Advancing non-rigid 3D/4D human mesh registration for ultra-personalization

Tajdari, F.

DOI
10.4233/uuid:300b839f-78fc-4d1b-a7d6-007d080e902e

Publication date
2023

Document Version
Final published version

Citation (APA)
Tajdari, F. (2023). *Advancing non-rigid 3D/4D human mesh registration for ultra-personalization*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:300b839f-78fc-4d1b-a7d6-007d080e902e

ADVANCING NON-RIGID 3D/4D HUMAN MESH REGISTRATION FOR ULTRA-PERSONALIZATION

ADVANCING NON-RIGID 3D/4D HUMAN MESH REGISTRATION FOR ULTRA-PERSONALIZATION

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Thursday 7 December 2023 at 12:30 o'clock

by

Farzam TAJDARI

Master of Science in Mechanical Engineering,
Amirkabir University of Technology, Iran,
born in Golpayegan, Iran.

This dissertation has been approved by the promotor(s).

promotor: Dr. Y. Song
copromotor: Dr. T. Huysmans

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Dr. Y. Song ,	Delft University of Technology
Dr. T. Huysmans,	Delft University of Technology
Independent members:	
Prof. dr. D. Aschenbrenner	Hochschule Aalen, Germany
Prof. dr. S. Verwulgen	U. of Antwerp, Belgium
Prof. ir. D.J. van Eijk	Delft University of Technology
Prof. dr. ir. R.H.M. Goossens	Delft University of Technology
Prof. dr. P. Vink	Delft University of Technology

Keywords: personalized design, human body shapes, 3D mesh registration, 4D scanning, computational design framework

Cover design F. Tajdari
Layout F. Tajdari

ISBN 978-94-6366-792-0

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

The most exciting phrase in science is not "Eureka!" but "That's funny..."

Isaac Asimov

CONTENTS

LIST OF FIGURES

LIST OF TABLES

SUMMARY

Personalized designs bring significant added value to the products and the users. However, they also pose challenges on the product design process. For instance, for products for personalized fit, each may differ subject to each user's body shape and preference. Presently, there exist knowledge and methods which support designing personalized products/services, with sample applications in the fields of medical products, shoes, clothing industry, etc. Meanwhile, the major steps in these methods are manual or semi-automated, thus designing Ultra Personalized Products and Services (UPPS) can be a tedious and time-consuming task. Furthermore, the design process is usually not optimized and most applications are employing ad-hoc approaches. Designers need a systematic approach to designing UPPS.

Project UPPS is targeting a complete human-centered approach that goes beyond the boundaries of traditional manufacturing industry: it requires adaptation to digital representations and computation-controlled production systems are essential. The workflow of developing UPPS is framed in several phases; data collection, data analysis, design, production, and testing. However, it not only encompasses the introduction of new production technology and organisation, but by its one-on-one customer relationship, it also requires a profound rethink of current design and market approaches. Furthermore, UPPS necessitates on-demand production, with little to no stock, creating opportunities for localised production and reshoring.

With UPPS, we need advanced tools to collect offline and online the needed data e.g., 3D point cloud, from the involved parties i.e., the user and the product. Next and using the data, there is a necessary need for more effective and efficient design tools/methods to understand the geometry of the parties in both 3D space and 4D space i.e., time series of the 3D space geometries. Finally, a holistic understanding of the constraints and abilities of the production line, and infrastructure of embedding the user in the framework loop for feed-backing the personalized products for possible future improvements complete the UPPS cycle requirements.

This Ph.D. project aims at developing a robust mesh registration framework that accounts for addressing the challenges, especially for human body shapes.

The first challenge is how to establish a framework that enables transferring meaningful features from one 3D mesh to another. In the area of 3D geometry processing, mesh registration comprises the fundamental knowledge that can be used to build the connections between two meshes, thus by understanding the features of one of the meshes, we can automatically understand the corresponded features in the other mesh treated as knowledge transfer between the meshes. One of the most well-known approaches to addressing knowledge transfer is the non-rigid Iterative Closest Point (ICP) registration method. Though the method was successfully applied in a vast number of applications, e.g. statistical shape modeling, computer vision, multimedia applications, 3D deformation of the human spinal column detection, and 3D human body analysis, the non-rigid registration is a non-trivial and ill-defined problem with a high number of Degrees-Of-Freedom (DOFs). This caused challenges, e.g. establishing meaningful robust correspondences in the registration process, ensuring the convergence towards the desired minimum in the optimization, and maintaining the quality of the mesh regarding the source surface.

The second challenge involves identifying and extracting sequential dynamic 3D features from a series of 3D meshes. Human movements usually result in considerable shape deformation of different body parts. 4D scan, which explains 3D geometric shapes over time, is a pioneer for understanding dynamic anthropometry in different applications. For example, 4D scans present important information about human body deformation during motion, conducting specific activities with dynamic workloads. This expresses the need for 4D scanners design and 4D scanning techniques. To acquire 4D scans, multiple (depth) cameras are usually employed. To do so, arrangement of several cameras is a necessary stage in any motion tracking technique. The arrangement should guarantee the free movement of the users to be able to track the targeted features. Apparently, the accuracy and the quality of the 3D reconstructed point-clouds rely on the number of the used cameras and their configuration. However, the number is competing with the cost and the complexity of the final product and partly the speed of the process. Those cameras can be synchronized for capturing continuous images at a given moment. However, it is challenging to balance the needed resolutions of the images, the needed time duration, the buffer of the depth cameras, the data transfer rate, the computing power, and the storage. Due to the challenge, an accumulative delay forms regarding each camera even if all the cameras are synchronised hardware-based. Therefore, this dissertation aims at finding an automatic method to optimally define the position and orientation of a set of cameras in a foot scanner, and achieve software-based synchronisation of the cameras.

Although 3D/4D scanning and parametric design are powerful techniques for the automation of UPPS design, it does not constitute a framework for explaining designerly freedom nor does it directly allow for handling all kinds of design and manufacturing constraints. For the final stage of this Ph.D. project, a new and generic data-enabled and data-driven design framework is explored in the context of the UPPS project, namely a computational design framework, that supports the designer in the creative process while also delivering robust automation.

In conclusion, this dissertation presents a pipeline that effectively automates the geometry-based designerly knowledge transformation whether in 3D space or 4D space. Firstly, software programs are developed to address fast and robust 3D point cloud registration as surface mesh or volumetric mesh. On the other hand, new hardware namely 4D scanners, are designed and manufactured capable to achieve high-speed 4D scanning of a moving object i.e., 30 frames per second. In this regard, a novel time-synchronisation software is implemented on the scanner to maintain time-wise alignment of the frames. Using the 3D point cloud registration methods, we present a comprehensive mesh-morphing 4D scan product from the software to the hardware. In the end, the integration of all the developed software and hardware to design is discussed.

SAMENVATTING

Gepersonaliseerde ontwerpen voegen aanzienlijke meerwaarde toe aan producten en hun gebruikers. Ze brengen echter ook uitdagingen met zich mee voor het productontwerpproces. Bijvoorbeeld, voor producten met een gepersonaliseerde pasvorm kan elk product verschillen afhankelijk van de lichaamsvorm en voorkeur van de gebruiker. Op dit moment zijn er reeds kennis en methoden beschikbaar die het ontwerpen van gepersonaliseerde producten/diensten ondersteunen, met voorbeelden in de medische sector, schoenen, kledingindustrie, enz. Echter zijn de belangrijkste stappen in deze methoden handmatig of slechts semi-geautomatiseerd, waardoor het ontwerpen van Ultra Personalized Products and Services (UPPS) een tijdrovende taak kan zijn. Bovendien is het ontwerpproces meestal niet geoptimaliseerd en maken de meeste toepassingen gebruik van ad-hoc benaderingen. Ontwerpers hebben een systematische aanpak nodig voor het ontwerpen van UPPS.

Het UPPS Project richt zich op een volledig op de mens gerichte benadering die de grenzen van de traditionele productie-industrie overstijgt: het vereist aanpassing aan digitale representaties en digitaal gestuurde productiesystemen zijn essentieel. De workflow voor het ontwikkelen van UPPS bestaat uit verschillende fasen: gegevensverzameling, gegevensanalyse, ontwerp, productie en gebruik. Het omvat echter niet alleen de introductie van nieuwe productietechnologie en organisatie, maar vereist ook een diepgaande heroverweging van huidige ontwerp en marktbenaderingen vanwege de persoonlijke relatie met de klant. Bovendien vereist UPPS on-demand productie, met weinig tot geen voorraad, wat mogelijkheden biedt voor lokale productie en reshoring.

Voor UPPS hebben we geavanceerde tools nodig om offline en online de benodigde gegevens te verzamelen, bijv. 3D puntenwolk, van zowel de betrokken gebruiker als het te produceren product. Vervolgens, met behulp van deze gegevens, is er een noodzaak voor effectievere en efficiëntere ontwerptools/methoden om de geometrie van deze modellen in zowel 3D als 4D (tijdreeksen van 3D geometrieën) beter te kunnen interpreteren. Ten slotte is er een holistisch begrip vereist van de beperkingen en mogelijkheden van de productietechnieken en moet er nagedacht worden over de aanpak van het integreren van de gebruiker in het UPPS raamwerk, ten behoeve van feedback over het gebruik van de gepersonaliseerde producten voor mogelijke toekomstige verbeteringen.

Dit Ph.D. project heeft tot doel een robuust mesh-registratieraamwerk te ontwikkelen dat rekening houdt met het aanpakken van uitdagingen, met name voor menselijke lichaamsvormen.

De eerste uitdaging is hoe een raamwerk op te zetten dat het overdragen van betekenisvolle kenmerken van de ene 3D-mesh naar de andere mogelijk maakt. In het domein van de 3D geometrie-verwerking omvat mesh-registratie de fundamentele techcniek die kan worden gebruikt om de correspondentie tussen twee meshes op te bouwen. Door de kenmerken van een van de meshes te begrijpen, kunnen we automatisch begrijpen welke overeenkomstige kenmerken er zijn in de andere mesh. Een van de meest bekende

aanpakken voor correspondentie is de niet-rigide Iterative Closest Point (ICP) registratiemethode. Hoewel de methode met succes is toegepast in een groot aantal applicaties, zoals statistische vormmodellering, computervisie, multimediatoepassingen, 3D deformatie van de menselijke wervelkolom en 3D analyse van het menselijk lichaam, is de niet-rigide registratie een uitdagend en slecht gedefinieerd probleem met een groot aantal vrijheidsgraden. Dit veroorzaakt uitdagingen, zoals het vaststellen van betekenisvolle robuuste overeenkomsten in het registratieproces, het waarborgen van de convergentie naar het gewenste minimum in de optimalisatie en het handhaven van de kwaliteit van de vervormde mesh ten opzichte van de kwaliteit van de mesh van het bronoppervlak.

