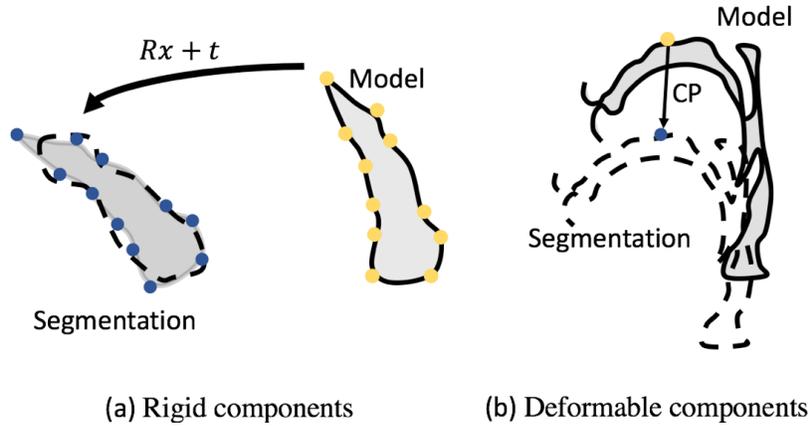


Figure 4.5: Methods of determining the target positions (blue points) for tracking markers (yellow points). (a) for tracking markers of rigid components, we first find the best-fit rigid transformation that maps the mesh of the rigid component to the corresponding segmentation; and then apply this rigid transformation to the tracking markers to generate their target positions. (b) for tracking markers of deformable components, we use their closest points on the corresponding segmentation as their target positions.

get positions. For tracking markers of deformable components, we use their closest points on the corresponding segmentation as their target positions in each image frame. Since the CT dataset does not capture the oral preparatory phase, we change the posture of the model to match with the first CT frame before the swallowing simulation.

4.3 Results

The data tracking simulation went through 21 frames in total. Figure 4.7-4.9 illustrate 8 frames in mid-sagittal, oblique and top views respectively. The muscle activations levels estimated by the inverse solver during swallowing are illustrated in Figure 4.10-4.12. As we notice the muscle activation patterns on both side (left and right) are similar, we show the left side only. The tracking performance for the FRANK components – the tracking errors and the standard deviations of the Euclidean distance between the target positions and the resulting marker positions – are shown in Figure 4.6. As we notice the tracking error around the corniculate cartilage is significantly larger than other part of the larynx, we illustrated their tracking performance separately.

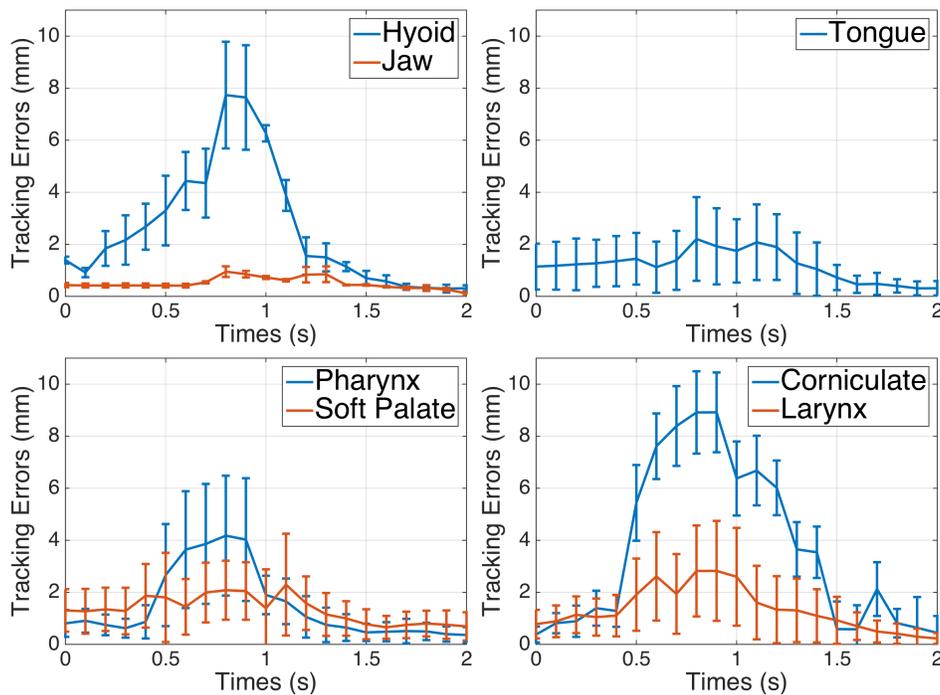


Figure 4.6: Tracking error and the standard deviation of each component during swallowing simulation, which was computed as the Euclidean distance between the target positions and the resulting marker positions.

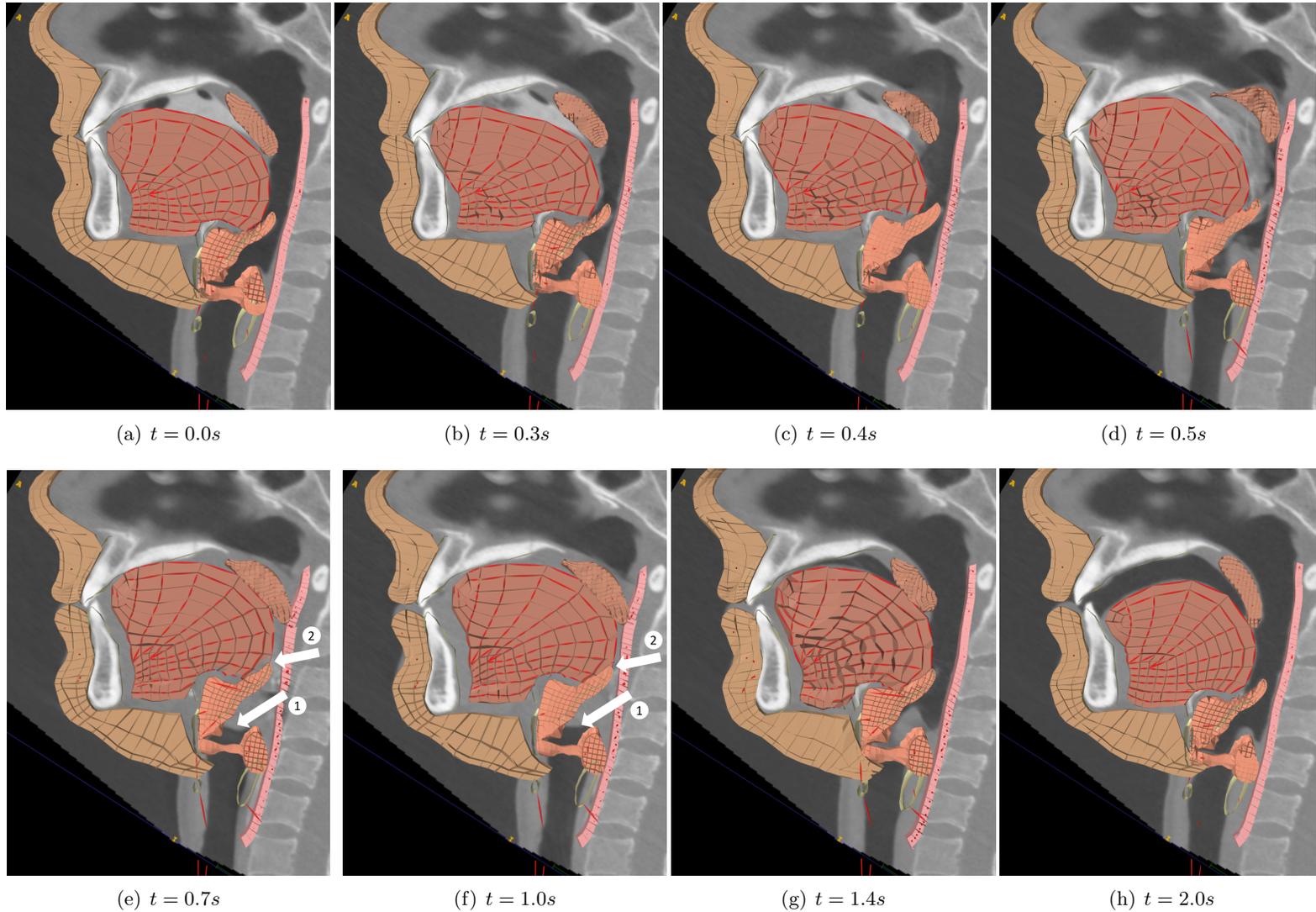


Figure 4.7: A sequence of video frames showing the mid-sagittal view of the inverse simulation of swallowing. The muscle activation pattern at these time frames are illustrated in Figure 4.10-4.12 at the vertical dash lines. Arrow 1 shows that the model fails to fully reproduce the laryngeal motion: The vocal fold of the model is much lower than its position shown in the CT images. Arrow 2 shows that the contact behaviour between tongue base and epiglottis – which is ignored in the simulation – may contribute to epiglottis inversion.

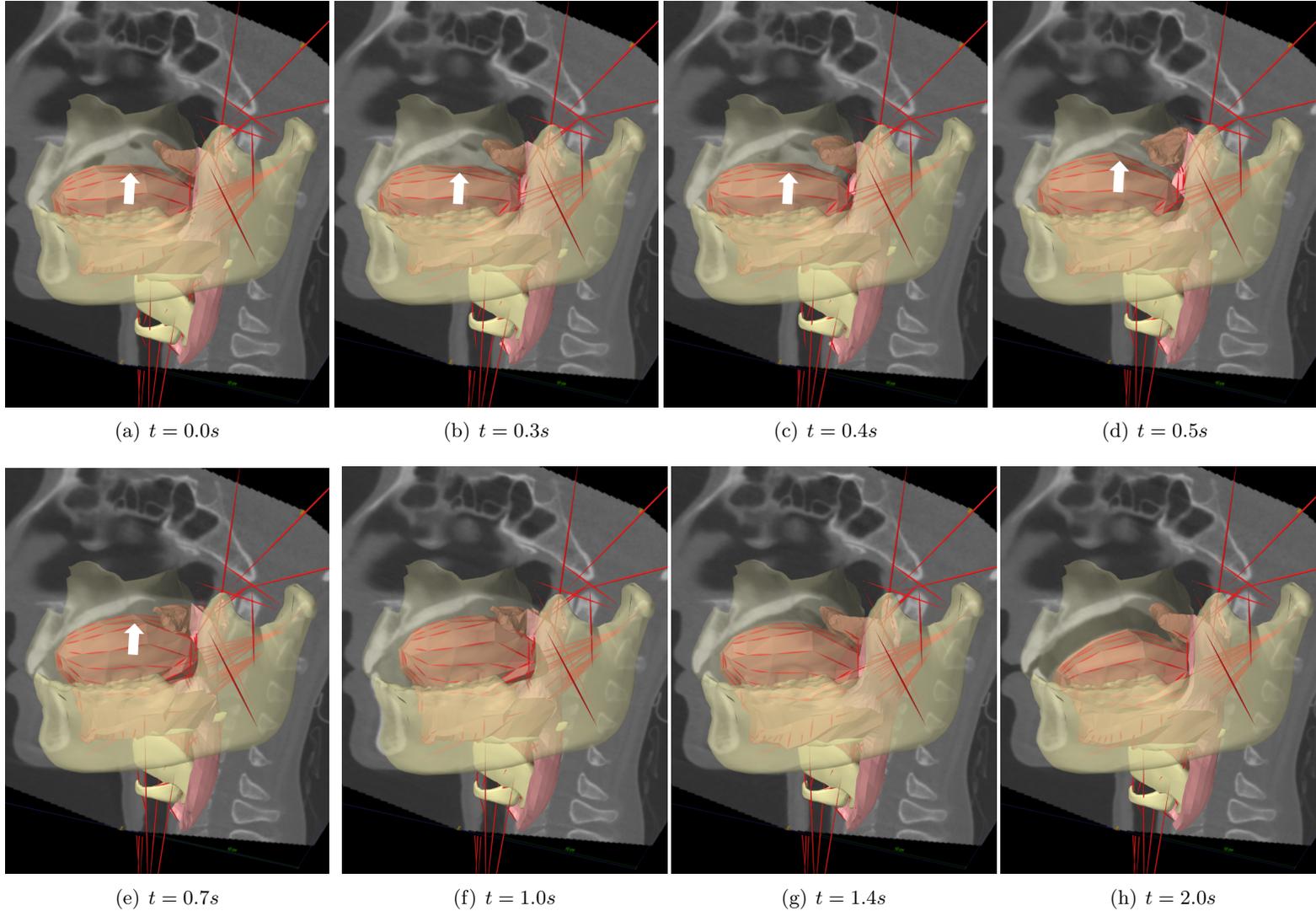


Figure 4.8: A sequence of video frames showing the oblique views of the inverse simulation of swallowing. Face model is hidden in the images, in order to illustrate the occluded structures. The muscle activation pattern at these time frames are illustrated in Figure 4.10-4.12 at the vertical dash lines. Arrows show that the wavelike motion of the tongue directs the food bolus backwards and downwards.

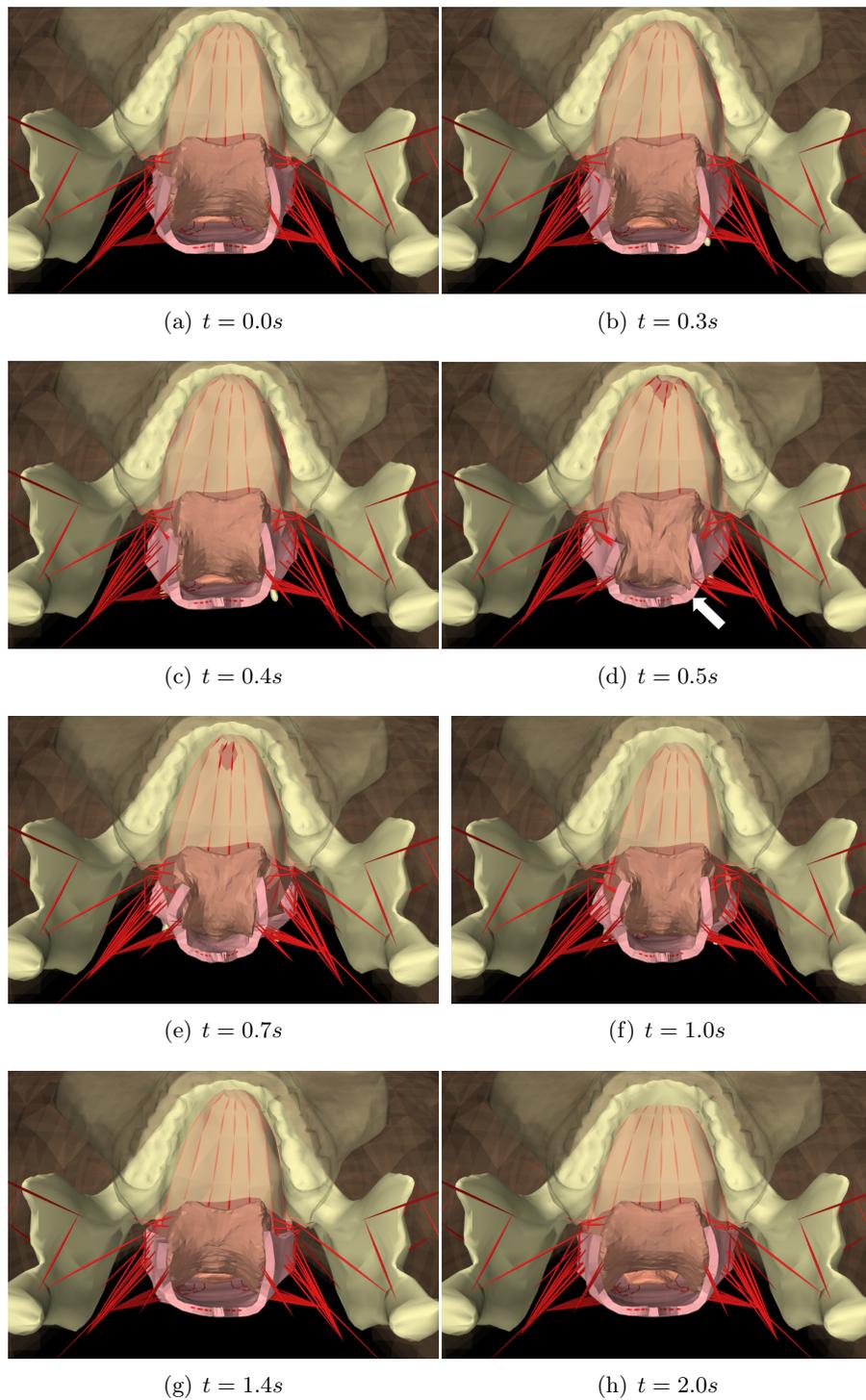


Figure 4.9: A sequence of video frames showing the top view of the inverse simulation of swallowing. The muscle activation pattern at these time frames are illustrated in Figure 4.10-4.12 at the vertical dash lines. Arrow shows velopharyngeal closure.

