3D biological tissue shape reconstruction dealing with motion and deformation based on medical images

Submitted To
Graduate School of Information Science and Technology
Osaka University

January 2025

Norio FUKUDA

Abstract

This dissertation focuses on methods for reconstructing 3D biological tissue shapes from medical images, addressing challenges associated with intra-individual dynamic deformation and inter-individual static variation. Dynamic deformation arises from movements caused by spontaneous motion or external forces, while static variation reflects individual anatomical differences. These challenges require distinct approaches to restore 3D tissue shapes accurately.

The first study proposes a method for reconstructing 3D ultrasound images during surgery. It uses a biplane probe to track tissue movement caused by compression. This approach enables real-time correction of non-cyclic motion and achieves high accuracy without relying on external position sensors, improving the system's practicality.

The second study develops a tool to efficiently generate labeled datasets by allowing operators to deform template shapes based on 3D medical images. This method reduces annotation time and enables non-specialist users to produce datasets suitable for training machine learning models. These datasets have been shown to enhance model performance compared to those generated by conventional expert-based methods.

The third study presents a statistical model for estimating muscle

attachment sites using patient-specific bone shapes. By incorporating anatomical variation observed across multiple specimens, this method enables accurate estimation of attachment sites, supporting applications such as personalized musculoskeletal modeling.

These studies provide practical solutions for addressing the effects of motion and deformation in 3D tissue shape reconstruction. The results improve the accuracy and efficiency of data generation and shape analysis in medical image processing, supporting diagnosis and treatment planning applications.

List of publications

As the first author

Journal articles

1 Fukuda, N., Otake, Y., Nakamoto, M., Ukimura, O., & Sato, Y. (2024). Position sensor-less reconstruction of 3D ultrasound image of the prostate using biplane transrectal ultrasound (TRUS) probe. Journal of Japan Society of Computer Aided Surgery, 26, (accepted for publication).

2 Fukuda, N., Konda, S., Umehara, J., & Hirashima, M. (2024). Efficient musculoskeletal annotation using free-form deformation. Scientific Reports, 14(1), 16077.

3 Fukuda, N., Otake, Y., Takao, M., Yokota, F., Ogawa, T., Uemura, K., ... & Sato, Y. (2017). Estimation of attachment regions of hip muscles in CT image using muscle attachment probabilistic atlas constructed from measurements in eight cadavers. International journal of computer assisted radiology and surgery, 12, 733-742.

International conference

4 Fukuda, N., Otake, Y., Takao, M., Yokota, F., Ogawa, T., Uemura, K., Nakaya, R., Tamura, K., Grupp, R. B., Farvardin, A., Armand, M., Sugano, N., & Sato, Y. "Statistical Estimation of Attachment of Hip Muscles Based on Measurement in Cadavers," Proceedings of CAOS International 2016, pp. 351-354.

5 Fukuda, N., Nakamoto, M., Okada, T., Yoshito, O., Shoji, S., Ukimura, O., Gill, S., & Yoshinobu, S. 2015. "Cadaver Evaluation of Image-Based Tracking Using Biplane Transrectal Ultrasound Probe for Focal Prostate Therapy," the 2015 Joint Conference of the International Workshop on Advanced Image Technology (IWAIT) and the International Forum on Medical Imaging in Asia (IFMIA), Tainan, Taiwan.

6 Fukuda, N., Chen, Y. W., Nakamoto, M., Okada, T., & Sato, Y. (2010). "A scope cylinder rotation tracking method for oblique-viewing endoscopes without attached sensing device," In Software Engineering and Data Mining (SEDM), 2010 2nd International Conference on 2010, pp. 684-687.

As a co-author

Journal articles

10 Umehara, J., Fukuda, N., Konda, S., & Hirashima, M. "Validity of Freehand 3-D Ultrasound System in Measurement of the 3-D Surface Shape of Shoulder Muscles," Ultrasound in Medicine & Biology 48.9 (2022), pp. 1966-1976.

11 Kaneko, M., Fukuda, N., Nagano, H., Yamada, K., Yamada, K., Konishi, E., ... & Ukimura, O. (2022). Artificial intelligence trained with integration of multiparametric MR-US imaging data and fusion biopsy trajectory-proven pathology data for 3D prediction of prostate cancer: A proof-of-concept study. The Prostate 82.7 (2022), pp. 793-803.

12 Takao, M., Otake, Y., Fukuda, N., Sato, Y., Armand, M., & Sugano, N. "The Posterior Capsular Ligamentous Complex Contributes to Hip Joint Stability in Distraction," The Journal of arthroplasty 33.3 (2018), pp. 919-924.

13 Matsugasumi, T., Baco, E., Palmer, S., Aron, M., Sato, Y., Fukuda, N., Suer, E., Bernhard, J.-C., Nakagawa, H., & Azhar, R. A. "Prostate Cancer Volume Estimation by Combining Magnetic Resonance Imaging and Targeted Biopsy Proven Cancer Core Length: Correlation with Cancer Volume," The Journal of urology 194.4 (2015), pp. 957-965.

14 Shoji, S., Hashimoto, A., Nakamoto, M., Fukuda, N., Fujikawa, H., Endo, K., Tomonaga, T., Nakano, M., Terachi, T., & Uchida, T. "Morphological Analysis of the Effects of Intraoperative Transrectal Compression of the Prostate During High-Intensity Focused Ultrasound for Localized Prostate Cancer," International Journal of Urology 22.6 (2015), pp. 563-571.

International conference

15 Enami, H., Takao, M., Fukuda, N., Umehara, J., Hirashima, M., Konda, S., Uemura, K., Hamada, H., Sugano, N. "Sliding portion at acetabulum of psoas major tendon - dynamic analysis using wide bore MRI," International Society for Computer Assisted Orthopaedic Surgery.

16 Otake, Y., Yokota, F., Fukuda, N., Takao, M., Takagi, S., Yamamura, N., O'Donnell, L. J., Westin, C.-F., Sugano, N., & Sato, Y. "Patient-Specific Skeletal Muscle Fiber Modeling from Structure Tensor Field of Clinical Ct Images," International Conference on Medical Image Computing and Computer-Assisted Intervention 2017, pp. 656-663.

17 Nishii, T., Kawakami, S., Fukuda, N., Otake, Y., & Sato, Y. 2016. "Virtual Arthroscopic Simulator for Estimating Optimal Portal Placements in Arthroscopic Hip Surgery," Journal of Hip Preservation Surgery, p. hnw030. 050.

18 Otake, Y., Yokota, F., Takao, M., Fukuda, N., Sugano, N., & Sato, Y. 2016. "Analysis of Muscle Fiber Structure Using Clinical Ct:

Preliminary Analysis Using Cadaveric Images," Proceedings of CAOS International 2016, pp. 343-346.

19 Shin, T., Ukimura, O., de Castro Abreu, A. L., Isotani, S., Fukuda, N., Matsugasumi, T., Aron, M., Mimata, H., & Gill, I. S. 2016. "Pd30-10 Three-Dimensional Printed Model of Prostate Anatomy and Targeted Biopsy-Proven Index Tumor to Facilitate Nerve-Sparing Prostatectomy," The Journal of Urology (195:4), pp. e715-e716.

20 Nakamoto, M., Ukimura, O., Fukuda, N., Shoji, S., Abreu, A. L., Leslie, S., Sato, Y., & Gill, I. S. 2012. "Image Fusion for Assessment of Therapeutic Adequacy in Targeted Cryotherapy for Prostate Cancer: 3d Visualization of the Preoperative Cancer Model and the Intraoperative Cryolesion Model," JOURNAL OF ENDOUROLOGY: MARY ANN LIEBERT INC 140 HUGUENOT STREET, 3RD FL, NEW ROCHELLE, NY 10801 USA, pp. A405-A406.

21 Ukimura, O., de Castro Abreu, A. L., Fukuda, N., Shoji, S., Sato, Y., Gill, I., & Nakamoto, M. 2012. "V1238 Real-Time Tracking of Biplane Trus Images by Image Registration Accelerated with Graphics Processing Unit," The Journal of Urology (187:4), pp. e500-e501.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Medical image analysis plays a crucial role in modern healthcare by precisely understanding patient-specific anatomical features, thereby enhancing the quality of diagnosis and treatment (Ma et al., 2024; Uhl et al., 2021; Kobatake and Masutani, 2017). Among the various applications, three-dimensional (3D) shape reconstruction is indispensable for preoperative planning (Hamid et al., 2023), intraoperative guidance (Malhotra et al., 2023), and postoperative assessment (Cui et al., 2024). This technology has widespread use in fields such as radiation therapy and robotic surgery (Yamout et al., 2023). However, accurate reconstruction of 3D shapes faces significant challenges, including the complexity of anatomical structures, dynamic deformations caused by motion or external forces, and static inter-individual variations in anatomy.

This study addresses the challenges of 3D shape reconstruction by proposing novel approaches to handle dynamic and static deformations. Specifically, it focuses on improving the accuracy and efficiency of shape reconstruction methods by targeting two fundamental problems:

(1) correcting dynamic deformations caused by fast dynamics such as motion from respiration, postural changes, or external forces during imaging, and (2) managing static deformations that account for inter-individual anatomical variations. It should be noted that dynamic deformation encompasses both fast dynamics, as mentioned above, and slow dynamics caused by gradual changes such as aging or disease progression. However, as slow dynamics cannot be captured through single imaging and require longitudinal databases, they are treated equivalently to static deformation in image processing. Therefore, this study intentionally focuses on fast dynamics and static deformation while excluding detailed discussions of slow dynamics.

1.2 Addressing Dynamic Deformations

Dynamic deformations in medical imaging often arise from periodic physiological processes such as respiration (Chen et al., 2023) or cardiac motion (Pan et al., 2024). Existing methods leverage this periodicity to achieve high-precision corrections. However, periodicity cannot be assumed in scenarios such as intraoperative imaging, where organ motion is induced by external forces or instrument manipulation. This research addresses the challenge of compensating for non-cyclic dynamic deformations.

The first objective of this dissertation is to develop a method for accurate 3D shape reconstruction in such cases. The proposed method utilizes a bi-plane ultrasound probe to acquire cross-sectional images in two orthogonal planes, enabling real-time compensation for motion without relying on external position sensors. This technique achieves precise reconstruction even in non-cyclic motion, enhancing its appli-

cability in clinical environments.

1.3 Addressing Static Deformations

Static deformations involve inter-individual anatomical variations and require different approaches for effective handling. This study focuses on two specific objectives under this category:

Template-to-Target Adaptation (One-to-one Shape Transformation)

One-to-one shape transformation involves adapting a template model to match a specific individual's anatomy. Traditional methods, such as multi-atlas approaches (Wang et al., 2012), align multiple templates to the target anatomy through optimization-based deformation. However, these methods often lack control over the deformation process, resulting in undesirable or anatomically inappropriate transformations. To overcome this limitation, this research introduces a fully controllable deformation interface that allows operators to manipulate the template model directly. The proposed tool enables non-experts to efficiently adapt templates to target images, generating high-quality annotated datasets in a fraction of the time required by traditional approaches. This significantly improves the efficiency and accuracy of template-to-target adaptation.

Statistical Modeling of Anatomical Variations (Many-to-one Shape Transformation)

Many-to-one transformations are essential for capturing and modeling anatomical variations across individuals. Previous methods, such as

Statistical Shape Models (SSM) (Heimann and Meinzer, 2009) and Active Appearance Models (AAM) (Cootes et al., 1998), have been widely used to analyze global shapes or pixel-based variations. However, these approaches are unsuited for handling localized structures, such as the closed surfaces representing muscle attachment regions on bone surfaces. This research bridges this gap by constructing a statistical model of muscle attachment regions using data from multiple specimens. The proposed method estimates patient-specific attachment regions based on bone shape, enabling high-precision reconstruction of these regions. This advancement facilitates a better understanding of inter-individual anatomical variations and their integration into 3D reconstruction processes.

1.4 Significance of This Dissertation

This dissertation addresses critical challenges in 3D biological tissue shape reconstruction by proposing methods to manage dynamic deformations caused by motion and static anatomical variations between individuals. The research introduces techniques for real-time shape recovery without reliance on external tracking devices, efficient dataset generation through template deformation, and statistical modeling of anatomical features.

These methods improve the accuracy and efficiency of 3D reconstruction and dataset creation, particularly in applications requiring patient-specific anatomical modeling. This work provides a basis for further advancements in medical image analysis and related computational methods by focusing on practical implementations and data-driven approaches.

1.5 Structure of the Dissertation

This dissertation is organized as follows. Chapter 2 addresses the challenge of non-cyclic dynamic deformations caused by external forces during surgery. It introduces a method for real-time 3D shape reconstruction using a biplane ultrasound imaging system. Chapter 3 focuses on efficient musculoskeletal shape generation through a novel deformation tool, enabling the creation of annotated datasets with high accuracy and reduced time. Chapter 4 presents a statistical modeling approach for estimating muscle attachment regions, incorporating inter-individual anatomical variations to enhance musculoskeletal simulations. Finally, Chapter 5 summarizes the findings and discusses future directions, including the integration of these methods into platforms for personalized medicine.

Chapter 2

3D Shape Reconstruction Addressing Displacement Caused by External Forces

Original title: Position sensor-less reconstruction of 3D ultrasound image of the prostate using biplane transrectal ultrasound (TRUS) probe

2.1 Overview

This study presents a novel method for three-dimensional ultrasound (3D-US) reconstruction using a bi-plane transrectal ultrasound (TRUS) probe designed for 3D guidance of focal therapy. The bi-plane TRUS probe simultaneously acquires two orthogonal cross-sectional images, enabling precise compensation for prostate motion induced by external pressure during probe sweeping. The proposed method addresses the challenge of non-cyclic, dynamic deformations in clinical scenarios. Importantly, the reliability and effectiveness of this method have been rigorously validated through phantom and cadaver experiments. These experiments have demonstrated that the reconstruction accuracy

of this method is equivalent to or exceeds that of position sensor-based approaches. These findings highlight the potential clinical utility of this method in scenarios requiring accurate 3D reconstruction without periodic motion assumptions.

2.2 Background

In recent years, focal therapy has emerged as a promising treatment approach for prostate cancer, aiming to selectively ablate cancerous tissue while preserving healthy prostate tissue, as opposed to performing a total prostatectomy (Javier-DesLoges et al., 2023; Fujihara and Ukimura, 2022). A definitive diagnosis of prostate cancer typically involves a biopsy procedure, where suspected areas and 12-16 systematic locations within the prostate are sampled under ultrasound (US) guidance. For focal therapy to be successful, accurate three-dimensional (3D) localization of the tumor boundaries and biopsy sites is essential, necessitating precise 3D mapping of tissue sampling points. To meet these needs, new systems have been developed to enhance the accuracy of biopsy targeting and tissue sampling.

One such system is the MR-US fusion biopsy system (Light et al., 2024; Falagario et al., 2024; Fletcher et al., 2023), which leverages the high sensitivity of magnetic resonance (MR) imaging to better identify potential tumor sites prior to biopsy. In this method, MR imaging is performed in advance to localize suspected cancerous regions, and these locations are targeted during the biopsy, with the trajectory of the biopsy needle monitored in real-time on the MR image. This setup enables the quantification of sampling sites within the MR coordinate system by mapping the needle trajectory to the MR image space, allowing for

accurate localization of biopsy sites.

The MR-US fusion biopsy system maps the MR and US images through registration. At the start of the procedure, a 3D-US image (referred to as the reference US image) that includes the prostate and surrounding area is reconstructed. The prostate contour is extracted from the reference US image and the MR image to establish a correspondence between the coordinate systems. During the biopsy, this correspondence is used to map the needle position in the MR coordinate system and display it on the MR image in real-time.

However, differences in prostate positioning between the initial reconstruction of the reference US image and the biopsy procedure can occur due to interventions like needle insertion. A 2D-3D registration is performed between the real-time 2D-US images and the 3D reference US image to account for these changes. This 2D-3D registration (where ”2D” refers to cross-sectional views rather than projections) determines the relative positioning in real-time, allowing adjustments to the needle position based on the updated prostate location (De Silva et al., 2017; Xu et al., 2008). The accuracy of this correction relies heavily on the precision of the reference US image.

Traditionally, systems such as Artemis (Bax et al., 2008)(Eigen, USA) and Biojet (Hadaschik et al., 2011) (DK Technologies, USA) have used 3D position sensors to acquire the reference US image. This process involves sweeping a tracked transrectal US (TRUS) probe to obtain a series of 2D-US images, which are then reconstructed into a 3D volume. Such methods assume that the prostate remains stationary during probe sweeping. However, Baumann et al. reported that the prostate can shift by an average of 13.8 mm due to TRUS probe pressure dur-

ing procedures (Baumann et al., 2012), suggesting that the prostate may move during the acquisition of multiple images, potentially compromising the accuracy of the reconstructed reference US image. This movement could limit the precision of needle positioning based on the corrected reference image.

To address this, we have developed a novel method for reconstructing a reference US image that accounts for prostate movement during probe sweeping. Our approach aims to enhance the accuracy of needle targeting under the assumption that prostate movement may occur. This study proposes a new reconstruction method that leverages a biplane TRUS probe, which enables simultaneous imaging of two orthogonal cross-sections in real time. Utilizing this feature, we acquire an image sequence from one cross-section (parallel to the probe's sweeping direction) that includes the prostate's midsagittal plane. This sequence generates an extended field-of-view (EFOV) US image (Scherzer and Schoisswohl, 2002; Weng et al., 1997).

The novelty of our approach lies in correcting prostate motion caused by probe pressure during EFOV-US image reconstruction. Additionally, we achieve 3D reconstruction without using position sensors by accurately aligning the orthogonal cross-section images. In experiments, we first compare the accuracy of our method to the position-sensor-based approach using phantom models under ideal conditions for position-sensor accuracy. Subsequently, we conduct cadaver studies simulating actual surgical conditions, where prostate movement is expected, to compare needle placement accuracy between reconstructions based on our method and those based on the position sensor. This study demonstrates the effectiveness of our method in improving needle targeting

precision for prostate biopsies.

2.3 Materials and Method

2.3.1 Biplane TRUS probe

In this study, we employ a biplane TRUS probe capable of simultaneously capturing two orthogonal cross-sections, as illustrated in Fig. 2.1. This configuration allows for the acquisition of a vertical cross-sectional image I_v (highlighted by the pink frame on the left side of Fig. 2.1), which is parallel to the probe's axis, as well as a horizontal cross-sectional image I_h (highlighted by the green frame) that is perpendicular to it. The operator can view these two cross-sectional images concurrently on the display, as shown on the right side of Fig. 2.1. Hereafter, the images captured at the same time t from the vertical and horizontal cross-sections will be denoted as $I_{v,t}$ and $I_{h,t}$, respectively.

2.3.2 3D-US reconstruction

Figure 2.2 provides an overview of the proposed reconstruction method. In this approach, the biplane TRUS probe is swept to maintain the vertical cross-sectional plane on the same plane (the prostate's midsagittal plane). Adjacent vertical cross-sectional images are rigidly aligned to create an extended field-of-view (EFOV) US image (Weng et al., 1997). For an experienced urologist, maintaining this midsagittal plane during sweeping is relatively straightforward. Following the aligned positions of each vertical cross-section, the orthogonal horizontal cross-sectional images are arranged within a 3D-US image space to reconstruct the reference US image.

Figure 2.3 illustrates the reconstruction process when the prostate shifts during probe sweeping. In Fig. 2.3(a), the positional relationship is shown when the prostate moves due to probe pressure, along with the acquired 2D-US images from the vertical cross-section. From t_1 to t_3, the probe shifts leftward while simultaneously compressing the prostate upwards. Figure 2.3(b) shows the positioning when using a location sensor to align images $I_{(v,t_1)}$ through $I_{(v,t_3)}$ based on the probe's measured position, while Fig. 2.3(c) shows the positioning achieved using the proposed method. The prostate movement remains uncorrected when using the location sensor, leading to a distorted reference US image reconstructed from the horizontal cross-sections. Conversely, in the proposed method, motion correction between adjacent vertical cross-sections ensures that horizontal cross-sections are appropriately arranged. This enables the proposed method to robustly reconstruct a reference US image, even when the prostate moves during sweeping.