De tweede uitdaging omvat het identificeren en extraheren van sequentiële dynamische 3D kenmerken uit een reeks 3D meshes. Menselijke bewegingen leiden meestal tot aanzienlijke vormvervorming van verschillende lichaamsdelen. 4D scans, die 3D-geometrische vormen in de tijd capteren, zijn baanbrekend voor het begrijpen van dynamische antropometrie in verschillende toepassingen. Zo geven 4D scans bijvoorbeeld belangrijke informatie over de vervorming van het menselijk lichaam tijdens beweging, bij het uitvoeren van specifieke activiteiten met dynamische werklasten. Hieruit volgt de behoefte van het ontwerp van 4D scanners en 4D scantechnieken. Om 4D scans te verkrijgen, worden vaak meerdere (diepte)camera's gebruikt. Daarbij is de relatieve opstelling van de verschillende camera's een belangrijk onderdeel van bewegingsregistratie. De opstelling moet de vrije beweging van de gebruikers toelaten om de relevante kenmerken te kunnen volgen. Het blijkt dat de nauwkeurigheid en de kwaliteit van de 3D gereconstrueerde puntenwolken afhankelijk zijn van het aantal gebruikte camera's en hun configuratie. De hoeveelheid camera's concurreert echter met de kost en complexiteit van het eindproduct en ook beinvloed ook deels de efficientie van het opname- en verwerkingsproces. Deze camera's kunnen worden gesynchroniseerd om continue beelden op een gegeven moment vast te leggen. Het is echter uitdagend om het vereiste niveau van resolutie van de beelden, de benodigde tijdspanne, de buffer van de dieptecamera's, de gegevensoverdrachtsnelheid, de rekenkracht en de opslag in evenwicht te brengen. Deze uitdaging resulteert in een opstapelende vertraging bij de verschillende camera's, zelfs als ze allen hardwarematig zijn gesynchroniseerd. Daarom heeft dit proefschrift tot doel een automatische methode te vinden om de positie en oriëntatie van een set camera's in een voetscanner optimaal te definiëren en de synchronisatie softwarematig aan te vliegen.

Hoewel 3D/4D scannen en parametrisch ontwerp krachtige technieken zijn voor de automatisering van UPPS ontwerp, vormt het geen kader voor ontwerpvrijheid, noch staat het direct toe om alle soorten ontwerp en productiebeperkingen aan te pakken. Voor het laatste stadium van dit Ph.D. project wordt een nieuw en generiek, op data gebaseerd en data-gestuurd, ontwerpraamwerk onderzocht in het kader van het UPPS-project. Met name een computationeel ontwerpraamwerk dat de ontwerper ondersteunt in het creatieve proces en tegelijkertijd robuuste automatisering oplevert.

In conclusie presenteert deze dissertatie een workflow die de op geometrie gebaseerde en ontwerpgerichte automatisering effectief uitvoert, zowel in 3D als 4D. Ten eerste worden algoritmes ontwikkeld voor snelle en robuuste 3D registratie van oppervlakte- of volumemeshes. Aan de andere kant worden nieuwe hardwareopstellingen, namelijk 4D scanners, ontworpen en vervaardigd om in staat te zijn tot snelle 4D-scanning van een bewegend object (tot 30 frames per seconde). In dit opzicht wordt een nieuwe tijdssynchronisatiesoftware geïmplementeerd op de scanner om de tijdsafstemming van de frames te bewerkstelligen. In conclusie presenteren we uiteindelijk een uitgebreid 4D scanproduct van de software tot de hardware. Tot slot wordt de integratie van alle ontwikkelde software en hardware voor het ontwerpen in een reeks van toepassingen besproken.

I

INTRODUCTION

1

INTRODUCTION

Personalized designs bring added value to the products and the users. Meanwhile, they also pose challenges to the product design process, with a focus on personalized fit, as each product may differ for each user's preference. The first step towards addressing the difficulties is understanding the geometry of both the user and the product. Thus, this thesis presents methods utilized to automatically extract 3D and 4D features from raw scanned data, an optimized 4D foot scanner, and a discussion on integrating these methods and tools into the personalized product design framework. Achieving the methods and the tools faces several challenges which are presented in three research questions that are explained in detail and solved in corresponding research cycles. This chapter ends with the organization of this dissertation.

IN the midst of mass-produced products, Ultra-Personalized Products and Services (UPPS) have emerged as a popular choice by embedding added values in the product/service for individuals. Among the different types of UPPS, personalized-fit products address the need for personalized products regarding the physical interactions between the product and the consumer, the environment, and/or other products that are used by the consumer. Physical characteristics of the product, such as shape, size, mass, area, quantity, color palette, etc., and the personalized interactions (e.g., comfort), present the added values of the personalized-fit products [173].

This dissertation aims at developing tools and methods for designing personalized-fit products under the influence of disruptive technology in the fourth industrial revolution [66], e.g., 3D scanning, 3D printing. Currently, there are technologies and knowledge that support designing personalized-fit products, and they already have many applications in the fields of medical products, shoes, clothing industry, etc. However, most of these technologies/methods are manual or semi-automated, therefore designing personalized-fit products is often a time-consuming and tedious job. In addition, the design process is often not optimized and most applications utilize ad-hoc solutions. Designers ask for a systematic approach for designing personalized-fit products and they also need more effective and efficient design tools/methods to support and accelerate such an approach. A generic framework is missing to systematically approach personalized-fit products design and production, taking into account the complexities of customer variety in shape and preferences, design and production constraints, certification standards, and mechanisms for continuous feedback throughout the process.

1.1. PERSONALIZATION

For customers, size and shape specifications are among the most important parameters in choosing a product, e.g. clothes, shoes, spectacles, and sports products. Amongst, standard sizing and shape systems helped industries and factories to be able to implement mass-production. However, there are some disadvantages with existing solutions. In mass production, a large number of identical or similar products are manufactured which may not be sold and often result in a huge waste of materials [283]. Moreover, rudimentary [283] sizing is another issue that causes a mismatch between customers' preferences and mass-produced product(s). Such discrepancies may reduce comfort, and can even result in not using or discarding the product. These problems reveal the missing gaps in the traditional mass-production and highlight the opportunities of developing a comprehensive methodology to satisfy the needs of the personalized-fit products in society, e.g. for groups such as obese people, amputees, people with non-standard body types, etc.

The accelerated development of technologies, especially digital technologies, brings more possibilities to personalized product design and production [79]. For instance, contrary to traditional manufacturing, advanced manufacturing methods, e.g. 3D printing, have undergone a revolution in the past decade. The potential for personalized mass production has motivated manufacturers to incorporate 3D printing technology in production, leading to potentially radical changes in the future of personalized product design and manufacturing [67].

To find a comprehensive approach to leverage these opportunities, an overarching view is necessary that encompasses 1) the transition from mass customization to mass personalization, incorporating the UPPS Model, and 2) the integration of semi/fully-automated production.

1.1.1. MASS CUSTOMIZATION TO MASS PERSONALIZATION

Wide applications of machinery in manufacturing have resulted in two different production principles [283]: First, the continuation of craft production, where machines are employed to increase the variety of products [124]. Second, mass production, where the use of machines signifies higher efficiency in production, leading to lower product prices [198]. However, mass production has its downsides. The concept of the mass production feedback loop recognizes that disruptive changes in technology can influence the acceptance of conventionally produced mass products [122]. To address the heterogeneity of customer needs in society, the concept of "mass customization" was introduced in [197], defined as follows: "Mass customization is the new frontier in business for both manufacturing and service industries. At its core is a tremendous increase in variety and customization without a corresponding increase in costs. At its limit, it is the mass production of individually customized goods and services. At its best, it provides strategic advantage and economic value."

The two aforementioned articles [198, 197] have been published in the 90s when no one could predict that the technological revolution would bring this fundamental consideration to a higher level. The different stages of production are visualized in Figure 1.1, taken from [19], and depict the evolution towards personalized production. The figure is extracted according to [125], which has been expanded upon in [19] for personalized production. Mentioning that there are ample attempts regarding personalization in the literature (as summarized in [129]). We depict this with several arrows. The difference between personalization and customization has been explained in [129] as: "We term personalization in a narrow sense as system-initiated personalization and customization as user-initiated personalization." In the figure, the UPPS approach is added as a distinct subcategory of personalized production. Targeting batch size one, it differs from other attempts, respectively.

The reason that UPPS does not end up at the same point as craft production does, is the most important significance of Figure 1.1. Accordingly, the reason is that personalizing a product is time-consuming with current methods, resulting in a lower number of personalized manufactured products. Moreover, the amount of technology usage should be investigated and the results can be appreciated as the third dimension. As mentioned in [19] “Thus, the Volume-variety diagram should be treated as a spiral with a third dimension, rather than as a circle.” Accordingly, the target of UPPS is resumed as “The development of radical new product propositions for the manufacturing industry through the innovative use of data and by making products fully customized” [235]. In this project, we want to emphasize that the visualization of the UPPS model known as NextUPPS presented in Figure 1.2 towards UPPS project is a broad field that includes multi-disciplinary knowledge of automation in production, design, and management.

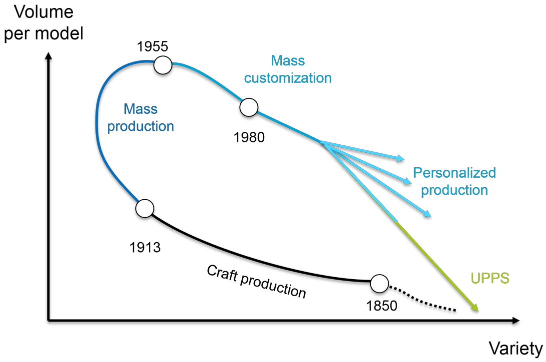


Figure 1.1: Volume-variety diagram in different stages of production includes craft production between 1850 and 1913, mass production between 1913-1955, mass customization between 1955-1980, personalized production, and UPPS from 1980 to the present [19].

1.2. UPPS PROJECT

In 2016, Delft University of Technology (TUD) as Work Package 1 (WP1), Eindhoven University of Technology (TU/E) as WP2, and the University of Twente (UT) as WP3, collaborated on the nextUPPS project, which was proposed to the Dutch Creative Industries. The UPPS theoretical model was initially introduced in the proposal and further refined through workshops involving the principal investigators (PIs) and Ph.D. students from the three universities. This model served as an approach to identifying UPPS challenges. Figure 1.2 provides an overview of the model, highlighting the Analysis, Design, Produce, and Test/Use stages within the UPPS product/service system [129].

The proposed UPPS model is an integrated concurrent approach toward UPPS product design and manufacturing. Starting from the market with the actual needs of the user, their preferences, sizing, and (working) context will be analyzed utilizing population models, resulting in digital twin representations of the customer and their activities and environment. Together with collected (anthropometric) data/models, manufacturing constraints, and legal requirements, an expert system will make an analysis and propose design preferences and parameters. These preferences and parameters will be the input of a design automation system, and the outcome will be a digital twin of the personalized product to be produced, which will be manufactured and certificated by "advanced" manufacturing methods. The final stage of the UPPS workflow is also the start of a new iteration in the UPPS iterative product and process where feedback from the product usage is used to improve the process and product in an iterative fashion.