4.3. Results

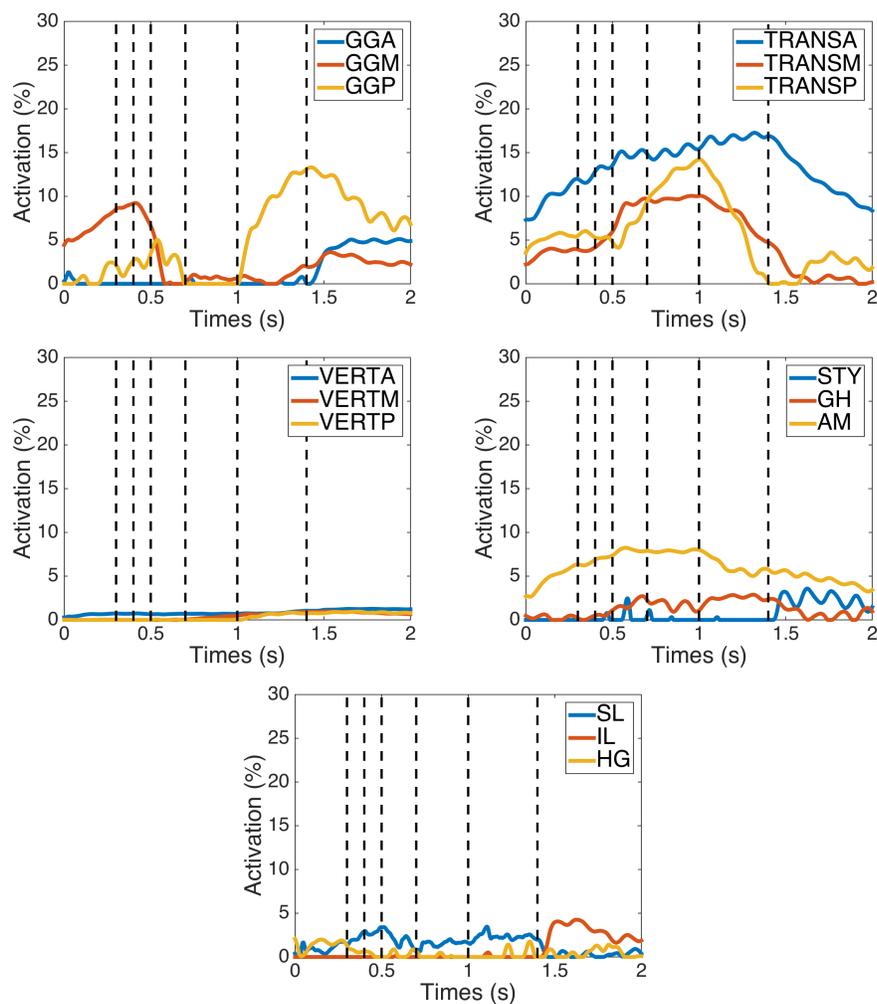


Figure 4.10: Activation levels of the extrinsic and intrinsic tongue muscles estimated by the inverse solver during the swallowing. The vertical dash lines label the time of $t = 0.3s, 0.4s, 0.5s, 1.0s, 1.4s$.

4.3. Results

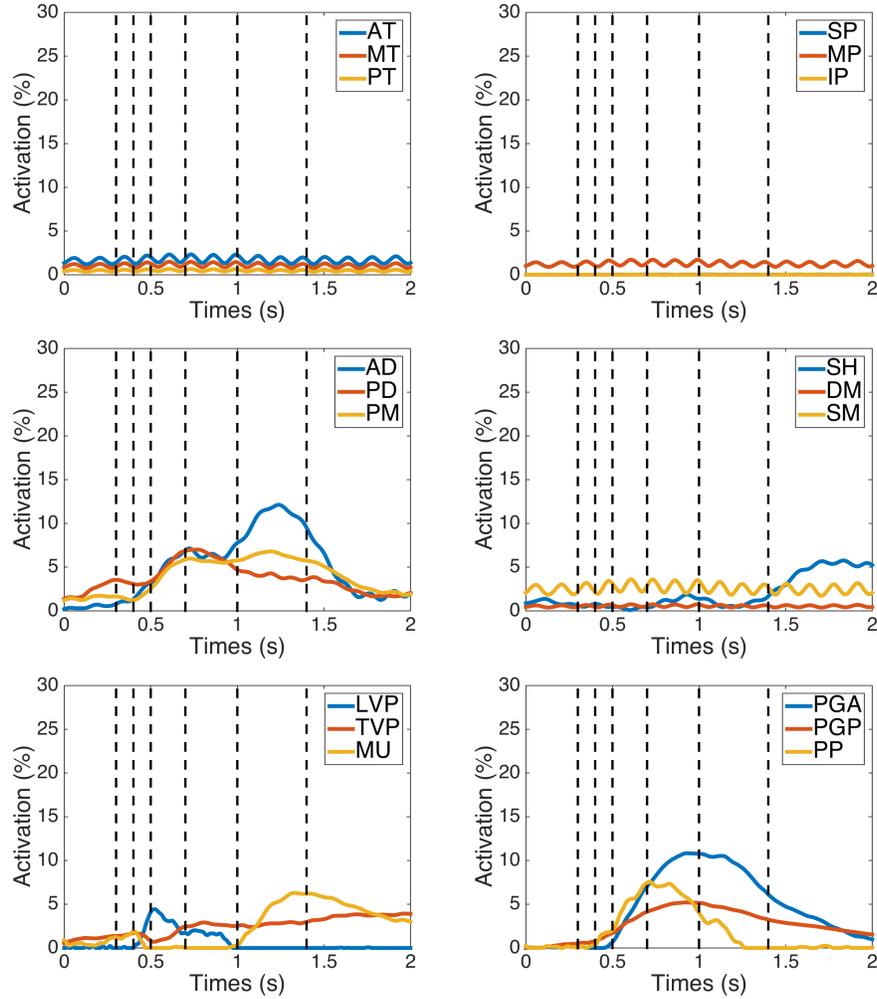


Figure 4.11: Activation levels of the suprahyoid muscles, mastication muscles and palate muscles estimated by the inverse solver during the swallowing. The vertical dash lines label the time of $t = 0.3s, 0.4s, 0.5s, 1.0s, 1.4s$.

4.3. Results

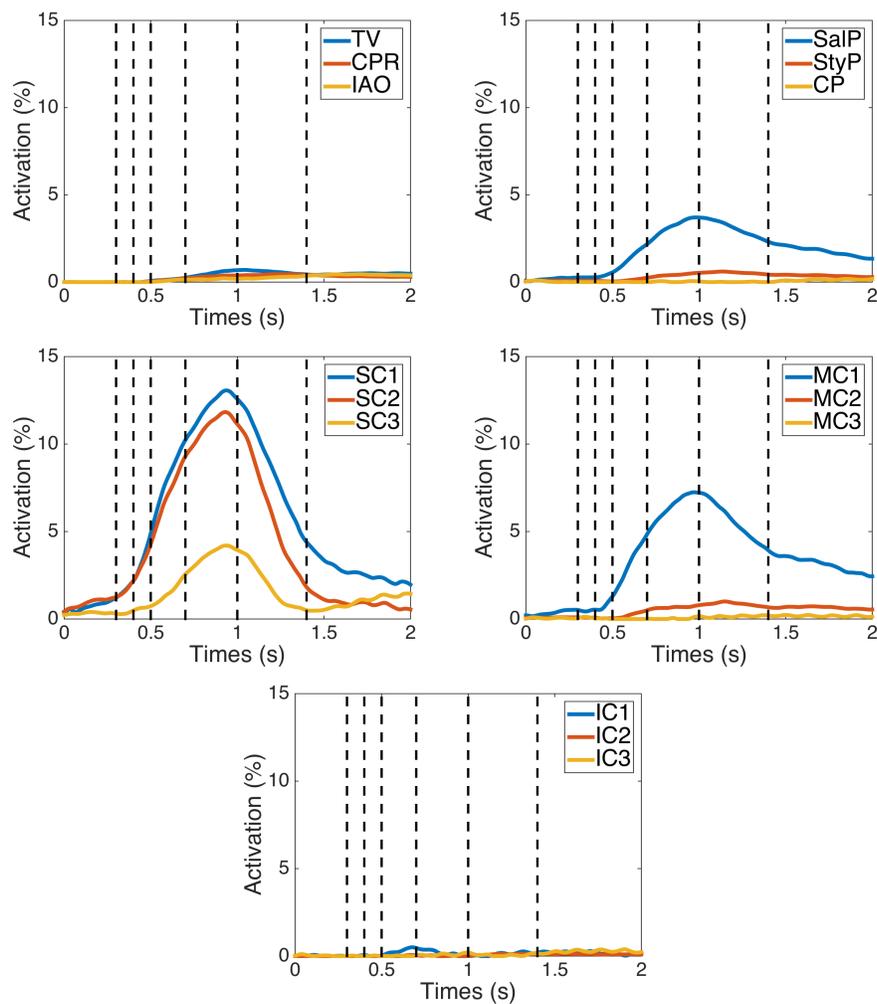


Figure 4.12: Activation levels of the laryngeal muscles and the pharyngeal muscles estimated by the inverse solver during the swallowing. Since only three laryngeal muscles (TV, CPR, IAO) are activated during the simulation, other laryngeal muscles are not shown in the figure. The vertical dash lines label the time of $t = 0.3s, 0.4s, 0.5s, 1.0s, 1.4s$.

4.3. Results

Hyoid	Jaw	Tongue	Pharynx
$2.67mm \pm 0.88$	$0.43mm \pm 0.08$	$1.12mm \pm 0.70$	$0.86mm \pm 0.63$
Soft-palate	Larynx	Corniculate	
$1.88mm \pm 1.23$	$1.35mm \pm 0.96$	$1.3781mm \pm 0.6689$	

Table 4.1: The peak values of the tracking errors of the FRANK components during oral transport phase (from $t = 0.0s$ to $t = 0.4s$).

Oral Transport Phase At the outset of the oral transport phase (at $t = 0.0s$), by contracting SL, GGM and TRANS muscles, the bolus is contained between the lingual dorsum and hard palate, as shown in Figure 4.10. Then, by contracting PD and AM, hyoid bone is pulled to move superiorly; in the meanwhile, via coordinating the intrinsic tongue muscles (including TRANS and SL) and the extrinsic tongue muscles (including AM, GGM, GGP), the tongue blade is pushed upwards and moves in a wavelike motion, which directs the bolus posteriorly towards the oropharynx (as shown in Figure 4.8). The mastication muscles (SP, MP, SM, DM, AT, MT and PT) are activated throughout the oral transport phase (from $t = 0.0s$ to $t = 0.4s$) to stabilize the jaw and tongue. During oral transport phase, except for the hyoid bone, the tracking errors of other tracking components are below 2.0mm (Table 4.3); the hyoid bone has the largest tracking error.

Pharyngeal Phase Pharyngeal phase lasts for approximate one second (from $t = 0.4s$ to $t = 1.4s$). As the oral transport phase ends, the bolus enters the oropharynx and crosses over the area of the anterior faucial pillars. This contact initiates the involuntary trigger of the pharyngeal phase [130]. TV and IAO contract to close the laryngeal inlet (from $t = 0.4s$). The tongue base is retracted towards the posterior pharyngeal wall. At the same time, LVP pulls the soft palate upwards and backwards; combined with SC contraction, the velopharyngeal closure is formed, which prevents the food bolus from entering the nasal cavity. During this motion, PP, PGA and PGP involve to stabilize the soft palate. As the velopharyngeal closure is formed (from $t = 0.5s$), the pharyngeal constrictors (SC, MC and IC) are sequentially activated from SC1 downward to IC3 (Figure 4.12), thus producing pharyngeal peristalsis from the oropharynx to the upper esophagus (Figure 4.7(d)-(f)). The movements of the tongue base and the pharyngeal wall are generated to provide the driven force for the pharyngeal bolus transit [130]. The contraction of the suprahyoid muscles (AM, PM,



Figure 4.13: The model fails to reproduce the target hyoid posture at $t = 0.9s$. The yellow points represent the tracking markers; the blue points represent their corresponding target positions.

GH, AD, PD) directs the hyoid bone superiorly and anteriorly (from $t = 0.4s$). The larynx – as it attaches to the hyoid rigid body – passively moves along with the hyoid bone. The superior and anterior movements of the hyoid bone direct the larynx under the tongue base, thereby preventing the bolus entering the laryngeal inlet, as shown in Figure 4.7(d)-(f).

Large tracking errors appear on the larynx and hyoid models during pharyngeal phase (Figure 4.6 and Table 4.3). At $t = 0.9s$, as shown Figure 4.13, the hyoid model fails to reproduce the target posture. In the meanwhile, the larynx fails to raise the corniculate cartilage. As shown in Figure 4.14, the tracking markers (yellow points) around the corniculate cartilage are substantially lower than their target positions (blue points) at $t = 0.9s$. As a consequence, the vocal fold of the larynx model in Figure 4.7(d)-(f) locates at the position of tracheal inlet of the CT images, leaving the laryngeal inlet open.

4.3. Results

Hyoid	Jaw	Tongue	Pharynx
$7.70mm \pm 0.95$	$0.95mm \pm 0.08$	$2.83mm \pm 0.80$	$4.31mm \pm 0.37$
Soft-palate	Larynx	Corniculate	
$2.32mm \pm 1.07$	$2.21mm \pm 0.91$	$8.94mm \pm 0.67$	

Table 4.2: The peak values of the tracking errors of the FRANK components during pharyngeal phase (from $t = 0.4s$ to $t = 1.4s$).

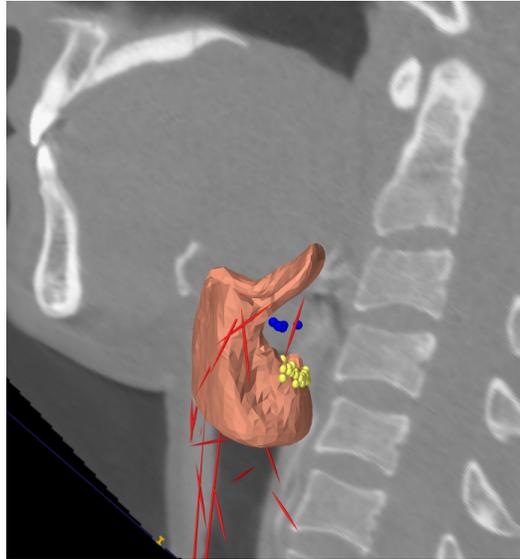


Figure 4.14: The model fails to raise the the corniculate cartilage at $t = 0.9s$, leaving the laryngeal inlet open during the pharyngeal phase. The yellow points represent the tracking markers; the blue points represent their corresponding target positions.

4.4 Discussion and Conclusion

As the results demonstrate, the model was able to achieve normal swallowing that closely resembles real tissue motion captured during the oral transport phase and the pharyngeal phase. In general, the tracking error in pharyngeal phase is greater than that in oral transport phase. Two model components suffered from large tracking error: the hyoid bone and the larynx. There are several possible explanations for their limited tracking performance. Firstly, the laryngeal complex model is over simplistic and does not provide the fidelity to describe the realistic biomechanics in that region. As an example, we model the laryngeal cartilages as rigid bodies embedded inside the larynx FE model, which makes the larynx model over stiff; in turn, the larynx model lacks the flexibility to allow the elevation of the corniculate cartilage and the inversion of the epiglottis.

In addition, the accuracy of the swallowing simulation is not only limited by the biomechanical model but also the resolution of the image data. The dynamic CT images used in this study do not describe the motion of a specific tissue point inside the pharyngeal wall. Besides, from $t = 0.6s$ to $t = 1.1s$, the CT images fail to depict the boundary between the tongue and the pharyngeal wall. These limitations make the vertical motion information of the pharyngeal wall unavailable. The elevation of the pharynx is essential for the elevation of larynx [130]; Pearson et al. [113, 114] suggest the long pharyngeal muscle group (SalP, StyP, PP) are activated during swallowing to assist hyolaryngeal elevation. However, SalP, StyP and PP were not activated from $t = 0.5s$ to $t = 1.0s$ in the swallowing simulation (Figure 4.12), which may affect the motion of both the hyoid bone and the larynx.