The following is the mathematical formulation of this method. Let the plane containing the EFOV-US image, precisely the midsagittal plane of the prostate during sweeping, be the x-y plane. For the EFOV-US image reconstruction, the positional relationship between adjacent vertical cross-sectional images $I_{(v,t-1)}$ and $I_{(v,t)}$ $(t = 1, 2, 3, \ldots, n)$ is obtained by performing rigid registration in three degrees of freedom within the x-y plane. Here, the three degrees of freedom in rigid alignment are determined by finding dx, dy, and $d\theta$ that minimize the cost C_t as shown below (Scherzer and Schoisswohl, 2002). The cost function C_t for aligning adjacent vertical cross-sectional images $I_{(v,t-1)}$ and $I_{(v,t)}$ is defined as follows:

$$C_t = \sum_{(x_v, y_v)} \left[I_{(v,t-1)}(x_v, y_v) - I_{(v,t)}(x'_v, y'_v) \right]^2 \tag{2.1}$$

where the transformed coordinates (x'_v, y'_v) are obtained by applying a rigid transformation to (x_v, y_v). This transformation is represented as:

$$\begin{pmatrix} x'_v \\ y'_v \\ 1 \end{pmatrix} = \begin{pmatrix} \cos d\theta & -\sin d\theta & dx \\ \sin d\theta & \cos d\theta & dy \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_v \\ y_v \\ 1 \end{pmatrix} \tag{2.2}$$

In this transformation, dx and dy are the translations in the x- and y-directions, respectively, and $d\theta$ is the rotation angle around the z-axis. This rigid transformation minimizes C_t by optimally aligning the two adjacent cross-sections. We use stochastic gradient descent (Qiao et al., 2014) to minimize the cost.

Next, the obtained displacement values dx, dy, and $d\theta$ in the 2D coordinate system for the vertical cross-section are converted into displacement values for the horizontal cross-section in the 3D coordinate system. Let Σ_v and Σ_h represent the 3D coordinate systems for the vertical and horizontal cross-sections, respectively, where $I_{(v,0)}$ and $I_{(h,0)}$ each have their origins on an axis z orthogonal to the x-y plane. The transformation matrix $dT_{(v,t)}$ for the relative position between adjacent vertical cross-sections $I_{(v,t-1)}$ and $I_{(v,t)}$ in the coordinate system Σ_v is expressed as follows using dx, dy, and $d\theta$:

$$dT_{(v,t)} = T(dx, dy, 0) \cdot R_z(d\theta) \tag{2.3}$$

where $T(x, y, z)$ denotes a 3D translation and $R_z(\theta)$ represents a rotation about the z-axis in homogeneous coordinates. Given the transformation matrix T_{vh} for mapping from Σ_v to Σ_h, the positional relationship $dT_{(h,t)}$ for adjacent horizontal cross-sections $I_{(h,t-1)}$ and $I_{(h,t)}$ in Σ_h is obtained by:

$$dT_{(h,t)} = T_{vh} \cdot dT_{(v,t)} \tag{2.4}$$

The transformation matrix $T_{(h,t)}$ at any given time t is then calculated as:

$$T_{(h,t)} = \prod_{n=1}^{t} dT_{(h,n)} \tag{2.5}$$

Using $T_{(h,t)}$ for each time point t, we reconstruct all horizontal cross-sectional images $I_{(h,t)}$ ($t = 1, 2, 3, \ldots, N$).

The transformation T_{vh} is determined through prior calibration. First, a position sensor marker is attached to the TRUS probe, and the transformation matrices T_{ph} and T_{pv} from the TRUS probe to Σ_h and Σ_v, respectively, are obtained via standard ultrasound calibration (Sato et al., 1998). Consequently, we derive T_{vh} as follows:

$$T_{vh} = T_{pv} \cdot T_{ph}^{-1} \tag{2.6}$$

2.3.3 Scanning Procedure

The proposed method assumes that prostate motion is restricted to in-plane translations and rotations within vertical cross-sections during the scanning process. To minimize unpredictable motion components resulting from external probe pressure, the scanning procedure is designed to withdraw the probe from within the body rather than applying pressure from outside. This approach stabilizes the prostate's movement and ensures that it predominantly adheres to in-plane motion, a prerequisite for accurate image alignment.

Furthermore, excessive probe sweeping speed is avoided to reduce the effects of motion blur in the acquired ultrasound images. Rapid probe movements can lead to image artifacts, compromising the extended field-of-view (EFOV) quality and the reconstructed 3D-US images. To ensure the optimal quality of the input data, operators are instructed to perform smooth and steady probe sweeps while maintaining the vertical cross-sectional plane on the midsagittal plane of the prostate.

These considerations are integral to the robustness and accuracy of the proposed reconstruction method, enabling effective correction of prostate motion during scanning and reliable arrangement of horizontal cross-sections in the 3D-US image space.

2.4 Experiments and Results

In the experiments, we first conducted a preliminary study using a prostate biopsy training phantom, Model 053 (CIRS, USA). This phantom is structurally designed to sweep the probe while visually confirming the prostate model without inducing prostate movement. Under these conditions, reconstruction using a position sensor can be performed with high accuracy since the prostate does not move. Therefore, this setup provides ideal conditions for validating the accuracy of the position sensor-based reconstruction method (hereafter referred to as the ”phantom experiment”).

Next, to simulate conditions closer to actual surgical procedures, we conducted experiments on two cadavers to evaluate the accuracy of 3D-US image reconstruction of the prostate region and subsequent 2D-3D

alignment for estimating needle placement accuracy (hereafter referred to as the ”cadaver experiment”). The cadaver experiment was approved by the Ethics Review Committee for Research Involving Human Subjects at Nara Institute of Science and Technology (Approval Number 2022-I-8).

For both experiments, we used the type 8808e biplane TRUS probe (BK Medical, Denmark) and the Polaris Spectra optical position sensor (Northern Digital, Canada).

2.4.1 Phantom experiments

Experimental setting

In the phantom experiment, we used a 3D-US image of the phantom captured by an MRI scanner as the ground truth data. We then compared the reconstruction results of the 3D-US images obtained using the position sensor and the proposed method. During the sweeping process, we ensured that no pressure was applied to the phantom with the TRUS probe, preventing any movement of the phantom.

The evaluation was conducted by comparing the prostate region manually extracted from each reconstructed 3D-US image to the prostate region derived from the ground truth data. This comparison was quantified using the Dice coefficient (Dice, 1945) to assess spatial overlap and the Average Symmetric Surface Distance (ASD) (Crum et al., 2006) to measure the mean surface distance error.

Experimental results

Figure 2.4 shows three cross-sectional views of the 3D-US images reconstructed using MR imaging, the proposed method, and the position sensor method. Table 2.1 presents the quantitative accuracy evaluation results, using the MR-derived region as the ground truth. A qualitative assessment of the three cross-sectional images revealed no significant differences between the reconstruction methods.

In the accuracy evaluation in Table 2.1, the ASD was 0.66 mm for the proposed method and 0.59 mm for the position sensor method, with a difference of less than 0.1 mm. The Dice coefficient was 93.67 % for the proposed method, slightly higher than the 93.32 % achieved by the position sensor method. Overall, both methods demonstrated comparable accuracy.

2.4.2 Cadaver experiments

Experimental setting

Figure 2.5(a) illustrates the setup for the cadaver experiment conducted at the University of Southern California Hospital. Fresh cadavers with natural prostate movement, similar to that of actual patients, were used. For the reference US image reconstruction, approximately 250 and 300 pairs of orthogonal cross-sections were obtained from two cadavers, with sweep lengths of approximately 5 cm and 6 cm, respectively. Due to the difficulty in comparing the reconstructed images to a ground truth shape in cadaver experiments, we assessed the needle placement estimation error by acquiring 2D-US images for simulated biopsy punctures after reconstructing the reference 3D-US image using both the position

sensor and the proposed method. The alignment cost for the 2D-3D registration (Xu et al., 2008) was computed as the sum of both cross-sections because of the biplane TRUS probe.

Before the experiment, several metal markers were embedded within the prostate of each cadaver. An example of a metal marker in the ultrasound image is shown in Fig. 2.5(b). 2D-US images were acquired at locations where both cross-sectional views captured metal markers to evaluate needle placement. Each pair of cross-sections contained 2-4 metal markers, and ten pairs of cross-sections were used for each cadaver. In total, 27 points were evaluated for cadaver 1 and 24 for cadaver 2.

Following 2D-3D alignment, corresponding metal markers in the reference US image and each cross-sectional pair were manually extracted, and the Euclidean distance between their centroids was calculated to determine the marker position error. To prevent high-intensity metal markers from influencing image alignment, the regions containing the metal markers were masked to exclude them from the alignment calculation.

Experimental results

Figure 2.6 shows typical sagittal plane images reconstructed from the two cadaver specimens. The sagittal plane includes the probe's sweeping direction and is parallel to the vertical cross-section (midsagittal plane of the prostate). When comparing sagittal images, the reconstructions using the proposed method appeared more curved than those using the position sensor, particularly in cadaver 2, where the difference was more pronounced.

Figure 2.7 presents a detailed comparison of the reconstruction results for cadaver 2, where significant differences were observed. In the sagittal plane indicated by the green line on the left, the cross-sectional planes are aligned in parallel when using the position sensor, whereas, in the proposed method, they are arranged in a curved fashion, closely resembling the schematic shown in Fig. 2.3. In comparing the prostate contour traced by the urologist, the contour of the sagittal plane in the 3D-US image reconstructed with the proposed method (green dashed line) more closely matched the original vertical cross-section contour (pink dashed line) compared to the position sensor contour (light blue dashed line). The average distance between the original and reconstructed image contours was 0.86 mm with the proposed method, compared to 1.55 mm with the position sensor.

The results of the marker position error are presented in Fig. 2.8 and Table 2.2. In cadaver 1, the average marker position error decreased from 2.72 mm with the position sensor method to 2.12 mm with the proposed method, though this difference was not statistically significant. In cadaver 2, the average error decreased from 5.10 mm to 2.70 mm, showing a statistically significant difference ($p < 0.01$) based on the Wilcoxon rank-sum test.

Discussion

This study relates to various methods developed for reconstructing 3D-US images from 2D-US images without position sensors using a free-hand approach. These methods include techniques based on speckle noise correlation(Gee et al., 2006), deep learning-based positioning and alignment(Prevost et al., 2017; Miura et al., 2023), and methods us-

ing inertial sensors. However, such approaches generally achieve a different level of accuracy than position sensors(Luo et al., 2023). They rely on a single cross-sectional image, which lacks reliable information to estimate out-of-plane motion. This leads to cumulative errors and instability in positional estimation. In contrast, the proposed method achieves accuracy comparable to position sensors, as demonstrated in the phantom experiment. This is likely due to its ability to directly measure out-of-plane motion using orthogonal 2D-US images, a significant advantage over prior methods. While the proposed method is not immune to cumulative errors–arising from its reliance on aligning adjacent frames–the errors remain minimal in practice. The phantom experiment revealed that even when cumulative errors occurred, they were negligible, resulting in reconstruction accuracy nearly identical to that of position sensor methods. This suggests that the proposed method's design inherently limits the accumulation of such errors under controlled conditions. Nonetheless, future improvements could involve exploring strategies to mitigate cumulative errors further, such as incorporating information from non-adjacent frames. However, when scanning dynamic tissues such as the prostate, using too-distant frames may lead to inconsistencies due to non-rigid tissue deformation, potentially compromising reconstruction accuracy. Thus, carefully selecting frames tailored to the specific characteristics of the scanned tissue is essential to maximize precision while minimizing adverse effects from non-rigid deformations.

The marker position errors in Fig. 2.8 and Table 2.2 represent alignment errors between the reference US image and real-time 2D-US images, reflecting the effectiveness of prostate motion correction during interventions, such as biopsy needle insertion. The observed differences

in marker position errors between the position sensor and our proposed method are attributed to variations in the reconstruction accuracy of the reference US image, given that both approaches use the same 2D-3D alignment algorithm for real-time imaging. Figure 2.7 shows that significant prostate displacement due to TRUS probe pressure was observed in cadaver 2, leading to lower reconstruction accuracy with the position sensor-based method. However, the proposed method demonstrated accurate reconstruction by correcting for this prostate movement, achieving better alignment in cases of substantial motion.

Concerning clinical accuracy requirements, Karnik et al. proposed that an alignment error of less than 2.5 mm is clinically significant for focal therapy, as treatable tumors typically have a radius of at least 5 mm (Karnik et al., 2010). While the result for cadaver two slightly exceeded this threshold at 2.7 mm, the proposed method nonetheless demonstrated significant accuracy improvements over the position sensor, suggesting clinical relevance. Although rigid alignment was employed for 2D-3D registration in this study, De Silva et al. noted that incorporating non-rigid alignment to account for prostate deformation can improve accuracy by approximately 0.8 - 1.6 mm over rigid alignment. Thus, introducing non-rigid alignment into the proposed method could achieve the desired clinical accuracy and enhance its application in a surgical setting.

A limitation of the proposed method is that it performs rigid alignment across the entire vertical cross-sectional image during EFOV-US image generation, which does not account for potential prostate deformation during probe sweeping. This approach may inadvertently capture the motion of surrounding tissues, potentially reducing reconstruc-

tion accuracy. An automated method for extracting the prostate region from 2D-US images, using deep learning-based recognition techniques (Anas et al., 2018; Jiang et al., 2023, 2024), could allow more precise alignment by focusing exclusively on prostate tissue, thereby improving reconstruction accuracy. Additionally, since deformation may occur during probe sweeping, automated prostate recognition could enable the selective use of minimally deformed 2D-US images or provide warnings when deformation is detected. Incorporating these automated contour extraction techniques into the reconstruction process remains an important direction for future research.

The cadavers used in this study were thawed frozen specimens, which better preserve the physical properties of living tissue than formalin-fixed specimens, thereby providing high reliability of results even with a single specimen (Ukimura and Matsugasumi, 2015). In this setting, the proposed method demonstrated comparable or superior accuracy to that of a position sensor across both cadavers, with significant improvement observed in one cadaver, indicating potential clinical utility for biopsy guidance and needle placement.

Another critical consideration is that the proposed method relies on EFOV-US imaging, which requires the 2D-US images to remain in a single plane during probe sweeping. For standard EFOV-US imaging, guidelines can be drawn directly on the body surface to assist in maintaining a single plane; however, this is more challenging with a TRUS probe inside the rectum. Nonetheless, the TRUS probe allows limited movement, mainly pivoting around the anus, and by visually monitoring the horizontal cross-section from the biplane TRUS probe, urologists can feasibly maintain the required conditions. In the cadaver experi-

ment, position sensor data analysis showed minimal rotation (average of 0.4°) and out-of-plane movement (average of 0.6 mm), confirming that sweeping was performed with high precision within the same plane. Furthermore, within the urologist's observation range, the prostate in the horizontal cross-sectional images did not deviate from the midsagittal line, ensuring consistency during the procedure. This indicates that similar conditions could be achieved in actual surgeries. However, the accuracy of the proposed method is not guaranteed if the scanning procedure deviates from these conditions. For example, excessive probe pressure or rapid sweeping speeds may introduce unpredictable motion components or image artifacts, reducing reconstruction accuracy. While the scanning procedure described in Section 2.3.3 was strictly followed in this study, we did not evaluate the performance of the reconstruction method under conditions that deviate from these guidelines. Investigating the method's robustness under non-ideal scanning conditions remains an important direction for future research.

2.5 Conclusion

In this study, we have developed a 3D-US reconstruction method that provides an accurate reference US image by compensating for prostate movement during probe sweeping. By utilizing the biplane TRUS probe's real-time, dual orthogonal cross-sectional imaging, the proposed method captures out-of-plane motion using the orthogonal image, enabling motion-corrected reconstruction of the prostate. In cadaver experiments using specimens with properties close to living tissues, the proposed method achieved accuracy comparable or superior to that of a position sensor-based approach. This success demonstrates the potential for clinical ap-

Table 2.1: Reconstruction accuracy of 3D US-image by the proposed method and optical position sensor using MRI as ground truth.

	Average Symmetric Surface Distance (ASD) [mm]	Dice Coefficient [%]
Proposed method	0.66	93.67
Position sensor	0.59	93.37

Table 2.2: Table of marker position error [mm] in cadaver experiments.

	Cadaver 1		Cadaver 2	
	Position Sensor	Proposed Method	Position Sensor	Proposed Method
Ave.	2.72	2.12	5.10	2.70
Max.	7.05	5.92	13.19	4.44
Std.	1.68	1.14	3.10	1.07

plication in sensor-free biopsy guidance and needle placement, offering a promising future for this method in the medical field.

The proposed method provides a sensor-free system that offers comparable accuracy to position sensors while significantly enhancing procedural efficiency. Position sensor-based systems can be cumbersome, as optical systems require careful avoidance of marker occlusion, and mechanical systems restrict the working area, complicating procedures. In contrast, our method streamlines the process without compromising accuracy, potentially allowing for more precise focal therapy and improving patients' quality of life post-treatment. This reassurance about the method's practicality can instill confidence in its potential use.

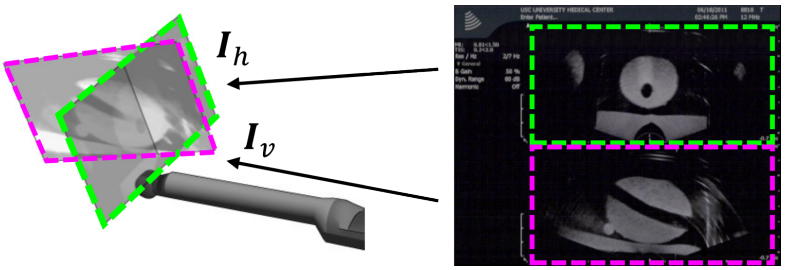


Figure 2.1: Biplane TRUS probe. Horizontal (green) and vertical (pink) cross-sectional images, which are perpendicular each other, are obtained simultaneously and displayed in real-time.

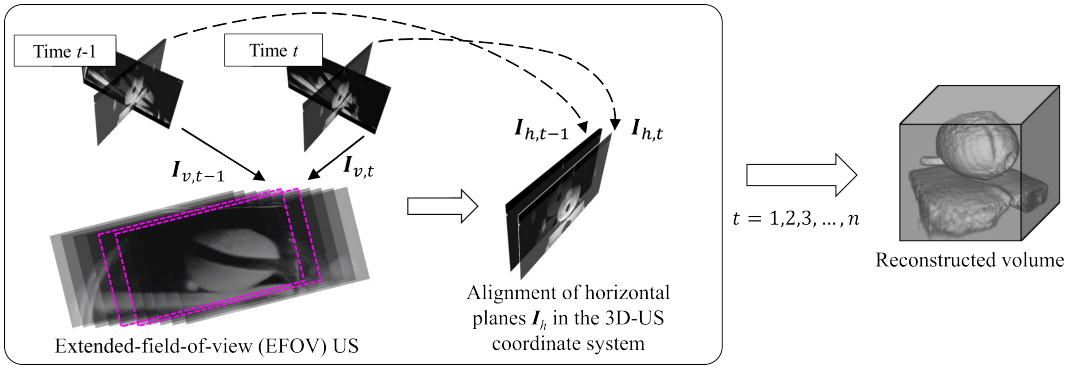


Figure 2.2: Overview of 3D-US reconstruction using biplane US probe. Temporally adjacent two vertical cross-sectional images are rigidly registered to reconstruct the EFOV-US image and their positional relations are used to localize the horizontal images and finally reconstruct 3D-US images. This procedure is performed under the assumption that the probe is swept so as to keep the vertical images on the mid-sagittal plane of the prostate. This assumption is quite realistic for experienced urologists.