Among the three involved knowledge institutes, TU/E focuses on the role of the user in the process of ultra-personalized product and service creation. They are researching how to facilitate the exchange of collaborative creation between the designer and the co-designer while minimizing the perceived and actual risk of sizing and customer comfort during sizing. UT is investigating: 1. The characteristics of personalization and possible classification of personalized products with the aim to weigh the level of personalization against required effort; 2. The gap between the flexibility of designs of personalized products and the flexibility of production processes; 3. A possible classification of pro-

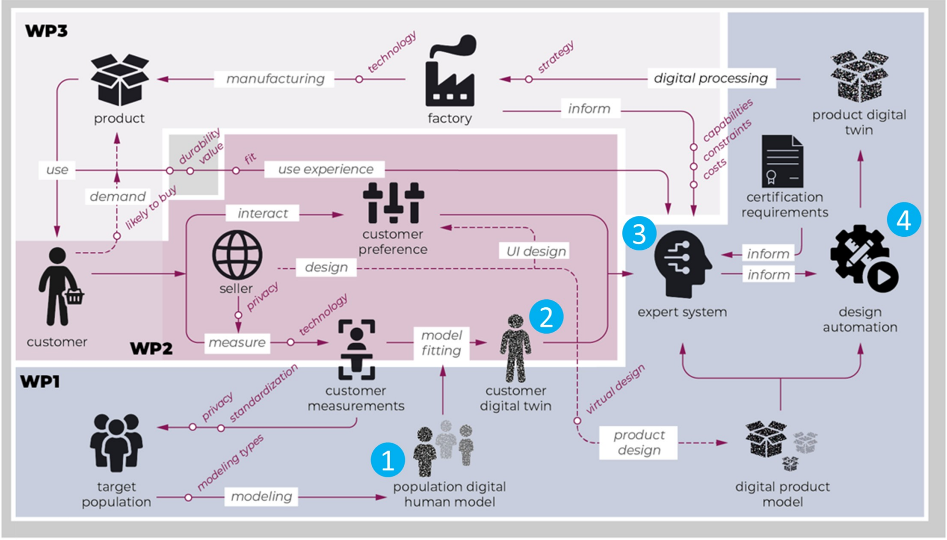


Figure 1.2: UPPS connectivity of different units including four WPs: WP1) Automating UPPS; WP2) Defines designer and the co-designer relationship; and WP3) Manufacturing considerations of the product digital twin.

cess steps for automation to increase insight into successful automated manufacturing of personalized products; and 4. Human-Robot collaboration in manufacturing and assembly in the context of personalized products to gain a better insight into using collaborative robots in reconfigurable manufacturing systems.

TUD is responsible for the overarching framework of the model and introducing tools and methods for designing UPPS. In detail, we illustrated the interactions between the involved units in UPPS concerning TUD (stages 1-4) in Figure 1.2, namely 1: Population modeling, 2: Digital twinning, 3: Decision-making system, and 4. Design automation/product. Accordingly, the UPPS cycle [235] at TUD begins with the analysis stage, where a 3D digital twin of the customer is generated from a population model based on some form of digitization (e.g., 3D scanning). In the second step, utilizing an expert system and considering customer preferences, activities, and production constraints, such as costs and certification, the product's digital twin is designed. Thirdly, the specific piece of the designed product (including both body shape and preferences) is produced in a production plant with a batch size of one. In the final steps, the product is delivered to the customer, who tries and starts using it. Meanwhile, there are possibilities to enhance the personalized product for each user in the future, based on feedback collected from the product's usage (e.g., sensors or customer responses).

1.2.1. AUTOMATION IN THE SEMI-AUTOMATED PERSONALIZATION

Semi-automated product personalization is a beginning step towards fully-automated personalization and involves techniques that bridge the knowledge of engineers and designers [181]. By utilizing digital data from scanning and, if necessary, engaging with customers, the physical human form and their preferences can be discerned through data analysis. Based on the physical form, designers can then proceed to design the products, select suitable materials, and determine appropriate digital fabrication techniques [84], such as 3D printing [170]. However, this process often proves to be time-consuming, requiring several iterations to achieve the desired outcome.

Fully-automated approaches have the potential to expedite this procedure. For example, in the clothing industry, Susan et al. [21] studied the process of automated pattern generation using software tools from full-body 3D scanning. As validation, they compared the fitting quality of the customized jackets prototype with a set of 10 ready-to-wear jackets, in which 70% of the participants were satisfied with the customized jacket prototype. In shoe personalisation, Nachtigall [181] introduced a procedure for creating fully functional shoes and presented a research prototype [183]. Through this work, he demonstrated that automated software solutions and an understanding of design practice are prerequisites that support the realization of Ultra Personalized 4D shoemaking and related digital crafts, thereby partially verifying the feasibility of UPPS goal. Accordingly, the goal of UPPS is to create a highly individualized and tailored experience for customers, resulting in increased satisfaction, loyalty, engagement, and ultimately, business success.

1.3. RESEARCH GAPS

Although many works have investigated UPPS, particularly in the context of personalized-fit products, there are several research gaps that hinder the realization of fully automated personalized-fit products. These research gaps include:

- There is a lack of essential methods to extract meaningful 3D features on human body shapes automatically.
- The robustness of the methods [13, 180, 102] in the presence of artefacts on the collected 3D data of the human body shapes is still not addressed comprehensively.
- In the area of 3D human body shapes processing, there are limited works investigating directly extracting time sequences of the internal structure of the human body from X-ray images. Namely, reconstructing a time-sequential 3D volumetric human spine from X-ray images to predict pediatric spinal deformity (PSD).
- 4D human body scanners are needed to collect movement and deformation of human bodies. However, there are limited studies about the optimal design of the scanners subject to maximum visibility of the dynamic body by each camera.
- 4D scanning of a deformable object is still challenging, extracting 4D features is rarely reported, as establishing a connection between the 3D extracted features in a time series of 4D dataset is still unsolved.
- The 4D feature extraction may get more challenging when the cameras of the 4D scanner are not time-wise synchronized.

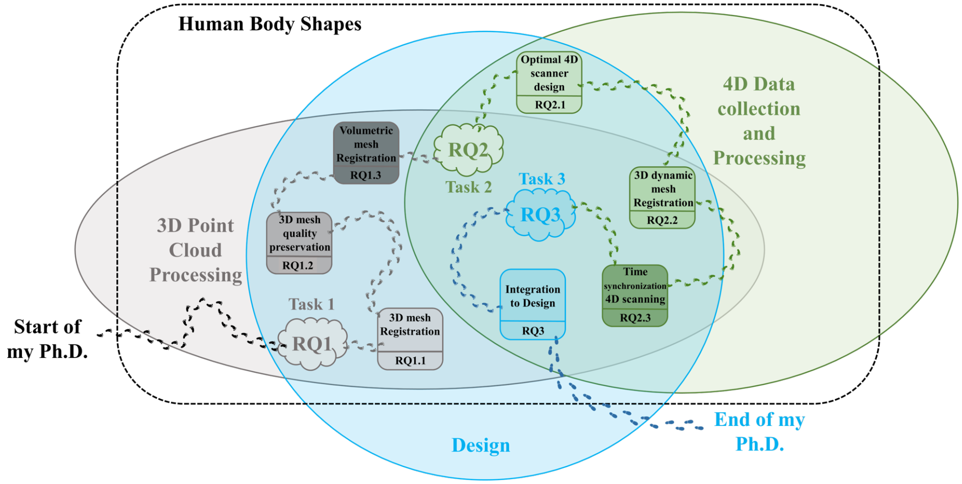


Figure 1.3: The connection between the different components of this thesis and the research questions in the area of human body shapes investigated in three areas of 3D point cloud processing, 4D data collection and processing, and design.

- Finally, a framework is lacking capable of integrating the solutions towards the aforementioned gaps together and expressing how to utilise them in the context of design automation.

1.4. RESEARCH QUESTIONS

To fill the scientific gaps for developing an integrated design method for designing personalized-fit products, we frame the Ph.D. research in three inter-connected areas with a set of research questions as shown in Figure 1.3. In the figure, the first research area is 3D human body shape processing, which forms the basis of the geometric fit of human body shape. The second area is 4D data collection and processing, addressing dynamic deformable body shapes. The third area is Design, which highlights the need for designers to formulate an integrated approach to designing personalised-fit products.

In the area of 3D human body shape processing, many tools and methods were developed, however, with different limitations, e.g. non-rigid Iterative closest points (ICP) method is only suitable to match scans of a fixed object, which minimizes the distance of the closest points from the source mesh to the target mesh, thus other features of 3D points clouds are neglected results in limited accuracy of the pattern matching between the two meshes. This limitation made it difficult to apply the method to create a statistical shape model of human which can approach different human body shapes based on a set of data. The first Research Question (**RQ1**) is therefore set as:

- **RQ1:** *What type of registration methods can contribute to establish 3D human shape model and their use as prior knowledge in fitting new subject scans?*

One of the most well-known approaches to addressing knowledge transfer is non-rigid registration methods [7, 71]. Here, non-rigidly registering a point cloud is treated as deforming the point cloud to have the best match to a different point cloud, where the point clouds in this thesis focus on 3D space. In the past decade, non-rigid registration has been widely applied in many applications such as motion analysis [96], shape analysis [68], and medical image registration [75]. Among different types of non-rigid surface registration methods, the non-rigid Iterative Closest Point (ICP) registration has attracted much attention, mainly due to its simplicity, efficiency, and effectiveness. Using the non-rigid ICP method, a source surface is registered to the target surface via two iterative steps, where the goal is to find meaningful correspondences, e.g. match anatomical equivalent zones/points. First, the correspondence of each vertex in the source surface to the target surface is established based on a metric, which is usually defined as a type of distance (e.g., distance between points). Then, a cost function is defined based on those correspondences and is minimized by finding a non-rigid transformation. In a practical application, these two steps iterate until a local minimum of the cost function is found or the iteration steps exceed a threshold.

Though it was successfully applied in many applications, such as gender scoring [89], statistical shape modeling [100], computer vision [138], multimedia applications [224], human-computer interactions [162], 3D deformation of the human spinal column detection [257], image face alignment [46], and 3D human body analysis [114], the non-rigid registration is a non-trivial and ill-defined problem with a high number of degrees-of-freedom (DOFs). According to Figure 1.3, sub-research question 1.1 refers to challenges for preserving features of the source surface in the design and implementation of a non-rigid ICP registration algorithm [268]. Here, features account for salient geometric features which form compound higher-level descriptors. A salient geometric feature, or in short, a salient feature, consists of a cluster of descriptors that locally describe a nontrivial region of the surface [83], such as curvature. These challenges include establishing meaningful robust correspondences in each step of the iteration [265], ensuring convergence towards the desired minimum in the optimization [143], and maintaining the quality of the mesh regarding the source surface. These challenges resulted in sub-research question (**RQ1.1**) as:

- **RQ1.1:** *How to integrate a shape descriptor to the nonrigid ICP algorithm to consider the shape characteristics?*

In the context of non-rigid registration, most available methods aim to non-rigidly align a point cloud to another point cloud, involving deformation. However, the interpretation of the ideal deformation in these methods remains ambiguous as the deformation is typically considered as minimizing the geometric error with respect to the corresponding points, as in the case of the non-rigid ICP algorithm. Nevertheless, such deformation may result in the loss of certain features in the source mesh and/or a reduction in the overall mesh quality. The primary challenge lies in the utilization of uniform parameter values for all points on the source mesh during registration, disregarding the fact that each point requires a detailed examination based on its specific conditions, such as the presence of holes or artifacts on the target mesh, in order to determine appropriate parameter values. This challenge gives rise to the formulation of the following sub-research question:

  - **RQ1.2:** *How to derive a nonrigid ICP-based algorithm that is robust to artefacts in a 3D shape?*

Finally, **RQ1.3** is formulated as follows, based on the challenges associated with the registration process and the considerations explained in **RQ1.1** and **RQ1.2**, while considering the utilization of volumetric meshes for human body shapes within the source and/or target meshes.