Furthermore, since unilateral constraints may lead to oscillating or sticking behaviour in the inverse simulation [136], we did not incorporate the contact behaviour in the swallowing simulation. The lack of contact may affect the swallowing biomechanics. For example, during the pharyngeal phase, the contact between the tongue base and the epiglottis may contribute to the epiglottis inversion (Figure 4.7(c)-(f)).

To summarize, our model is capable of reproducing oropharyngeal motion of normal swallowing. However, it shows limited ability to track hyolaryngeal movements. The superior and anterior movements of the hyoid and the larynx are essential to swallowing for two reasons. Firstly, it will direct the laryngeal inlet under the tongue base and seal off the airway to cease respiration and to prevent any misdirected food or liquid from entering the trachea. Secondly, these movements will create a biomechanical force to pull the cricoid cartilage up and away from the posterior pharynx.

geal wall, thus opening the esophagus; the opening of the UES will create an additional source of negative pressure (or suction force) in the upper esophagus, which would greatly enhance the efficiency of pharyngeal bolus transit [130]. Therefore, reproducing of bolus transit biomechanics of normal swallowing requires a functional hyolaryngeal complex. Future work is needed to increase the fidelity of hyolaryngeal complex model in the upper-airway template. Moreover, applying different image modalities, such as tagged MRI, to investigate swallowing motion may provide more accurate kinematic information and help with reduction of the ambiguity of tissue motion.

Although the model failed to reproduce the hyolaryngeal motion, the proposed subject-specific modelling and simulation method can be applied to assist with the prediction and assessment of the postoperative swallowing biomechanics before intralaryngeal prosthesis implantation. Intralaryngeal prosthesis implantation is a treatment option for the patient suffering from aspiration due to an impairment of the pharyngeal phase (or larynx). Such application does not require a fully functional larynx. Our model demonstrates the ability to reproduce the swallowing motion in the oropharynx and velopharynx. After insertion of a virtual intralaryngeal prosthesis, such as the prosthesis introduced in [123], coupled with certain fluid simulation methods, the model can help with estimating post-implantation bolus-transit biomechanics. In future work, efficient fluid-structure interaction methods for swallowing simulation will be investigated. Besides, as the contact behaviours between organs may influence the biomechanics of the swallowing, we are investigating methods to incorporate unilateral constrains in the inverse simulation.

Chapter 5

Data-Driven Speech Simulation: Towards Speaker-Specific Articulatory Synthesis

Speech production is an essential part of daily life for most people. Speech synthesis has long been a vital assistive technology tool which allows environmental barriers to be removed for people with a wide range of disabilities. The physical patterns carrying linguistic information in speech are generated by movements of the articulatory apparatus, such as jaw and tongue. Hence, a careful analysis of the physical properties of the speech apparatus, including an anatomical, biomechanical and aerodynamic characterization should contribute to the understanding of the structures of languages and of their evolutions [117]. Articulatory synthesis refers to computational techniques for synthesizing speech based on models of the human vocal tract (controlled by speech articulators, such as the tongue and jaw) and the articulation processes occurring there. This technique targets at simulating specific voices, speaking styles, and emotions for an arbitrary speaker [13]; it is widely considered as a valuable computational aid for the analysis and assessment of speech disorders of specific individuals.

Following Perkell [116], who initiated pioneer works in physiological modelling of tongue, more sophisticated and more realistic orofacial models [17, 42, 43, 57, 96, 117, 131, 154] have been developed to enable comprehensive understanding of the articulatory characters of the upper-airway organs in speech production. For example, Buchaillard et al. [17] use a 3D biomechanical model of tongue and oral cavity (shown in Figure 5) to study the motor control patterns of the tongue during the production of French cardinal vowels. Their model includes two dynamic organs, i.e. a FE tongue and a rigid-body-represented hyoid bone; the other fixed structures that shape the oral cavity are linked with the tongue and hyoid using point-

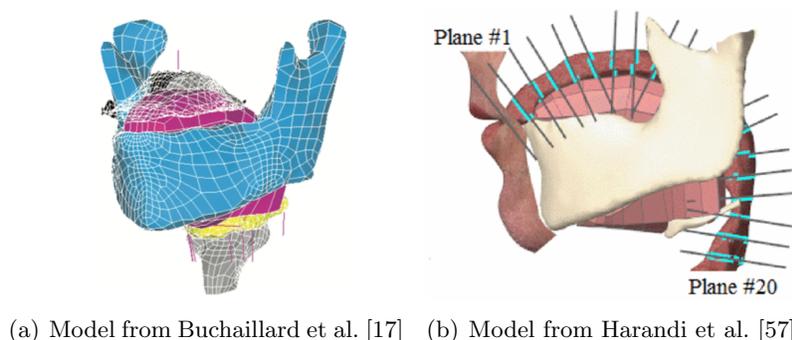


Figure 5.1: Illustration of two biomechanical models for speech simulation reported in previous studies. (a) biomechanical model of upper-airway complex from Buchaillard et al [17]. This model include two dynamic components: the FE tongue (in magenta) and the rigid-body-represented hyoid bone (in yellow). Other structures shaping the the oral cavity are fixed in space to provide boundary conditions. © Acoustical Society of America (2009), Buchaillard et al. [17], adapted with permission. (b) The template biomechanical model of oropharynx reported by Harandi et al. [57]. The vocal tract shape is controlled by the motions of the tongue and jaw. The Planes – used to extract the area function of the vocal tract – are orthogonal to the vocal tract center line and evenly distributed from the lip (#1) to the epiglottis (#20). © IEEE (2015), Harandi et al. [57], adapted with permission.

to-point muscles, and restrict the tongue motion via contact behaviours. Stavness et al. [137] develop a jaw-tongue-hyoid model using dynamic FE method combined with rigid-body dynamics; this model is able simulate co-articulation effects in speech production. Based on Stavness’s jaw-tongue-hyoid model, Harandi et al. [57] introduce a speaker-specific articulatory synthesis framework: They first register the jaw-tongue-hyoid model to a MR data set (as shown in Figure 5) and simulate the tongue speech motion of the utterance a-geese (/ə-gis/) based on tagged MR images; then the area function of the deformed vocal tract is extracted from the model as the input of an 1D implementation of the Navier-Stockes equations [150]. As the several important speech articulators (e.g. the face, soft-palate, larynx and pharynx) are missing in Harandi’s model, only inside of the oral cavity the vocal tract is deformable.

In recent years, computational orofacial models have been applied to pre-

dict and assess the impact of head-and-neck surgery on speech production. Surgical treatment of head-and-neck cancer may cause severe consequences on mobility of the articulators and even strong impairments of speech production. Fujita et al. [38] construct a FE tongue model based on MRI data to predict and verify the changes in the movements of the tongue with a tumor before and after partial glossectomy. Similarly, Buchaillard et al. [16] build an oropharyngeal model based on MR and CT scans of the patient to predict the consequences of the tongue surgery on tongue movements, according to the size and location of the tissue loss.

Currently, modelling of coupled articulatory biomechanical system still attracts much less attention than modelling of isolated articulators. The vocal tract shape of the speakers can provide vital geometric information for articulatory synthesis. In this chapter, we apply the proposed subject-specific modelling methods to create coupled upper-airway systems for two speakers. Similar to Harandi et al. [57], we simulate the speech movements based on the kinematic information of the tongue and jaw extracted from tagged MR images; the other speech articulators that lack tracking data act as functional regularizers moving along with them (the tongue and jaw).

5.1 Medical Dataset

Our MRI data captures two normal Caucasian American English speakers in the supine position, who repeated the utterances a-geese (/ə-gis/) in time with a metronome. Cine MRI is able to depict the motion of soft-tissue boundaries, but it usually fails to distinguish internal tissues. In contrast, tagged MR is capable of providing kinematic information of internal tissues, but it blurs organ boundaries. Therefore, we base our simulation on synchronized 3D tagged and cine MRI data, which capture the motion of tissues in the tongue and chin. Both the cine and tagged MRI data were acquired using a Siemens 3.0T Time-Trio MRI scanner with a 12-channel head and a 4-channel neck coil. The in-plane image resolution is $1.875mm \times 1.875mm$ with a slice thickness of $6mm$. Other sequence parameters include the following: TR $36ms$, TE $1.47ms$, flip angle 6° , and turbo factor 11. The axial, sagittal and coronal stacks of cine MRI slices are combined to form isotropic super-resolution volumes for 26 TFs, using a maximum posterior-Markov Random Field method with an edge-preserving regularization scheme [158].

Table 5.1 summarizes the information of each individual speaker. The phonemes of interest (/ə/, /g/, /i/ and /s/) are identified in specific Time Frames (TFs) in the utterance. Each vowel is identified at the TF before the

5.1. Medical Dataset

tongue begins to move toward the next consonant, i.e. that is the maximum vowel position. Each consonant is identified at the TF when the tongue first contacts the palate, i.e. the initial frame, rather than the maximum consonant position. The mid-sagittal slices of the cine MR images in the chosen TFs are illustrated in Figure 5.2.

Speaker index	Sex (M/F)	Age (years)	TF # in /ə-gis/			
			ə	g	i	s
A	F	43	8	10	14	23
B	M	22	6	10	18	20

Table 5.1: Speaker information in this study: Sex, age, and time frames associated with individual sounds in the utterance a-geese (/ə-gis/). The corresponding mid-sagittal slices of the cine MR images in these TFs are illustrated in Figure 5.2.

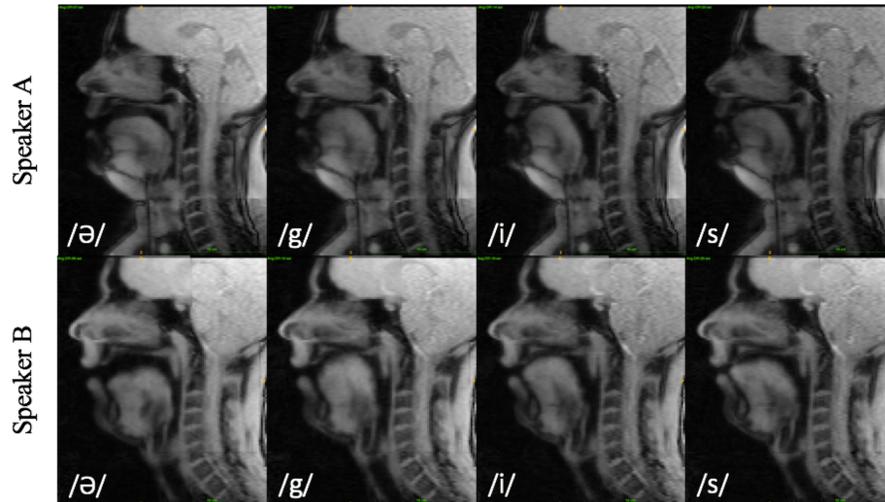


Figure 5.2: Mid-sagittal slice of cine MR images at the specific TFs of the phonemes of interest (/ə/, /g/, /i/ and /s/). The top and the bottom rows show the speaker A and speaker B respectively.

5.2. Speaker-Specific Modelling

Organs (RC)	Speaker	Method	Geometry Result
Jaw	A & B	Manual	Partial
Tongue	A & B	Manual	Complete
Lip & Chin	A & B	Auto	Partial
Airway	A & B	Auto	-

Table 5.2: Summary of segmentation.

Speaker	Metric	Jaw	Lip & Chin	Tongue	Airway
A	$D_{mean}(mm)$	0.65	0.69	0.58	1.08
	$D_{max}(mm)$	3.17	3.59	3.89	4.12
	$\sigma(mm)$	0.59	0.70	0.63	0.88
B	$D_{mean}(mm)$	0.51	0.99	0.61	1.07
	$D_{max}(mm)$	2.33	4.59	3.90	4.05
	$\sigma(mm)$	0.45	0.69	0.63	0.83

Table 5.3: Registration errors within the two speaker-specific models of upper-airway complex. D_{mean} , D_{max} and σ are the mean surface distance, the maximum surface distances and the standard deviation.

5.2 Speaker-Specific Modelling

Using the model registration methods introduced in Chapter 3, we created two speaker-specific upper-airway models based on the first TF of the cine MR data. We extracted the segmentation for Reference Component (RC) from the cine MR images. The segmentation information is summarized in Table 5.2.

Figure 5.3 illustrates our 3D upper-airway models for speakers A and B. Registration errors are shown in Table 5.2. Note the mean values of surface-representation error fall in the range of image resolution. Larger registration errors are localized around ill-defined features (e.g. the laryngeal inlet), or the around poor-quality elements (e.g. the lip and laryngeal inlet.)

Figure 5.4 compares average quality and the number of poor-quality elements of FE meshes in the template and our speaker models. Note that our registration method improves the average mesh quality in the tongue and face, while maintaining the quality about the same level for the pharynx, soft palate and larynx. Besides, the numbers of elements where the mean-ratio values are smaller than 0.1 were reduced for all the FE meshes.

5.2. Speaker-Specific Modelling

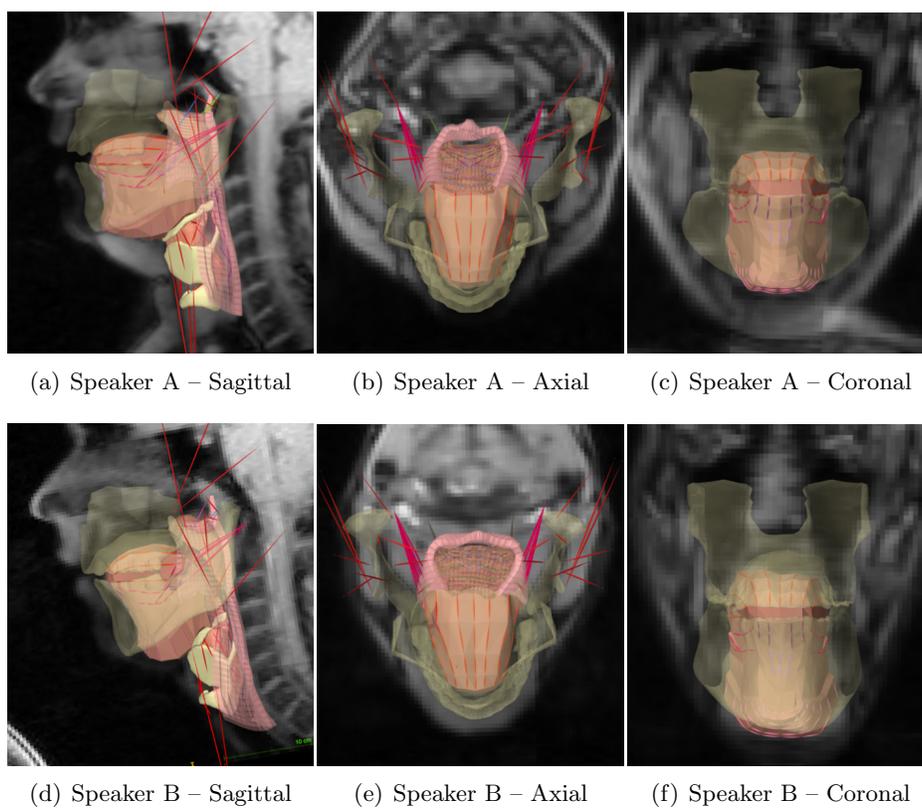
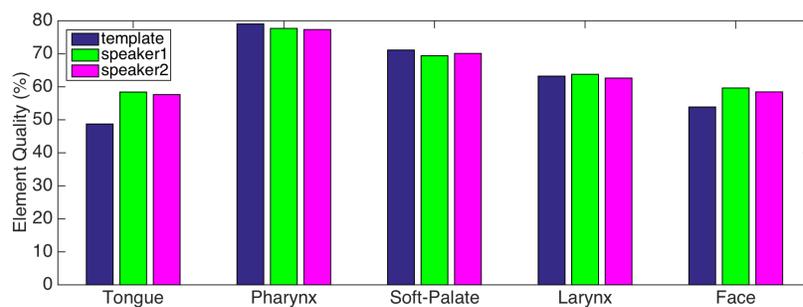
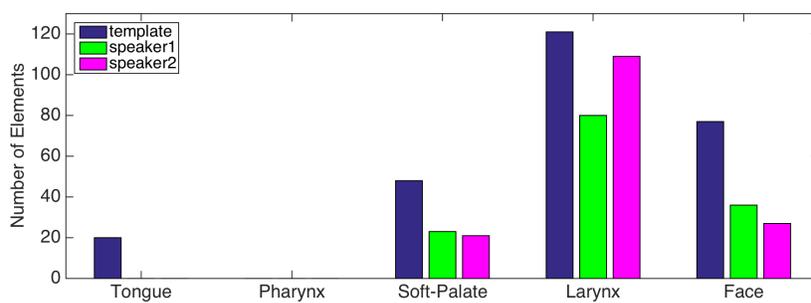


Figure 5.3: Speaker-specific models of upper-airways complex.