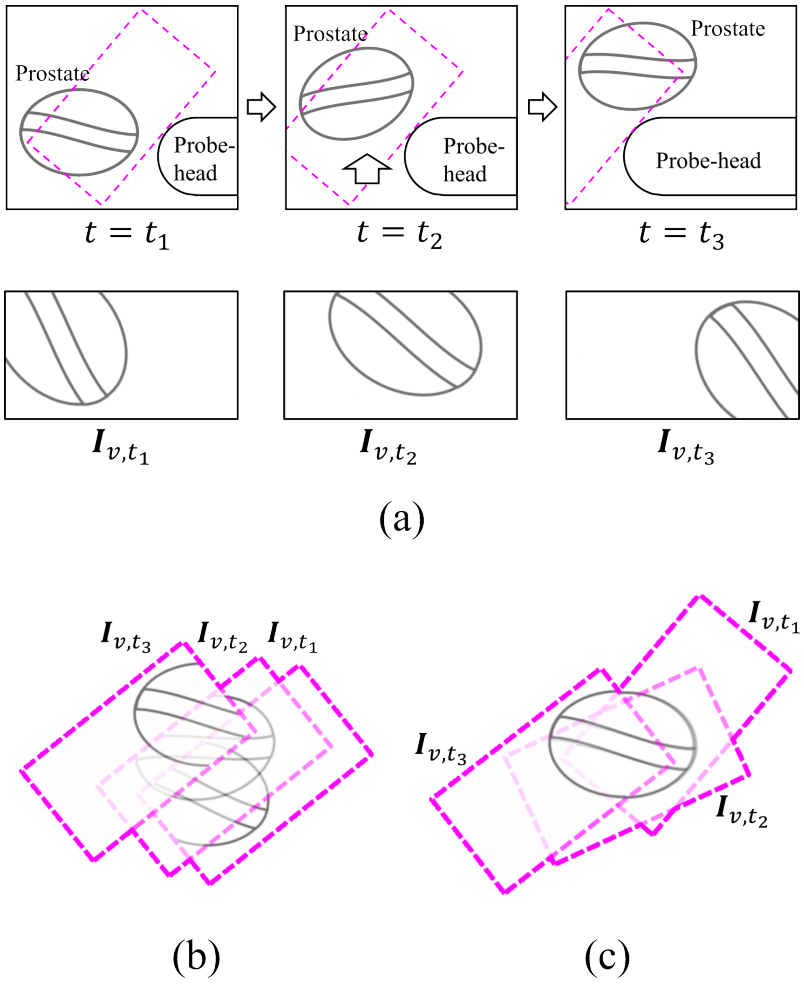


Figure 2.3: Schematic diagrams of compensation of unwanted effects of prostate motion. (a) Typical motion of the prostate caused by pressure of probe motion. (b) Spatial arrangement of vertical cross-sections in the position sensor coordinate system. (c) Spatial arrangement of vertical cross-sections in the proposed method, in which the prostate motion is corrected.

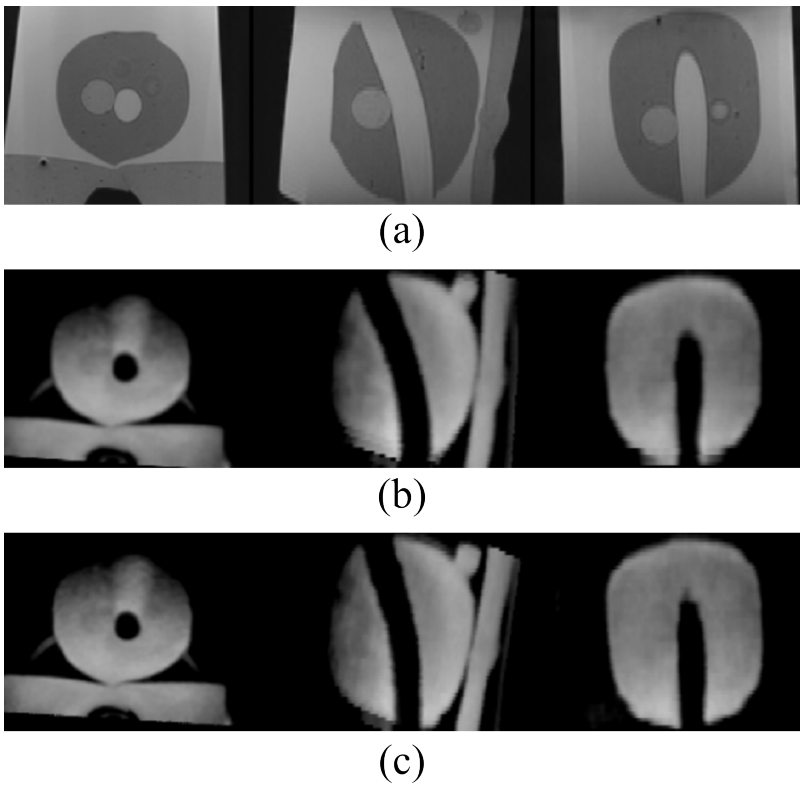
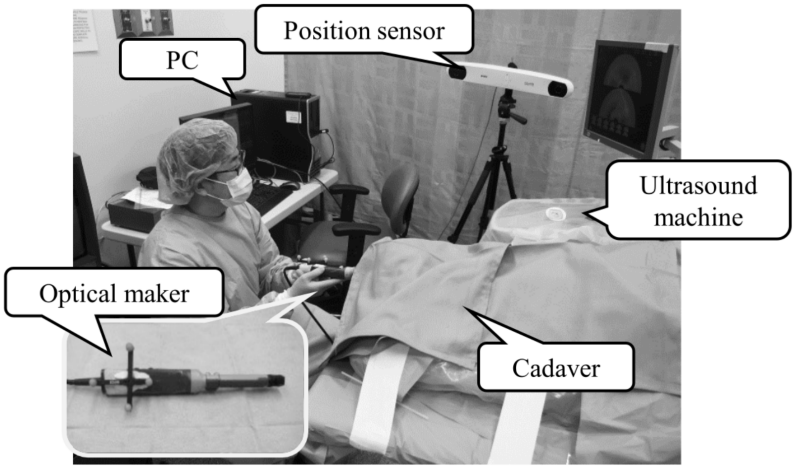
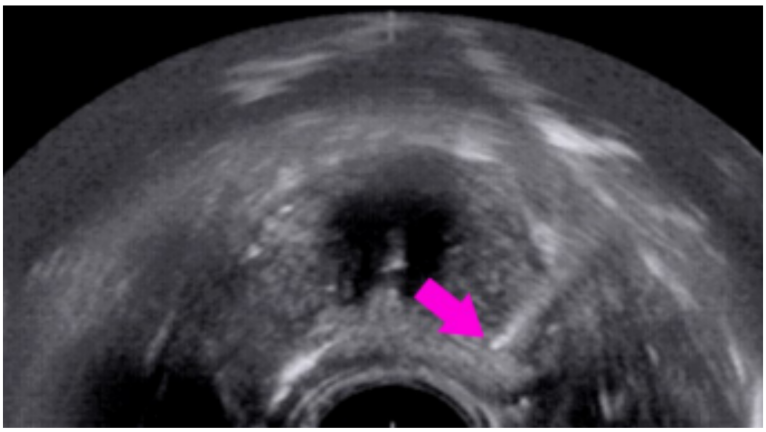


Figure 2.4: Comparison of reconstructed 3D-US images by phantom experiments. Three orthogonal cross-sectional images are shown. (a) MR images used as ground truth. (b) Reconstruction by the proposed method. (c) Reconstruction using optical position sensor.



(a)



(b)

Figure 2.5: Cadaver experiment. (a) Experimental setup. Optical markers were attached to the biplane probe. The optical camera of the position sensor was arranged so as to minimize occlusions of the markers. (b) Metal marker (shown by arrow) imaged in US image. Metal makers were embedded in the cadaver prostate.

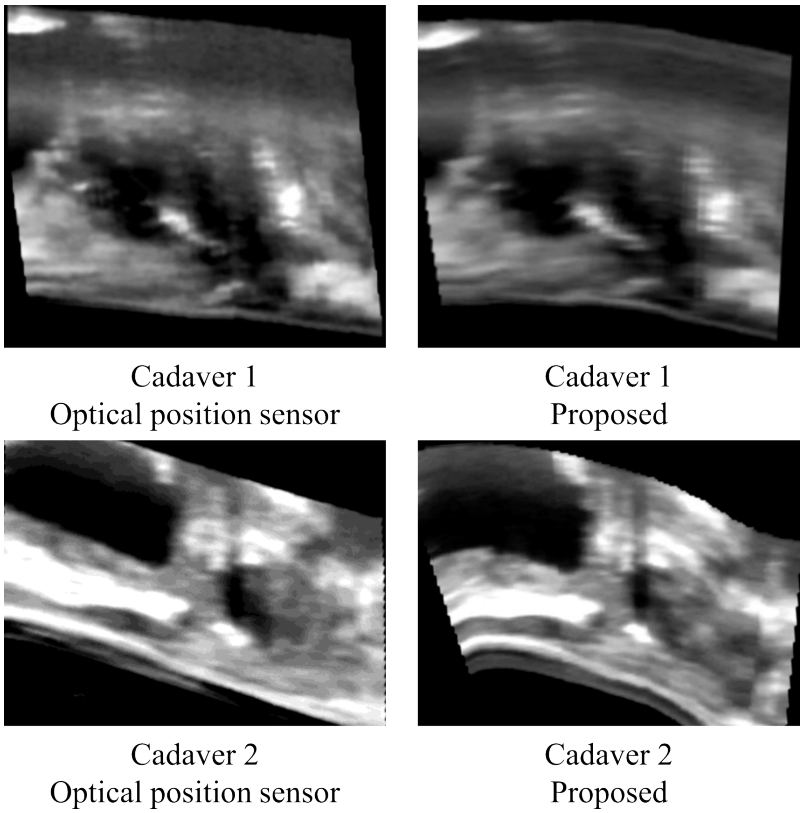


Figure 2.6: Comparison of reconstructed 3D-US images on the sagittal plane between proposed method and optical position sensor. See Fig. 7 for the prostate region in the reconstructed images. Curved parts are indicated by arrows.

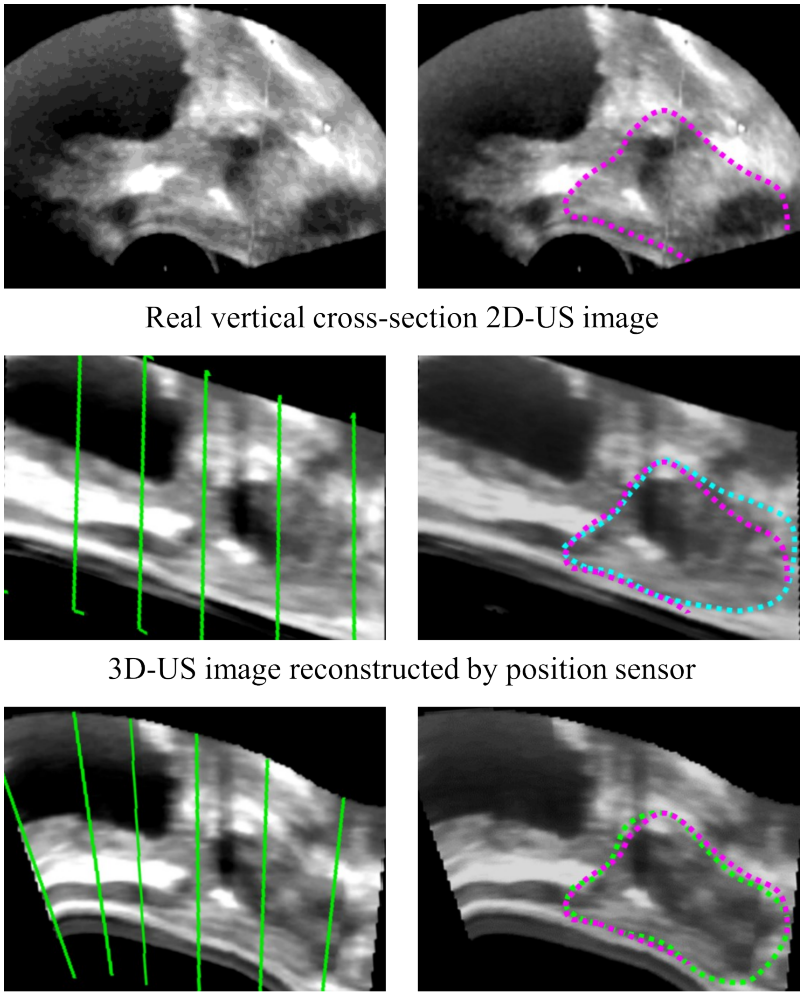


Figure 2.7: Comparison of sagittal sections of reconstructed 3D-US images and original vertical cross-sectional image of cadaver 2. Upper: Original vertical image and superimposed expert trace of the prostate on it. Middle: Position sensor-based reconstruction and comparison of prostate contours. Bottom: Proposed reconstruction and comparison of prostate contours. See texts for explatation of dotted contours in the cross-sectional images.

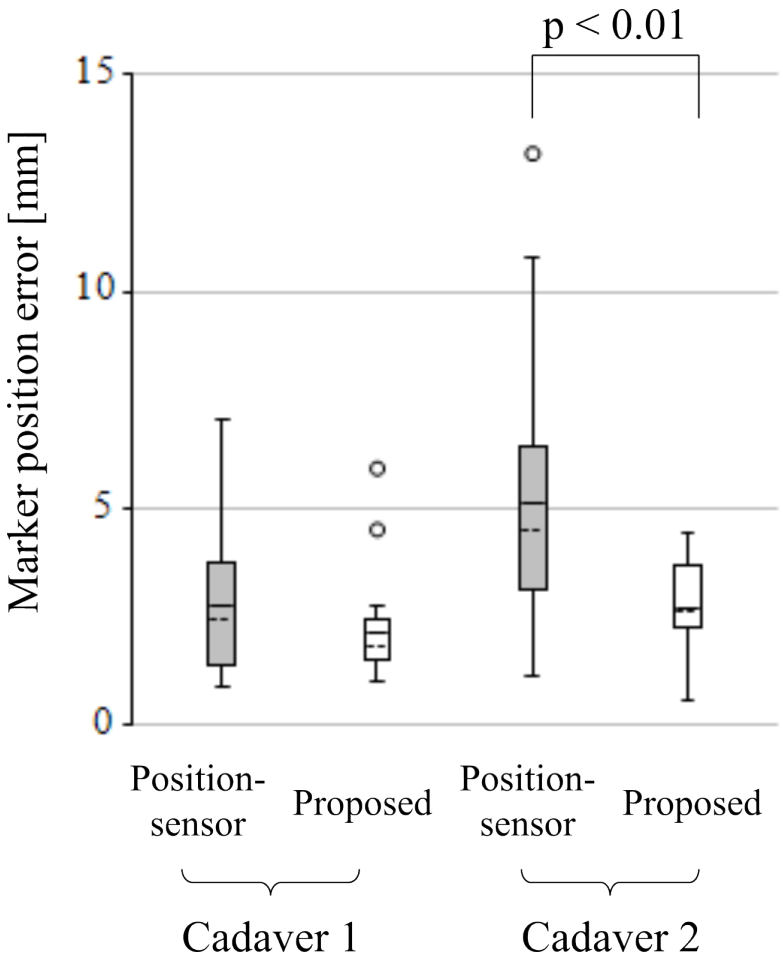


Figure 2.8: Box plots of marker position error in the cadaver experiments and results of statistical significance test.

Chapter 3

Efficient Musculoskeletal Shape Generation Through Controlled Deformation

Original title: Efficient Musculoskeletal Annotation Using Free-Form Deformation: A Novel Approach for Dataset Creation

3.1 Overview

Constructing training datasets for automatic muscle segmentation traditionally requires skilled operators, leading to high costs and limited scalability. To overcome these challenges, we developed a novel tool that allows non-experts to efficiently adapt a template three-dimensional (3D) anatomical model to target magnetic resonance images (MRI) using a controllable free-form deformation interface. This approach simplifies the annotation process by enabling independent adjustments in axial, sagittal, and coronal directions, allowing for simultaneous adaptation of all muscle structures. Experiments revealed that tool-assisted annotations by non-experts achieved a Dice coefficient exceeding 0.75

compared to expert annotations, with no significant errors such as mislabeling or omissions, demonstrating the high accuracy of our tool. Additionally, datasets generated with this tool enabled automatic segmentation networks to achieve performance comparable to or surpassing those trained on expert-annotated datasets. This method significantly enhances the efficiency and accessibility of high-quality dataset generation, addressing the need for scalable solutions in medical image analysis.

3.2 Background

Because the volume of muscle is an important determinant of its force-producing capacity, quantification of individual muscle volume in vivo is of considerable interest in the fields of health, medicine, and sports (Fukunaga et al., 2001; Arokoski et al., 2002). Muscle volume or cross-sectional area has been assessed by segmenting muscles using magnetic-resonance (MR) images (Fukunaga et al., 2001; Arokoski et al., 2002; Tracy et al., 1999; Barnouin et al., 2014; Pons et al., 2018; Ogier et al., 2020) or computed-tomography images(Momose et al., 2017; Liu et al., 2012; Rasch et al., 2007). Segmentation is typically performed manually by experts. However, manual segmentation (the task of tracing the contour of each muscle slice-by-slice) is time-consuming–particularly when the number of muscles is large.

Recently, automatic segmentation techniques based on deep neural networks and artificial intelligence (AI) have been revolutionized (Badrinarayanan et al., 2017; Ronneberger et al., 2015; Minaee et al., 2022). These technique have been applied to muscle segmentation: once the network is trained with a dataset of several dozen cases, it

can instantaneously recognize individual muscles from medical images with similar accuracy to an expert (Dice coefficient of 0.8 – 0.9) (Ni et al., 2019; Hiasa et al., 2020; Agosti et al., 2022; Cheng et al., 2022; Guo et al., 2021; Conze et al., 2020; Riem et al., 2023; Alipour et al., 2024; Wesselink et al., 2022). Such AI-based methods are promising; however, to train a network for different body parts or subject groups, manual segmentation or annotation tasks are necessary to create a training dataset. Thus, there is a growing need for techniques that assist in the annotation and efficient creation of datasets.

Several annotation-assistance methods, including classical techniques and recently developed AI-assisted methods, have been proposed for image processing. Classical techniques such as "graph cuts" and "snakes" help operators to trace the boundaries of objects of interest by utilizing the local intensity features of the image(Boykov and Jolly, 2001; Kass et al., 1988). However, these approaches often struggle to effectively annotate structures with unclear boundaries, such as muscles. In contrast, AI-assisted techniques provide an interactive framework where AI iteratively suggests segmentations based on expert corrections through simple scribbling or clicking, continuing until the expert is satisfied(Liu et al., 2022; Diaz-Pinto et al., 2022; Philbrick et al., 2019). This increases the efficiency of the annotation process but typically requires substantial pre-training and may not provide satisfactory results unless the operational conditions of the trained model (such as imaging modality and acquisition area) are satisfied.

In the aforementioned studies, researchers assumed that datasets should be created according to expert judgment. However, the reliance on experts inevitably incurs time and human-resource costs. Therefore, in

the present study, we developed a system that allows segmentation even by non-experts. A key innovation is allowing non-experts to annotate muscles by deforming a template 3D anatomical model that includes full musculoskeletal elements to fit the target MR image. If the template model can be deformed while maintaining the topology of internal structures, it can serve as an alternative to the anatomical knowledge typically provided by experts.

The manipulation of a complex object shape has been performed with free-form deformation (FFD) (Hsu et al., 1992), which is a powerful modeling tool for describing deformation by manipulating the locations of the control points. FFD manipulation is implemented in several modeling software programs, where users can intuitively deform an object (2D image or 3D surface data) by dragging control points with a mouse.

Although a single set of control points is normally established in FFD, handling these control points across three cross-sectional (axial, coronal, and sagittal) interfaces, which are commonly used for reading 3D medical images, is challenging (see Fig. 3.1). For example, when a user moves a control point on an axial plane (as indicated by the red arrows in Figs. 3.1a and b), the moved control point disappears from the planes in the other directions (as indicated by the cyan arrows on the coronal and sagittal planes in Fig. 3.1b), where it was originally located, making it difficult to track the control point.

We address this challenge by creating three independent sets of control points corresponding to the axial (red), sagittal (yellow), and coronal (green) directions (Fig. 3.2d). The entire deformation field is calculated by multiplying the three deformation fields. With respect to

a specific direction (e.g., axial), only the corresponding set of control points (e.g., red control points) is displayed and can be manipulated within that plane (Fig. 3.2a). In this setting, manipulating a control point on a certain plane (indicated by the red arrow in Fig. 3.2b) does not affect the control points associated with other planes (indicated by the yellow and green arrows in Fig. 3.2b). This approach allows the intricate deformation of the template through manipulation across all three planes (as demonstrated by the red, yellow, and green grids in Fig. 3.2c) without losing track of the control points.