  - **RQ1.3:** *How to implement the nonrigid-ICP registration method for registering volumetric meshes on each other?*

Human movements often result in significant shape deformations of various body parts. The use of 4D scanning, which captures 3D geometric shapes over time, has garnered considerable attention in gaining a better understanding of dynamic anthropometry across different applications [35, 39]. For instance, 4D scans provide valuable insights into human body deformations during motion, particularly when engaging in specific activities with dynamic workloads. In summary, this technology aids in establishing a fundamental understanding of human movements and finds applications in various domains, including building virtual avatars, ergonomics, computer games, personal protective equipment, workwear, sportswear, and other practical garments [34].

However, the transition from 3D point cloud processing to 4D scanning is accompanied by several hardware and software limitations. As a result, the following Research Question (**RQ2**) is formulated:

- **RQ2:** *What type of tools and registration methods can contribute to establish 4D human shape model and their use as prior knowledge in fitting new subject scans?*

To acquire 4D scans, the utilization of multiple depth cameras is often necessary. Therefore, the arrangement of these cameras plays a crucial role in any motion-tracking technique, ensuring unrestricted user movement for effective tracking of targeted features.

The identification of feature points, which can be detected using either passive or active landmarks, relies on the camera outputs. Through triangulation between the outputs of multiple cameras, the 3D positions of the landmarks are determined [74, 161]. The accuracy and quality of the resulting 3D reconstructed point clouds [92] are inherently dependent on the number of cameras employed and their configuration. However, this number presents a trade-off between the cost of the final product and the processing speed. The investigation into the complexity of designing an optimal 4D scanner is addressed in Research Question **RQ2.1**:

  - **RQ2.1:** *How to optimally define the position and orientation of a set of cameras in a 4D foot scanner?*

In 4D scanning, each camera generates a set of point clouds corresponding to each frame, and there are no inherent connections or associations between these point clouds. Consequently, the absence of point correspondence between frames hinders the ability to capture and analyze dynamic features across frames. This leads to the formulation of Research Question **RQ2.2**:

  - **RQ2.2:** *How to establish meaningful dynamic features between frames of a camera which guarantees to find the proper corresponded points in a limited number of iterations from a Source mesh to a Target mesh?*

The synchronization of multiple cameras in order to capture continuous images at a specific moment is possible. However, striking a balance between various factors such as image resolution, time duration, buffer capacity of the depth cameras, data transfer rate, computing power, and storage poses significant challenges [33]. For example, capturing depth images of size 640x576 using six cameras at 30 frames per second would require a bandwidth of approximately 2 Gb/s. This presents difficulties in the design of a 4D scanning system, particularly for low-cost implementations. As a result, "dropped frames" can often be observed in the captured data. Moreover, due to the substantial data transfer involved, a nonlinear accumulative delay occurs for each camera, even when they are hardware-based synchronized [171]. The challenge related to this accumulated delay gives rise to Research Question **RQ2.3**:

  - **RQ2.3:** *How to design a higher level of software-based synchronisation method between cameras in a 4D foot scanner?*

3D/4D scanning and parametric design are powerful techniques for the automation of UPPS design; however, it does not constitute a framework for expressing designerly freedom nor does it directly allow for handling all kinds of design and manufacturing constraints. Thus, a new and generic data-enabled and data-driven design framework in the context of the UPPS project, namely a computational design framework, is necessary that supports the designer in the creative process while also delivering robust automation. Expert systems may play a central role here. To this end, bridging from the tools and methods to the designers' knowledge guide us to the following third research equation:

- **RQ3:** *How can we use the 3D/4D human data in the design automation for UPPS ?*

We answer these research questions in the following chapters where the structure is explained in Section 1.5.

1.5. THESIS STRUCTURE

The main body of this Ph.D. thesis consists of seven Chapters that answer the corresponding research questions (RQs) subject to the sub-research questions that were defined in Section 1.4 within the UPPS framework. Figure 1.4 presents an overview of the thesis. Each **RQ** is investigated in an individual part, and each sub-research question is presented and discussed in a separate chapter. The thesis ends with a discussion and recommendation. Parts II- IV includes several chapters that present the introduction, related work, theory, results, and conclusion. Short descriptions of each Chapter are presented in the following paragraphs.

Chapter 2: In Chapter 2, and to address **RQ1.1**, a novel non-rigid registration method is presented, making use of a newly defined semi-curvature concept, which draws inspiration from the definition of Gaussian curvature. In the process of establishing point correspondences, a dynamic weighted criterion based on the distance and semi-curvature

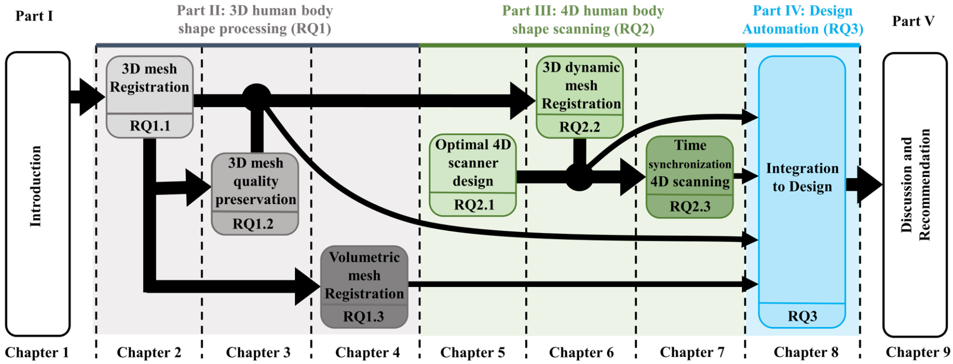


Figure 1.4: An overview of the research cycles and the organization of the thesis, where gray color belongs to the domain of 3D human body shape processing, green color belongs to 4D human body shape scanning, and blue color belongs to design automation.

is employed to select the corresponding point on the target surface for each point on the source surface. The cost function is reformulated to incorporate the semi-curvature, stiffness, and distance terms, effectively penalizing errors in both the distance and semi-curvature components within a guaranteed stable region. To facilitate a robust and efficient optimization process, the semi-curvature term is linearized, defining a region of attraction and establishing the stability of the approach. Experimental results demonstrate that the proposed method better preserves features of local areas on the original surface with higher curvature values, outperforming conventional methods.

Chapter 3: Although Chapter 2 addresses a promising registration method for preserving the features of the source mesh, the challenge outlined in **RQ1.2** concerning the balance between feature preservation and deformation in the presence of mesh artifacts remains. Therefore, in Chapter 3, a non-rigid ICP algorithm is presented, which approaches the challenge as a control problem. An adaptive feedback control scheme is derived, ensuring global asymptotic stability, to control the stiffness ratio and achieve maximum feature preservation and minimum mesh quality loss during the registration process. A cost function is formulated, incorporating the distance term and the stiffness term. The initial stiffness ratio value is determined using an Adaptive Neuro-Fuzzy Inference System (ANFIS)-based predictor, which takes into account the source mesh and target mesh topology, as well as the distances between the correspondences. Throughout the registration process, the stiffness ratio of each vertex is continuously adjusted based on intrinsic information obtained from shape descriptors of the surrounding surface, as well as the progress made in the registration steps. Additionally, the estimated process-dependent stiffness ratios are utilized as dynamic weights for establishing correspondences in each registration step. Experimental results on both simple geometric shapes and 3D scanning datasets demonstrate the superiority of the proposed approach over existing methodologies, particularly in regions where features are not prominent or where interferences among features exist. This is attributed to its capability to incorporate the inherent properties of the surface during the mesh registration process.

Chapter 4: Predicting Pediatric Spinal Deformity (PSD) based on X-ray images obtained during a patient's initial visit poses a significant challenge. In response to **RQ1.3**, Chapter 4 investigates this issue by proposing a volumetric geometry-based bone growth model that can be applied in various applications to enhance the bio-informed mechanistic machine learning framework. The primary focus is on examining and predicting spine curvature in PSD cases, specifically adolescent idiopathic scoliosis. The proposed technique leverages a segmented 3D volumetric geometry of the human spine obtained from 2D X-ray images. To achieve this, the anteroposterior and lateral views of the X-ray images are segmented using an active contour model based on gradient vector flow snakes. These 2D contours surrounding each vertebra are extracted with minimal user input, and the snake parameters are calibrated and automatically computed across the dataset, enabling fast image segmentation and data collection. The 2D segmented outlines of each vertebra are then transformed into a 3D image segmentation result. To establish a mesh morphing approach and create a 3D atlas spine model, the Iterative Closest Point mesh registration technique is employed. This process facilitates the extraction of spinal geometry data from the comprehensive 3D volumetric model, which can be utilized as inputs to the mechanistic machine learning network. Furthermore, the proposed bio-informed deep learning network, incorporating the modified bone growth model, achieves competitive or even superior performance compared to other state-of-the-art learning-based methods.

Chapter 5: Optical motion capture involves the estimation of three-dimensional (3D) positions of points by employing triangulation with multiple depth cameras. With regards to **RQ2.1**, the success of this technique heavily relies on the visibility of points from different cameras and the overlap of captured meshes between them. The accuracy of the position estimation is largely influenced by the camera parameters, such as their location and orientation. Consequently, the configuration of the camera network plays a crucial role in the quality of the resulting mesh. In Chapter 5, an optimal approach for camera placement is proposed based on the characteristics of the D435i depth camera from Intel RealSense. The optimization problem involves a cost function that consists of various terms aimed at minimizing or maximizing specific parameters. The minimization terms include the distance of the cameras to the center of the scanning object, resolution error, and sparsity. On the other hand, the maximization terms encompass the distance between pairs of cameras, the percentage of captured points from the object, and the level of overlap between cameras. The accuracy and robustness of the proposed algorithms are evaluated through experimental measurements, and the sensitivity to the number of cameras is also investigated. This analysis provides insights into the performance of the camera placement optimization approach and its effectiveness in achieving accurate and reliable results in optical motion capture applications.

Chapter 6: In accordance with **RQ2.2**, 4D scans, which refer to continuous 3D scans captured at different timestamps, are crucial for the development of garments/apparel in 4D that can accommodate body movements. In Chapter 6, a semi-automatic workflow is proposed for constructing 4D scans of body parts, with a particular focus on aligning and registering noisy scans at specific timestamps. The process begins by capturing continuous 3D scans of the moving body parts from various depth cameras positioned at different angles. At a given timestamp, the collected 3D scans are roughly aligned to a template using the rigid ICP algorithm. Subsequently, these scans are further registered using a newly proposed non-rigid Iterative Closest-Farthest Points (ICFP) algorithm. The ICFP algorithm establishes correspondences between the source and target scans by utilizing either the closest or farthest points, based on a newly defined logical distance concept and probability theory. This approach aims to improve the accuracy of registration while being robust against noise. Experimental results demonstrate that the ICFP method exhibits strong resilience to noise and achieves acceptable scanning accuracy. Overall, the proposed workflow and the ICFP algorithm contribute to the construction of reliable and accurate 4D scans of body parts, facilitating the design and development of garments/apparel that can adapt to the dynamic movements of the human body.