5.2. Speaker-Specific Modelling



(a)



(b)

Figure 5.4: Comparison of the element quality of the speaker-specific models and the template models. Figure (a) illustrates the average element qualities (η_{mean}) of the each FE components. Figure (b) illustrates the number of elements whose mean-ratio values are smaller than 0.1 ($\#(\eta < 0.1)$).

5.2.1 Inverse Simulation

Tongue protrusion is an important speech motion of the utterance /ə-gis/. To demonstrate the tongue-protrusion ability of our models (including the template, Speaker A and Speaker B), we performed an inverse simulation to produce a tongue-protrusion posture: The models targeted at keeping the jaw at the neutral position (i.e. the initial position), while forcing the tongue blade and tongue center to touch the hard palate. As shown in Figure 5.5, we used six tracking markers (three at the tongue blade, the other three at the tongue center) for the tongue and three for the jaw. In the first 0.10s of the simulations, each model was allowed to settle under gravity. After that, the tracking markers of the jaw stayed at their original positions, while the tracking markers of the tongue gradually went upwards until met with the hard palate at 0.17s. The models settled again without any further changes on the target positions until 0.20s. The tracking weights (w_m), regularization weight (w_a) and damping weight (w_d) were set to 1.0, 0.05, 0.05 respectively. Since the target motion is bilateral symmetry, we enforce the left and right muscles to be activated together and with the same intensity. The target muscles used in the inverse simulation are listed below:

1. Muscles of the tongue: genioglossus [GG], hyoglossus [HG], styloglossus [STY], verticalis [VERT], transversus [TRANS], geniohyoid [GH], anterior mylohyoid [MH], and longitudinal (inferior [IL], superior [SL]). GG, VERT and TRANS were further divided into three individual segments (anterior [A], middle [M], posterior [P]).
2. Muscles of the Jaw and Hyoid: jaw openers including digastric (anterior [AD], posterior [PD]) and stylo-hyoid (SH); jaw closers including temporal (anterior [AT], middle [MT], posterior [PT]), masseter (superficial [SM], deep [DM]), and medial pterygoid (MP); other jaw and hyoid muscles including pterygoid (superior-lateral [SP], inferior-lateral [IP]).

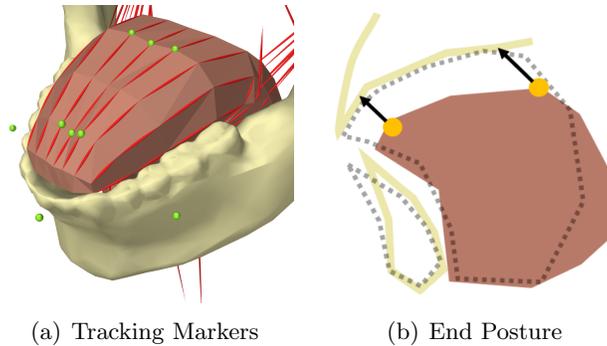


Figure 5.5: Illustration of the inverse simulation. (a) The yellow points represent the tracking markers; we use six tracking markers for the tongue, three tracking markers for the jaw. (b) the models target at keeping the jaw at the neutral position, while forcing the tongue to touch the hard palate. The dash lines represent the potential end postures of the jaw and tongue after the inverse simulation.

The end postures and the inverse-predicted muscle activations of the three models are illustrated in Figure 5.6. The muscles – whose activation levels were greater than 1% – were considered as active muscles; other muscles (including GGA, GGM, IL, IP and SP) are not shown in Figure 5.6. As shown in Figure 5.6, three models used the same muscles to achieve the tongue protrusion postures. TRANS narrows the tongue and VERT widens and grooves it; when these muscles are activated together, they protrude the tongue. From 0.1s onwards, Speaker A and Speaker B used TRANS in conjunction with VERTM and VERTP to protrude the tongue. However, VERT in the template showed a low level of activity throughout, which suggests the template had relatively less tongue protrusion. Other muscles of the three models were activated following a similar pattern during the simulations. GGP pulled the tongue root forwards. STY pulled the tongue dorsal backwards to stabilize the tongue; in the meanwhile, this muscle assisted the tongue to move upwards. HG, GH and MH stabilized the elevated the hyoid to assist the upward and forward motion of the tongue. Jaw openers and closers were activated simultaneously to stabilize the jaw.

The tracking errors – average distances between the tracking markers and their corresponding target positions – of the template, Speaker A and Speaker B are summarized in Table 5.2.1. Note that three models are ca-

5.2. Speaker-Specific Modelling

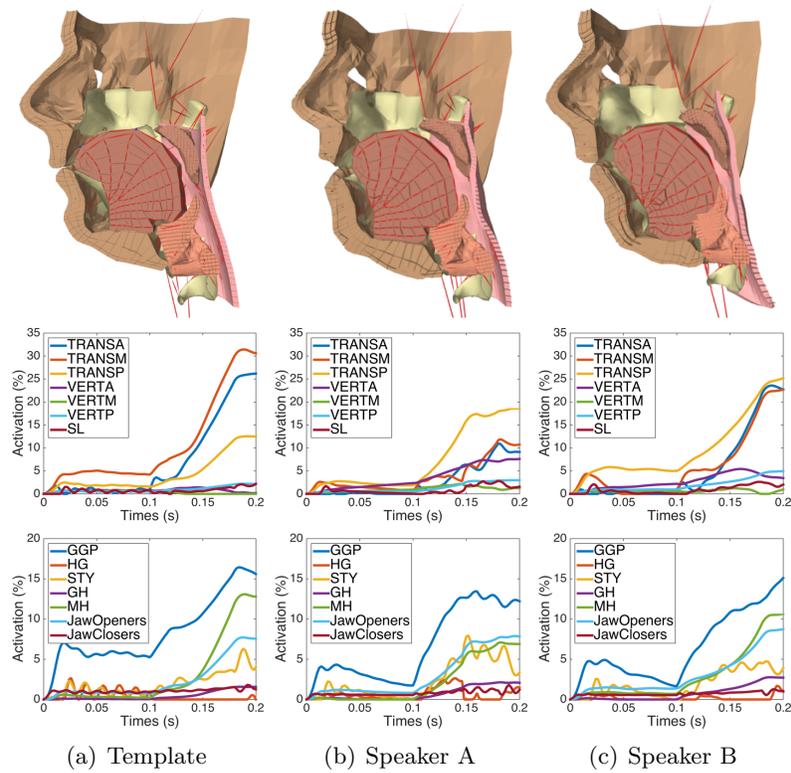


Figure 5.6: The end postures and the inverse-predicted muscle activations of the template, Speaker A and Speaker B.

pable of tongue protrusion and, in the meanwhile, keeping the jaw at the neutral position: For both the tongue and jaw, three models achieve sub-millimeter tracking accuracy.

	Template	Speaker A	Speaker B
Tongue	0.37 ± 0.11 (mm)	0.17 ± 0.04 (mm)	0.61 ± 0.37 (mm)
Jaw	0.01 ± 0.00 (mm)	0.02 ± 0.00 (mm)	0.01 ± 0.00 (mm)

Table 5.4: Tracking errors of the template, Speaker A and Speaker B at the end posture.

5.3 Speech Simulation

Models track the soft-tissue in the tongue, and lower jaws motion, based on tagged MRI trajectories available for them. The tracking weights (w_m), regularization weight (w_a) and damping weight (w_d) were set to 1.0, 0.05, 0.05 respectively. To reduce the computational cost of the inverse problem, we consider the speech motion as bilateral symmetry; the left and right muscles were activated together and with the same intensity. The target muscles included in the speech simulation are the same as the muscles listed in Section 5.2.1.

5.3.1 Definition of the Trajectories

We perform speech simulations based on the tissue points motion estimated from the tagged MR and cine MR images. The two dimensional motion of the tongue and lower-chin tissue-points was estimated from tagged MR image slices using the HARmonic Phase (HARP) algorithm [111]. We further applied the Enhanced Incompressible Deformation Estimation Algorithm (E-IDEA) to combine the 2D motion data and make a 3D deformation field [160]. E-IDEA imposed a smooth, divergence-free, vector spline to interpolate velocity fields across the target organs. In HARP, the displacement field at each TF was calculated with reference to the first TF when the tags were initially applied. We calculate displacements between successive TFs – from the n th to the $(n + 1)$ th TF – by using the following process:

$$T_{n \rightarrow n+1} = T_{n \rightarrow 1} \circ T_{1 \rightarrow n+1}$$

where $T_{i \rightarrow j}$ denotes the displacement field from the i th to the j th TF. The $T_{n \rightarrow 1}$ is computed by inverting the E-IDEA displacement field $T_{1 \rightarrow n}$ using a simple fixed-point algorithm [25].

To reduce the influence of the image noise on the estimated motion, we discarded the tissue points that are close to the surfaces (in the range of 2mm). The displacement vectors were averaged in a spherical region of predefined radius (2mm) around each point of interest. As we noticed the estimated motions of isolated tissue points at the lower chin were not reliable, we calculated their best-fit rigid transformations as the motion trajectory of the jaw model for our inverse simulations. For each speaker, we tracked about 40 tissue points in the tongue, and the rigid motion of the jaw that represented by 20 tracking markers attached to it. A linear interpolation was performed between successive TFs to calculate the intermediate displacements.

5.3.2 Results

Figure 5.7 demonstrates the cross section views of the Speaker A and Speaker B in 5 key time frames. Figure 5.8 illustrates the estimated muscle activations. Motor equivalence is evident in our dataset; especially in amount of activation used by each speaker, considerable similarity is also visible. Comparing like colors for TRANS and VERT in Figure 5, red is anterior, blue middle, and yellow/green posterior, we can see similarity, in that neither speaker activates the TRANS and VERT simultaneously. This indicates a lack of protrusion, despite the forward motion of the tongue from /g/ to /i/ to /s/. The exception is a short pulse before /i/ when Speaker A did activate TRANS simultaneously with VERT. Looking at the muscles separately, both speakers used a low-level activation pattern for TRANS throughout the speech task, possibly to keep the tongue from spreading too widely. During /gi/ they used VERTP in conjunction with GGP (row 2), to pull the tongue root forward. Interestingly, other than this one instance, GG was not used by either speakers, possibly because as the largest tongue muscle [141] it is used for larger tongue motions. In this speech task, the jaw was quite closed and only small motions were needed to execute the sounds. Differences appear as well. Speaker A activated the entire VERT muscle (row 1) for /i/, to keep the tongue surface flat/grooved, while Speaker B achieved that shape with VERTA, but not VERTM. The IL muscle (row 2), on the other hand, shows greater motor differences; Speaker A, but not B, used IL during /i/.

When looking at the jaw and hyoid muscles in row 3, the speakers were again quite similar. The jaw closers (red) showed a low level of activity throughout, since the jaw was quite closed during this speech task. Activation was seen during /g/ and /s/ for the jaw openers and GH, which pulled the hyoid forwards. These activity peaks may relate to the switch from voiceless /g/ to the voiced /i/, or they may elevate or stabilize the tongue root during the consonants. Stabilization and elevation of the tongue root and hyoid would assist the tongue to make stable contact with the palate to precisely produce and direct the airflow for the consonants.

The average tracking error – the mean values and standard deviations of the Euclidean distance between the target positions and the corresponding marker positions – at the key TFs is summarized in Table 5.5 and Table 5.6. Note that the error is in the range of MRI resolution. However, compared with Speaker A, Speaker B has larger tracking error in both the tongue and jaw.

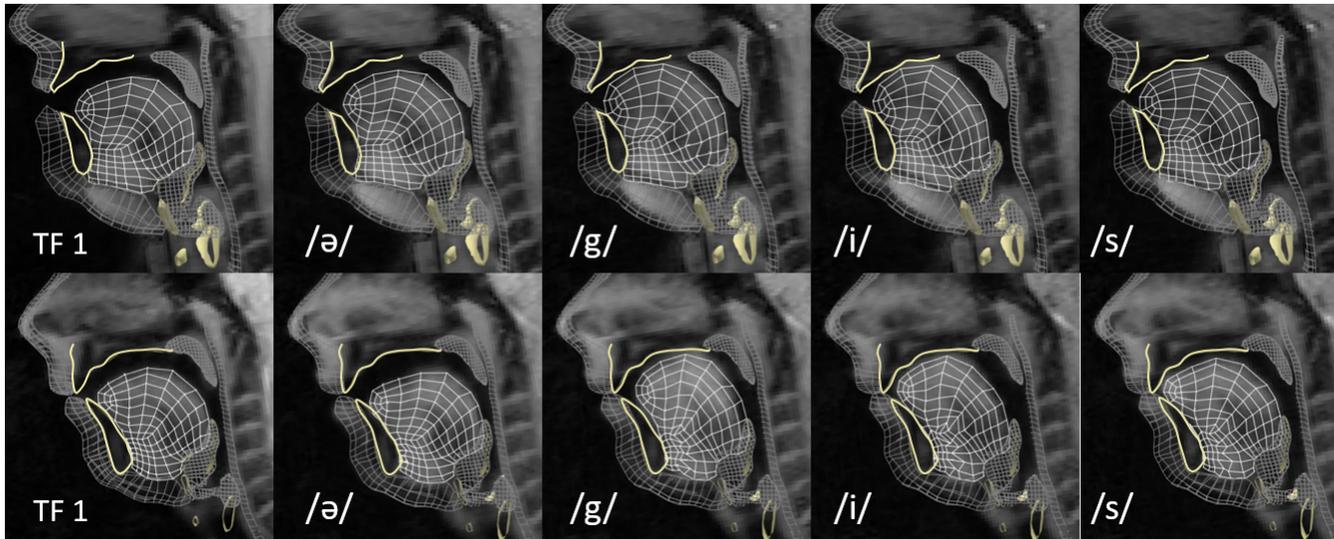


Figure 5.7: Simulation results for Speaker A (top row) v.s. Speaker B (bottom row) overlaid on the mid-sagittal slice of cine MRI. Meshes of the FE soft-tissues models are shown in white (tongue) and gray (others), while the bony structures are shown in cream.

5.3. Speech Simulation

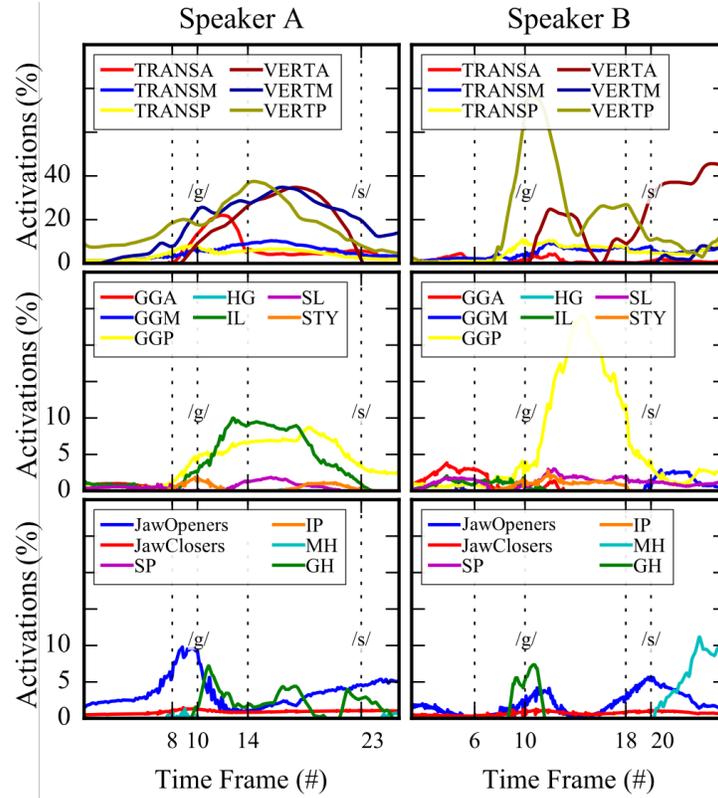


Figure 5.8: Estimated muscle activations

Speaker	ə	g	i	s
A	0.74±0.31	1.22±0.45	1.27±0.44	0.81±0.29
B	1.33±0.82	1.87±0.80	1.87±0.81	1.72±0.59

Table 5.5: Tongue tracking error (mm) over 40 control points.