The template model can be deformed by the entire deformation field, presented interactively to the user, and superimposed on the target MR image. The user is asked to deform the model by manipulating the control points to ensure that the model accurately fit the MR images. Although it is unrealistic to expect non-experts without anatomical knowledge to manually segment each muscle individually, graphically fitting the template model to the target MR image can be intuitively accomplished, even by non-experts. Importantly, once this task is completed, annotation of all the muscles in the template model is simultaneously achieved, making the process highly efficient.

In this study, we conducted several key tasks to ensure the effectiveness of our system. First, non-experts without specialized anatomical knowledge were engaged in segmentation of dozens of MR image data, and the system's effectiveness was verified. Second, we trained a deep neural network using the dataset obtained in this study and assessed its ability to perform automatic segmentation. Furthermore, we compared the performance of this network with that of a network traditionally trained using a dataset annotated by individuals with anatomical knowl-

edge. The objective of these investigations was to demonstrate how our system can streamline the creation of training datasets for AI-based automatic segmentation, optimizing both the time-efficiency and cost-effectiveness of the annotation process.

3.3 Materials and Methods

Image data collection

We collected MR images of the upper-extremity musculoskeletal system from 88 healthy male participants (age 22.7 ± 2.6 years, height 172.0 ± 5.2 cm, weight 64.0 ± 9.1 kg). The experiment was approved by the Ethics Committee of the National Institute of Information and Communications Technology (Approval Number: B160051702, Approval Date: January 18, 2018) and was conducted in accordance with the Declaration of Helsinki. All the participants provided written informed consent before participating in the study.

The imaging protocol utilized a 3-T MR system (MAGNETOM Vida, Siemens Healthineers) with a T1-weighted VIBE (Volumetric Interpolated Breath-hold Examination) Dixon sequence (acquisition plane: coronal; field of view: 448 mm × 448 mm; slice thickness: 1.5 mm; resolution: 1.0 mm × 1.0 mm × 1.5 mm; repetition time: 5.88 ms, echo time: 2.46 ms for in-phase image and 1.23 ms for opposed-phase image; flip angle: 5°). The scanning sequence was repeated three or four times by moving the bed from the neck to the wrist to cover the target region of the left upper limb. To reduce the distortion around the boundary of the field of view, the center of the image was moved by 196 or 248 mm with image overlaps of 250 or 200 mm, respectively. Three

or four images were reconstructed as having a large field of view using the composing software installed in the console application (Syngo MR XA10, Siemens Healthineers). Opposed-phase images were used in this study because they provide better visibility of the anatomical structures around the muscle attachment area.

3.3.1 Template model

We devised a novel segmentation support tool that utilizes the FFD of a template human musculoskeletal model. Two template models were used in this study: one for the left upper arm (Fig. 3.3a) and another for the left shoulder (Fig. 3.3b). Fig. 3.3c presents the 24 elements comprising the template models. The models were created by an expert who manually segmented the muscles and bones from a participant's MR image in a conventional manner using the 3D Slicer platform (Fedorov et al., 2012) over a period of > 50 h. They comprised 1-mm-grid 3D volume data, with each voxel having one of the 25 labels (24 elements and background).

3.3.2 Deformation of template model

Our system overlays the template model and the MR image to be segmented in each view of the three cross-sectional (axial, coronal, and sagittal) interfaces (Fig. 3.2). To allow the user to manipulate the control points without losing sight of them, we introduced three sets of control points (red, yellow, and green) corresponding to the axial, sagittal, and coronal directions, respectively (Fig. 3.2). Each set of control points (e.g., red points) is visible only in view of the corresponding plane (e.g., axial) and is allowed to move within the plane. This con-

straint ensures that all the control points are on the prescribed planes, preventing users from losing sight. The system interactively displays deformations in response to the user actions conducted thus far in all three directions. During this process, the template model is deformed by the entire deformation field, which is calculated by multiplying the three individual deformation fields (Fig. 3.2d). Each field is determined by the current locations of the control points in each direction. Notably, the FFD_{whole} is a function provided by the Visualization Toolkit (VTK) library(Schroeder et al., 2006), where three FFDs are concatenated. Internally, this is equivalent to sequentially applying each FFD in a cascade manner.

3.3.3 Implementation and workflow

We developed a segmentation tool (Seg Muscle) as a plugin for the widely used 3D Slicer platform (Fedorov et al., 2012) – a powerful medical image processing software program. Thus, users can leverage the extensive functionality of 3D Slicer while benefiting from the unique features of our plugin. The typical workflow for the segmentation task using our plugin in 3D Slicer is as follows:

1. The user loads the target medical image into 3D Slicer.
2. The user selects our plugin, chooses the appropriate template anatomical model, and then positions it over the target image using intuitive mouse operations.
3. The user manipulates the control points on the template model to fit the target image.

4. As the user adjusts the control points, the software updates the shape of the template model in real time, providing immediate feedback on the quality of the fit.

5. Once satisfied with the fit, the user saves the results, and the process is complete. If not satisfied, the user can employ the current deformation result as the initial value and then finely adjust the fit using denser control points in the subsequent step.

3.3.4 Tool-assisted segmentation

Non-experts performed segmentation tasks on the data from 87 participants using our tool. These non-experts were four individuals experienced in using general office applications on a personal computer but with limited anatomical knowledge. They were instructed to initially fit the template model in a certain direction (e.g., axial), then adjust it in the next direction (e.g., coronal), and finally adjust it in the remaining direction (e.g., sagittal). They were asked to perform at least two rounds of this process for each participant's data. Deforming the template model while preserving its topology has the advantage of allowing non-experts to effectively perform segmentation. However, fine-tuning each muscle in our system is difficult because manipulating a control point simultaneously changes the shape of multiple muscles. Therefore, to avoid wasting time, we informed the operators that fine-tuning the individual muscles was not necessary. No time restrictions were imposed. On average, segmentation took approximately 2 h per participant, with the shortest time being approximately 30 min and the longest approximately 4 h.

3.3.5 Refinement of segmentation

We also created conventional manual segmentation data. These data were created by medical students, physical therapists, or equivalents with a basic understanding of musculoskeletal anatomy and experience in reading medical images. This process refined the results of the tool-assisted segmentation described above. They manually modified the boundary of each muscle individually using the 3D Slicer platform, without our plugin. They were instructed to complete the task within 1 h per element. On average, about 20 hours were spent per participant's data. After this process, an expert with approximately eight years of experience in manually segmenting muscles from medical images reviewed and further adjusted the data. In total, data for 20 participants were created in accordance with a previous study in which a deep neural network with a dataset of 20 participants exhibited good classification performance (Hiasa et al., 2020).

3.3.6 Dataset for network training

The dataset used for training the network consisted of MR image data and corresponding label data. The MR image data comprised 400 – 450 axial images per participant, each having dimensions of 448 mm × 264 mm, with a 1-mm gap. These images covered all the elements listed in Fig. 3.3c. The label data consisted of 1-mm-grid 3D volume data, where each voxel was assigned one of 23 labels (22 elements and background), in which the three parts of the pectoralis major were combined. The following three datasets were used.

1. Expert–20: This is a conventional manual segmentation dataset

consisting of the refined label data of 20 participants. It served as the ground truth because of the involvement of an expert in the process.

2. Tool–20: This dataset consists of tool-assisted segmentation labels for the same 20 participants as the Expert–20 dataset. It allowed comparison between the manual and tool-assisted segmentation methods for the same set of participants.

3. Tool–87: The third dataset consisted of tool-assisted segmentation labels for all 87 participants. This dataset provided insights into the effects of applying the tool-assisted method to a larger number of participants.

3.3.7 Automated segmentation network

We employed a 2D U-Net with an added dice loss function, which is widely used in deep learning(Milletari et al., 2016; Isensee et al., 2021). The 2D U-Net was chosen because it is one of the most widely adopted architectures for medical image segmentation problems. This selection allows us to highlight the performance differences attributable to the quality of the datasets (Expert–20, Tool–20, and Tool–87), using a widely recognized and standardized architecture. For implementation, we used the PyTorch framework. In addition, we applied data augmentation during training using the albumentation library.

The training parameters were as follows: a learning rate of 0.0001, the Adam optimization algorithm, a batch size of 8, 30 epochs, mixed-precision training enabled, learning rate scheduling with ReduceLROnPlateau, and early stopping if there was no improvement in the vali-

dation metric for five consecutive evaluations conducted 10 times per epoch across epochs. The experiments were performed on a computer with the following specifications: CPU, Intel(R) Xeon(R) E5-2698 v4 @ 2.20GHz; GPU, Tesla V100 DGXS 32GB; RAM, 256GB; OS, Ubuntu 18.04.

We trained the network using one of three datasets to obtain a classifier capable of classifying every voxel into 23 labels (22 elements and background) for a given set of MR image data. To evaluate the performance of the classifier, we employed 5-fold validation. For the Expert-20 and Tool–20 datasets, the classifier was trained using data from 16 participants in each fold. The classification results for the data of the remaining four participants in each fold were compared with the ground truth. The procedure for the Tool–87 dataset was similar to that for the Tool–20 dataset. However, for the Tool–87 dataset, the classifier was trained using the data from 83 participants in each fold. This included the same 16 participants as the Tool–20 dataset and 67 additional participants. The agreement between the classification results and ground truth was quantified using the Dice coefficient.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, we used the GPT-4 model developed by OpenAI to improve readability and language. After using this tool, the authors reviewed and edited the content as required. We take full responsibility for the content of the manuscript.

3.4 Results

Quality of tool-assisted segmentation

Figure 3.4 shows representative axial cross-sectional images comparing the initial template labels (Fig. 3.4a), tool-assisted labels from non-experts (Fig. 3.4b), and ground-truth labels from experts (Fig. 3.4c) for the same participant data. The initial template labels (Fig. 3.4a) indicate a state in which only the location is roughly aligned with the MR image. Because no deformation was performed yet, the agreement between the initial template labels and ground-truth labels was poor; quantitative analysis using 20 participants' data revealed that Dice coefficient was 52.6% ± 6.53% (Fig. 3.4e). Figure 3.4b shows the state after segmentation by a non-expert using our tool. Importantly, tool-assisted labels do not include beginners' mistakes, such as adjacent muscles being labeled as opposite to each other, which is attributed to the tool's ability to deform the template while maintaining the topology. In the visual inspection, the tool-assisted labels (Fig. 3.4b) exhibited good agreement with the ground-truth labels (Fig. 3.4c) for most of the elements, although there were mismatches for a few muscles, e.g., the pectoralis minor (red arrow in the upper panel of Fig. 3.4b) and the medial part of the triceps brachii (red arrow in the lower panel of Fig. 3.4b). Quantitative analysis of data from 20 participants indicated that the Dice coefficient increased to 78.2% ± 4.38% with tool-assisted segmentation by non-experts (Fig. 3.4e).

3.4.1 Network training and classification result

The network was trained using each of the three datasets to obtain the classifier. The training time for N = 20 and N = 87 datasets was approximately 1.5 and 9 h, respectively. The obtained classifier could classify muscles and bones from the MR images of one participant within approximately 30 – 40 s (35 ms per slice), excluding file loading. Figure 3.5 shows representative classification results for each classifier. Consistent with previous studies, the classifier trained with the Expert–20 dataset exhibited good performance (compare Figs. 3.5a and d), although there was an underestimation for a few muscles (see the red arrow in Fig. 3.5a). The classifier trained with the Tool–20 dataset exhibited discrepancies for several muscles (red arrows in Fig. 3.5b), likely owing to its relatively coarse segmentation compared with the Expert–20 dataset. Interestingly, training with a larger dataset (i.e., Tool–87) mitigated these discrepancies (see Fig. 3.5c).

Figure 3.6 quantitatively illustrates the differences in classification performance among the three classifiers. Figure 3.6a shows the Dice coefficient between the classification results and ground truth for the 22 musculoskeletal elements averaged from the data of 20 participants. Although there were differences in the Dice coefficient among different muscles, it was consistently smaller for the Tool–20 classifier than for the Expert–20 classifier; however, it increased from Tool–20 to Tool–87. The data averaged over 22 elements are presented as box plots in Fig. 3.6b. The average and standard deviation of the Dice coefficient for the Expert–20, Tool–20, and Tool–87 classifiers were 78.6% ± 6.99%, 72.1% ± 9.86%, and 80.7% ± 2.78%, respectively. Wilcoxon signed-rank tests with Bonferroni correction indicated that the Dice co-

efficient for the Tool–20 classifier was significantly smaller than that for the Expert-20 classifier (Z = 3.92, corrected P < 0.001) and that it increased from Tool–20 to Tool–87 (Z = 3.92, corrected P < 0.001), surpassing that of Expert–20 (Z = 2.54, corrected P $= 0.033$).

3.5 Discussion

Effectiveness of the proposed tool

Our study highlights the effectiveness of the proposed tool in allowing non-experts to perform segmentation. It is important to note that non-experts can rapidly generate accurate labels without conspicuous errors such as mislabeling of adjacent muscles or omission of musculature. Although the segmentation quality was relatively coarse compared with the precision achieved by experts, the quality (Dice coefficient of approximately 78%) reached a level similar to that of the network trained with Expert–20. Asking non-experts to manually segment each muscle individually is unrealistic; however, our results suggest that they can effectively fit the template model to the target MR image graphically, which serves as a form of segmentation. Another advantage of this template-model approach is that once the deformation is complete, the segmentation of all the muscles is accomplished simultaneously. Typically, the time required for manual segmentation is proportional to the number of elements involved; however, with the proposed tool, the number of muscles ceases to be a limiting factor. Consequently, our tool reduces labor costs and enhances scalability.

However, it is also important to consider the limitations regarding the generalizability of the template. While a single template successfully

facilitated segmentation across 87 cases in this study, this was due to the relatively uniform characteristics of the subject group. In scenarios where there is greater variability among subjects, such as differences in disease conditions or age groups, it may be necessary to prepare multiple templates to achieve accurate segmentation. Furthermore, while this study addressed variations in limb positioning by handling the shoulder and arm separately, more extreme variations in limb positioning might require the creation of new, dedicated templates. Even in such cases, the overall labor cost with the tool remains lower compared to manual segmentation without the tool.

3.5.1 Effect of training data accuracy and quantity on classification performance

Although the proposed tool has many advantages, it is not without limitations. Because the design is aimed at preserving the topology of the template, the generated labels may not perfectly capture the detailed shape of an individual. In fact, there was a tendency for more under- or over-segmentation for the tool-assisted labels than for the ground-truth labels. Accordingly, when using an equivalent amount of participant data, the Tool–20 classifier exhibited significantly inferior performance to the Expert–20 classifier. However, as demonstrated by the Tool–87 classifier, increasing the number of participants can increase the accuracy beyond that of the Expert–20 classifier. This outcome implies that with regard to network performance, accumulating more data can compensate for issues in data quality. Of course, gathering a large amount of high-quality data is ideal. However, considering that this is expensive and time-consuming, our findings suggest that there is also a viable strategy for acquiring a large amount of data that can be easily col-

lected, even if the quality of the data is slightly inferior. Although it is an estimate, Expert-20 required ≥ 400 h of expert time, whereas Tool–87 needed approximately 178 h of non-expert time. This indicates that tool-assisted segmentation is a viable option.

However, there is no assurance that gathering large amounts of relatively coarse data will always be beneficial. If the errors in the coarse dataset are biased in a certain direction relative to the true values, the network trained with this dataset could mistakenly recognize the muscle in that direction. Thus, the increased accuracy observed in this study suggests that the errors in tool-assisted labels are uniformly distributed around the true value. This uniform error distribution may be due to the tool's constraints. The limited manipulability from the coarse grid and the design that deforms all labels simultaneously likely reduce arbitrary bias. Overestimation in one area tends to be balanced by underestimation in an adjacent area, resulting in a more evenly spread error. While these tool characteristics might explain the observed error distribution, this hypothesis requires further investigation. Although verifying this is beyond the scope of this study, researchers should be aware of these potential benefits when using the tool. Future work could explore this hypothesis further to provide more definitive guidelines and detection mechanisms.

3.5.2 Opportunities for the tool usage

Although networks for automatic muscle segmentation have been developed, an important consideration is the need for a segmentation dataset of body parts. To model the entire body, segmentation data for muscles throughout the body are essential. To date, most studies have focused on

the lower limbs (Ni et al., 2019; Hiasa et al., 2020; Agosti et al., 2022; Cheng et al., 2022; Guo et al., 2021), and a few studies have focused on the upper limb (Conze et al., 2020; Riem et al., 2023; Alipour et al., 2024) and trunk (Wesselink et al., 2022). Thus, to realize the automatic recognition of the entire body, we still need to create a segmentation dataset for these parts. Additionally, the accuracy of recognition can be influenced by not only the body part but also the attributes of the individual, such as age, sex, and pathology(Ogier et al., 2021; Conze et al., 2020). Hence, to build a universal network, a dataset encompassing a wide range of attributes is necessary, whereas to build a specialized network targeting specific groups, datasets tailored to these attributes are needed. Therefore, manual segmentation remains a critical task in many instances, for which our tool can be effectively utilized.

3.5.3 Modality-free approach

An AI-assisted segmentation tool is also available for dataset creation(Liu et al., 2022; Diaz-Pinto et al., 2022; Philbrick et al., 2019). With the assistance of AI, users can interactively perform segmentation using simple operations. However, the AI must be trained on the modalities and body parts of the data being analyzed. In contrast, our tool offers a modality-free approach for both the template model and segmentation target image. As the template model is merely labeled volume data, where each voxel is assigned one of the musculoskeletal element labels, it can be created from any image source. For instance, a template can be generated from a specialized source such as the Visible Korean Human dataset(Park et al., 2006), which contains high-resolution and specialized images. Once created, regardless of the modality used in

its development, a template can be used in our tool to segment various 3D medical images, including MRI, CT, and 3D-US, further broadening the applicability of the tool. Although we used templates created from MR images to segment the MR images in this study, our tool is flexible, allowing it to accommodate a wide range of medical imaging scenarios and requirements.

Considering that the dual-modality-free nature of this tool significantly broadens its range of applications and enhances its value to the medical and image-processing communities, we have provided the tool as a plugin for 3D Slicer. It can be used free of charge for academic or non-commercial purposes.

3.5.4 Comparative evaluation of tool-assisted vs. manual segmentation

While the proposed tool has demonstrated its utility in enabling non-experts to create segmentation labels, the scope of comparisons performed in this study is limited. Specifically, we compared tool-assisted segmentation by non-experts with manual segmentation by experts but did not evaluate two additional potential groups: tool-assisted segmentation by experts and manual segmentation by non-experts. This limits our ability to independently assess the effects of tool usage and expertise level on segmentation quality.

Notably, non-experts' manual segmentation of individual muscles is virtually impossible without extensive training. To perform such a task, a non-expert would require prolonged and intensive guidance from a skilled expert, effectively elevating their proficiency to at least an intermediate level. At that point, the individual could no longer be clas-

sified as a non-expert. Consequently, segmentation labels created by non-experts without using the tool do not exist, making experiments involving this group infeasible. Conversely, based on observations during preliminary experiments, expert tool-assisted segmentation is expected to produce higher-quality labels in less time than non-expert tool-assisted segmentation. However, when used to generate training datasets for deep learning, our findings suggest that the performance of networks trained with tool-assisted labels created by non-experts already surpasses those trained with manually segmented labels created by experts. This raises questions about whether the additional cost of involving experts with the tool would result in sufficient performance gains to justify the effort.

Future research should address these gaps by systematically evaluating the effects of expertise and tool usage on segmentation quality and network performance. Additionally, cost-effectiveness should be a central consideration, as the current results suggest that the proposed tool offers significant advantages in terms of scalability and efficiency, even when used by non-experts.

3.5.5 Future research directions

Future research should focus on refining the algorithm to increase the label accuracy while maintaining the modality-free nature and cost-effectiveness of the tool. Integrating the proposed tool with other segmentation techniques and evaluating its performance on diverse medical imaging datasets can verify its efficacy and broaden its applicability. In addition, developing a more advanced network that leverages the characteristics of the datasets created by our tool, such as the error variance,

is crucial for achieving more precise recognition.