Chapter 7: Although this chapter is mainly focusing on **RQ2.3**, it uses promoted version of methods and tools extracted form Chapter 5 and Chapter 6, namely, next generation of 4D foot scanner, and mesh-morphing technique. Thus, individually in Chapter 7, we introduce a generic framework that is able to 1) synchronise scans captured by asynchronous cameras through a novel deep-learning-based network, which is capable of aligning 3D scans captured by different cameras to the timeline of a specific camera; and 2) register a high-quality template to synchronised scans at each timestamp to form a high-quality 3D mesh model using a non-rigid registration method. Using a newly developed 4D foot scanner, we validate the framework and create the first open-access data-set, namely the 4D feet.

Chapter 8: This chapter addresses **RQ3** and focuses on the feasibility of integrating the outputs from the previous chapters (Chapter 2 to Chapter 7) into the design process. The chapter presents a comprehensive framework that aims to bridge the gap between personalized design and design practice. First, the general workflow of personalized product design is introduced. Then, different steps in the workflow, such as human data/parameters acquisition, computational design, design for digital fabrication, and product evaluation, are presented. Tools and methods that are often used in different steps in the process are also outlined. Human data acquisition, 3D scanning, and digital human models are addressed in the context of data acquisition. For computational design, the use of computational thinking tools such as abstraction, decomposition, pattern recognition, and algorithms is discussed. In design for digital fabrication, additive manufacturing methods (e.g., Fused Deposition Modeling (FDM)) and their design requirements are highlighted. Regarding product evaluation, both functional evaluation and usability evaluation are considered, and the evaluation results can serve as the starting point for the next design iteration. Finally, several case studies are presented to enhance the understanding of the workflow, emphasize the importance of different steps in the workflow, and explore variations in the approach based on different contexts. In conclusion, the aim is to provide designers with a holistic view of the design process for personalized products and to assist practitioners in generating innovations at each step of the process.

II

3D HUMAN BODY SHAPE PROCESSING

2

3D FEATURE-BASED NON-RIGID REGISTRATION METHOD

Preserving features of a surface as characteristic local shape properties captured e.g. by curvature, during non-rigid registration is always difficult where finding meaningful correspondences, assuring the robustness and the convergence of the algorithm while maintaining the quality of mesh are often challenges due to the high degrees of freedom and the sensitivity to features of the source surface. In this chapter, we present a non-rigid registration method utilizing a newly defined semi-curvature, which is inspired by the definition of the Gaussian curvature. In the procedure of establishing the correspondences, for each point on the source surface, a corresponding point on the target surface is selected using a dynamic weighted criterion defined on the distance and the semi-curvature. We reformulate the cost function as a combination of the semi-curvature, the stiffness, and the distance terms, and ensure to penalize errors of both the distance and the semi-curvature terms in a guaranteed stable region. For a robust and efficient optimization process, we linearize the semi-curvature term, where the region of attraction is defined and the stability of the approach is proven. Experimental results show that features of the local areas on the original surface with higher curvature values are better preserved in comparison with the conventional methods. In comparison with the other methods, this leads to, on average, 75%, 8% and 82% improvement in terms of quality of correspondences selection, quality of surface after registration, and time spent of the convergence process respectively, mainly due to that the semi-curvature term logically increases the constraints and dependency of each point on the neighboring vertices based on the point's degree of curvature.

Parts of this chapter have been published in:
Tajdari F, Huysmans T, Yang Y, Song Y. Feature preserving non-rigid iterative weighted closest point and semi-curvature registration. *IEEE Transactions on Image Processing*. 2022 Feb 9;31:1841-56 [251].

2.1. INTRODUCTION

IN the past decade, non-rigid registration is widely applied in many applications such as motion analysis [96], shape analysis [68], and medical image registration [75]. Among different types of non-rigid surface registration methods, the non-rigid Iterative Closest Point (ICP) registration attracted much attention, mainly due to its simplicity, efficiency and effectiveness [7]. Using the non-rigid ICP method, a source surface is registered to the target surface via two iterative steps. First, the correspondence of each vertex in the source surface to the target surface is established based on a metric, which usually is defined as a type of distance (e.g. distance between points). Then, a cost function is defined based on those correspondences, and is minimized by finding a non-rigid transformation [71]. In a practical application, these two steps iterate until a local minimum of the cost function is found or the iteration steps exceed a threshold.

Though it was successfully applied in many applications, e.g. gender scoring [89], statistical shape modeling [100], computer vision [138], multimedia applications [224], human-computer interactions [162], 3D deformation of the human spinal column detection[257], image face alignment [46], and 3D human body analysis [114], the non-rigid registration is a non-trivial and ill-defined problem with a high number of degrees-of-freedom (DOFs). Accordingly, there are many challenges for preserving features of the source surface in the design and implementation of a non-rigid ICP registration algorithm[268]. Here features account for salient geometric features which form compound higher-level descriptors. A salient geometric feature, or in short, a salient feature, consists of a cluster of descriptors that locally describe a nontrivial region of the surface [83] i.e. curvature. Those challenges are, e.g. establishing meaningful robust correspondences in each step of the iteration [265], ensuring the convergence towards the desired minimum in the optimization [143], and maintaining the quality of the mesh regarding the source surface. A typical example is that in the minimization of distances between closest points from the target to the source surfaces, the correspondences between feature points may change when the stiffness term, or the weight of it, is not large enough. This is especially true when the feature is not prominent, or when there is interference between/among features.

In the past decades, researchers made considerable progress in tackling those challenges, for instance, using landmarks (LMs) to improve the accuracy of the correspondences [109, 175, 107, 217], introducing prior knowledge regarding the geometric shapes[104], integrating more terms in the cost functions [147]. However, problems are not fully solved. For instance, in matching 3D scans of human feet, it is often found that some toes are bonded together and/or part of a toe is hidden behind another. Though researchers tried to introduce LMs, either manually or automatically, the limited number of LMs does not always contribute to establishing the desired correspondences between source and target surfaces, neither lead to a correct registration result. A better measure, which synthesizes the intrinsic properties of features of both the source and the target surface, is needed in establishing the meaningful correspondences as well as defining the cost function.

In this chapter, we propose a new asymptotically robust approach of non-rigid ICP by integrating a newly defined semi-curvature in the definition of the metric for establishing the correspondences as well as the cost function for finding the non-rigid transformation. Our scientific contributions are:

1. We introduce the semi-curvature, which is monotonically related to the Gaussian curvature with similar geometric meaning, and has unique mathematical characteristics of the capability of being linearized;

2. We integrate the semi-curvature in the cost function for establishing corresponding point pairs in two surfaces using a dynamic weighting factor to address intrinsic properties across the complete surfaces;

3. We linearize the cost function with guaranteed regions of attraction in the minimization process for a robust registration process as well as a registration result with high quality mesh.

The remainder of the chapter is arranged as follows: first recent developments in the non-rigid ICP registration method regarding the aforementioned three challenges are presented. Then we propose our approach where the definition and the characteristics of the semi-curvature are addressed. In Section 7.4, the setup of the experiments is introduced where in Section 3.5, experimental results on the comparison of the proposed approach and other non-rigid registration methods are presented. Finally, a short conclusion is drawn and future research directions are highlighted as well.

2.2. RELATED WORK

2.2.1. CORRESPONDENCES

For a better preservation of the features on the source surface, establishing meaningful dense correspondences between the source and the target surface throughout the registration process is key. Using LMs is a common method to improve the accuracy of the correspondences in the registration process. Besides manual specification of corresponding points on the two surfaces, pattern recognition algorithms are often deployed to find those LMs automatically, e.g. [109, 175, 107, 217]. However, those LMs are often located at the extrema of certain intrinsic properties and they are sparsely distributed around the surface. The full spectrum of the intrinsic properties of the surface is often not completely used in the process of establishing correspondences. To embed more shape information in the registration process, researchers introduced different intrinsic properties in establishing the correspondence. For instance, Li [141] employed a single ℓ_2-norm optimization framework utilizing confidence weights to improve robustness. Dai et al. [58] presented an iterative registration method that combines ICP with Coherent Point Drift (CPD) to achieve a more stable correspondence establishment. As a recent improvement, a Bayesian Coherent Point Drift (BCPD) approach was presented in [102]. The method utilizes the coherent drift in the variational Bayesian inference theory, while keeping the fundamental features of the CPD algorithm. Recently, a rigid ICP based registration algorithm was presented in [313] which uses curvature feature similarity to find more accurate correspondences. However, the method is sensitive to noise, and the exponential growth of the computing time regarding the number of points in the surfaces also prevents its wide application to more complicated geometries. Besides shape intrinsic properties, prior knowledge can also contribute to establish meaningful correspondence. For instance, Hontani et al. [104] used a reference statistical shape model (SSM) to find outliers in the non-rigid registration. Guo et al. [97] proposed an

ℓ_0 model for establishing correspondences between deformed body shapes, and this method had higher accuracy and robustness in dynamic 3D reconstruction and tracking. However, the construction of an SSM [310] is always an expensive task regarding both the manpower and the computing time, and it is not always possible to build an SSM a priori regarding the shape to be registered.

2.2.2. CONVERGENCE IN THE MINIMIZATION

With meaningful correspondences between the source and the target surfaces, a robust minimization strategy that is used in each step of the registration is essential for a successful registration. A few works addressed this issue regrading the metrics in the cost function, the initial conditions and the non-rigid transformation. For instance, Sharp et al. [222] used the Euclidean and topological metrics to reduce the probability of instability and possible deviation from the global minimum. The effect of initial condition was investigated in [227], where a better initial condition was proposed using a novel evaluation method based on the genetic algorithm. Regarding the rigid transformation, the transformation of the surface was restricted to the rotation with the normal vector of the faces in [134] to guarantee the robustness only for rotation components. To extend the domain of robustness (following global robustness) of the rotation angles, constraints of the rotations are proposed in [316] to limit the unnecessary rotational transformation that leads to instability. Recently, Uttaran et al. [30] presented a fast and locally robust SE-based (Special Euclidean) methodology that optimizes a cost function based on motion estimation [93]. Haris et al. [24] introduced the local minimum escape ICP algorithm, improving the conventional ICP method by proposing local minimum estimation and escape mechanisms. However, the robustness of the minimization strategy was not yet fully guaranteed.