Speaker	ə	g	i	s
A	0.89±0.37	1.24±0.32	0.94±0.15	1.32±0.14
B	1.23±0.13	1.89±0.25	1.62±0.39	2.04±0.84

Table 5.6: Jaw tracking error (mm) over 20 control points.

5.4 Discussion and Conclusion

This chapter demonstrates the potential of our proposed model registration technique for creating speaker-specific biomechanical models of upper-airway complex that facilitate personalized analysis of the speech production. Based on the first TF of the cine MR images, we created upper-airway models for two specific speakers. We enabled an inverse simulation to test the ability of our models with respect to an important speech movement, i.e. tongue-protrusion. As a result, the speaker-specific models are capable of protruding their tongue and, in the meanwhile, keeping their jaw at the neutral position.

Based on tagged and cine MRI of a simple speech motion, the speaker-specific models reproduced the speech motion of the utterance /ə-gis/. Our results suggest that the inverse solver tracks well and predicts functional movements effectively at the areas where tracking data exist: The tracking errors are in the range of MRI resolution. The inverse-estimated muscle activations make good agreement with the expectation of speech researchers. We speculate that several factors may have contributed to larger tracking error measured for Speaker B. First, Speaker B has a slightly larger tongue compared to Speaker A, with 103.17cm^2 vs. 95.61cm^2 volume. This means that on average each element in the FE model of Speaker B's tongue needs to account for deformations of a larger region of tissue. Second, Speaker B shows larger local deformations as well. For example, at the 10th TF (/g/), the average deviation of motion, measured based on displacement of our 40 control points, is about 5.84mm for Speaker B vs. 3.25mm for Speaker A. These factors suggest that Speaker B may need a higher FE tongue resolution to track the data more accurately. Finally, in this study the left and right muscles are activated together as one exciter, since the speech motion is believed to be bilateral. However, for example at the 10th TF (/g/) Speaker B shows an average of 1.82mm unilateral tongue motion, compared to just 0.15mm for Speaker A. We suspect that such large unilateral motion also contributed to the tracking error.

In this study, we track the motion of the tongue and lower jaw; the other FRANK components act as biomechanical regularizers in the simulations. To allow high-fidelity articulatory synthesis, the speech motion data for other articulators, such as soft palate or larynx, need to be incorporated into the speech simulation. However, currently, tagged MRI still suffers from noise and inaccuracy, especially near edges, due to blurring of the tag patterns caused by the HARP bandpass filter and tag fading [159]. Figure 5.9 shows an example of erroneous tracking that happens near the tongue surface. This limitation hampers the use of this technique for tracking small

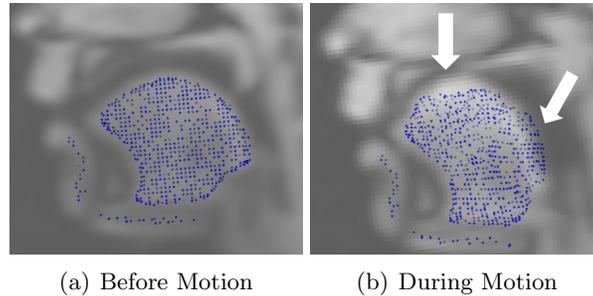


Figure 5.9: Illustration of erroneous tracking of tissue points near tongue surface using tagged MRI. The blue points represent the tissue points computed from the tagged MR data set (in the mid-sagittal slice), which is synchronized with the cine MRI shown in the figure. As demonstrated by the arrows, the computed tissue points fail to track the motions of tongue blade and tongue root.

and thin organs, such as soft palate or lips.

To summarize, this study demonstrates the potential of our proposed model registration technique for creating speaker-specific biomechanical models of upper-airway complex that facilitate personalized analysis of the speech production. Our method regularizes the functional movement using the morphic generic geometry and muscle attachments for areas where motion tracking data is missing to accommodate multi-modal datasets. Our results suggest the mesh resolution of FE models plays an important role in determining how detailed the tissue motions can be accurately tracked. Therefore, in order to balance the computational cost and tracking accuracy, the subject-specific methods need to tailor the mesh resolution of the resulting model to the target anatomical or kinematic details. Finally, we expect to incorporate complete motion data of the upper-airway system in our simulation. Acquiring accurate tissue motion from medical data remains to be an important subject of ongoing research.

Chapter 6

Conclusion

6.1 Summary

Subject-specific modelling enables the transition from generic understanding of biomechanical phenomena to addressing biomechanics of a particular individual. The goal of this thesis is to demonstrate the feasibility of using *in-vivo* medical images to create comprehensive upper-airway models for particular subjects and to simulate coupled upper-airway behaviours, such as swallowing and speech. Creating a biomechanical model of upper-airway complex relies heavily on expert interaction; the slow process of model creation prevents us from simulating a large number of individual cases. To ease the modelling efforts, we explore the use of model-registration techniques that register a predefined comprehensive upper-airway model (FRANK) to subject-specific medical data. The contributions of this thesis are summarized below:

Identified a model-registration strategy for creating comprehensive subject-specific models of upper-airway complex. Model-registration based subject-specific modelling methods should maintain the numerical stability and accuracy of the comprehensive biomechanical template, and should approximate the motor control behaviours of the target subject. To resolve the ambiguity and sparsity of the subject data, we regularize the registration by minimizing the morphological deviation from the template. This method targets at preserving three types of regularity. **1. inter-component regularity:** Maintain the spatial relationship (including connectivity, topology, relative postures and size) between model components. **2. intra-component regularity:** Preserve the underlying discretization structures of every subcomponents. **3. functional regularity:** Keep the functional information (including coupling attachments between components, muscle attachments and biomechanical properties) similar to the template but relevant to the new model geometry.

Developed a multi-structure registration technique, which can preserve both the inter- and intra-component regularity. This registration technique generates a homeomorphic mapping function, which superimposes a collection of template meshes (with different types of geometric discretization) onto certain target dataset (surfaces or point cloud), while minimizing the spatial distortion. For each individual mesh component, the mapping function is a similarity transformation, i.e. the local transformations are combinations of scaling and rotation; hence the generated deformation field preserves the underlying discretization structures of the template components.

Developed a model-registration workflow for generation of subject-specific upper-airway complex models. This workflow is based on our proposed multi-structure registration technique. According to correspondences established between a template model and subject data, we transfer a functional model from the generic space into a specific subject space by enforcing the correspondence constraints and minimizing the morphological deviation from the template. A penalty-based element-quality controlling method is applied to prevent poor-quality elements (in the template meshes) from further degrading during registration. To preserve the functional regularity, all the functional information, including muscles, coupling attachments between components and biomechanical properties, is updated to stay relevant to the registered subject-specific meshes.

Demonstrated the feasibility of our template-based subject-specific modelling methods by creating two comprehensive upper-airway models for the particular subjects, and demonstrated the model functionality in a set of biomechanical simulations Based on geometries extracted from two medical image volumes, we created two subject-specific models of upper-airway complex. To demonstrate the functionality of our models, we performed activation-driven simulations to model speech production on three vowels /**a**/, /**i**/ and /**u**/. Our models can maintain their stability during the simulation and properly respond to the muscle activations and external loadings (such as gravity). The kinematics obtained from the forward simulations of /**a**/ were used to drive the inverse simulations. Compared with the template model, our models showed comparable tracking performance (smaller tracking errors) and yielded similar muscle-activation predictions.

Demonstrated the potential of the proposed subject-specific modelling methods for personalized analysis of swallowing biomechanics. The sequential segmentation of the moving upper-airway, jaw and hyoid bones were used to drive the inverse simulation of normal swallowing using a subject-specific model. The model simulated the swallowing motion from the outset of the oral transit phase to the early stage of the esophagus phase. The model demonstrated the ability to reproduce the tissue motion happening in oral cavity, oropharynx and velopharynx. However, it showed limited ability to reproduce the hyolaryngeal motion during pharyngeal phase.

Demonstrated the potential of the proposed subject-specific modelling methods for personalized analysis of speech production. Using the proposed model-registration methods, we created two speaker-specific models. We performed inverse simulations to test the tongue-protrusion ability of our models. The result shows the speaker-specific models are capable of protruding their tongue and, in the meanwhile, keeping their jaw at the neutral position. Then, we enabled personalized speech simulations of the utterance /ə-gis/. The models reproduced the speech motion of the tongue and jaw, based on speaker-specific tagged and cine MRI data, with sub-voxel tracking error; the other upper-airway components acted as functional regularizers to move with the tongue and jaw in the simulations. The inverse-estimated muscle activations made good agreement with the speech-expert knowledge; these results suggest our methods may facilitate the investigation of the speech motor-control mechanisms.

6.2 Future Directions

Subject-specific biomechanical models bridge the gap between the human knowledge and the clinical recordings and measurements of a specific individual. This thesis presents a modelling and simulation framework, which facilitates personalized analysis of the coupled upper-airway system. This framework is illustrated in Figure 6.1.

6.2. Future Directions

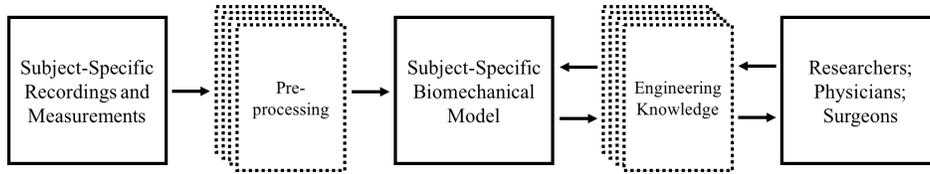


Figure 6.1: Current subject-specific modelling framework. The subject data, the biomechanical models and the human knowledge are loosely coupled.

Our methods are based on a comprehensive anatomical template, which is created through continuous collaboration of a team of interdisciplinary researchers. Based on the structural and dynamic information extracted from medical data, we map the comprehensive template from the generic space to the subject domain and then perform functional analysis on the corresponding human behaviours. However, this framework requires pre-processing on the subject data, which can be time-consuming, and may bring in additional artifacts into the results. Moreover, currently, interaction with biomechanical model still requires considerable engineering knowledge, which hinders not only the improvement of the model but also its applications in research and biomedicine.

We expect to continually narrow the gap between the subject data, the biomechanical models, and human knowledge. As shown in Figure 6.2, on the one hand, future work should minimize the pre-processing efforts on the subject data; biomechanical models should act as functional anatomical regularizers to reduce the ambiguity and assist interpretation of medical recordings and measurements; in turn, the medical data can provide the information to calibrate the biomechanical models and drive clinically-relevant analysis. On the other hand, the subject-specific modelling platform should allow experts, who do not have engineering background, to input the human knowledge into the models, and also to retrieve outcomes from them.

We will continue building the subjects-specific modelling platform that would assist accelerating the iterations in Figure 6.2, and ensures the human knowledge, the anatomical models and the medical recordings converging to the reality. In the following paragraphs, we would like to highlight a few directions in detail.

6.2. Future Directions

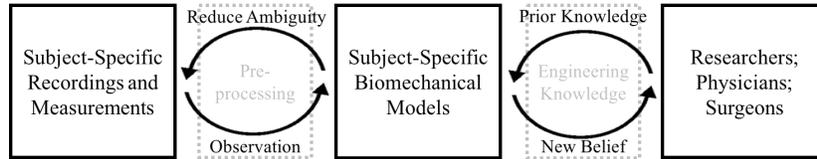


Figure 6.2: Target subject-specific modelling framework. The subject data, the biomechanical models and the human knowledge inform each other.

Improve Geometric Representation of the Biomechanical Template The template model (FRANK) has three limitations in terms of its geometric representation. First, since hexahedron mesh can void volumetric locking and has relatively low computational cost, all the deformable components in FRANK are represented as hexahedron or hexahedron-dominant FE models. However, generation of good-quality hexahedron mesh for a given geometry is still an ill-posed problem. Due to the complexity of organ geometry, the template model contains poor mesh-quality FE components. Second, the rigid body representation is over simplistic for modelling cartilages. This representation fails to model certain important behaviours, such as epiglottic inversion. Third, there is no systematic method available that allows the functional resolution (e.g. mesh resolution of the FE models, DoF of the bony structures) of the template model to be tailored to the specific applications and the data resolutions.

Future work is needed to find alternative geometric representations for deformable organs. For example, ANP-based tetrahedron FE model can be an alternative representation for organs that have complex geometry; cartilages can be modelled as thin-shell structures, which is flexible and has relatively low computational cost. In addition, systematic methods should be explored to adjust the functional resolution of the biomechanical models based on the needs of the users; such improvement would allow the models to be used in a wide range of application scenarios.

Statistical Atlas In order to handle the missing data, we regularize the registration in the way of minimizing the morphological deviation from the template. This regularization can effectively avoid undesirable distortions, but it may not be anatomically relevant. Our subject-specific modelling methods can benefit from replacing the deterministic template with a statistical atlas. A statistical atlas can not only represent the average information of human anatomy but also describe the inter-subject variations. This

anatomical prior knowledge can effectively reduce the morphological uncertainty and make anatomically-relevant interpretation of the subject data. Thus, it would increase our confidence on the registration results when the subject organ geometry is partially unknown in the medical data. As an example, Wang et al.[152] propose a statistical atlas based subject-specific FE modelling framework. Future work should explore a more generalized framework to accommodate different geometric representations.

Improve the Model Registration Technique First, the proposed structure-preserving FFD can be further improved by adjusting the resolution of the deformation grid according to the spatial distribution of the registration error. Such improvement can significantly reduce computational cost and increase the registration accuracy.

Second, although our method is able to maintain the overall mesh quality of the template, it does not guarantee minimum quality for elements. Mesh-Match-and-Repair (MMRep) [18] uses a post-registration mesh-repair method to improve the quality of the bad elements to the minimum value (0.03 JR); however, this minimal value is insufficient for maintaining the numerical stability and accuracy of FE models when large deformation involves in biomechanical simulations; moreover, this method can fail to untangle the mesh when excessive distortion is generated during mesh-matching process. In future work, we expect to blend our method into the MMRep framework, which would maintain the overall mesh quality of the template and also guarantee a minimum quality for individual elements. However, an extra care needs to be taken for preserving the FE-FE attachments. One solution is to replace the local Gauss-Seidel mesh-repair method with appropriate global mesh optimization methods [92] that allow displacement constraints at the attachment areas.

Third, it is useful to replace the points-based correspondence with certain intensity-based metric. Establishment of correspondence in the image space can help with removing the additional pre-processing steps (as shown in Figure 6.1 and Figure 6.2), such as segmentation, from the registration workflow.

Fourth, we assume the template and the recorded subject are in the same pose. However, this assumption may not be true in certain cases. Particularly, when articulated structures (such as cervical spine) involves in the registration, the results can be sensitive to pose difference between the model and the subject. Moreover, our assumption also ignores the muscle contraction of the subject. Future work should couple our registration

method with certain articulated registration techniques. An important question for future studies is to determine the initial muscle-contraction levels of the subject-specific models.

Pre-stressing Configuration Currently, we assume that the subject organ geometries extracted from the medical images are in stress-free configurations. Such assumption does not hold in most *in vivo* conditions, due to the external forces (such as gravity) and tetanic contractions. Future study is needed to determine the pre-stressing reference configurations of subject-specific models. A good approximation of the pre-stressing configuration can avoid bias of the modelled biomechanics. Vavouraki et al. [151] propose an inverse analysis method to derive the pressure-free configuration of human aortas, and the gravity-free shape of the female breast. However, their studies only focus on isolated FE models. Further study is needed to determine the unloading reference configuration of hybrid models (such as FRANK).

Subject-Specific Muscle Configuration In this study, we assume the subject has similar muscular configurations and forces with the template. To obtain more accurate motor control patterns of the specific subjects, we expect to extract personalized muscular information (e.g. the fibre paths, pennation angles and PCSA) from *in-vivo* imaging techniques, such as Diffusion Tensor MRI (DTI).