Moreover, the development of techniques for constructing personalized musculoskeletal models based on the shapes recognized by networks is vital. This is because the diagnosis of motor functions and recommendations for physical training programs rely on not only recognized shapes but also the ability of the musculoskeletal model to simulate individuals' movements. A pipeline was developed to create personalized skeletal models from automatically recognized bone geometries (Modenese and Renault, 2021) using OpenSim (Delp et al., 2007). Efforts are also being made to create finite-element muscle models (Blemker and Delp, 2005) and fiber models (Modenese and Kohout, 2020) from the three-dimensional muscle geometries. In the future, rather than developing recognition technologies and model construction techniques separately, it will be necessary to develop them considering that both should function complementarily.

3.6 Conclusions

We addressed the challenges of medical image annotation — particularly in musculoskeletal structures — by developing an innovative tool that allows non-experts to efficiently generate accurate labels. Our results indicated the effectiveness of the proposed system for reducing the time and cost associated with the annotation process while maintaining adequate quality for deep learning-based automatic segmentation. Furthermore, the versatility and adaptability of our system across different imaging modalities and anatomical structures were demonstrated, highlighting its potential to contribute to the advancement of medical image analysis. By providing a practical solution to the challenges of annota-

tion, our research can support the development of systems for evaluating individual musculoskeletal functions by constructing subject-specific musculoskeletal models from medical images.

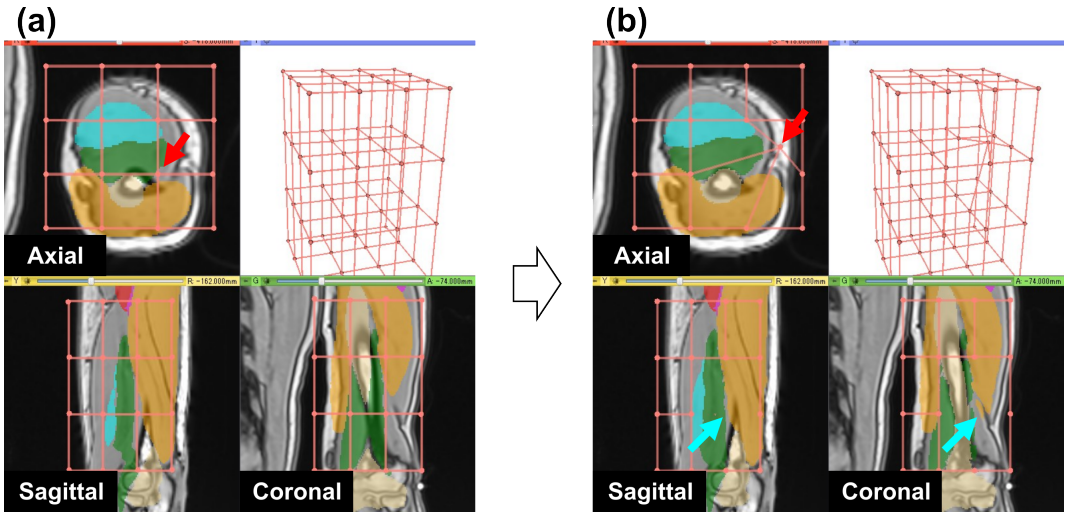


Figure 3.1: Difficulty in FFD with one set of control points. (a) One set of control points is shared for axial, sagittal, and coronal views. (b) If a user moves a control point in a certain plane (red arrow), that point disappears from the other plane (cyan arrows).

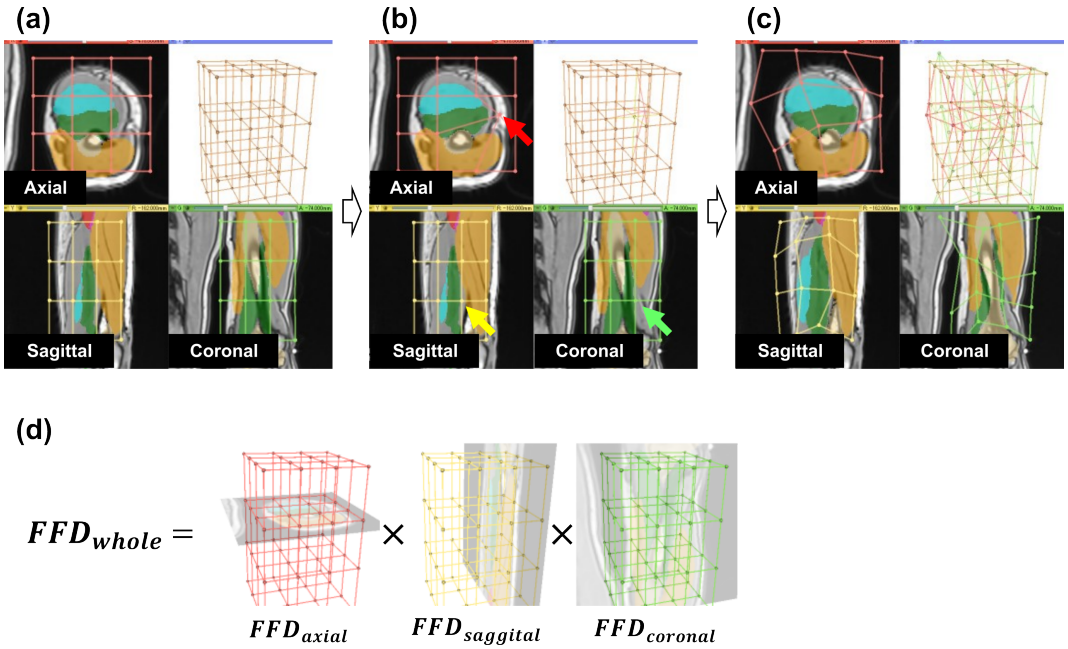
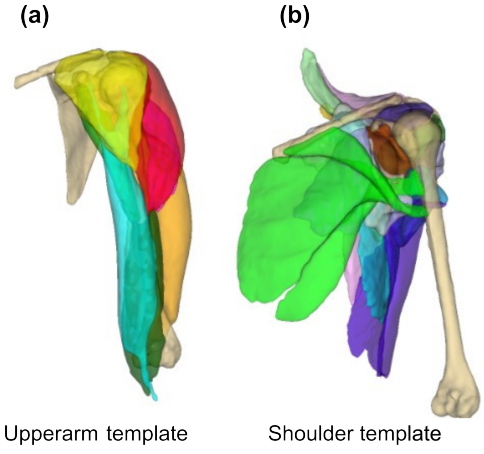


Figure 3.2: FFD based on three independent sets of control points. (a) Three different sets of control points are set for the template model. (b) Manipulating a control point on a certain plane (red arrow) does not affect the control points associated with other planes (yellow and green arrows). (c) Users can deform the template from the other planes also (yellow and green) without losing sight of the control points. (d) Whole deformation field is calculated by multiplying the three deformation fields.



(c)

Element name	Colors	Upperarm template	Shoulder template
Humerus		✓	✓
Clavicle		✓	✓
Scapula		✓	✓
Biceps Brachii		✓	
Brachialis		✓	
Coracobrachialis		✓	
Deltoid Acromial		✓	
Deltoid Clavicular		✓	
Deltoid Spinal		✓	
Triceps Brachii		✓	
Infraspinatus			✓
PectoralisMinor			✓
Subscapularis			✓
Supraspinatus			✓
Teres Minor			✓
Teres Major			✓
Trapezius			✓
Latissimus Dorsi			✓
Serratus Anterior			✓
Pectoralis Major (Abdominal)			✓
Pectoralis Major (Sternum)			✓
Pectoralis Major (Clavicle)			✓
Rhomboid			✓
Elevator Scapulae			✓

Figure 3.3: Template model for the upper-extremity musculoskeletal system: (a) left upper arm; (b) left shoulder; (c) name and color of each element.

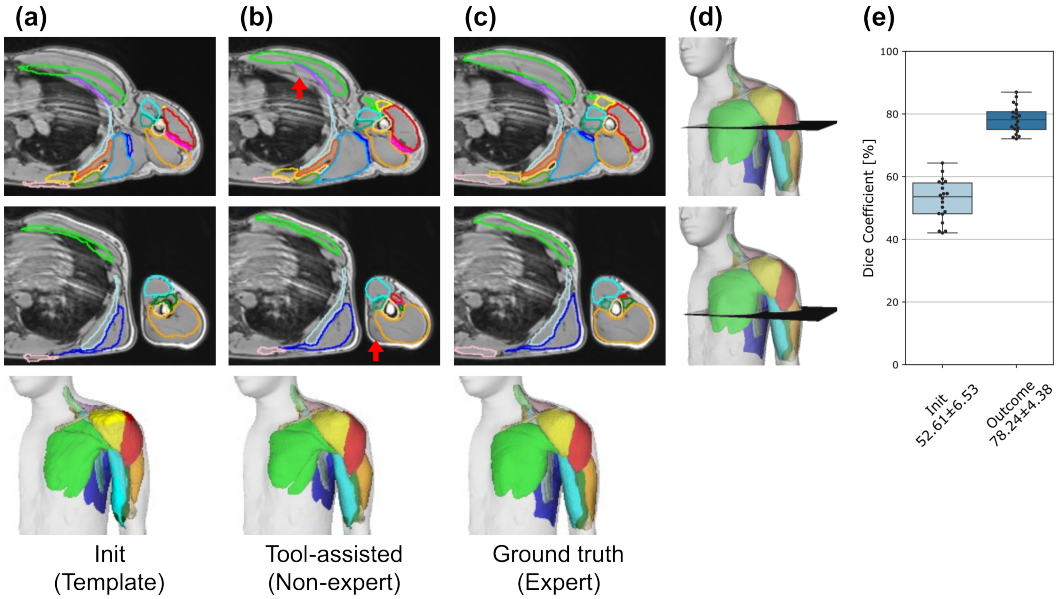


Figure 3.4: Figure 4. Results of tool-assisted segmentation. Representative axial cross-sectional images with (a) initial template labels, (b) tool-assisted labels, and (c) ground-truth labels for the same MR image. (d) Locations of the axial cross-sectional images. (e) Dice coefficient comparing initial template labels and tool-assisted labels with ground-truth labels.

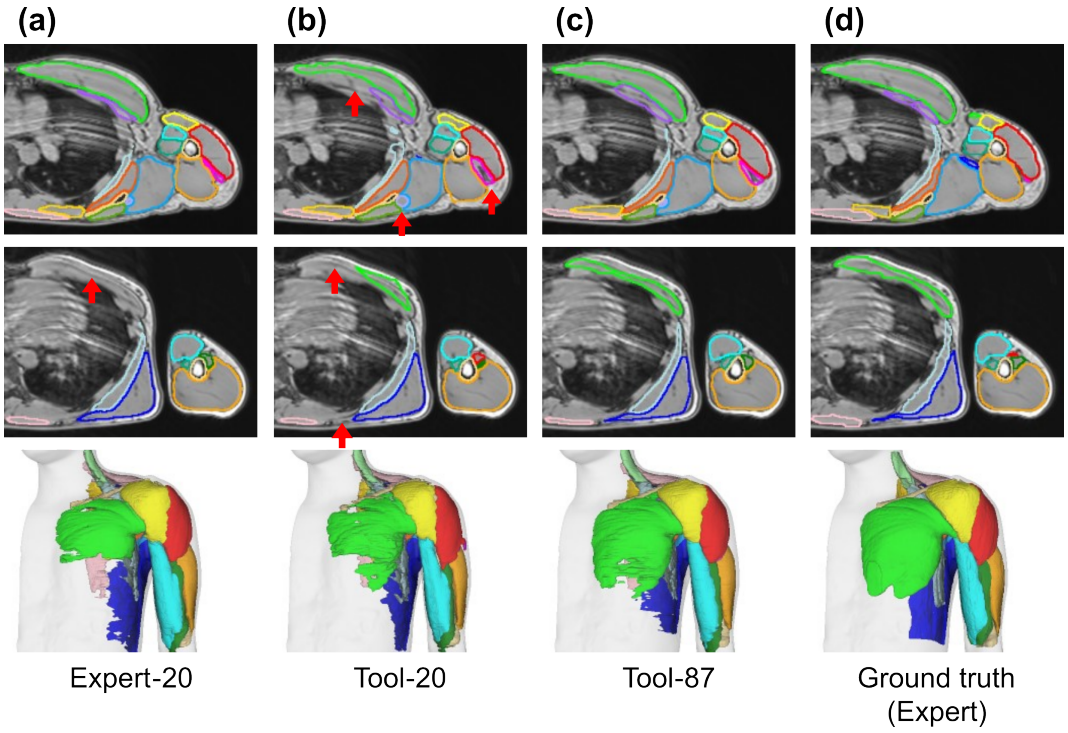


Figure 3.5: Segmentation results for three classifiers. Representative axial cross-sectional images comparing the classification results for the (a) Expert–20, (b) Tool–20, and (c) Tool–87 datasets and (d) the ground truth for the same MR image.

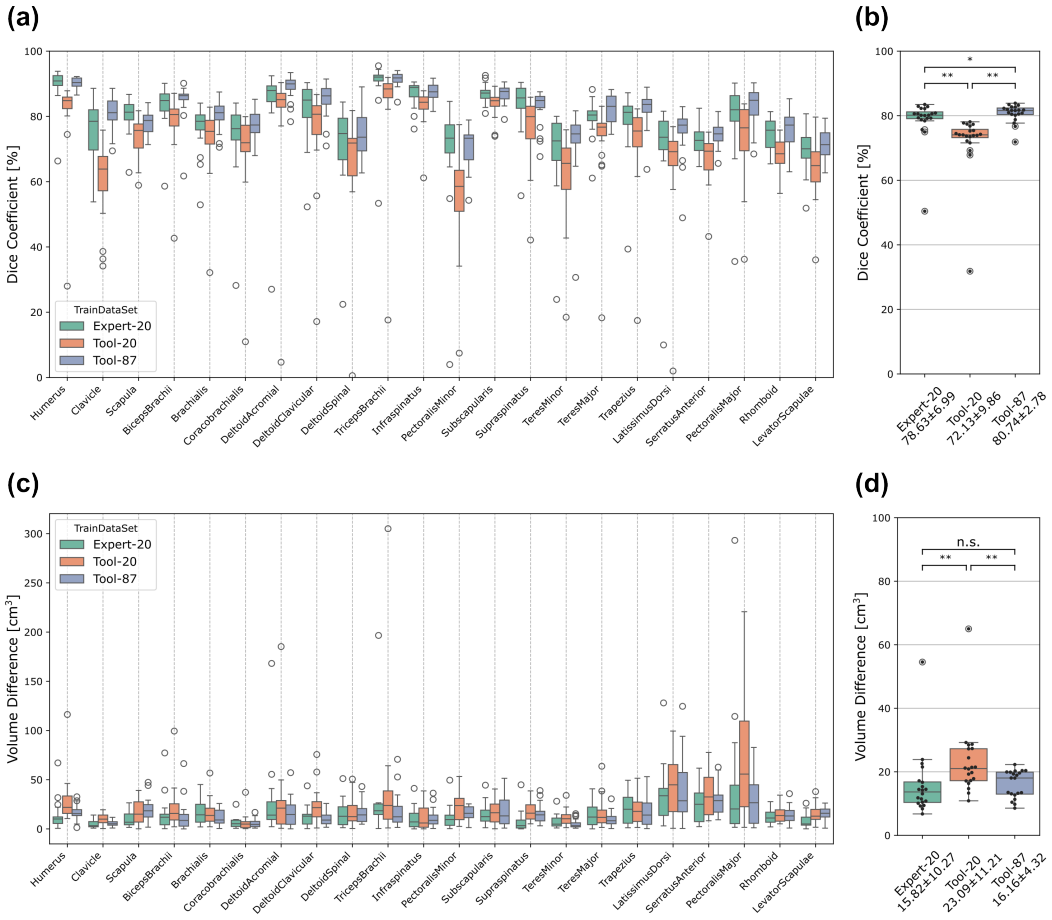


Figure 3.6: Classification performance for the Expert–20, Tool–20, and Tool–87 classifiers. (a) Dice coefficient between the classification result and the ground truth for the 22 musculoskeletal elements averaged among the 20 participants' data. (b) Each dot corresponds to individual participant data averaged over 22 elements. Boxes denote the 1st and 3rd quartiles, and the median is marked with a horizontal line in each box.

Chapter 4

Statistical Modeling of Anatomical Variations for Personalized Muscle Attachment Estimation

Original title: Estimation of attachment regions of hip muscles in CT image using muscle attachment probabilistic atlas constructed from measurements in eight cadavers

4.1 Overview

Patient-specific musculoskeletal modeling is critical for preoperative planning and postoperative orthopedic surgery and rehabilitation evaluation. However, identifying muscle attachment regions, often invisible in CT and MRI, remains challenging. This study introduces a method to estimate patient-specific muscle attachment regions using statistical modeling based on cadaver data. Muscle attachment probabilistic atlases (PAs) were constructed by integrating attachment data from eight cadavers onto an average bone surface via non-rigid registration. These atlases enable the estimation of patient-specific attachment regions by

mapping average regions to patient bone shapes derived from CT scans. Experimental results showed that the proposed method improved Dice similarity coefficients by over 10 % and reduced boundary distance errors by an average of 1.1 mm compared to previous methods. By effectively capturing inter-individual variations and enabling high-precision estimation of attachment regions, this method advances the understanding and reconstruction of localized anatomical features, addressing critical challenges in musculoskeletal modeling.

4.2 Background

One of the important application areas of musculoskeletal modeling is biomechanical simulation for assessment of surgical outcome in orthopaedic surgery and rehabilitation medicine (Delp et al., 1994; Piazza and Delp, 2001; Reinbolt et al., 2009). A string muscle model has been widely utilized such as "OpenSim" by Stanford Univ (Delp et al., 2007) or AnyBody Technology (Damsgaard et al., 2006), etc. However, these models employ a generic anatomy model and do not reflect the patient specific anatomy. In order to obtain a patient specific musculoskeletal model that allows biomechanical simulation, it is crucial to locate the regions that each muscle attaches to the bones.

A difficulty in application of the patient-specific biomechanical simulation comes from the fact that the muscle attachment regions are typically invisible in CT and MR images. Several previous studies conducted cadaver experiments to directly record the attachment regions using a three-dimensional (3D) position sensor during dissection (Van der Helm et al., 1992), (Carbone et al., 2015), (Horsman et al., 2007). Carbone et al reported TLEM 2.0 which is a dataset for mus-

culoskeletal modeling of lower extremity including the attachment sites of 55 muscles and bone surfaces measured from one fresh male cadaver (Carbone et al., 2015). Ito et al. measured the attachment regions of the short external rotators on twenty hips and analyzed inter-subject variations via viewpoint normalized photograph of specimens (Ito et al., 2012). They reported that attachment sites of the conjoined tendon and piriformis have considerably larger variability among individuals than previously thought. However, estimation of the attachment sites from medical images was not addressed.

In current imaging technologies, in-vivo direct measurement of the patient-specific muscle attachment regions is not reliable from medical images. In several studies, therefore, cadaver experiments were conducted for their measurement, and then the measurement results were used to estimate patient-specific attachment regions of an unseen subject. Kaptein et al estimated the attachment regions with a morphing-based method using landmark registration of implanted metal screws in three cadavers (Kaptein and Van der Helm, 2004). Pellikaan et al employed the same approach using registration between the vertices on bone surface in two cadavers (Pellikaan et al., 2014). They performed pair-wise cross validation experiments between two specimens, and their focus did not include inter-subject variations of the muscle attachment regions.

We propose a framework for statistical modeling of the muscle attachment regions from measurement of multiple cadavers and its application to estimating patient-specific muscle attachment regions using bone shapes derived from a patient CT image. In order to model the inter-subject variations in 3D, we use probabilistic atlas (PA), which

has been typically used for representing variations of organ locations and shapes. (Mazziotta et al., 2001), (Park et al., 2006) The attachment regions of each cadaver specimen are measured in 3D and normalized to a surface of an average bone shape by using a non-rigid registration so that the PA of muscle attachment regions is represented on an average bone surface. We evaluate the estimation accuracy of the attachment areas using PA.