2.2.3. MESH QUALITY

Preserving mesh quality can be attributed to part of the feature preservation in a local scale, e.g. details of the shape intrinsic properties might be lost if the resulting surface of the non-rigid registration is too smooth. Knupp [121] indicated that the mesh quality can be quantified as "*an element quality metric is a scalar function of node positions that measures some geometric property of the element*". Therefore, to preserve the mesh quality of the surface after registration, embedding regularization terms, e.g. ℓ_2-norm in a least-squares sense [238], in the formulation of the cost function is a popular choice. For instance, Amberg et al. [13] introduced an ℓ_2-norm cost function where a stiffness term is embedded to have more logical deformation considering the similarity of the surfaces. Sumner et al. [240] employed the as-rigid-as-possible term introduced by [233] to the cost function. In 2009, Liao et al. [150] described a Thin Plate Spline (TPS) based terms [52], and combined it with graduated assignment algorithm to formulate smoothness constraints. Rouhani et al. [214] also integrated locally rigid transformations to formulate the non-rigid deformation optimization problem. A sparse non-rigid registration method using an ℓ_1-norm cost function is employed by Yang et al. [306]. However, the position constraints (e.g., near piece-wise rigid deformation) were not sufficiently well embedded in the model to establish the connectivity of the piece-wise rigid deformation. Li et al. [143] recruited local affine transformations and orthogonality constraint

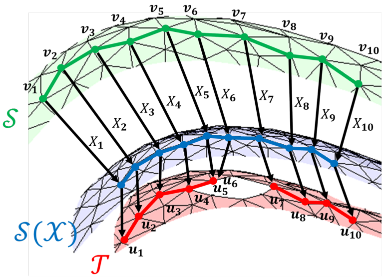


Figure 2.1: Match the source surface to the target surface.

together to capture surface details for preserving local shape in the registration [301, 297]. Different densities in the source and the target meshes also pose challenges to preserving mesh quality in the non-rigid registration. For a better match of similar surfaces in the source and target surfaces with different point densities, Tazir et al. [264] proposed the idea of matching points representing the local regions of the source cloud with the points representing the corresponding local regions in the target. Recently, Ayan [43] presented a registration method based on an energy function combining the strength of local and global geometry along with an intermediate level representation of the point cloud. The method is practical to address uniform deformation and preserve the original mesh quality; however, it is very sensitive to the point density, high deformation and detailed geometry areas.

2.3. METHODOLOGY

This section introduces the proposed non-rigid ICP approach. First, a concise description of the approach is given based on the conventional ICP algorithm [13]. This is followed by the introduction of the semi-curvature term, the improved corresponding points search approach, the cost function and the optimization process.

2.3.1. THE APPROACH

In the registration process of the non-rigid ICP, the source surface $\mathscr{S} = (\mathscr{V}, \mathscr{E})$, consisting of n vertices in $\mathscr{V}$ and m edges in $\mathscr{E}$, is registered to the target surface $\mathscr{T}$ step by step. Figure 2.1 illustrates a step of the registration process. In the figure, the meshes are assumed to be triangular meshes, and the vertices are labeled by numbers. In this step, first, the correspondences between vertices v_i in the source surface $\mathscr{S}$ (green) and vertices u_i in the target surface $\mathscr{T}$ (red) are established. Then v_i is transformed by locally affine transformations (X_i) towards the target surface $\mathscr{T}$ (red). The transformed source surface is $\mathscr{S}(X)$ (blue). This procedure iterates till an optimal stable state is obtained.

Based on the basic concept introduced in Figure 2.1, the flowchart in Figure 3.4 describes the proposed iterative approach by finding optimal X_i to transform v_i to u_i (where u_i is comparable with $X_i v_i$) in each iteration. It is worth mentioning that in the approach of Figure 3.4 there are two nested loops. In the outer loop (where k is updated), the weights in the metric for establishing correspondences are automatically updated. In

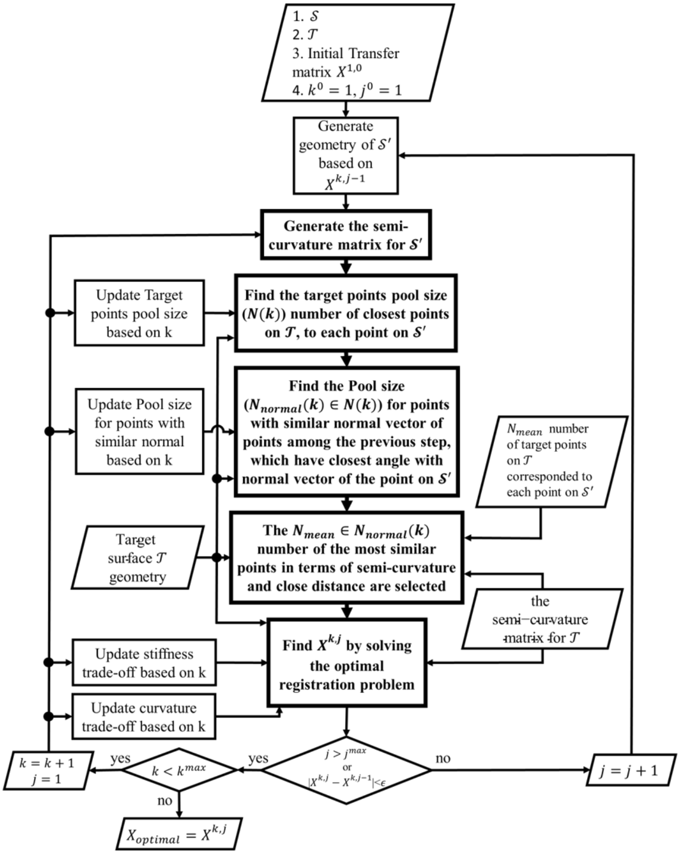


Figure 2.2: The proposed registration process.

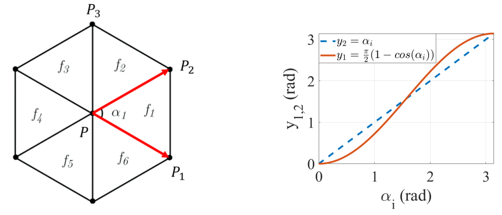
the inner loop (where j is updated), the cost function for finding the optimal transformation matrix is continuously being minimized based on the weights till the change in X is less than a small value of ϵ or $j > j^{\max}$. The two nested loops in each of their iterations account for addressing different aspects of the measures, e.g. the Euclidean distance and the difference of semi-curvatures of two corresponding points.

As a final step in the registration, each vertex of $\mathcal{V}(X)$ is projected onto the target surface along the normal vectors of the transformed surface. Keeping the original topology of the source mesh, the final transformed vertices represent the registered surface of the original source.

2.3.2. SEMI-CURVATURE

Using the notation from Figure 2.3(a), the Gaussian curvature of a point P on a triangular mesh can be approximated [304] as:

$$K_G(P) = \frac{3\left(2\pi - \sum \alpha_i\right)}{\sum \mathcal{A}(f_i)} \tag{2.1}$$



(a) (b)

Figure 2.3: (a) Curvature parameters of point P. (b) Relations between α_i and $\frac{\pi}{2}(1 - cos(\alpha_i))$.

where, α_i is the angle of the triangle formed by P and its 1-ring neighboring vertices. It can be calculated as Eq. (2.2).

$$\alpha_i = cos^{-1}\left(\frac{\left(\overrightarrow{P}_i - \overrightarrow{P}\right).\left(\overrightarrow{P}_{i+1} - \overrightarrow{P}\right)}{\| \overrightarrow{P}_i - \overrightarrow{P} \| \| \overrightarrow{P}_{i+1} - \overrightarrow{P} \|}\right) \tag{2.2}$$

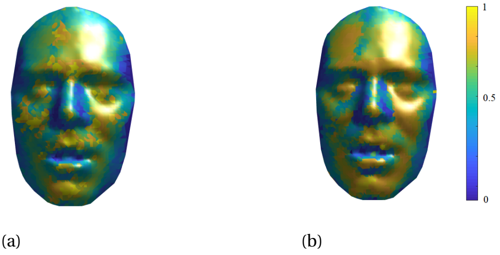
$\mathscr{A}$ in Eq. (2.1) is a function to calculate the area of each triangle f_i. It can be approached as:

$$\mathscr{A}(f_i) = 0.5 \| \left(\overrightarrow{P}_i - \overrightarrow{P}\right) \times \left(\overrightarrow{P}_{i+1} - \overrightarrow{P}\right) \| \tag{2.3}$$

As the function cos^{-1} in Eq. (2.2) makes the relationship between Gaussian curvature values and vectors of P to P_i nonlinear, we replace α_i in Eq. (2.1) by $\frac{\pi}{2}(1 - cos(\alpha_i))$. For a triangle mesh, $0 \le \alpha_i \le \pi$, therefore the domains of α_i is the same as $\frac{\pi}{2}(1 - cos(\alpha_i))$ for all vertices. Regarding the values, these two expressions equal to each other at $\alpha_i = 0, \frac{\pi}{2}$, and π, however, slightly different in other α_i as y_1 and y_2 is shown in Figure 2.3(b). Replacing α_i by $\frac{\pi}{2}(1 - cos(\alpha_i))$ in Eq. (2.1), we define the semi-curvature of vertex P as:

$$K(P) = \frac{3(2\pi - \sum \frac{\pi}{2}(1 - cos(\alpha_i)))}{\sum \mathscr{A}(f_i)} \tag{2.4}$$

The newly defined semi-curvature in Eq. (2.4) and the Gaussian curvature suggested by Eq. (2.1) are monotonically related as Figure 2.3(b). For acute α_i, we have $y_1 < y_2$ and when α_i is obtuse, $y_1 > y_2$ holds. According to Eq. (2.4), if a part of a mesh consists only of acute triangles, the value of semi-curvature on each vertex is larger than Gaussian curvature and for obtuse triangles, the value is smaller. Therefore, compared to the Gaussian curvature, semi-curvature might have a larger value when the valence of the vertex, i.e. the number of one-ring vertices, is large, e.g. ≥5. This transformation from the Gaussian curvature Eq. (2.4) to the semi-curvature Eq. (2.1), as regulator of curvature, is instigated by increasing the chance of visibility of the features (suggested by semi-curvature function versus curvature function on saddle points) as it is found, complex models' features usually are indexed to their saddle points [147]. Thus, enlarging the deviation spectrum between high curved and low curved points can improve corresponding points selection between the two meshes. To clarify, the Gaussian curvature



(a) (b)

Figure 2.4: Visualisation of normalised curvature values where yellow color indicates the highest curved area and blue color indicates the lowest curved area. (a) Using Gaussian curvature. (b) Using the proposed semi-curvature.

region was further addressed as depicted in Figure 2.4. In the figure, it can be observed that using the semi-curvature, the yellow region on the face (Figure 2.4(b)) is larger compared to using the Gaussian curvature (Figure 2.4(a)). Generally, the saddle points come into view because of exaggeration in the semi-curvature function. It is expected that the function would attempt to keep larger deviation of the semi-curvature value differences to provide a higher chance of detecting the corresponding saddles. This logical deviation gives more freedom to the deformation with less confliction between several correspondences. Moreover, increasing the valence of a vertex results in an increase of the semi-curvature value, the registration will be enhanced to logically exert different stiffness values for the areas with more features. This concludes to have $\mathscr{S}$ with different stiffness in different parts based on the the degree of pronouncedness of feature that exists in the area.

Thus, the new curvature formulation satisfies the goal of this chapter, to emphasize more on highly curved areas. As an example in Figure 2.4, the location of high curved (yellow), and low curved (blue) areas are the same comparing Figure 2.4(a) with Figure 2.4(b), while the area of the domains is different.