Subject-Specific Tissue Properties Our methods do not personalize the elastic properties of soft-tissues. Tissue properties may influence the muscular coordinating patterns of a specific physiological task and are needed for reliable subject-specific modelling. *In vivo* measurement tissue properties is still an ill-posed problem. Elastography provides a non-invasive way for such measurement. For example, Cheng et al. [26] apply magnetic resonance elastography to measuring the viscoelastic properties of tongue and soft palate. However, as soft tissues exhibit non-linear stress-strain behaviour, tissue properties at larger deformations cannot be inferred from elastography alone [155]. Future work should further the investigation of *in vivo* tissue-property measurement.

Neural Control Currently, our knowledge of the upper-airway system – the relationship between the motor commands the upper-airway sub-structures receiving and the activity they performing – is still limited. The

6.3. Concluding Remarks

muscular structures and the organization of motor control are both essential for determining the range of the motion that is anatomically and physiologically feasible. Unfortunately, due to the anatomical complexity and ethical reasons, careful anatomical data about the compartmental organization of human upper-airway system are not yet available. Nevertheless, with the advancement of *in-vivo* imaging techniques, increasing computational power, and further improvement of the model fidelity, inverse modelling will eventually allow inference to the organization of motor control. Future study will continue increasing the realism of the neurophysiological control mechanism.

Multi-Physics Simulation Human behaviours, such as swallowing, breathing and speech production, involve multiple physical phenomena. Comprehensive analysis of these behaviours needs coupling of different computational simulation methods to allow integration of solid mechanics, hydrodynamics, aerodynamics, etc. Currently, most biomechanical studies focus on structural analysis of human bodies; multi-physics simulation methods attract much less attention and research effort, which limits the application of subject-specific biomechanical models. Future study should explore efficient and reliable multi-physics simulation methods.

Human-Model Interaction Currently, interaction with biomechanical models requires considerable engineering knowledge and programming efforts. It would be beneficial to explore interaction methods that can ease the human interventions in modelling and simulation processes.

6.3 Concluding Remarks

To conclude, this thesis has presented new modelling and simulation techniques that enable transition of the upper-airway research from generic understanding of isolated biomechanical phenomena to analyzing comprehensive biomechanics of a specific individual. We applied these techniques to personalized analysis of complex upper-airway activities, including swallowing and speech production. This work provides a starting point for developing an accurate, efficient, interactive, subject-specific modelling platform, which will eventually assist diagnosis and treatment of a wide range of upper-airway dysfunctions.

Bibliography

- [1] Hidalgo D. A and Pusic A. L. Free-flap mandibular reconstruction: a 10-year follow-up study. *Plastic and Reconstructive Surgery*, 110(2):438–49; discussion 450, 2002.
- [2] Hossam Abdelmunim and Aly A. Farag. Elastic shape registration using an incremental free form deformation approach with the icp algorithm. In *2011 18th IEEE International Conference on Image Processing*, pages 729–732, 2011.
- [3] Dicko Ali-Hamadi, Tiantian Liu, Benjamin Gilles, Ladislav Kavan, Francois Faure, Olivier Palombi, and Marie-Paule Cani. Anatomy transfer. *ACM Transactions on Graphics (TOG)*, 32(6):1–8, 2013.
- [4] Brett Allen, Brian Curless, and Zoran Popovi. Articulated body deformation from range scan data. pages 612–619. ACM, 2002.
- [5] Brett Allen, Brian Curless, and Zoran Popovi. The space of human body shapes: reconstruction and parameterization from range scans. *ACM Transactions on Graphics*, 22(3):587, 2003.
- [6] B. Amberg, S. Romdhani, and T. Vetter. Optimal step nonrigid icp algorithms for surface registration. pages 1–8. IEEE, 2007.
- [7] Peter Anderson, Negar M. Harandi, Scott Moisiuk, Ian Stavness, and Sidney Fels. A comprehensive 3d biomechanically-driven vocal tract model including inverse dynamics for speech research. In *INTER-SPEECH*, 2015.
- [8] Dragomir Anguelov, Praveen Srinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers, and James Davis. Scape: shape completion and animation of people. *ACM Transactions on Graphics (TOG)*, 24(3):408–416, 2005.
- [9] K. S. Arun, T. S. Huang, and S. D. Blostein. Least-squares fitting of two 3-d point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(5):698–700, 1987.

- [10] A. Aubel and D. Thalmann. Interactive modeling of the human musculature. pages 167–255. IEEE, 2001.
- [11] Thomas Baer, J. Peter Alfonso, and Kiyoshi Honda. Electromyography of the tongue muscles during vowels in /apvp/ environment. *Ann Bull RILP*, 22:7–19, 1988.
- [12] P. J. Besl and H. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.
- [13] P. Birkholz. Modeling consonant-vowel coarticulation for articulatory speech synthesis. *PLOS ONE*, 8(4):e60603, 2013.
- [14] J. Bonet and A. J. Burton. A simple average nodal pressure tetrahedral element for incompressible and nearly incompressible dynamic explicit applications. *Communications in Numerical Methods in Engineering*, 14(5):437–449, 1998.
- [15] Alexander M. Bronstein, Michael M. Bronstein, and Ron Kimmel. Topology-invariant similarity of nonrigid shapes. *International Journal of Computer Vision*, 81(3):281–301, 2009.
- [16] Stphanie Buchaillard, Muriel Brix, Pascal Perrier, and Yohan Payan. Simulations of the consequences of tongue surgery on tongue mobility: implications for speech production in postsurgery conditions. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 3(3):252–261, 2007.
- [17] Stphanie Buchaillard, Pascal Perrier, and Yohan Payan. A biomechanical model of cardinal vowel production: Muscle activations and the impact of gravity on tongue positioning. *The Journal of the Acoustical Society of America*, 126(4):2033–2051, 2009.
- [18] Marek Bucki, Claudio Lobos, and Yohan Payan. A fast and robust patient specific finite element mesh registration technique: Application to 60 clinical cases. *Medical Image Analysis*, 14(3):303–317, 2010.
- [19] Marek Bucki, Claudio Lobos, Yohan Payan, and Nancy Hirschfeld. Jacobian-based repair method for finite element meshes after registration. *Engineering with Computers*, 27(3):285–297, 2011.

- [20] Cynthia Burke, Karen Patrias, National Library of Medicine (U.S.), Reference, and Web Services Section. Visible human project: January 1987 through march 2007 : 912 citations. Technical Report 2007-1, U.S. Dept. of Health and Human Services, Public Health Service, National Library of Medicine, Reference and Web Services Section, 2007.
- [21] TA Burkhart, DM Andrews, and CE Dunning. Finite element modeling mesh quality, energy balance and validation methods: A review with recommendations associated with the modeling of bone tissue. *JOURNAL OF BIOMECHANICS*, 46(9):1477–1488, 2013.
- [22] JQ Campbell and AJ Petrella. Automated finite element modeling of the lumbar spine: Using a statistical shape model to generate a virtual population of models. *JOURNAL OF BIOMECHANICS*, 49(13):2593–2599, 2016.
- [23] Matthieu Chabanas, Vincent Luboz, and Yohan Payan. Patient specific finite element model of the face soft tissues for computer-assisted maxillofacial surgery. *Medical Image Analysis*, 7(2):131–151, 2003.
- [24] Hui Chen, Sidney Fels, Tricia Pang, Ling Tsou, Fernanda R. de Almeida, and Alan A. Lowe. Three-dimensional reconstruction of soft palate modeling from subject-specific magnetic resonance imaging data. *Sleep and Breathing*, 16(4):1113–1119, 2012.
- [25] Mingli Chen, Weiguo Lu, Quan Chen, Kenneth J. Ruchala, and Gustavo H. Olivera. A simple fixed-point approach to invert a deformation field. *Medical Physics*, 35(1):81–88, 2008.
- [26] S. Cheng, S.C. Gandevia, M. Green, R. Sinkus, and L.E. Bilston. Viscoelastic properties of the tongue and soft palate using {MR} elastography. *Journal of Biomechanics*, 44(3):450 – 454, 2011.
- [27] Yongchoel Choi and Seungyong Lee. Injectivity conditions of 2d and 3d uniform cubic b-spline functions. *Graphical Models*, 62(6):411–427, 2000.
- [28] Haili Chui and Anand Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2):114–141, 2003.

- [29] S. Corazza, E. Gambaretto, L. Mundermann, and T. P. Andriacchi. Automatic generation of a subject-specific model for accurate markerless motion capture and biomechanical applications. *IEEE Transactions on Biomedical Engineering*, 57(4):806–812, 2010.
- [30] Batrice Couteau, Yohan Payan, and Stphane Lavalle. The mesh-matching algorithm: an automatic 3d mesh generator for finite element structures. *Journal of Biomechanics*, 33(8):1005–1009, 2000.
- [31] Francis A. Duck. *Physical Properties of Tissue: A Comprehensive Reference Book*. Academic Press, 1990.
- [32] Philip J. Edwards, Derek L. G. Hill, John A. Little, and David J. Hawkes. A three-component deformation model for image-guided surgery. *Medical Image Analysis*, 2(4):355–367, 1998.
- [33] Cumhuri Ertekin, Gaye Eryaar, Nevin Grgr, ehnaz Arc, Yaprak Secil, and Tlay Kurt. Orbicularis oculi muscle activation during swallowing in humans. *Experimental Brain Research*, 224(1):79–91, 2013.
- [34] Sidney Fels and Bryan Gick. Interdisciplinary approaches for advancing articulatory speech theory and synthesis. *Canadian Acoustics*, 44(3), 2016.
- [35] Cormac Flynn, Ian Stavness, John Lloyd, and Sidney Fels. A finite element model of the face including an orthotropic skin model under in vivo tension. *Computer Methods in Biomechanics and Biomedical Engineering*, 18(6):571, 2015.
- [36] LA Freitag and P. Plassmann. Local optimization-based simplicial mesh untangling and improvement. *INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING*, 49(1-2):109–125, 2000.
- [37] Naoko Fujii, Yoko Inamoto, Eiichi Saitoh, Mikoto Baba, Sumiko Okada, Satoshi Yoshioka, Toshiaki Nakai, Yoshihiro Ida, Kazuhiro Katada, and Jeffrey B. Palmer. Evaluation of swallowing using 320-detector-row multislice ct. part i: Single- and multiphase volume scanning for three-dimensional morphological and kinematic analysis. *Dysphagia*, 26(2):99–107, 2010.
- [38] Satoru Fujita, Jianwu Dang, Noriko Suzuki, and Kiyoshi Honda. A computational tongue model and its clinical application. *Oral Science International*, 4(2):97–109, 2007.

- [39] Terry A. Gaige, Thomas Benner, Ruopeng Wang, Van J. Wedeen, and Richard J. Gilbert. Three dimensional myoarchitecture of the human tongue determined in vivo by diffusion tensor imaging with tractography. *Journal of Magnetic Resonance Imaging*, 26(3):654–661, 2007.
- [40] Amit Gefen. *Patient-specific modeling in tomorrow's medicine*, volume 9. Springer, New York;Heidelberg;, 2012 edition, 2012.
- [41] Michle Gentil and Thomas Gay. Neuromuscular specialization of the mandibular motor system: Speech versus non-speech movements. *Speech Communication*, 5(1):69 – 82, 1986.
- [42] Jean-Michel Gerard, Pascal Perrier, and Yohan Payan. *3D biomechanical tongue modeling to study speech production*. Speech Production: Models, Phonetic Processes, and Techniques, Psychology Press, 2006.
- [43] Jean-Michel Gerard, Reiner Wilhelms-Tricarico, Pascal Perrier, and Yohan Payan. A 3d dynamical biomechanical tongue model to study speech motor control. *Research Developments in Biomechanics*, 1:49–64, 2003.
- [44] J.M. Gerard, J. Ohayon, V. Luboz, P. Perrier, and Y. Payan. Non-linear elastic properties of the lingual and facial tissues assessed by indentation technique: Application to the biomechanics of speech production. *Medical Engineering & Physics*, 27(10):884 – 892, 2005. Advances in the finite element modelling of soft tissue deformation.
- [45] Bryan Gick, Peter Anderson, Hui Chen, Chenhao Chiu, Ho B. Kwon, Ian Stavness, Ling Tsou, and Sidney Fels. Speech function of the oropharyngeal isthmus: a modelling study. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2(4):217–222, 2014.
- [46] Bryan Gick, Ian Wilson, and Donald Derrick. *Articulatory phonetics*. Wiley-Blackwell, Malden, MA;Chichester, West Sussex;, 2013.
- [47] B. Gilles, L. Reveret, and DK Pai. Creating and animating subject-specific anatomical models. *COMPUTER GRAPHICS FORUM*, 29(8):2340–2351, 2010.
- [48] Benjamin Gilles, Laurent Moccozet, and Nadia Magnenat-Thalmann. *Anatomical Modelling of the Musculoskeletal System from MRI*, pages 289–296. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

- [49] Evgeny Gladilin and Alexander Ivanov. Computational modelling and optimisation of soft tissue outcome in cranio-maxillofacial surgery planning. *Computer Methods in Biomechanics and Biomedical Engineering*, 12(3):305–318, 2009.
- [50] Leslie Glupker, Katherine Kula, Edwin Parks, William Babler, Kelton Stewart, and Ahmed Ghoneima. Three-dimensional computed tomography analysis of airway volume changes between open and closed jaw positions. *American journal of orthodontics and dentofacial orthopedics : official publication of the American Association of Orthodontists, its constituent societies, and the American Board of Orthodontics*, 147(4):426, 2015.
- [51] S. Gold and A. Rangarajan. A graduated assignment algorithm for graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):377–388, 1996.
- [52] Manuel Gonzalez Hidalgo, Arnau M. Torres, Javier Varona Gmez, and SpringerLink ebooks Engineering. *Deformation models: tracking, animation, and applications*, volume 7;7.:. Springer, Dordrecht;New York;, 2013.
- [53] Dan Grauer, Lucia S. H. Cevidanes, Martin A. Styner, James L. Ackerman, and William R. Proffit. Pharyngeal airway volume and shape from cone-beam computed tomography: Relationship to facial morphology. *American Journal of Orthodontics & Dentofacial Orthopedics*, 136(6):805–814, 2009.
- [54] Nicole M. Grosland, Ritesh Bafna, and Vincent A. Magnotta. Automated hexahedral meshing of anatomic structures using deformable registration. *Computer Methods in Biomechanics and Biomedical Engineering*, 12(1):35–43, 2009.
- [55] A. G. Hannam, I. Stavness, J. E. Lloyd, and S. Fels. A dynamic model of jaw and hyoid biomechanics during chewing. *Journal of Biomechanics*, 41(5):1069–1076, 2008.
- [56] Alan G. Hannam, Ian K. Stavness, John E. Lloyd, S. S. Fels, Arthur J. Miller, and Donald A. Curtis. A comparison of simulated jaw dynamics in models of segmental mandibular resection versus resection with alloplastic reconstruction. *The Journal of Prosthetic Dentistry*, 104(3):191–198, 2010.

- [57] N. M. Harandi, J. Woo, M. R. Farazi, L. Stavness, M. Stone, S. Fels, and R. Abugharbieh. Subject-specific biomechanical modelling of the oropharynx with application to speech production. pages 1389–1392. IEEE, 2015.
- [58] Negar Harandi Mohaghegh. 3d subject-specific biomechanical modeling and simulation of the oral region and airway with application to speech production, 2016.
- [59] Christian Head, Daniel Alam, Joel A. Sercarz, Jivianne T. Lee, Jeffrey D. Rawnsley, Gerald S. Berke, and Keith E. Blackwell. Microvascular flap reconstruction of the mandible: a comparison of bone grafts and bridging plates for restoration of mandibular continuity. *Otolaryngology - Head and Neck Surgery*, 129(1):48–54, 2003.
- [60] AM A. Heemskerk, GJ G. Strijkers, A. A. Vilanova, MR M. st, and K. K. Nicolaij. Determination of mouse skeletal muscle architecture using three dimensional diffusion tensor imaging. *Magnetic Resonance in Medicine*, 53(6):1333–1340, 2005.
- [61] A. V. Hill. The mechanics of active muscle. *Proceedings of the Royal Society of London B: Biological Sciences*, 141(902):104–117, 1953.
- [62] Andrew Ho, Mark Nicosia, Angela Dietsch, William Pearson, Jana Rieger, Nancy Solomon, Maureen Stone, Yoko Inamoto, Eiichi Saitoh, Sheldon Green, and Sidney Fels. 3d dynamic visualization of swallowing from multi-slice computed tomography. pages 1–1. ACM, 2014.
- [63] Andrew K. Ho, Ling Tsou, Sheldon Green, and Sidney Fels. A 3d swallowing simulation using smoothed particle hydrodynamics. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2(4):237–244, 2014.
- [64] QX Huang, B. Adams, M. Wicke, and LJ Guibas. Non-rigid registration under isometric deformations. *COMPUTER GRAPHICS FORUM*, 27(5):1449–1457, 2008.
- [65] Thomas J. R. Hughes. *The finite element method: linear static and dynamic finite element analysis*. Dover Publications, Mineola, NY, 2000.
- [66] Y. Inamoto, E. Saitoh, S. Okada, H. Kagaya, S. Shibata, M. Baba, K. Onogi, S. Hashimoto, K. Katada, P. Wattanapan, and J. B. Palmer.