Contributions of this paper are as follows; 1) construction of PA of the muscle attachment regions, 2) application of the PA to estimation of patient-specific attachment regions from medical images, and 3) evaluation of the proposed framework using the data physically measured in eight cadavers. We plan to contribute the PA and the average bone shape to registered users in order to allow patient-specific biomechanical simulation for a wider user-base.

4.3 Materials and Method

4.3.1 Data measurement

Figure 4.2 shows a workflow of the process of collecting the ground truth data of muscle attachment regions. Before dissection, the bone regions in CT images were extracted using our automated segmentation method (Yokota et al., 2009) to reconstruct the bone shape models . During dissection, locations of each muscle attachment region were recorded using an in-house measurement system similar to a surgical navigation system. Then, the measured attachment regions were normalized by warping the bone shape of specimen to an average bone shape.

Data conditions

Eight fresh cadaver specimens of the lower extremity including three males and five females with the age at death of 67.5 $\pm$ 14 y.o. were used in the experiments. We acquired CT images of all the specimens prior to the experiment with the slice interval of 0.3 mm (in-plane pixel size varies between 0.68 mm to 0.85 mm). Figure 4.1 shows volume renderings of the CT images of the cadaver specimens. Variations in the muscle shape in the cadaver specimens can be observed. The left leg of all specimens and right leg of three specimens were recorded (see Fig. 4.1).

Recording cadaver data

For the recording of the muscle attachment regions, we developed a measurement system as a plug-in module of 3D Slicer (https://www.slicer.org/) (Fedorov et al., 2012). The system was connected to an optical tracker (Polaris, Northern Digital Inc., Canada) with OpenIGTLink (Tokuda et al., 2009). 3D Slicer is an open source platform for medical image processing, and OpenIGTLink helps to connect trackers and imaging devices.

The recording was performed as follows. First, as shown in Fig. 4.3 (a), the optical markers were attached to the pelvis and femur with a fixation device used in a real surgical setup. Next, the coordinate system of the bone shape model was registered to the optical tracker coordinate system by paired-point registration using anatomical landmarks followed by iterative closest point (ICP) registration, which was performed using a module included in Slicer IGT (Ungi et al., 2016). Figure 4.3 (b) and (c) shows the point set for these registrations. The

average fiducial registration error (FRE) of the rough registration was 4.37±1.87 mm and of ICP was 0.56±0.15 mm. Then, dissection of individual muscles and tracing of the boundary of the muscle attachment region were repeated by the orthopedic surgeons.

The recorded points of muscle attachment boundaries were manually refined to remove outlier measurements due to tracking noise and project the points onto the bone surface in order to fill the gap between the measured points and the bone surface that occurred due to soft tissues remained on the bone.

Geometric normalization

For the purpose of geometric normalization, the measured muscle attachment regions of each cadaver specimen were mapped to the average bone shape. Surface-based non-rigid registration was performed between the average and each specimen's bone shapes and the computed deformation field was applied to map the vertices of the muscle attachment region to the average bone shape coordinate system. The average bone shape was constructed from clinical CT images, 200 female (Yokota et al., 2009). Note that the dataset for average model constructiondid not include the cadaver specimens used in the data recording. We used image registration tool kit (IRTK) provided by Rueckert et. al. (https://biomedia.doc.ic.ac.uk/software/irtk/) (Rueckert et al., 1999) for the non-rigid registration.

4.3.2 Data analysis

Constructing muscle attachment PA

Figure 4.4 shows the workflow of construction of muscle attachment PA and average attachment regions, and their utilization for estimation of the attachment region of a target patient. In Fig. 4.4, $\boldsymbol{R}_m(m = 1, 2, \ldots, N)$ denotes a muscle attachment region on a bone shape S_m, where N is the number of learning data. $\boldsymbol{R}_m$ is a binary vector whose length is the number of vertices included in bone shape S_m, indicating whether each vertex in S_m is inside (1) or outside (0) of an attachment region. S_{ave} is the average bone shape from clinical CTs and $T_{S_m \to S_{ave}}$ indicates a non-rigid transformation from S_m to S_{ave}. $\boldsymbol{R}'_m$ denotes a geometrically normalized attachment region of $\boldsymbol{R}_m$. In the learning phase, as shown in the upper row of Fig. 4.4, the muscle attachment PA was constructed by averaging the normalized attachment region $\{\boldsymbol{R}'_1, \boldsymbol{R}'_2, \ldots, \boldsymbol{R}'_N\}$ as follow:

$$\boldsymbol{P}_0(x) = \frac{1}{N} \sum_{m=1}^{N} \boldsymbol{R}'_m \tag{4.1}$$

As shown in Fig.4.5 (a), in order to avoid discontinuities occurring in a simple averaging method with a small number of training data, we smoothed the PA by convolution of Gaussian kernel (Fig. 4.5 (b)). We applied Gaussian kernel of standard deviation $\sigma = 2.0$ mm to $\boldsymbol{P}_0(x)$ and finally obtained muscle attachment PA $\boldsymbol{P}(x)$.

Obtaining and utilizing average muscle attachment regions

The average muscle attachment region $\boldsymbol{R_{ave}}$ was calculated by thresholding the muscle attachment PA. The threshold value was determined

so that the area of the average attachment region was equal to the average area, which is indicated in the third column in Table 4.1. The areas of the attachment regions obtained by thresholding the muscle attachment PA by all possible threshold values (with 0.01 step) were computed and the threshold value that yielded the area closest to the average area was selected.

The obtained average muscle attachment regions were utilized for estimating the attachment region of a specific patient. As shown in the lower row of Fig. 4.4, the average muscle attachment region was mapped to the target bone shape S_t, which was reconstructed from a CT image of the unseen patient, using the deformation field computed by non-rigid registration between the average and target bone shapes.

4.3.3 Experimental conditions

Target muscles and shape model resolution

Ten muscle attachments were measured, which were insertion and origin of the gluteus maximus, gluteus medius and gluteus minimus, insertion of piriformis, obturator internus and quadratus femoris, and origin of vastus lateralis. Where the insertion of obturator internus was an attachment of a conjoined tendon consisting of obturator internus, gemellus superior and gemellus inferior. Two measurements of the insertion of quadratus femoris and one of origin of vastus lateralis was missed in recording due to mechanical trouble. The reference bone of the origin of the gluteus maximus, gluteus medius and gluteus minimus was the pelvis, and the other attachments were on the femur. The average bone shapes had 244,526 and 266,648 vertices in the femoral and pelvis models, respectively. The ground-truth and the estimated muscle

attachment region are represented as a binary value (zero: back ground, one: muscle attachment) at each vertex on the average bone shape, and the muscle attachment PA is represented as a continuous-value from 0.0 to 1.0 indicating the probability. The spatial resolution of the femur model was about 4.9 vertices/mm^2 and the pelvis model was about 5.5 vertices/mm^2. The density of vertices was sufficiently high for describing details of the muscle attachment region because the average area of the smallest muscle attachment (insertion of obturator internus muscle) was more than 50 mm^2.

Evaluation methods

The accuracy of the proposed estimation method of the patient specific muscle attachment region was evaluated by leave-one-out cross-validation. We compared our proposed method to a previous method that uses pairwise morphing proposed by Pellikaan et al(Pellikaan et al., 2014). In order to evaluate the previous method, all combinations of pairwise morphing between the reference and target of the muscle attachment measurement data were tested. Given eight left and three flipped right cadavers, 55 ($\frac{11\times10}{2}$) combinations were tested. Two evaluation metrics were used to evaluate accuracy of the estimation: 1) Dice coefficient (Dice, 1945) for evaluating overlap of the estimated and ground truth region, 2) Average boundary distance which was a mean of the symmetric absolute distance of closest points on the boundary (Styner et al., 2008).

4.4 Results

Table 4.1 shows the average and standard deviation of the area and boundary length of each attachment region of cadaver specimens. The measurements on the right side of the specimen were mirror transformed so that it was dealt with in the same manner as the left one. Most attachment regions were measured in eight left and three right hips, however the origin of vastus lateralis and two insertions of quadratus femoris were missed due to a system trouble at recording.

Figure 4.6 shows the results of construction of the muscle attachment PA. The probability value is color-coded in Fig. 4.6. Red area indicates the area with high probability, green middle and blue low. The origin of the gluteus medius muscle has a large red area, that indicates sharing a large common region among patients. On the other hand, there are no red area in the insertion of the gluteus minimus and the obturator internus, indicating relatively large variations.

Table 4.2 shows comparison between the proposed estimation method and the previous Pellikaan's method. Average of Dice coefficient of the proposed method was more than 10 % larger than Pellikaan's method in most cases except for the origin of the gluteus medius (6.1 % larger) and the origin of the gluteus minimus (6.3 % larger), and the average symmetrical boundary distance of the proposed method was 1.4 mm smaller than Pellikaan's method in average.

Figure 4.7 shows boxplot of Dice coefficient (upper) and average boundary distance (lower) on each muscle attachment; in these figures, asterisks indicates results of Wilcoxon rank sum test; the single asterisk (*) indicates that the p-value was less than 0.05 and the double aster-

isk (**) indicates that the p-value was less than 0.01. As shown in Fig. 4.7, the accuracy improved significantly on eight muscle attachments in Dice coefficient and six muscle attachments in average boundary distance out of ten attachments.

Two muscle attachments (origin of Vastus Lateralis and insertion of Quadratus Femoris) did not show significant differences on the both error metrics, however the average Dice similarity coefficient of these attachment regions were, respectively, 13.4 % and 12.2 % higher than previous method and the average symmetric boundary distance were 1.2 mm and 0.8 mm smaller as shown in Table 4.2.

Figure 4.8 shows results of a representative case which had an accuracy near the mean value among all of muscle attachments. The previous pairwise method produced different estimation results depending on the choice of the reference subject. The Dice coefficient was 70.1 % for the best subject (left upper) and 11.5 % for the worst (left lower), and the average was 47.3 % for the worst (right), while the proposed method resulted in 55.4 %. The proposed method produced a better result than the average produced by the previous method.

4.5 Discussion and Conclusion

In this paper, we conducted cadaver experiments to measure 10 attachment regions of hip muscles and built a probabilistic atlas of the muscle attachment regions, as shown in Fig. 4.6, which is the main contribution of this study. The muscle attachment PA clarified the variations of the location of the muscle attachments. To the best of our knowledge, there is no study analyzing the inter-subject variations of the muscle

attachment in the normalized bone space. Moreover, we proposed a framework to statistically estimate the muscle attachment region using the muscle attachment PA. Our goal is to estimate patient-specific muscle attachment region from CT image. Comparison with the previous method (Pellikaan et al., 2014) suggested that the muscle attachment PA is useful to estimate patient-specific muscle attachment regions based on bone shape models reconstructed from CT image.

As observed in the comparison results with the previous pairwise estimation method, the proposed method had better stability and higher accuracy than the previous method. As shown in Fig. 4.8, the result of the previous method greatly depend on the choice of the target-source pair for the registration, while the proposed method stably estimated in all cases. Even though increasing the accuracy, the process for the estimation of the proposed method is basically the same as the previous method once the muscle attachment PA is built.

The proposed method has limitation for the muscle attachment with large variation. Two muscle attachments which did not show significant differences in both error metrics had the first and third largest standard deviation of the surface area (see Table 4.1). There is a possibility that an estimation by simple thresholding of muscle attachment PA with the average area was not stable in these cases.

We believe that the outer surface of the muscle body or the directions of muscle fibers would become a better predictor for estimating these attachment regions with larger variability, which is in the scope of our ongoing work. An automated method for segmentation of individual muscles (Yokota et al., 2018) and estimation of muscle fiber orientations from CT images (Otake et al., 2016) that we developed should help to

improve the estimation accuracy.

Table 4.1: Numerical values of measured attachment regions.

	Number of recorded subjects	Surface area [mm^2]	Boundary length [mm]
Attachment region on Femur			
Insertion of Gluteus Maximus	11	809.9 ± 276.8	196.6 ± 51.9
Insertion of Gluteus Medius	11	689.0 ± 261.2	144.4 ± 33.1
Insertion of Gluteus Minimus	11	357.5 ± 237.6	94.9 ± 30.3
Insertion of Gluteus Piriformis	11	103.7 ± 36.9	45.4 ± 7.4
Insertion of Gluteus Obturator Internus	11	98.3 ± 33.8	46.0 ± 8.0
Origin of Vastus Lateralis	9	1045.8 ± 588.4	192.0 ± 33.1
Insertion of Quadratus Femoris	10	499.9 ± 285.5	116.4 ± 26.7
Attachment region on Pelvis			
Origin of Gluteus Maximus	11	1051.9 ± 394.9	191.4 ± 50.7
Origin of Gluteus Medius	11	5705.4 ± 993.5	462.9 ± 64.9
Origin of Gluteus Minimus	11	4314.9 ± 874.6	320.2 ± 42.0

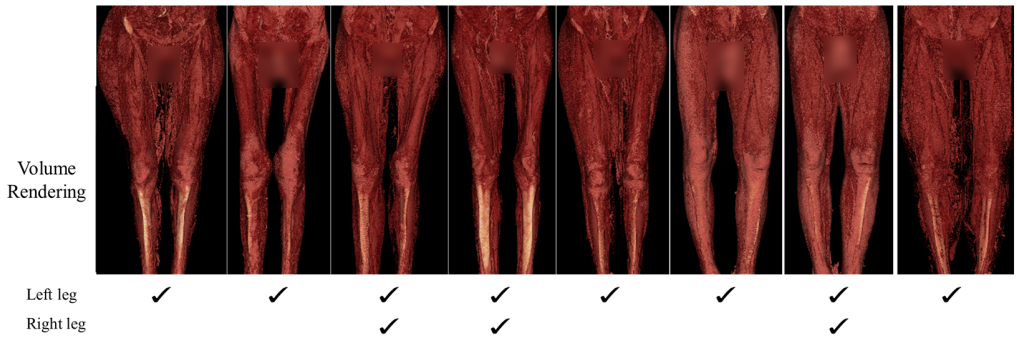


Figure 4.1: Volume rendering images of CT volume of the cadaver specimens used in the experiments. The check marks at the bottom indicate the side for which the measurement of muscle attachment regions was conducted.

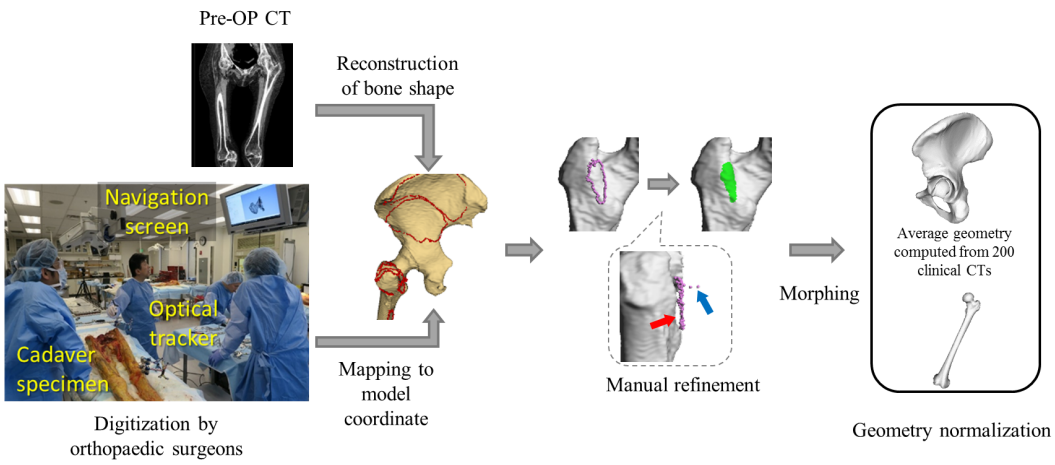


Figure 4.2: Workflow of collecting the ground truth data of muscle attachment regions. The red arrow in the middle box (manual refinement) indicates the gap between the bone surface and measured points due to soft tissues that could not be completely removed. The blue arrow indicates outlier points caused by the measurement noise.

Table 4.2: Average and standard deviation of Dice coefficient [%] and average symmetrical boundary distance [mm] of the proposed and previous methods.

	Dice coefficient [%]	Average symmetrical boundary distance [mm]
All muscle attachments		
Proposed	58.0 ± 20.5	4.5 ± 2.5
Previous (Pairwise)	46.3 ± 24.2	5.6 ± 3.1
Insertion of Gluteus Maximus		
Proposed	59.9 ± 15.1	5.4 ± 3.8
Previous (Pairwise)	47.9 ± 17.6	7.6 ± 5.2
Insertion of Gluteus Medius		
Proposed	62.9 ± 14.4	3.5 ± 0.9
Previous (Pairwise)	49.8 ± 17.6	4.4 ± 1.1
Insertion of Gluteus Minimus		
Proposed	50.6 ± 23.4	3.6 ± 1.7
Previous (Pairwise)	34.5 ± 23.6	4.9 ± 2.4
Insertion of Piriformis		
Proposed	49.6 ± 20.1	2.0 ± 0.7
Previous (Pairwise)	32.1 ± 20.6	2.7 ± 0.9
Insertion of Obturator Internus		
Proposed	26.1 ± 18.8	3.9 ± 3.3
Previous (Pairwise)	14.8 ± 22.1	5.8 ± 4.4
Origin of Vastus Lateralis		
Proposed	57.7 ± 12.6	4.3 ± 2.7
Previous (Pairwise)	44.3 ± 18.9	5.1 ± 1.7
Insertion of Quadratus Femoris		
Proposed	59.3 ± 11.5	3.7 ± 0.9
Previous (Pairwise)	47.1 ± 15.1	4.5 ± 1.4
Origin of Gluteus Maximus		
Proposed	61.2 ± 12.2	4.9 ± 1.6
Previous (Pairwise)	49.3 ± 17.2	6.8 ± 2.4
Origin of Gluteus Medius		
Proposed	77.7 ± 5.6	6.3 ± 2.0
Previous (Pairwise)	74.3 ± 6.0	6.0 ± 1.8
Origin of Gluteus Minimus		
Proposed	75.4 ± 6.3	7.0 ± 1.6
Previous (Pairwise)	68.8 ± 7.3	7.7 ± 1.2

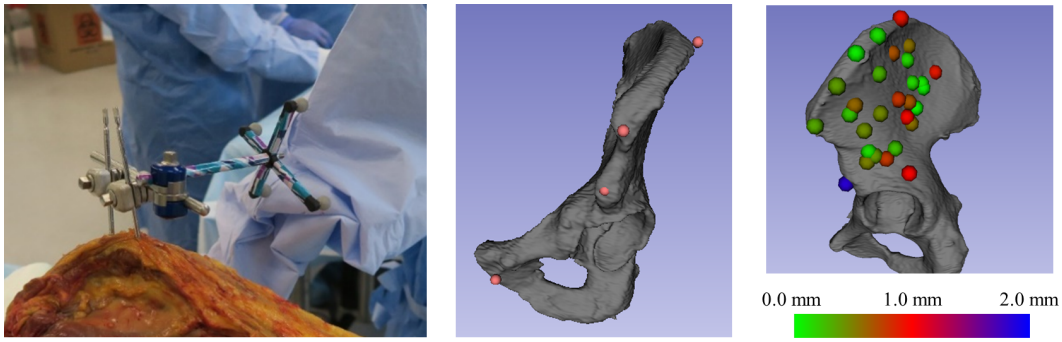


Figure 4.3: Registration between the optical tracker coordinate system and the bone shape model coordinate system. (a) Attachment of the optical marker to the pelvic bone. (b) Anatomical landmark points for initial paired point registration. (c) Result of ICP registration with 30 points on the pelvic bone. The color of each sphere indicates the fiducial registration error at each point. The radius of the sphere is 2.0 mm.