2.3.3. ESTABLISH CORRESPONDENCES

In the use of a conventional ICP method, given a point on $\mathscr{S}$, the closest point on $\mathscr{T}$ is considered as its corresponding point. As only the Euclidean distance is used in establishing the correspondences, the intrinsic properties of the surface are not embedded. To avoid the loss of intrinsic properties, we introduce a new criterion as Eq. (3.22), which combines the Euclidean distance and semi-curvature to find the candidate point on $\mathscr{T}$.

$$H = \zeta^{\eta} H_d + (1 - \zeta^{\eta}) H_c \tag{2.5}$$

Where H_d is the distance term composed of h_{d_i} related to P_i, and H_c is the differential semi-curvature term composed of h_{c_i} related to P_i. Moreover, in Eq. (3.22), ζ is a linearly increasing coefficient by k in range of $[0,1]$, and η is a constant parameter. The higher the value over 1 for η, the larger the effect of semi-curvature values is in driving the corresponding points selection by H. Accordingly, the smaller the η is than 1, the lower the effect of semi-curvature values are to select the correspondences by H. In the case that η is equal to 1, both distance (H_d) and semi-curvature (H_c) have equal effect in terms of the H for correspondences selection.

In practice, sparse similarities prevent us to search all points in $\mathscr{T}$ for establishing correspondences. Therefore, we only consider candidate correspondence points in a pre-processed region including N_{normal} number of points in $\mathscr{T}$. Then, H_d and H_c in Eq. (3.22) can be elaborated as:

$$\left[h_{c_i}\right]_{N_{normal}\times 1} = \frac{\dfrac{K(P_i)}{max(K)} - \dfrac{\left[K_{target}\right]_{N_{normal}\times 1}}{max(K_{target})}}{\max\left(\dfrac{K(P_i)}{max(K)} - \dfrac{\left[K_{target}\right]_{N_{normal}\times 1}}{max(K_{target})}\right)} \tag{2.6}$$

$$\left[h_{d_i}\right]_{N_{normal}\times 1} = \frac{v_i - [U]_{N_{normal}\times 1}}{\max\left(v_i - [U]_{N_{normal}\times 1}\right)} \tag{2.7}$$

where K_{target} consists of the semi-curvature values of the points on the $\mathscr{T}$ and arranged based on correspondences, H_d is a N_{normal} by 1 matrix, consisting of distances of the N_{normal} points from $\mathscr{T}$ to P_i on $\mathscr{S}$. Similarly, H_c is a N_{normal} by 1 matrix, consisting of deviation of the N_{normal} points on $\mathscr{T}$ in terms of the semi-curvature value (Eq. (2.4)) from P_i on $\mathscr{S}$. The pre-processing procedure to find the region with N_{normal} points has several steps as follows.

- Briefly, N (lowering value, from $N_r n$ to N_{mean}) number of closest points on $\mathscr{T}$ to each point on $\mathscr{S}$ are chosen, where N_r is a ratio in [0 1] and n is the number of vertices on target surface.

- Then, N_{normal} (lowering value, from M_{normal} to N_{mean}, number of vertices with most similar angle (between normal vector of the vertices on $\mathscr{T}$ and current point on $\mathscr{S}$) are selected.

- Finally, N_{mean} number of the candidate vertices with a lower value of H are picked among the previous step in which the final corresponding point is the average of these N_{mean} vertices.

Once the correspondences are selected from the target, we check the number of aligned vertices from the source to each of the points of the target. For the ones aligned to more than N_{mean} number of source vertices, we only select the N_{mean} of them with minimum H. At the end, for the source vertices that have less than N_{mean} connections, we consider the corresponded point as the transformed point of previous step. This may avoid conflicts between the correspondences and preserve the original mesh detail (shape of the surface's faces).

2.3.4. THE COST FUNCTION & MINIMIZATION

In this section, based on the established correspondences (v_i, u_i), a cost function consisting of different terms is defined and then minimized with guaranteed stability, convergence, and robustness. In the following we introduce each term in the cost function first, then we describe the optimization process based on the linearised cost function.

Conventional terms from Amberg [13]

For a non-rigid registration, the distance of the deformed source and the target should be minimized. Thus, a distance term is selected as the first component of the cost function to be minimised as,

$$E_d = \sum_{v_i \in \mathcal{V}} w_i \| X_i v_i - u_i \|^2, \tag{2.8}$$

where, w_i is the weight of the distance term, X describes a set of transformations of displaced source vertices $\mathcal{V}(X)$. The transformation matrix X_i for each vertex in the source is a 3×4 transformation matrix as:

$$X_i = \begin{bmatrix} r_{xx} & r_{xy} & r_{xz} & d_x \\ r_{yx} & r_{yy} & r_{yx} & d_y \\ r_{zx} & r_{zy} & r_{zz} & d_z \end{bmatrix}, \tag{2.9}$$

where r, and d define all afine transformations. The transformation matrix X of all vertices is described in a $4n \times 3$ matrix as $X = \left[X_1 \cdots X_n\right]^T$.

A canonical form of Eq. (2.8) is addressed in Eq. (2.10), introduced by swapping the position of transformation matrix, and correspondences (v_i, u_i). The sparse matrix D is formed to facilitate the transformation of the source vertices with the individual transformations contained in X via matrix multiplication, and denoted as $D = diag(v_1^T, v_2^T, \ldots, v_n^T)$. The corresponding points are also arranged as $U = \left[u_1 \cdots u_n\right]^T$ and the distance term can be derived as:

$$E_d = \| W \left(DX - U\right) \|_F^2 \tag{2.10}$$

where W is a diagonal matrix consisting of weights w_i. To regularise the deformation, an additional stiffness term is employed. Using the Frobenius norm $\|.\|_F$, the stiffness term penalizes difference of the transformations of neighboring vertices, through a weighting matrix $G = diag(1, 1, 1, \gamma)$.

$$E_s = \sum_{i,j \in \mathcal{E}} \| \left(X_i - X_j\right) G \|_F^2 \tag{2.11}$$

During the deformation, γ is a parameter to stress differences in the skew and rotational part against the translational part of the deformation. The value of γ can be specified based on data units and the types of deformation [13].

Addressing the function of the stiffness term to penalise differences of transformation matrices of the neighboring vertices, the node-arc incidence matrix M (e.g. Dekker [61]) of the template mesh topology is employed to convert the stiffness term functional into a matrix form. As the matrix is fixed for directed graphs, the construction is one row for each edge of the mesh and one column per vertex. To establish the node-arc incidence matrix of the source topology, the indices (i.e. the subscripts) of edges and vertices are addressed, for any edge of r which is connected to vertices (i, j), in r^{th} row of M, and the nonzero entries are $M_{ri} = -1$ and $M_{rj} = 1$. Therefore, we formulate the stiffness term as

$$E_s = \| \left(M \otimes G\right) X \|_F^2 \tag{2.12}$$

SEMI-CURVATURE TERM

Using distance term only does not fully utilize the intrinsic properties of the surfaces, e.g. high curvature regions were not addressed, this may lead to suboptimal solutions of the registration. Moreover, it may cause conflicts between different parts of a shape, e.g. selecting the same parts on the target as correspondences for different parts on the source, which are far from the correct corresponding (group of) vertices on the source surface. Thus, considering an extra term, which is independent of the distance term, in the cost function to preserve features is necessary to avoid such problems. In the proposed approach, we embed a new term defined on the semi-curvature to eliminate the conflations, and to address the intrinsic properties across the surface. It contains the differences of semi-curvature values of the points on $\mathscr{S}$ with the correspondences on $\mathscr{T}$ as follows.

$$E_c = \|K - K_{target}\|_F^2 \tag{2.13}$$

To reveal the connection of $K(P)$ in Eq. (2.4) with X, here we extend $cos(\alpha_i)$ based on the transfer matrix

$$\begin{aligned} cos(\alpha_i) &= \frac{\left(\overrightarrow{P}_i - \overrightarrow{P}\right).\left(\overrightarrow{P}_{i+1} - \overrightarrow{P}\right)}{\|\overrightarrow{P}_i - \overrightarrow{P}\|\|\overrightarrow{P}_{i+1} - \overrightarrow{P}\|} \\ &= \frac{\left(D_{P_i} - D_p\right) X X^T \left(D_{P_{i+1}} - D_p\right)^T}{\|\overrightarrow{P}_i - \overrightarrow{P}\|\|\overrightarrow{P}_{i+1} - \overrightarrow{P}\|} \end{aligned} \tag{2.14}$$

where we use $D_{P_{i+1}}$, D_{P_i}, and D_P to denote the corresponding row of D to the points P_{i+1}, P_i, and P, respectively. Generally, $D_{P_{i+1}}$ is a $1 \times 4n$ row in D with the same index that P_{i+1} has in $\mathscr{V}$. If we define

$$\begin{aligned} a_{f_i} &= \left[\frac{(D_{P_i} - D_p)}{\|\overrightarrow{P}_i - \overrightarrow{P}\|\|\overrightarrow{P}_{i+1} - \overrightarrow{P}\|}\right]_{1\times 4n}, \\ c_{f_i} &= \left[\left(D_{P_{i+1}} - D_p\right)^T\right]_{4n\times 1}, \end{aligned}$$

Eq. (2.14) turns to be

$$cos(\alpha_i) = \left[a_{f_i}\right]_{1\times 4n} X_{4n\times 3} X^T_{3\times 4n} \left[c_{f_i}\right]_{4n\times 1} \tag{2.15}$$

Using Eq. (2.15) to replace the α_i in Eq. (2.4), the semi-curvature can be calculated as:

$$K(P) = \left(\frac{3\pi}{2}\frac{\sum a_{f_i}}{\sum \mathscr{A}(f_i)}\right) X X^T \left(\sum c_{f_i}\right) - \left(\frac{-6\pi + \frac{3\pi}{2}\sum 1}{\sum \mathscr{A}(f_i)}\right). \tag{2.16}$$

Let's consider,

$$\underline{a}_P = \frac{3\pi}{2}\frac{\sum a_{f_i}}{\sum \mathscr{A}(f_i)} \tag{2.17}$$

$$\underline{c}_P = \sum c_{f_i} \tag{2.18}$$

$$\underline{b}_P = \frac{-6\pi + \frac{3\pi}{2}\sum 1}{\sum \mathscr{A}(f_i)} \tag{2.19}$$

The semi-curvature matrix on $\mathscr{S}$ is defined as:

$$K_{n\times n} = \left[\underline{A}_{n\times 4n} X_{4n\times 3} X^T_{3\times 4n} \underline{C}_{4n\times n}\right]_{n\times n} - \underline{B}_{n\times n} \tag{2.20}$$

where $\underline{A} \in \mathbb{R}^{n\times 4n}$, which is a sparse matrix composed of elements $\underline{a}_P$. Similarly, $\underline{C} \in \mathbb{R}^{4n\times n}$, which is a sparse matrix composed of elements $\underline{c}_P$. $\underline{B} \in \mathbb{R}^{n\times n}$ is a diagonal matrix composed of elements $\underline{b}_P$. Similarly, $K \in \mathbb{R}^{n\times n}$, is a diagonal matrix composed of the semi-curvature values of vertices on S. Thus, to minimize the difference of the semi-curvature between a point on the source and the corresponding point on the target, the semi-curvature term, in the cost function as E_c is presented as follows

$$\begin{aligned} E_c(X) &= \|K - K_{target}\|_F^2 \\ &= \|\underline{A}XX^T\underline{C} - (\underline{B} + K_{target})\|_F^2 \\ &= \|\underline{A}X^2 - (\underline{B} + K_{target})\underline{C}^{-1}\|_F^2 \ \|\underline{C}\|_F^2 \end{aligned} \tag{2.21}$$

where, K_{target} consists of the semi-curvature values of vertices on $\mathscr{T}$, calculated through Eq. (2.4) and arranged based on correspondence points. As $\|\underline{C}\|_F^2$ is a constant matrix in all iterations, it is considered as part of the weight matrix of W_c, or can be neglected. Also, note that in Eq. (2.21), $\underline{C}^{-1}$ is the Moore-Penrose pseudoinverse of matrix $\underline{C}$.