- Anatomy of the larynx and pharynx: effects of age, gender and height revealed by multidetector computed tomography. *Journal of Oral Rehabilitation*, 42(9):670–677, 2015.
- [67] Yoko Inamoto, Naoko Fujii, Eiichi Saitoh, Mikoto Baba, Sumiko Okada, Kazuhiro Katada, Yasunori Ozeki, Daisuke Kanamori, and Jeffrey B. Palmer. Evaluation of swallowing using 320-detector-row multislice ct. part ii: Kinematic analysis of laryngeal closure during normal swallowing. *Dysphagia*, 26(3):209–217, 2011.
- [68] Haruhi Inokuchi, Marls Gonzlez-Fernndez, Koichiro Matsuo, Martin B. Brodsky, Mitsumasa Yoda, Hiroshige Taniguchi, Hideto Okazaki, Takashi Hiraoka, and Jeffrey B. Palmer. Electromyography of swallowing with fine wire intramuscular electrodes in healthy human: Amplitude difference of selected hyoid muscles. *Dysphagia*, 31(1):33–40, 2015.
- [69] Bing Jian and B. C. Vemuri. Robust point set registration using gaussian mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1633–1645, 2011.
- [70] Barry Joe. Shape measures for quadrilaterals, pyramids, wedges, and hexahedra. Technical report, Zhou Computing Services Inc., 2008.
- [71] Grand R. Joldes, Adam Wittek, and Karol Miller. Non-locking tetrahedral finite element for surgical simulation. *Communications in Numerical Methods in Engineering*, 25(7):827–836, 2009.
- [72] Takahiro Kikuchi, Yukihiro Michiwaki, Tetsu Kamiya, Yoshio Toyama, Tasuku Tamai, and Seiichi Koshizuka. Human swallowing simulation based on videofluorography images using hamiltonian mps method. *Computational Particle Mechanics*, 2(3):247–260, 2015.
- [73] Yoon-Chul Kim, R. M. Lebel, Ziyue Wu, Sally L. D. Ward, Michael C. K. Khoo, and Krishna S. Nayak. Real-time 3d magnetic resonance imaging of the pharyngeal airway in sleep apnea: Real-time 3d mri of the pharyngeal airway. *Magnetic Resonance in Medicine*, 71(4):1501–1510, 2014.
- [74] YoonChul Kim, Cecil E. Hayes, Shrikanth S. Narayanan, and Krishna S. Nayak. Novel 16channel receive coil array for accelerated upper airway mri at 3 tesla. *Magnetic Resonance in Medicine*, 65(6):1711–1717, 2011.

- [75] Patrick M. Knupp. On the invertibility of the isoparametric map. *Computer Methods in Applied Mechanics and Engineering*, 78(3):313–329, 1990.
- [76] Patrick M Knupp. Hexahedral and tetrahedral mesh untangling. *Engineering with Computers*, 17(3):261–268, 2001.
- [77] PM Knupp. Achieving finite element mesh quality via optimization of the jacobian matrix norm and associated quantities. part ii - a framework for volume mesh optimization and the condition number of the jacobian matrix. *INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING*, 48(8):1165–1185, 2000.
- [78] Cornelius Lanczos. *The variational principles of mechanics*. Dover Publications, 1986.
- [79] David J. Larkman and Rita G. Nunes. Parallel magnetic resonance imaging. *Physics in Medicine and Biology*, 52(7):R15–R55, 2007.
- [80] Sung-Hee Lee, Eftychios Sifakis, and Demetri Terzopoulos. Comprehensive biomechanical modeling and simulation of the upper body. *ACM Transactions on Graphics (TOG)*, 28(4):1–17, 2009.
- [81] Z. Levi and C. Gotsman. D-snake: Image registration by as-similar-as-possible template deformation. *IEEE Transactions on Visualization and Computer Graphics*, 19(2):331–343, 2013.
- [82] J. P. Lewis, Matt Cordner, and Nickson Fong. Pose space deformation: A unified approach to shape interpolation and skeleton-driven deformation. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '00, pages 165–172, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.
- [83] Hao Li, Robert W. Sumner, and Mark Pauly. Global correspondence optimization for non-rigid registration of depth scans. *Computer Graphics Forum*, 27(5):1421–1430, 2008.
- [84] Zhi Li, Jeremy P. M. Mogk, Dongwoon Lee, Jacobo Bibliowicz, and Anne M. Agur. Development of an architecturally comprehensive database of forearm flexors and extensors from a single cadaveric specimen. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 3(1):3–12, 2015.

- [85] Miao Liao, Qing Zhang, Huamin Wang, Ruigang Yang, and Minglun Gong. Modeling deformable objects from a single depth camera. pages 167–174, 2009.
- [86] Sajan G. Lingala, Brad P. Sutton, Marc E. Miquel, and Krishna S. Nayak. Recommendations for realtime speech mri. *Journal of Magnetic Resonance Imaging*, 43(1):28–44, 2016.
- [87] Y. Lipman and T. Funkhouser. Mobius voting for surface correspondence. *ACM TRANSACTIONS ON GRAPHICS*, 28(3):1, 2009.
- [88] A. Liu and B. Joe. Relationship between tetrahedron shape measures. *BIT*, 34(2):268–287, 1994.
- [89] Anwei Liu and Barry Joe. On the shape of tetrahedra from bisection. *Mathematics of Computation*, 63(207):141–154, 1994.
- [90] Yonghuai Liu. Automatic 3d free form shape matching using the graduated assignment algorithm. *Pattern Recognition*, 38(10):1615–1631, 2005.
- [91] Yonghuai Liu. A mean field annealing approach to accurate free form shape matching. *Pattern Recognition*, 40(9):2418–2436, 2007.
- [92] Marco Livesu, Alla Sheffer, Nicholas Vining, and Marco Tarini. Practical hex-mesh optimization via edge-cone rectification. *ACM Transactions on Graphics (TOG)*, 34(4):1–11, 2015.
- [93] John E Lloyd, Ian Stavness, and Sidney Fels. Artisynt: A fast interactive biomechanical modeling toolkit combining multibody and finite element simulation. In *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*, pages 355–394. Springer, 2012.
- [94] Jeri A. Logemann. *Evaluation and treatment of swallowing disorders*. College-Hill Press, San Diego, CA, 1983.
- [95] V. Luboz, Y. Payan, and B Couteau. Automatic 3d finite element mesh generation: an atlas fitted onto patient data. In *Proceedings of the Fifth International Symposium on Computer Methods in Biomechanics and Biomedical Engineering*, 2001.
- [96] Jorge C. Lucero and Kevin G. Munhall. A model of facial biomechanics for speech production. *The Journal of the Acoustical Society of America*, 106(5):2834–2842, 1999.

- [97] DONNA S. LUNDY, CHRISTINE SMITH, LAURA COLANGELO, PAULA A. SULLIVAN, JERILYN A. LOGEMANN, CATHY L. LAZARUS, LISA A. NEWMAN, TOM MURRY, LORI LOMBARD, and JOY GAZIANO. Aspiration: Cause and implications. *Otolaryngology - Head and Neck Surgery*, 120(4):474–478, 1999.
- [98] Peter F. Mac Neilage. Electromyographic study of the tongue during vowel production. *The Journal of the Acoustical Society of America*, 35(5):783, 1963.
- [99] Massimo R. Mannarino, Francesco Di Filippo, and Matteo Pirro. Obstructive sleep apnea syndrome. *European journal of internal medicine*, 23(7):586, 2012.
- [100] D. Mateus, R. Horaud, D. Knossow, F. Cuzzolin, and E. Boyer. Articulated shape matching using laplacian eigenfunctions and unsupervised point registration. pages 1–8. IEEE, 2008.
- [101] Martin J. McKeown, Dana C. Torpey, and Wendy C. Gehm. Non-invasive monitoring of functionally distinct muscle activations during swallowing. *Clinical Neurophysiology*, 113(3):354–366, 2002.
- [102] Biren Mehta and Stefan Schaal. Forward models in visuomotor control. *Journal of Neurophysiology*, 88(2):942–953, 2002.
- [103] M. Meyer, M. Desbrun, P. Schroder, and A.H. Barr. *Discrete Differential-Geometry Operators for Triangulated 2-Manifolds*. Visualization and Mathematics III, 2003.
- [104] Srboljub M. Mijailovich, Boban Stojanovic, Milos Kojic, Alvin Liang, Van J. Wedeen, and Richard J. Gilbert. Derivation of a finite-element model of lingual deformation during swallowing from the mechanics of mesoscale myofiber tracts obtained by mri. *Journal of Applied Physiology*, 109(5):1500–1514, 2010.
- [105] A. C. Miracle and S. K. Mukherji. Conebeam ct of the head and neck, part 1: Physical principles. *American Journal of Neuroradiology*, 30(6):1088, 2009.
- [106] H. Mizunuma, M. Sonomura, K. Shimokasa, H. Ogoshi, S. Nakamura, and N. Tayama. Numerical modelling and simulation on the swallowing of jelly. *Journal of Texture Studies*, 40(4):406–426, 2009.

- [107] Scott Moisik and Bryan Gick. The quantal larynx revisited. *The Journal of the Acoustical Society of America*, 133(5):3522, 2013.
- [108] W. Mollemans, F. Schutyser, N. Nadjmi, F. Maes, and P. Suetens. Predicting soft tissue deformations for a maxillofacial surgery planning system: From computational strategies to a complete clinical validation. *Medical Image Analysis*, 11(3):282–301, 2007.
- [109] Andriy Myronenko and Xubo Song. Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2262–2275, 2010.
- [110] Mohammad A. Nazari, Pascal Perrier, Matthieu Chabanas, and Yohan Payan. Simulation of dynamic orofacial movements using a constitutive law varying with muscle activation. *Computer Methods in Biomechanics and Biomedical Engineering*, 13(4):469–482, 2010.
- [111] N. F. Osman, E. R. McVeigh, and J. L. Prince. Imaging heart motion using harmonic phase mri. *IEEE Transactions on Medical Imaging*, 19(3):186–202, 2000.
- [112] C. Papazov and D. Burschka. Deformable 3d shape registration based on local similarity transforms. *Computer Graphics Forum*, 30(5):1493–1502, 2011.
- [113] William G. Pearson Jr, Susan E. Langmore, Louis B. Yu, and Ann C. Zumwalt. Structural analysis of muscles elevating the hyolaryngeal complex. *Dysphagia*, 27(4):445–451, 2012.
- [114] William G. Pearson Jr, Sonja M. Molfenter, Zachary M. Smith, and Catriona M. Steele. Image-based measurement of post-swallow residue: The normalized residue ratio scale. *Dysphagia*, 28(2):167–177, 2013.
- [115] C. C. Peck, G. E. J. Langenbach, and A. G. Hannam. Dynamic simulation of muscle and articular properties during human wide jaw opening. *Archives of Oral Biology*, 45(11):963–982, 2000.
- [116] Joseph S. Perkell. A physiologically-oriented model of tongue activity in speech production. Thesis (Ph. D.)Massachusetts Institute of Technology, Dept. of Electrical Engineering, 1974.

- [117] Pascal Perrier, Yohan Payan, Stphanie Buchaillard, Mohammad A. Nazari, and Matthieu Chabanas. Biomechanical models to study speech. *Faits de Langues*, 37:155–171, 2011.
- [118] Igor Sazonov Perumal Nithiarasu and Si Yong Yeo. *Scan-Based Flow Modelling in Human Upper Airways*, chapter 0, pages 241 – 280. Springer, 2012.
- [119] Guy J. Petruzzelli, Kelly Cunningham, and Darl Vandevender. Impact of mandibular condyle preservation on patterns of failure in head and neck cancer. *Otolaryngology - Head and Neck Surgery*, 137(5):717–721, 2007.
- [120] Les A. Piegl and Wayne Tiller. *The NURBS book*. Springer, New York;Berlin;, 2nd edition, 1997.
- [121] U. Pinkall and K. Polthier. Computing discrete minimal surfaces and their conjugates. *Exper Math*, 2(1):15–36, 1993.
- [122] Knupp P.M. *Proceedings of the 45th AIAA Aerospace Sciences Meeting and Exhibit*, 2007.
- [123] Thibaut Raguin, Jean Carvalho, Sophie Riehm, Catherine Takeda, and Agns Dupret-Bories. Method for dealing with severe aspiration using a new concept of intralaryngeal prosthesis: A case report: New concept of intralaryngeal prosthesis. *Head & Neck*, 38(10):E2504–E2507, 2016.
- [124] Dylan F. Roden and Kenneth W. Altman. Causes of dysphagia among different age groups: a systematic review of the literature. *Otolaryngologic clinics of North America*, 46(6):965, 2013.
- [125] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.
- [126] Daniel Rueckert, Paul Aljabar, Rolf A. Heckemann, Joseph V. Hajnal, and Alexander Hammers. *Diffeomorphic Registration Using B-Splines*, volume 4191, pages 702–709. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

- [127] H. Schliephake, R. Schmelzeisen, R. Schnweiler, T. Schneller, and C. Altenbernd. Speech, deglutition and life quality after intraoral tumour resection. a prospective study. *International journal of oral and maxillofacial surgery*, 27(2):99, 1998.
- [128] Julia A. Schnabel, Daniel Rueckert, Marcel Quist, Jane M. Blackall, Andy D. Castellano-Smith, Thomas Hartkens, Graeme P. Penney, Walter A. Hall, Haiying Liu, Charles L. Truwit, Frans A. Gerritsen, Derek L. G. Hill, and David J. Hawkes. *A Generic Framework for Non-rigid Registration Based on Non-uniform Multi-level Free-Form Deformations*, pages 573–581. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.
- [129] Thomas Sederberg and Scott Parry. Free-form deformation of solid geometric models. pages 151–160. ACM, 1986.
- [130] Stephanie M. Shaw and Rosemary Martino. The normal swallow: muscular and neurophysiological control. *Otolaryngologic clinics of North America*, 46(6):937, 2013.
- [131] Eftychios Sifakis, Andrew Selle, Avram Robinson-Mosher, and Ronald Fedkiw. Simulating speech with a physics-based facial muscle model. pages 261–270. Eurographics Association, 2006.
- [132] Karan Singh and Evangelos Kokkevis. Skinning characters using surface oriented free-form deformations. pages 35–42. Proceedings of the Graphics Interface 2000 Conference, 2000.
- [133] Y. Song, S. Wang, M. Chan, B. Chandra, A. Dhawan, and Y. Song. Femoral fracture risk assessment after intensity modulated radiation therapy (imrt) for the treatment of soft tissue sarcoma using a novel mathematical model. pages 95–98. IEEE, 2006.
- [134] M. Sonomura, H. Mizunuma, T. Numamor, H. Michiwaki, and K. Nishinari. Numerical simulation of the swallowing of liquid bolus. *Journal of Texture Studies*, 42(3):203–211, 2011.
- [135] Olga Sorkine and Marc Alexa. As-rigid-as-possible surface modeling. In *Proceedings of the Fifth Eurographics Symposium on Geometry Processing*, SGP '07, pages 109–116, Aire-la-Ville, Switzerland, Switzerland, 2007. Eurographics Association.