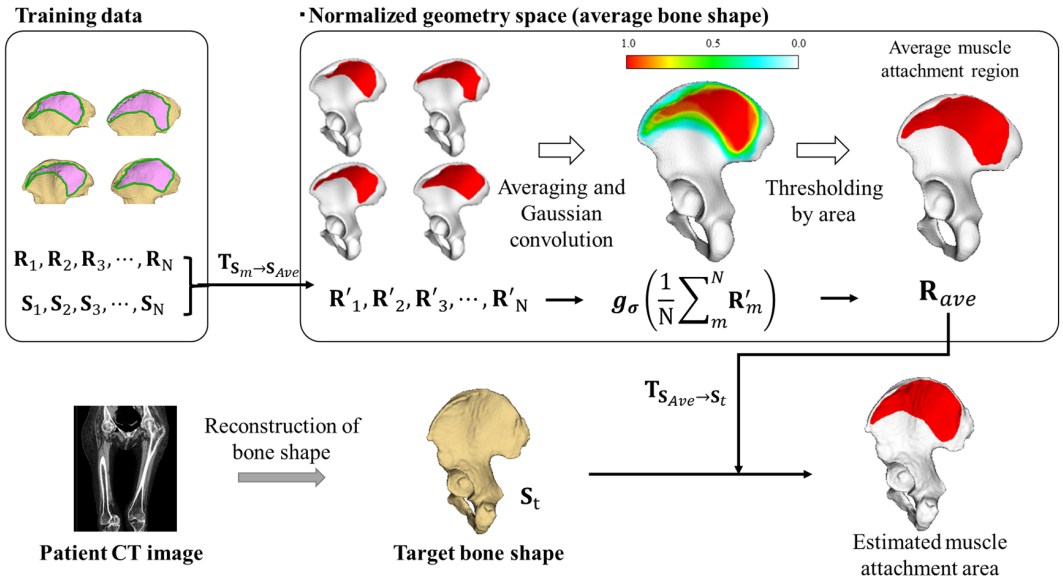
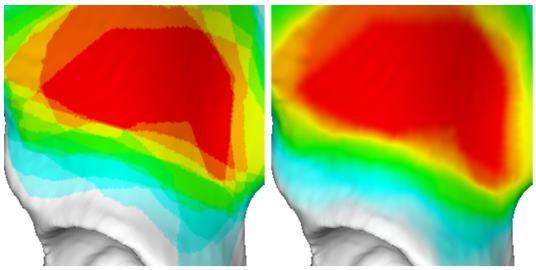


Figure 4.4: Workflow of estimation of muscle attachment region using probabilistic atlas. See texts for detailed explanations.



(a) **(b)**

Figure 4.5: Effect of the Gaussian convolution in building the PA of muscle attachment. (a) shows the result of simple averaging, (b) shows the result after applying smoothing with the Gaussian kernel of σ=2.0mm.

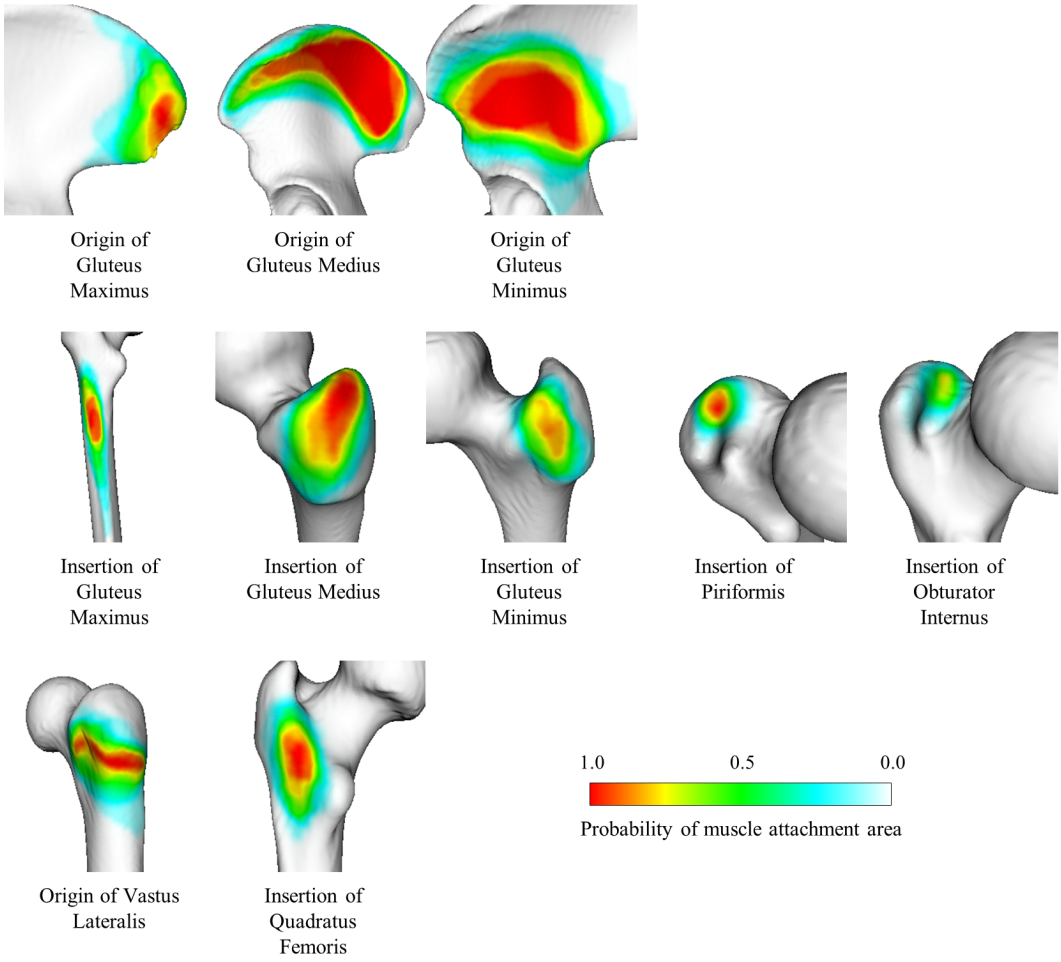


Figure 4.6: Probabilistic atlas of muscle attachment regions. Red area shows the area with high probability, blue low, and green middle.

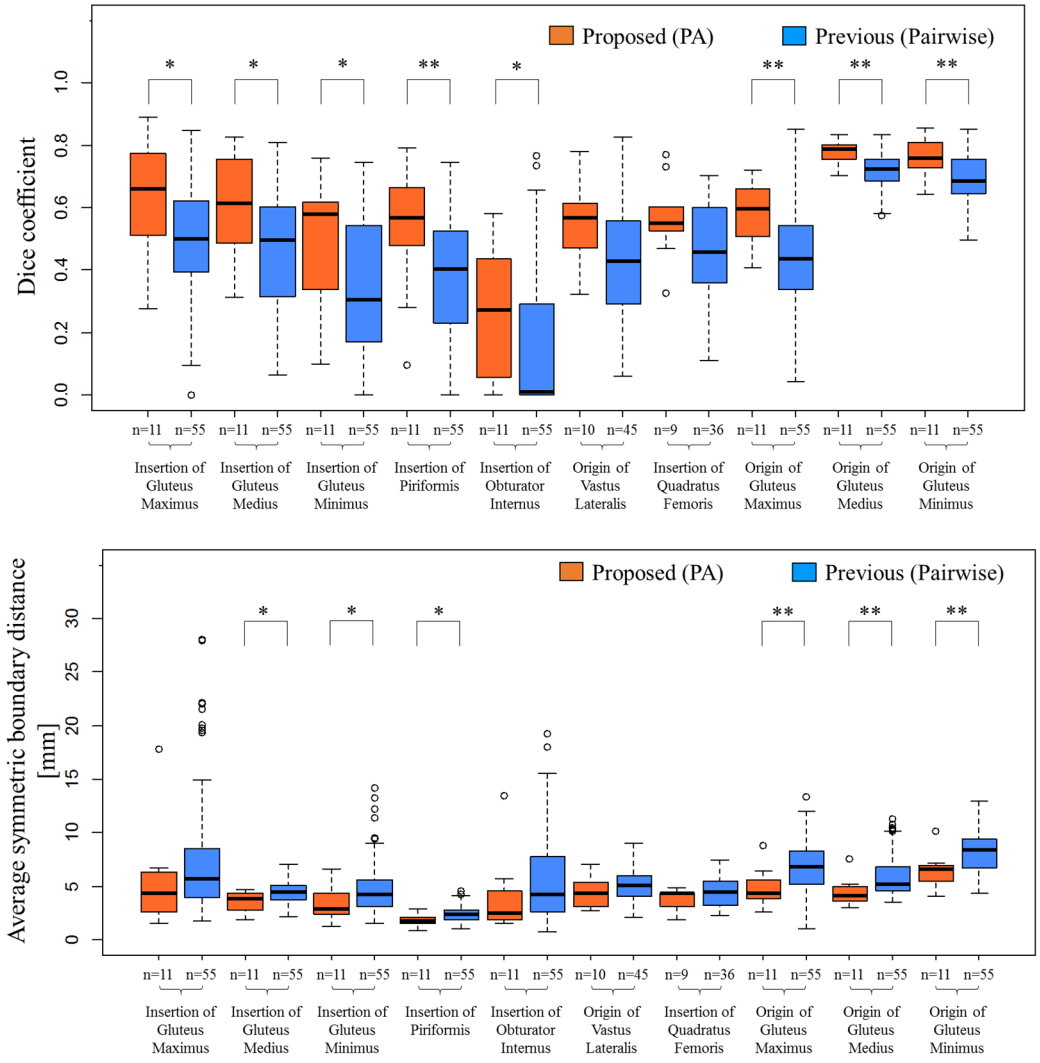


Figure 4.7: Evaluation results of Dice coefficient and average boundary distance. The asterisks (*) and (**) indicate significant differences from Wilcoxon rank sum test. $* : p < 0.05, ** : p < 0.01$

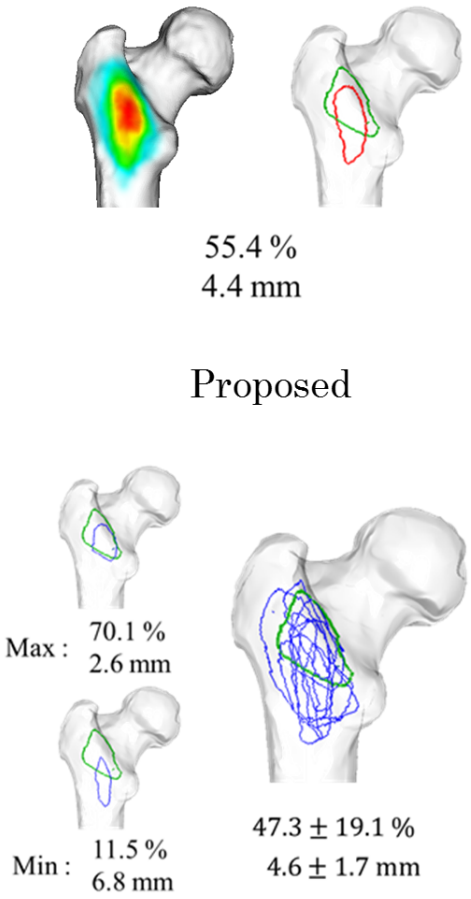


Figure 4.8: Representative case which has accuracy close to average among all attachments. Green line shows the ground truth, red line the estimation result by the proposed method, and blue lines the results by the previous pairwise method.

Chapter 5

Summary and future directions

5.1 Summary

In this paper, we have developed techniques for dynamic and static deformations through research on restoring 3D biological tissue shapes that correspond to deformations and movements based on medical images and have demonstrated their usefulness. The results of this research focus on correcting dynamic displacement deformations within individuals, statistical analysis of shape differences between individuals, and the efficient generation and application of data using these techniques.

Dealing with dynamic deformation

Our novel method for 3D shape recovery stands out for its ability to correct dynamic displacement deformations of organs during surgery. By utilizing a biplane ultrasound probe, we achieved real-time, high-precision recovery without the need for a position sensor, significantly enhancing operability compared to traditional sensor-dependent methods. This breakthrough is expected to pave the way for sensor-less

biopsy guidance systems, more precise needle insertion, and ultimately, improved patient outcomes.

Handling static deformation

Our data generation tool is a game-changer, enabling the efficient creation of high-quality annotation data even by non-experts. This tool's intuitive interface and quick operation have significantly improved the speed and quality of label data generation in medical image analysis. We have also demonstrated its ability to enhance automatic segmentation performance using deep learning. With its versatility across anatomical structures and imaging modalities, this tool is poised to revolutionize medical image analysis.

Statistical analysis of anatomical variations

Third, we constructed a statistical model of muscle attachment sites from multiple specimens and accurately estimated muscle attachment positions based on the patient's unique bone shape. This model was confirmed to be more valuable than conventional methods for estimating patient-specific anatomical shapes and was shown to be capable of highly accurate estimation that absorbs anatomical variations.

Significance of this research

The results of this research contribute to improving the accuracy and efficiency of medical image analysis by developing 3D shape recovery technology that can handle both dynamic and static deformation. This is expected to promote personalized medicine and enhance diagnostic

and treatment support. In addition, shape recovery and data generation technology utilizing statistical models is necessary to promote new developments in medical image analysis.

5.2 Future Directions

5.2.1 Construction of an integrated platform for the realization of personalized medicine

Our research has not only addressed individual challenges such as dynamic deformation correction technology, efficient annotation tools, and statistical analysis of anatomical variations, but also set the stage for an integrated platform that can revolutionize personalized medicine. This platform, which will enable real-time analysis of patient-specific dynamic and static shape information, is designed to provide highly personalized surgical support and treatment planning simulations based on each individual's anatomical characteristics.

5.2.2 Advancing Personalized Musculoskeletal Simulation

Building on the findings of Studies 2 and 3, we have developed a novel pipeline for automatically constructing individualized whole-body musculoskeletal simulation models. This pipeline, presented at domestic conferences, although unfortunately has not written a journal paper, utilizes AI to automatically extract skeletal and muscle belly shapes from MRI data and extrapolate the structure up to the attachment sites. Unlike conventional pipelines, which assume simplified string-based muscle models (Modenese and Kohout, 2020; Fernandez et al., 2023), our approach faithfully captures individual-specific musculoskeletal geom-

etry and muscle paths, offering a new generation of high-fidelity simulation models. We have successfully constructed musculoskeletal models for nearly 90 individuals using this pipeline. Leveraging these data, we created a statistical shape model of the musculoskeletal system, enabling the approximation of musculoskeletal models without requiring MRI imaging. This approach achieves a volume error of approximately 5%, offering a practical and efficient method for generating personalized models while reducing the cost and time associated with MRI acquisition. However, the simulation performance of these statistically predicted models compared to those generated directly from MRI data still needs to be verified. Future work will focus on evaluating the simulation accuracy of the approximate models and improving the predictive method to achieve higher simulation fidelity. These advancements are expected to enhance further the applicability of musculoskeletal simulation models in both research and clinical settings.

Acknowledgments

In the course of conducting this research, I have been fortunate to receive invaluable guidance and support from many individuals, both during my time as a student and at the research institution where I am currently affiliated. I would like to take this opportunity to express my heartfelt gratitude to all of them.

First and foremost, I would like to thank Professor Fumihiko Ino at Osaka University Graduate School of Information Science and Technology, who allowed me to be part of his laboratory despite the extra work they had to have because of me. If it were not for his long-term support, I would not be able to continue my work. Next, I would like to thank Professor Yasushi Yagi at The Institute of Scientific and Industrial Research, Osaka University, and Hajime Nagahara at Osaka University Institute for Datability Science for their time and effort in reviewing this thesis. Moreover, I would like to express my deepest gratitude to Professor Yoshinobu Sato at Graduate School of Information Science, Nara Institute of Technology and Science (NAIST), who guided the overall direction of my research during my doctoral studies. His remarkable insights and advice were instrumental in laying the foundation for my research.

I am also deeply grateful to Mr. Masahiko Nakamoto and Associate

Professor Yoshito Otake at NAIST, who directly supervised my daily research activities during my student years. From them, I gained not only specific advice on experimental and analytical methods but also valuable lessons on the attitude and dedication required of a researcher. I firmly believe that I could not have achieved the results presented in this thesis without their support.

At the Center for Information and Neural Networks (CiNet) of the National Institute of Information and Communications Technology (NICT), where I am currently affiliated, I have received considerable guidance and advice from Dr. Masaya Hirashima. His support and precise guidance have greatly contributed to the further development of my previous research, and I am deeply appreciative of his efforts.

I am also indebted to my collaborators, who provided crucial support and guidance in various research fields. In the field of urological surgery, I would like to express my deepest gratitude to Professor Osamu Ukimura at Kyoto Prefectural University of Medicine and Professor Inderbir S. Gill at University of Southern California for sharing their expertise and valuable experimental data. In the field of orthopedic surgery, I am extremely grateful to Professor Nobuhiko Sugano at Osaka University and Dr. Masaki Takao at Ehime University for providing expert knowledge and experimental data. In the fields of sports science and musculoskeletal science, I received invaluable advice from Dr. Shoji Konda at Osaka University and Dr. Jun Umehara at Kansai Medical University. The cooperation and contributions of these distinguished professionals were essential in establishing the foundation of this research.

I would also like to extend my gratitude to those who helped prepare

the experimental environment. In particular, Professor Mehran Armand at Johns Hopkins University generously provided access to essential experimental equipment and offered insightful advice. The CiNet MRI staff offered their expertise and meticulous support in operating and managing the experimental apparatus, for which I am deeply grateful. I also wish to thank Robert B. Grupp and Naoko Katagiri for their cooperation in setting up and managing the experiments.

Next, I would like to express my heartfelt gratitude to my family for their unwavering support. I am deeply thankful to my parents, who offered me immense help and encouragement during my student years. Although they have passed away, their support and teachings have been the cornerstone of my ability to continue my research. I extend my deepest respect, gratitude, and condolences to their memory.

I would also like to thank my wife, Nanase Fukuda, who has supported me throughout my daily life. Her understanding and dedicated support have enabled me to bring this research to its completion. I am truly grateful for her tireless efforts to support our family and my pursuits despite her own busy life.

Finally, I wish to express my sincere appreciation to everyone who has supported and contributed to this research at every stage.

References

Agosti, A., Shaqiri, E., Paoletti, M., Solazzo, F., Bergsland, N., Colelli, G., Savini, G., Muzic, S. I., Santini, F., Deligianni, X., Diamanti, L., Monforte, M., Tasca, G., Ricci, E., Bastianello, S., and Pichiecchio, A. (2022). Deep learning for automatic segmentation of thigh and leg muscles. *MAGMA*, 35(3):467–483.

Alipour, E., Chalian, M., Pooyan, A., Azhideh, A., Shomal Zadeh, F., and Jahanian, H. (2024). Automatic mri-based rotator cuff muscle segmentation using u-nets. *Skeletal Radiol*, 53(3):537–545.

Anas, E. M. A., Mousavi, P., and Abolmaesumi, P. (2018). A deep learning approach for real time prostate segmentation in freehand ultrasound guided biopsy. *Medical image analysis*, 48:107–116.

Arokoski, M. H., Arokoski, J. P. A., Haara, M., Kankaanpaa, M., Vesterinen, M., Niemitukia, L. H., and Helminen, H. J. (2002). Hip muscle strength and muscle cross sectional area in men with and without hip osteoarthritis. *Journal of Rheumatology*, 29(10):2185–2195.

Badrinarayanan, V., Kendall, A., and Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495.

Barnouin, Y., Butler-Browne, G., Voit, T., Reversat, D., Azzabou, N., Leroux, G., Behin, A., McPhee, J. S., Carlier, P. G., and Hogrel, J. Y. (2014). Manual segmentation of individual muscles of the quadriceps femoris using mri: a reappraisal. *J Magn Reson Imaging*, 40(1):239–47.

Baumann, M., Mozer, P., Daanen, V., and Troccaz, J. (2012). Prostate biopsy tracking with deformation estimation. *Medical image analysis*, 16(3):562–576.

Bax, J., Cool, D., Gardi, L., Knight, K., Smith, D., Montreuil, J., Sherebrin, S., Romagnoli, C., and Fenster, A. (2008). Mechanically assisted 3d ultrasound guided prostate biopsy system. *Medical physics*, 35(12):5397–5410.

Blemker, S. S. and Delp, S. L. (2005). Three-dimensional representation of complex muscle architectures and geometries. *Ann Biomed Eng*, 33(5):661–73.

Boykov, Y. Y. and Jolly, M.-P. (2001). Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In *Proceedings eighth IEEE international conference on computer vision. ICCV 2001*, volume 1, pages 105–112. IEEE.