- [136] Ian Stavness, Sidney Fels, John E. Lloyd, and Alan G. Hannam. Predicting muscle patterns for hemimandibulectomy models. *Computer methods in biomechanics and biomedical engineering*, 13(4):483–491, 2010.
- [137] Ian Stavness, John E. Lloyd, Yohan Payan, and Sidney Fels. Coupled hardsoft tissue simulation with contact and constraints applied to jaw-tonguehyoid dynamics. *International Journal for Numerical Methods in Biomedical Engineering*, 27(3):367–390, 2011.
- [138] Ian Stavness, Christy L. Ludlow, Bethany Chung, and Sidney Fels. Hyolaryngeal biomechanics modeling with intramuscular stimulation data. In *1st International Workshop on Dynamic Modeling of the Oral, Pharyngeal and Laryngeal Complex for Biomedical Applications (OPAL-2009)*, pages 73–78, Vancouver, BC, Canada, 2009.
- [139] Ian Stavness, Mohammad A. Nazari, Pascal Perrier, Didier Demolin, and Yohan Payan. A biomechanical modeling study of the effects of the orbicularis oris muscle and jaw posture on lip shape. *Journal of speech, language, and hearing research : JSLHR*, 56(3):878–890, 2013.
- [140] Ian Stavness, C. Snchez, John Lloyd, Andrew Ho, Johnty Wang, Sidney Fels, and Danny Huang. Unified skinning of rigid and deformable models for anatomical simulations. pages 1–4. ACM, 2014.
- [141] Maureen Stone, Jonghye Woo, Junghoon Lee, Tera Poole, Amy Seagraves, Michael Chung, Eric Kim, Emi Z Murano, Jerry L Prince, and Silvia S Blemker. Structure and variability in human tongue muscle anatomy. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, pages 1–9, 2016.
- [142] S. Sueda, A. Kaufman, and DK Pai. Musculotendon simulation for hand animation. *ACM TRANSACTIONS ON GRAPHICS*, 27(3):1, 2008.
- [143] Jochen Smuth, Marco Winter, and Gnther Greiner. Reconstructing animated meshes from time-varying point clouds. *Computer Graphics Forum*, 27(5):1469–1476, 2008.
- [144] Gary A. Thibodeau and Kevin T. Patton. *Anatomy & physiology*. Mosby, St. Louis, Mo, 5th edition, 2003.

- [145] Y. Tsin and T. Kanade. *A correlation-based approach to robust point set registration*, volume 3023, pages 558–569. SPRINGER-VERLAG BERLIN, BERLIN, 2004.
- [146] Ling Tsou. The effects of muscle aging on hyoid motion during swallowing : a study using a 3d biomechanical model, 2012.
- [147] Betty Tuller, KS Harris, and Bob Gross. Electromyographic study of the jaw muscles during speech. *J Phon.*, 9:175–188, 1981.
- [148] Martin Uecker, Shuo Zhang, Dirk Voit, Alexander Karaus, KlausDietmar Merboldt, and Jens Frahm. Realtime mri at a resolution of 20 ms. *NMR in Biomedicine*, 23(8):986–994, 2010.
- [149] Michael Vaiman, Ephraim Eviatar, and Samuel Segal. Surface electromyographic studies of swallowing in normal subjects: A review of 440 adults. report 1. quantitative data: Timing measures, 2004.
- [150] K. van den Doel and U. M. Ascher. Real-time numerical solution of webster’s equation on a nonuniform grid. *IEEE Transactions on Audio, Speech, and Language Processing*, 16(6):1163–1172, Aug 2008.
- [151] Vasileios Vavourakis, John H. Hipwell, and David J. Hawkes. An inverse finite element u/p-formulation to predict the unloaded state of in vivo biological soft tissues. *Annals of Biomedical Engineering*, 44(1):187–201, 2016.
- [152] XL Wang and XP Qian. A statistical atlas based approach to automated subject-specific fe modeling. *COMPUTER-AIDED DESIGN*, 70:67–77, 2016.
- [153] Anders Westermarck, Stefan Zachow, and Barry L. Eppley. Three-dimensional osteotomy planning in maxillofacial surgery including soft tissue prediction. *The Journal of craniofacial surgery*, 16(1):100–104, 2005.
- [154] Reiner Wilhelms-Tricarico. Physiological modeling of speech production: Methods for modeling soft-tissue articulators. *The Journal of the Acoustical Society of America*, 97(5 Pt 1):3085–3098, 1995.
- [155] Adam Wittek, Nicole M. Grosland, Grand R. Joldes, Vincent Magnotta, and Karol Miller. From finite element meshes to clouds of

- points: A review of methods for generation of computational biomechanics models for patient-specific applications. *Annals of Biomedical Engineering*, 44(1):3–15, 2016.
- [156] Wojciech Wolański, Bożena Gzik-Zroska, Edyta Kawlewska, Marek Gzik, Dawid Larysz, Józef Dzielicki, and Adam Rudnik. *Preoperative Planning of Surgical Treatment with the Use of 3D Visualization and Finite Element Method*, pages 139–163. Springer International Publishing, Cham, 2015.
- [157] Jonghye Woo, Junghoon Lee, Emi Z. Murano, Fangxu Xing, Meena Al-Talib, Maureen Stone, and Jerry L. Prince. A high-resolution atlas and statistical model of the vocal tract from structural mri. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 3(1):47–60, 2015.
- [158] Jonghye Woo, E. Z. Murano, M. Stone, and J. L. Prince. Reconstruction of high-resolution tongue volumes from mri. *IEEE Transactions on Biomedical Engineering*, 59(12):3511–3524, 2012.
- [159] Jonghye Woo, Maureen Stone, Yuanming Suo, Emi Z. Murano, and Jerry L. Prince. Tissue-point motion tracking in the tongue from cine mri and tagged mri. *Journal of Speech, Language, and Hearing Research*, 57(2):S626, 2014.
- [160] Fangxu Xing, Jonghye Woo, Emi Z. Murano, Junghoon Lee, Maureen Stone, and Jerry L. Prince. 3d tongue motion from tagged and cine mr images. *Medical image computing and computer-assisted intervention : MICCAI . International Conference on Medical Image Computing and Computer-Assisted Intervention*, 16(Pt 3):41, 2013.
- [161] Shuntaro Yamazaki, Satoshi Kagami, and Masaaki Mochimaru. Non-rigid shape registration using similarity-invariant differential coordinates. pages 191–198. IEEE, 2013.
- [162] Yusuke Yoshiyasu, Wan-Chun Ma, Eiichi Yoshida, and Fumio Kanehiro. As-conformal-as-possible surface registration. *Computer Graphics Forum*, 33(5):257–267, 2014.
- [163] Weimin Yu, Moritz Tannast, and Guoyan Zheng. Non-rigid free-form 2d3d registration using a b-spline-based statistical deformation model. *Pattern Recognition*, 2015.

- [164] Paul A. Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C. Gee, and Guido Gerig. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128, 2006.

Appendix A

Muscles and Ligaments in FRANK

This appendix summarizes all of the muscles and ligaments included in the FRANK model.

A.1 Muscles

This appendix summarizes all of the muscles included in the FRANK model. Most muscles are associated with a soft tissue component, but these associations should be considered model organization choices rather than anatomically-based groupings. The muscles are presented in the following tables according to their associations, which include: the jaw and hyoid region (Table A.1), the face (Table A.2), the tongue (Table A.3), the soft palate (Table A.4), the pharynx (Table A.5), and the larynx (Table A.6).

A.1. Muscles

Name	Abb.	F_{max}	ℓ_0	L_{opt}	L_{max}	T.R.
Anterior Temporal	AT	158.00	9.16	0.97	1.24	0.50
Middle Temporal	MT	95.60	8.47	0.95	1.34	0.48
Posterior Temporal	PT	75.60	6.72	0.93	1.22	0.51
Superficial Masseter	SM	190.40	5.70	0.96	1.25	0.46
Deep Masseter	DM	81.60	3.43	0.97	1.50	0.29
Medial Pterygoid	MP	174.80	5.38	0.98	1.23	0.64
Superior Lateral Ptery- goid	SP	28.67	2.20	1.11	1.51	0.00
Inferior Lateral Ptery- goid	IP	66.90	2.96	1.05	1.38	0.00
Anterior Digastric	AD	40.00	4.16	1.18	1.51	0.00
Posterior Digastric	PD	40.00	8.72	1.00	1.28	0.00
Posterior Mylohyoid	PM	35.40	3.39	1.10	1.41	0.00
Stylohyoid	SH	15.60	10.32	1.00	1.28	0.00

Table A.1: Jaw, hyoid muscles in FRANK. The table includes the muscle name, the abbreviated name (Abb.), maximum force (F_{max}) in Newtons, muscle rest length ℓ_0 in cm (for muscles symmetrical through the midsagittal plane, ℓ is the length on right half of the model), optimal length as a ratio of the rest length (L_{opt}), maximum length as a ratio of the rest length (L_{max}), and tendon ratio (T.R.). All muscles in this table use a Peck muscle model [115], with a passive fraction of 0.015 and damping of 0.001.

A.1. Muscles

Name	Abb.	F_{max}	ℓ_0	L_{opt}	L_{max}	T.R.
Depressor Anguli Oris	DAO	1.00	2.69	1.00	2.00	0.00
Buccinator	BUC	1.00	3.74	1.00	2.00	0.00
Depressor Labii Inferioris	DLI	1.00	1.85	1.00	2.00	0.00
Mentalis	MENT	1.00	1.62	1.00	2.00	0.00
Obicularis Oris Middle	OOM	1.00	7.44	1.00	2.00	0.00
Obicularis Oris Periphe- ral	OOP	1.00	8.51	1.00	2.00	0.00
Levator Labii Superioris Alaeque Nasi	LLSAN	1.00	3.25	1.00	2.00	0.00
Levator Anguli Oris	LAO	1.00	3.49	1.00	2.00	0.00
Risorius	RIS	1.00	5.30	1.00	2.00	0.00
Zygomaticus	ZYG	1.00	5.57	1.00	2.00	0.00
Levator Labii Superioris	LLS	1.00	3.51	1.00	2.00	0.00

Table A.2: Face muscles in FRANK. Table headings are the same as Table A.1. All muscles in this table use a Constant muscle model (muscle force is the product of F_{max} and the activation level), with a passive fraction of 0 and damping of 0.

A.1. Muscles

Name	Abb.	F_{max}	ℓ_0	L_{opt}	L_{max}	T.R.
Genioglossus Anterior	GGA	32.8	3.55	1.00	2.00	0.00
Genioglossus Middle	GGM	22.0	4.59	1.00	2.00	0.00
Genioglossus Posterior	GGP	67.2	4.80	1.00	2.00	0.00
Geniohyoid	GH	32.0	3.08	1.00	2.00	0.00
Hyoglossus	HG	118.0	3.85	1.00	2.00	0.00
Anterior Mylohyoid	AM	46.8	2.41	1.00	2.00	0.00
Styloglossus	STY	43.6	9.42	1.00	2.00	0.00
Transversus Anterior	TRANSA	90.8	1.37	1.00	2.00	0.00
Transversus Middle	TRANSM	90.8	1.37	1.00	2.00	0.00
Transversus Posterior	TRANSP	90.8	1.37	1.00	2.00	0.00
Verticalis Anterior	VERTA	36.4	1.64	1.00	2.00	0.00
Verticalis Middle	VERTM	36.4	1.64	1.00	2.00	0.00
Verticalis Posterior	VERTP	36.4	1.64	1.00	2.00	0.00
Inferior Longitudinal	ILA	16.4	8.61	1.00	2.00	0.00
Superior Longitudinal	SLA	34.4	9.49	1.00	2.00	0.00

Table A.3: Tongue muscles in FRANK. Table headings are the same as Table A.1. All muscles in this table use a Peck muscle model [115], with a passive fraction of 0 and damping of 0.

Name	Abb.	F_{max}	ℓ_0	L_{opt}	L_{max}	T.R.
Levator Veli Palitini	LVP	10.8	4.20	1.00	2.00	0.00
Musculus Uvulae	MU	5.56	3.31	1.00	2.00	0.00
Palatoglossus Anterior	PGA	3.22	5.60	1.00	2.00	0.00
Palatoglossus Posterior	PGP	3.22	5.24	1.00	2.00	0.00
Palatopharyngeus	PP	12.8	10.05	1.00	2.00	0.00
Tensor Veli Palitini	TVP	4.71	3.66	1.00	2.00	0.00

Table A.4: Soft palate muscles in FRANK. Table headings are the same as Table A.1. All muscles in this table use a Peck muscle model [115], with a passive fraction of 0.5 and damping of 0.

A.1. Muscles

Name	Abb.	F_{max}	ℓ_0	L_{opt}	L_{max}	T.R.
Inferior Constrictor 1	IC1	3.0	5.09	1.00	2.00	0.00
Inferior Constrictor 2	IC2	3.0	4.18	1.00	2.00	0.00
Inferior Constrictor 3	IC3	3.0	3.65	1.00	2.00	0.00
Middle Constrictor 1	MC1	3.0	6.36	1.00	2.00	0.00
Middle Constrictor 2	MC2	3.0	3.36	1.00	2.00	0.00
Middle Constrictor 3	MC3	3.0	3.21	1.00	2.00	0.00
Superior Constrictor 1	SC1	3.0	4.01	1.00	2.00	0.00
Superior Constrictor 2	SC2	3.0	4.18	1.00	2.00	0.00
Superior Constrictor 3	SC3	3.0	4.55	1.00	2.00	0.00
Crico Pharyngeal	CP	3.0	2.61	1.00	2.00	0.00
Salpingo Pharyngeus	SalP	3.0	9.31	1.00	2.00	0.00
Stylo Pharyngeus	StyP	3.0	8.19	1.00	2.00	0.00

Table A.5: Pharynx muscles in FRANK. Table headings are the same as Table A.1. All muscles in this table use a Peck muscle model [115], with a passive fraction of 0 and damping of 0.

Name	Abb.	F_{max}	ℓ_0	L_{opt}	L_{max}	T.R.
Criothyroid Pars Recta	CPR	1.00	1.64	0.95	1.15	0.10
Criothyroid Pars Oblique	CPO	1.00	2.14	0.95	1.15	0.10
Interarytenoid Transverse	IAT	1.00	1.78	0.95	1.15	0.10
Interarytenoid Oblique	IAO	1.00	1.84	0.95	1.15	0.10
Thyrohyoid Superior	TS	1.00	1.96	0.95	1.15	0.10
Thyrohyoid Inferior	TI	1.00	2.66	0.95	1.15	0.10
Thyroarytenoid External	TE	0.60	1.92	0.95	1.15	0.10
Thyroarytenoid Vocalis	TV	1.00	1.76	0.95	1.15	0.10
Lateral Cricothytenoid	LC	1.50	0.91	0.95	1.15	0.10
Posterior Cricothytenoid Oblique	PCO	1.00	1.92	0.95	1.15	0.10
Sternothyroid	ST	0.75	8.81	0.95	1.15	0.10
Sternohyoid	SteH	0.50	10.32	0.95	1.15	0.10

Table A.6: Larynx muscles in FRANK. Table headings are the same as Table A.1. All muscles in this table use a Peck muscle model [115], with a passive fraction of 0.2 and damping of 0.

A.2 Ligaments

The ligaments included in the FRANK model are given in Table A.7.

Name	Abb.	Stiffness	Damping
Cricotracheal	CTL	100.00	0.001
Cricotracheal Posterior	CTPL	100.00	0.001
Cricotracheal Anterior	CTAL	100.00	0.001
Cricoarytenoid Inferior	CAIL	500.00	0.001
Cricoarytenoid Medial	CAML	100.00	0.001
Cricoarytenoid Lateral	CALL	200.00	0.001
Cricoarytenoid Anterior	CAAL	100.00	0.001
Cricoarytenoid Posterior	CAPL	100.00	0.001
Thyroepiglottic	TEL	50.00	0.001
Thyrohyoid Lateral	TLL	5.00	0.001
Thyrohyoid Median	TML	5.00	0.001
Cricothyroid Lateral	CLL	5000.00	0.001
Cricothyroid Median	CML	50.00	0.001
Vocal Ligament	VL	100.00	0.001
Hyoepiglottic Median	HML	100.00	0.001
Hyoepiglottic Lateral	HLL	100.00	0.001
Thyrohyoid Anterosuperior Membrane 1	TAM1	5.00	0.001
Thyrohyoid Anterosuperior Membrane 2	TAM2	5.00	0.001
Thyrohyoid Posterosuperior Membrane 1	TPM1	5.00	0.001
Thyrohyoid Posterosuperior Membrane 2	TPM2	5.00	0.001

Table A.7: List of all ligaments in FRANK, including full and abbreviated (Abb.) name. Ligaments are modeled as springs with a stiffness (N/m) and damping. Some membranes, included in this table, are also modeled as having a spring-like contribution.