Carbone, V., Fluit, R., Pellikaan, P., Van Der Krogt, M., Janssen, D., Damsgaard, M., Vigneron, L., Feilkas, T., Koopman, H. F., and Verdonschot, N. (2015). Tlem 2.0–a comprehensive musculoskeletal geometry dataset for subject-specific modeling of lower extremity. *Journal of biomechanics*, 48(5):734–741.

Chen, S., Eldeniz, C., Fraum, T. J., Ludwig, D. R., Gan, W., Liu, J., Kamilov, U. S., Yang, D., Gach, H. M., and An, H. (2023). Respiratory motion management using a single rapid mri scan for a 0.35 t mri-linac system. *Medical physics*, 50(10):6163–6176.

Cheng, R., Crouzier, M., Hug, F., Tucker, K., Juneau, P., McCreedy, E., Gandler, W., McAuliffe, M. J., and Sheehan, F. T. (2022). Automatic quadriceps and patellae segmentation of mri with cascaded u(2) -net and sassnet deep learning model. *Med Phys*, 49(1):443–460.

Conze, P. H., Brochard, S., Burdin, V., Sheehan, F. T., and Pons, C. (2020). Healthy versus pathological learning transferability in shoulder muscle mri segmentation using deep convolutional encoder-decoders. *Comput Med Imaging Graph*, 83:101733.

Cootes, T. F., Edwards, G. J., and Taylor, C. J. (1998). Active appearance models. In *Computer Vision—ECCV'98: 5th European Conference on Computer Vision Freiburg, Germany, June 2–6, 1998 Proceedings, Volume II 5*, pages 484–498. Springer.

Crum, W. R., Camara, O., and Hill, D. L. (2006). Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE transactions on medical imaging*, 25(11):1451–1461.

Cui, W., Liu, X., Zhao, Z., Feng, Z., and Meng, X. (2024). Accuracy and postoperative assessment of robot-assisted placement of pedicle screws during scoliosis surgery compared with conventional freehand technique: a systematic review and meta-analysis. *Journal of Orthopaedic Surgery and Research*, 19(1):365.

Damsgaard, M., Rasmussen, J., Christensen, S. T., Surma, E., and De Zee, M. (2006). Analysis of musculoskeletal systems in the anybody modeling system. *Simulation Modelling Practice and Theory*, 14(8):1100–1111.

De Silva, T., Cool, D. W., Yuan, J., Romagnoli, C., Samarabandu, J., Fenster, A., and Ward, A. D. (2017). Robust 2-d–3-d registration optimization for motion compensation during 3-d trus-guided biopsy using learned prostate motion data. *IEEE transactions on medical imaging*, 36(10):2010–2020.

Delp, S. L., Anderson, F. C., Arnold, A. S., Loan, P., Habib, A., John, C. T., Guendelman, E., and Thelen, D. G. (2007). Opensim: open-source software to create and analyze dynamic simulations of movement. *IEEE transactions on biomedical engineering*, 54(11):1940–1950.

Delp, S. L., Ringwelski, D. A., and Carroll, N. C. (1994). Transfer of the rectus femoris: effects of transfer site on moment arms about the knee and hip. *Journal of Biomechanics*, 27(10):1201–1211.

Diaz-Pinto, A., Alle, S., Ihsani, A., Asad, M., Nath, V., Pérez-García, F., Mehta, P., Li, W., Roth, H. R., and Vercauteren, T. (2022). Monai

label: A framework for ai-assisted interactive labeling of 3d medical images. *arXiv preprint arXiv:2203.12362*.

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302.

Falagario, U. G., Pellegrino, F., Fanelli, A., Guzzi, F., Bartoletti, R., Cash, H., Pavlovich, C., Emberton, M., Carrieri, G., and Giannarini, G. (2024). Prostate cancer detection and complications of mri-targeted prostate biopsy using cognitive registration, software-assisted image fusion or in-bore guidance: a systematic review and meta-analysis of comparative studies. *Prostate Cancer and Prostatic Diseases*, pages 1–10.

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., and Sonka, M. (2012). 3d slicer as an image computing platform for the quantitative imaging network. *Magnetic resonance imaging*, 30(9):1323–1341.

Fernandez, J., Shim, V., Schneider, M., Choisne, J., Handsfield, G., Yeung, T., Zhang, J., Hunter, P., and Besier, T. (2023). A narrative review of personalized musculoskeletal modeling using the physiome and musculoskeletal atlas projects. *Journal of Applied Biomechanics*, 39(5):304–317.

Fletcher, P., De Santis, M., Ippoliti, S., Orecchia, L., Charlesworth, P., Barrett, T., and Kastner, C. (2023). Vector prostate biopsy: a novel magnetic resonance imaging/ultrasound image fusion transperineal biopsy technique using electromagnetic needle tracking under local anaesthesia. *European Urology*, 83(3):249–256.

Fujihara, A. and Ukimura, O. (2022). Focal therapy of localized prostate cancer. *International Journal of Urology*, 29(11):1254–1263.

Fukunaga, T., Miyatani, M., Tachi, M., Kouzaki, M., Kawakami, Y., and Kanehisa, H. (2001). Muscle volume is a major determinant of joint torque in humans. *Acta Physiologica Scandinavica*, 172(4):249–255.

Gee, A. H., Housden, R. J., Hassenpflug, P., Treece, G. M., and Prager, R. W. (2006). Sensorless freehand 3d ultrasound in real tissue: speckle decorrelation without fully developed speckle. *Medical image analysis*, 10(2):137–149.

Guo, Z., Zhang, H., Chen, Z., van der Plas, E., Gutmann, L., Thedens, D., Nopoulos, P., and Sonka, M. (2021). Fully automated 3d segmentation of mr-imaged calf muscle compartments: Neighborhood relationship enhanced fully convolutional network. *Comput Med Imaging Graph*, 87:101835.

Hadaschik, B. A., Kuru, T. H., Tulea, C., Rieker, P., Popeneciu, I. V., Simpfendorfer, T., Huber, J., Zogal, P., Teber, D., Pahernik, S., and Others (2011). A novel stereotactic prostate biopsy system integrating pre-interventional magnetic resonance imaging and live ultrasound fusion. *The Journal of urology*, 186(6):2214–2220.

Hamid, N., Portnoy, J. M., and Pandya, A. (2023). Computer-assisted clinical diagnosis and treatment. *Current allergy and asthma reports*, 23(9):509–517.

Heimann, T. and Meinzer, H.-P. (2009). Statistical shape models for 3d medical image segmentation: a review. *Medical image analysis*, 13(4):543–563.

Hiasa, Y., Otake, Y., Takao, M., Ogawa, T., Sugano, N., and Sato, Y. (2020). Automated muscle segmentation from clinical ct using bayesian u-net for personalized musculoskeletal modeling. *IEEE Trans Med Imaging*, 39(4):1030–1040.

Horsman, M. K., Koopman, H. F., van der Helm, F. C., Prosé, L. P., and Veeger, H. (2007). Morphological muscle and joint parameters for musculoskeletal modelling of the lower extremity. *Clinical biomechanics*, 22(2):239–247.

Hsu, W. M., Hughes, J. F., and Kaufman, H. (1992). Direct manipulation of free-form deformations. *ACM Siggraph Computer Graphics*, 26(2):177–184.

Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., and Maier-Hein, K. H. (2021). nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2):203–211.

Ito, Y., Matsushita, I., Watanabe, H., and Kimura, T. (2012). Anatomic mapping of short external rotators shows the limit of their preservation during total hip arthroplasty. *Clinical Orthopaedics and Related Research®*, 470(6):1690–1695.

Javier-DesLoges, J., Dall'Era, M. A., Brisbane, W., Chamie, K., Washington III, S. L., Chandrasekar, T., Marks, L. S., Nguyen, H., Daneshvar, M., and Gin, G. (2023). The state of focal therapy in the treatment of prostate cancer: the university of california collaborative (uc-squared) consensus statement. *Prostate Cancer and Prostatic Diseases*, pages 1–3.

Jiang, H., Imran, M., Muralidharan, P., Patel, A., Pensa, J., Liang, M., Benidir, T., Grajo, J. R., Joseph, J. P., and Terry, R. (2024). Microsegnet: a deep learning approach for prostate segmentation on micro-ultrasound images. *Computerized Medical Imaging and Graphics*, 112:102326.

Jiang, J., Guo, Y., Bi, Z., Huang, Z., Yu, G., and Wang, J. (2023). Segmentation of prostate ultrasound images: the state of the art and the future directions of segmentation algorithms. *Artificial Intelligence Review*, 56(1):615–651.

Kaptein, B. and Van der Helm, F. (2004). Estimating muscle attachment contours by transforming geometrical bone models. *Journal of Biomechanics*, 37(3):263–273.

Karnik, V. V., Fenster, A., Bax, J., Cool, D. W., Gardi, L., Gyacskov, I., Romagnoli, C., and Ward, A. D. (2010). Assessment of image registration accuracy in three-dimensional transrectal ultrasound guided prostate biopsy. *Medical physics*, 37(2):802–813.

Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: Active

contour models. *International journal of computer vision*, 1(4):321–331.

Kobatake, H. and Masutani, Y. (2017). Computational anatomy based on whole body imaging. *Springer*, 10:978–4.

Light, A., Mayor, N., Cullen, E., Kirkham, A., Padhani, A. R., Arya, M., Bomers, J. G. R., Dudderidge, T., Ehdaie, B., Freeman, A., Guillaumier, S., Hindley, R., Lakhani, A., Pendse, D., Punwani, S., Rastinehad, A. R., Rouviere, O., Sanchez-Salas, R., Schoots, I. G., Sokhi, H. K., Tam, H., Tempany, C. M., Valerio, M., Verma, S., Villeirs, G., van der Meulen, J., Ahmed, H. U., and Shah, T. T. (2024). The transatlantic recommendations for prostate gland evaluation with magnetic resonance imaging after focal therapy (target): A systematic review and international consensus recommendations. *Eur Urol*, 85(5):466–482.

Liu, Q., Xu, Z., Jiao, Y., and Niethammer, M. (2022). isegformer: Interactive segmentation via transformers with application to 3d knee mr images. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part V*, pages 464–474. Springer.

Liu, R., Wen, X., Tong, Z., Wang, K., and Wang, C. (2012). Changes of gluteus medius muscle in the adult patients with unilateral developmental dysplasia of the hip. *BMC Musculoskelet Disord*, 13:101.

Luo, M., Yang, X., Yan, Z., Li, J., Zhang, Y., Chen, J., Hu, X., Qian, J., Cheng, J., and Ni, D. (2023). Multi-imu with online self-consistency for freehand 3d ultrasound reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 342–351. Springer.

Ma, J., He, Y., Li, F., Han, L., You, C., and Wang, B. (2024). Segment anything in medical images. *Nature Communications*, 15(1):654.

Malhotra, S., Halabi, O., Dakua, S. P., Padhan, J., Paul, S., and Palliyali, W. (2023). Augmented reality in surgical navigation: a review of evaluation and validation metrics. *Applied Sciences*, 13(3):1629.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., et al. (2001). A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (icbm). *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 356(1412):1293–1322.

Milletari, F., Navab, N., and Ahmadi, S.-A. (2016). V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pages 565–571. Ieee.

Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., and Terzopoulos, D. (2022). Image segmentation using deep learning: A survey. *IEEE Trans Pattern Anal Mach Intell*, 44(7):3523–3542.

Miura, K., Ito, K., Aoki, T., and Ohmiya, J. (2023). Pose estimation of ultrasound probe using cnn and rnn with image reconstruction loss. In *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 1–4. IEEE.

Modenese, L. and Kohout, J. (2020). Automated generation of three-dimensional complex muscle geometries for use in personalised musculoskeletal models. *Annals of Biomedical Engineering*, 48(6):1793–1804.

Modenese, L. and Renault, J. B. (2021). Automatic generation of personalised skeletal models of the lower limb from three-dimensional bone geometries. *J Biomech*, 116:110186.

Momose, T., Inaba, Y., Choe, H., Kobayashi, N., Tezuka, T., and Saito, T. (2017). Ct-based analysis of muscle volume and degeneration of gluteus medius in patients with unilateral hip osteoarthritis. *BMC Musculoskelet Disord*, 18(1):457.

Ni, R., Meyer, C. H., Blemker, S. S., Hart, J. M., and Feng, X. (2019). Automatic segmentation of all lower limb muscles from high-resolution magnetic resonance imaging using a cascaded three-dimensional deep convolutional neural network. *J Med Imaging (Bellingham)*, 6(4):044009.

Ogier, A. C., Heskamp, L., Michel, C. P., Foure, A., Bellemare, M., Le Troter, A., Heerschap, A., and Bendahan, D. (2020). A novel segmentation framework dedicated to the follow-up of fat infiltration in individual muscles of patients with neuromuscular disorders. *Magnetic resonance in medicine*, 83(5):1825–1836.

Ogier, A. C., Hostin, M. A., Bellemare, M. E., and Bendahan, D. (2021). Overview of mr image segmentation strategies in neuromuscular disorders. *Front Neurol*, 12:625308.

Otake, Y. et al. (2016). Analysis of muscle fiber structure using clinical ct: preliminary analysis with cadaveric images. In *proceedings of the 16th annual meeting of the international society for computer assisted orthopaedic surgery, Osaka*.

Pan, J., Huang, W., Rueckert, D., Küstner, T., and Hammernik, K. (2024). Reconstruction-driven motion estimation for motion-compensated mr cine imaging. *IEEE Transactions on Medical Imaging*.

Park, J. S., Chung, M. S., Hwang, S. B., Shin, B., and Park, H. S. (2006). Visible korean human: its techniques and applications. *Clinical Anatomy: The Official Journal of the American Association of Clinical Anatomists and the British Association of Clinical Anatomists*, 19(3):216–224.

Pellikaan, P., Van Der Krogt, M., Carbone, V., Fluit, R., Vigneron, L., Van Deun, J., Verdonschot, N., and Koopman, H. (2014). Evaluation of a morphing based method to estimate muscle attachment sites of the lower extremity. *Journal of biomechanics*, 47(5):1144–1150.

Philbrick, K. A., Weston, A. D., Akkus, Z., Kline, T. L., Korfiatis, P., Sakinis, T., Kostandy, P., Boonrod, A., Zeinoddini, A., and Takahashi, N. (2019). Ril-contour: a medical imaging dataset annotation tool for and with deep learning. *Journal of digital imaging*, 32(4):571–581.

Piazza, S. J. and Delp, S. L. (2001). Three-dimensional dynamic simulation of total knee replacement motion during a step-up task. *J. Biomech. Eng.*, 123(6):599–606.

Pons, C., Borotikar, B., Garetier, M., Burdin, V., Ben Salem, D., Lempereur, M., and Brochard, S. (2018). Quantifying skeletal muscle volume and shape in humans using mri: A systematic review of validity and reliability. *PLoS One*, 13(11):e0207847.

Prevost, R., Salehi, M., Sprung, J., Ladikos, A., Bauer, R., and Wein, W. (2017). Deep learning for sensorless 3d freehand ultrasound imaging.

Qiao, Y., Lelieveldt, B. P. F., and Staring, M. (2014). Fast automatic estimation of the optimization step size for nonrigid image registration.

Rasch, A., Bystrom, A. H., Dalen, N., and Berg, H. E. (2007). Reduced muscle radiological density, cross-sectional area, and strength of major hip and knee muscles in 22 patients with hip osteoarthritis. *Acta Orthop*, 78(4):505–10.

Reinbolt, J. A., Fox, M. D., Schwartz, M. H., and Delp, S. L. (2009). Predicting outcomes of rectus femoris transfer surgery. *Gait & posture*, 30(1):100–105.

Riem, L., Feng, X., Cousins, M., DuCharme, O., Leitch, E. B., Werner, B. C., Sheean, A. J., Hart, J., Antosh, I. J., and Blemker, S. S. (2023). A deep learning algorithm for automatic 3d segmentation of rotator cuff muscle and fat from clinical mri scans. *Radiol Artif Intell*, 5(2):e220132.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015:*

18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, pages 234–241. Springer.

Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L., Leach, M. O., and Hawkes, D. J. (1999). Nonrigid registration using free-form deformations: application to breast mr images. *IEEE transactions on medical imaging*, 18(8):712–721.

Sato, Y., Nakamoto, M., Tamaki, Y., Sasama, T., Sakita, I., Nakajima, Y., Monden, M., and Tamura, S. (1998). Image guidance of breast cancer surgery using 3-d ultrasound images and augmented reality visualization. *IEEE Transactions on Medical Imaging*, 17(5):681–693.

Scherzer, O. and Schoisswohl, A. (2002). A fast and robust algorithm for 2d/3d panorama ultrasound data. *Real-Time Imaging*, 8(1):53–60.

Schroeder, W., Martin, K., and Lorensen, B. (2006). *The Visualization Toolkit (4th ed.)*. Kitware.

Styner, M., Lee, J., Chin, B., Chin, M., Commowick, O., Tran, H., Markovic-Plese, S., Jewells, V., and Warfield, S. (2008). 3d segmentation in the clinic: A grand challenge ii: Ms lesion segmentation. *MIDAS journal*, 2008:1–6.

Tokuda, J., Fischer, G. S., Papademetris, X., Yaniv, Z., Ibanez, L., Cheng, P., Liu, H., Blevins, J., Arata, J., Golby, A. J., et al. (2009). Openigtlink: an open network protocol for image-guided therapy environment. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 5(4):423–434.

Tracy, B. L., Ivey, F. M., Hurlbut, D., Martel, G. F., Lemmer, J. T., Siegel, E. L., Metter, E. J., Fozard, J. L., Fleg, J. L., and Hurley, B. F. (1999). Muscle quality. ii. effects of strength training in 65- to 75-yr-old men and women. *J Appl Physiol (1985)*, 86(1):195–201.

Uhl, J.-F., Jorge, J., Lopes, D. S., and Campos, P. F. (2021). Digital anatomy. *Cham: Springer International Publishing*.

Ukimura, O. and Matsugasumi, T. (2015). *Role and Technique of Transrectal Ultrasound for Focal Therapy*, pages 39–53. Springer.

Ungi, T., Lasso, A., and Fichtinger, G. (2016). Open-source platforms for navigated image-guided interventions. *Medical image analysis*, 33:181–186.

Van der Helm, F. C., Veeger, H., Pronk, G., Van der Woude, L., and Rozendal, R. (1992). Geometry parameters for musculoskeletal modelling of the shoulder system. *Journal of biomechanics*, 25(2):129–144.

Wang, H., Suh, J. W., Das, S. R., Pluta, J. B., Craige, C., and Yushkevich, P. A. (2012). Multi-atlas segmentation with joint label fusion. *IEEE transactions on pattern analysis and machine intelligence*, 35(3):611–623.

Weng, L., Tirumalai, A. P., Lowery, C. M., Nock, L. F., Gustafson, D. E., Von Behren, P. L., and Kim, J. H. (1997). Us extended-field-of-view imaging technology. *Radiology*, 203(3):877–880.

Wesselink, E. O., Elliott, J. M., Coppieters, M. W., Hancock, M. J., Cronin, B., Pool-Goudzwaard, A., and Weber Ii, K. A. (2022). Convolutional neural networks for the automatic segmentation of lumbar paraspinal muscles in people with low back pain. *Sci Rep*, 12(1):13485.

Xu, S., Kruecker, J., Turkbey, B., Glossop, N., Singh, A. K., Choyke, P., Pinto, P., and Wood, B. J. (2008). Real-time mri-trus fusion for guidance of targeted prostate biopsies. *Computer Aided Surgery*, 13(5):255–264.

Yamout, T., Orosz, L. D., Good, C. R., Jazini, E., Allen, B., and Gum, J. L. (2023). Technological advances in spine surgery: navigation, robotics, and augmented reality. *Orthopedic Clinics*, 54(2):237–246.

Yokota, F., Okada, T., Takao, M., Sugano, N., Tada, Y., and Sato, Y. (2009). Automated segmentation of the femur and pelvis from 3d

ct data of diseased hip using hierarchical statistical shape model of joint structure. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009: 12th International Conference, London, UK, September 20-24, 2009, Proceedings, Part II 12*, pages 811–818. Springer.

Yokota, F., Otake, Y., Takao, M., Ogawa, T., Okada, T., Sugano, N., and Sato, Y. (2018). Automated muscle segmentation from ct images of the hip and thigh using a hierarchical multi-atlas method. *International journal of computer assisted radiology and surgery*, 13:977–986.