2.3.5. LINEARISE E_c

In order to integrate the semi-curvature term in the cost functions for optimization, it is necessary to linearise the semi-curvature term toward the forms of other terms, e.g. Eq. (2.8) . To linearise the semi-curvature term, we assume

$$f(X) = \underline{A}X^2 - (\underline{B} + K_{target})\underline{C}^{-1}. \tag{2.22}$$

If we consider linear format of $f(X)$ as $f_l(X)$, then

$$f_l(X) = A_l X + B_l\left(-(\underline{B} + K_{target})\underline{C}^{-1}\right). \tag{2.23}$$

From [113], A_l, and B_l are

$$\begin{aligned} A_l &= \frac{\partial(\underline{A}X)}{\partial X} = \underline{A} \\ B_l &= \frac{\partial\left(-(\underline{B} + K_{target})\underline{C}^{-1}\right)}{\partial\left((\underline{B} + K_{target})\underline{C}^{-1}\right)} = -1 \end{aligned}$$

As A_l and B_l are independent to X, the linear system is valid around any point. The final linearised semi-curvature term is

$$f_l(X) = \underline{A}X - \left((\underline{B} + K_{target})\underline{C}^{-1}\right) \tag{2.24}$$

Considering, $E_c(X) = \|f_l(X)\|_F^2$, $A_c = \underline{A}$, and $B_c = (\underline{B} + K_{target})\underline{C}^{-1}$, the semi-curvature term can be denoted as:

$$E_c(X) = \|W_c\left(A_c X - B_c\right)\|_F^2 \tag{2.25}$$

Stability analysis: As we linearised the curvature term in Eq. (2.21), it is possible that X, which minimises $\|A_c X - B_c\|^2$, cannot minimise $\|A_c X^2 - B_c\|^2$. Thus, here we investigate the impact of $X^o = \frac{A_c^T B_c}{{A_c}^T A_c}$ on the nonlinear cost function. From Khalil [113, 249], for any system, if $E(X) > 0$ and $\frac{\partial E}{\partial t} \leq 0$, the system is globally stable [246, 248] (in our formulation, the number of intervals are equal to the effect of time in [113], i.e. $t = k$). For the system E_c is always positive, thus we only need to study the effect of X^o on $\frac{\partial E_c}{\partial k}$

$$\begin{aligned}
\frac{\partial E_c}{\partial k}(X^o) =& \frac{\partial E_c}{\partial X}(X^o)\frac{\partial X}{\partial k}(X^o) \\
=& 2\left(\underline{A}_k X_k^{o2} - B_{c_k}\right)\left(2\underline{A}_k X_k^o\right)^T \frac{X_k^o - X_{k-1}^o}{\Delta k} \leq 0
\end{aligned} \tag{2.26}$$

where $\Delta k = 1$, $X_k^o = \frac{\underline{A}_k^T B_{c_k}}{\underline{A_k}^T \underline{A}_k}$, and $X_{k-1}^o = \frac{\underline{A}_{k-1}^T B_{c_{k-1}}}{\underline{A}_{k-1}{}^T \underline{A}_{k-1}}$ then

$$\begin{aligned}
\frac{\partial E_c}{\partial k}(X^o) =& 2\left(\underline{A}_k\left(\frac{\underline{A}_k^T B_{c_k}}{\underline{A}_k{}^T \underline{A}_k}\right)^2 - B_{c_k}\right) \\
& \left(2\underline{A}_k\frac{\underline{A}_k^T B_{c_k}}{\underline{A}_k{}^T \underline{A}_k}\right)^T\left(\frac{\underline{A}_k^T B_{c_k}}{\underline{A}_k{}^T \underline{A}_k} - \frac{\underline{A}_{k-1}^T B_{c_{k-1}}}{\underline{A}_{k-1}{}^T \underline{A}_{k-1}}\right) \leq 0 \\
=& 4\left(\frac{{B_{c_k}}^2}{\underline{A}_k} - B_{c_k}\right) {B_{c_k}}^T\left(\frac{B_{c_k}}{\underline{A}_k} - \frac{B_{c_{k-1}}}{\underline{A}_{k-1}}\right) \leq 0 \\
=& 4\left(\frac{B_{c_k}}{\underline{A}_k} - I\right)\left(B_{c_k}{B_{c_k}}^T\right)\left(\frac{B_{c_k}}{\underline{A}_k} - \frac{B_{c_{k-1}}}{\underline{A}_{k-1}}\right) \leq 0
\end{aligned} \tag{2.27}$$

Here $\left(B_{c_k}{B_{c_k}}^T\right)$ is always positive, thus if $\left(\frac{B_{c_k}}{\underline{A}_k} - I\right)\left(\frac{B_{c_k}}{\underline{A}_k} - \frac{B_{c_{k-1}}}{\underline{A}_{k-1}}\right) \leq 0$, the system is stable. In this case, we have two scenarios for $\left(\frac{B_{c_k}}{\underline{A}_k}\right)$ and $\left(\frac{B_{c_{k-1}}}{\underline{A}_{k-1}}\right)$: **Scenario 1**:

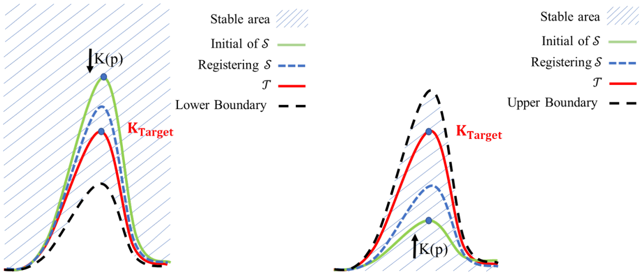
$$\frac{B_{c_{k-1}}}{\underline{A}_{k-1}} \leq \frac{B_{c_k}}{\underline{A}_k} \leq I \tag{2.28}$$

and **Scenario 2**:

$$I \leq \frac{B_{c_k}}{\underline{A}_k} \leq \frac{B_{c_{k-1}}}{\underline{A}_{k-1}} \tag{2.29}$$

Discussion: For an arbitrary point P, from (2.14), relative $\underline{A}\ \ \underline{C}$ can be considered as following, while $\frac{B_c}{\underline{A}} = \frac{\left(\underline{B}+K_{target}\right)}{\underline{A}\ \underline{C}}$:

$$\left(\underline{A}\ \underline{C}\right)_p = \frac{3\pi}{2}\frac{\sum\frac{(D_{P_i}-D_p)(D_{P_{i+1}}-D_p)}{\|\overrightarrow{P}_i-\overrightarrow{P}\|\|\overrightarrow{P}_{i+1}-\overrightarrow{P}\|}}{\sum\mathscr{A}(f_i)} \tag{2.30}$$



(a) (b)

Figure 2.5: Stability boundaries. (a) Lower boundary for Scenario 1. (b) Upper boundary for Scenario 2.

Suppose

$$r_p = \frac{\|\overrightarrow{P}_i^0 - \overrightarrow{P}^0\|\|\overrightarrow{P}_{i+1}^0 - \overrightarrow{P}^0\|}{\|\overrightarrow{P}_i - \overrightarrow{P}\|\|\overrightarrow{P}_{i+1} - \overrightarrow{P}\|} \tag{2.31}$$

where $\overrightarrow{P}^0$, $\overrightarrow{P}_i^0$, and $\overrightarrow{P}_{i+1}^0$ belong to $\mathscr{S}$ before registration, Eq. (2.30) can be denoted as:

$$\begin{aligned}\left(\underline{A}\ \underline{C}\right)_P &= \frac{3\pi}{2}\frac{\sum \frac{(D_{P_i}-D_p)(D_{P_{i+1}}-D_p)}{\|\overrightarrow{P}_i^0-\overrightarrow{P}^0\|\|\overrightarrow{P}_{i+1}^0-\overrightarrow{P}^0\|}\ r_p}{\sum \mathscr{A}(f_i)} \\ &= \frac{3\pi}{2}\frac{\sum cos(\alpha_i^0)\ r_p}{\sum \mathscr{A}(f_i)}\end{aligned} \tag{2.32}$$

where α_i^0 belongs to f_i in Figure 2.3(a) before registration. Considering $\left(\underline{B}\right)_P$ as in Eq. (2.19)

$$\frac{\underline{B_c}}{\underline{A}} = \frac{\frac{-6\pi+\frac{3\pi}{2}\sum 1}{\sum \mathscr{A}(f_i)} + K_{target_P}}{\frac{3\pi}{2}\frac{\sum cos(\alpha_i^0)\ r_p}{\sum \mathscr{A}(f_i)}} \tag{2.33}$$

where K_{target_P} is the semi-curvature value of the corresponding point for P on $\mathscr{T}$. Thus,

$$\Delta\left(\frac{\underline{B_c}}{\underline{A}}\right) \propto -\Delta K \tag{2.34}$$

Scenario 1: The proposed condition in Eq. (2.28) happens only when $\frac{\underline{B_{c_{k-1}}}}{\underline{A_{k-1}}} \leq \frac{\underline{B_{c_k}}}{\underline{A_k}}$, which means $K_{k-1}(P) \geq K_k(P)$ from Eq. (2.34). And, curvature in point P is decreasing, according to Figure 2.5(a). Replacing Eq. (2.33) by Eq. (2.28), we will have

$$\frac{-6\pi + \frac{3\pi}{2}\sum 1}{\sum \mathscr{A}(f_i)} + K_{target_P} \leq \frac{3\pi}{2}\frac{\sum cos(\alpha_i^0)\ r_p}{\sum \mathscr{A}(f_i)} \tag{2.35}$$

Adding $-\frac{3\pi}{2}\frac{\sum cos(\alpha_i)}{\sum \mathscr{A}(f_i)}$ to both sides of the above equation, and from Eq. (2.4)

$$K(P) - K_{target_p} \geq \frac{3\pi}{2}\frac{\sum cos(\alpha_i) - cos(\alpha_i^0)\ r_p}{\sum \mathscr{A}(f_i)} \tag{2.36}$$

This formula indicates that if the semi-curvature in point P is decreasing to reach the value of the semi-curvature of the corresponding point on $\mathscr{T}$, there is a lower limit for stability condition which is less than K_{target_P}.

Scenario 2: Using the same procedure as Scenario 1 and replacing Eq. (2.33) by Eq. (2.29), we will have

$$K(P) - K_{target_p} \leq \frac{3\pi}{2} \frac{\sum cos(\alpha_i) - cos(\alpha_i^0)\ r_p}{\sum \mathscr{A}(f_i)} \tag{2.37}$$

This formula means that if the semi-curvature in point P increases to reach the value of the semi-curvature of the corresponding point on $\mathscr{T}$, as shown in Figure 2.5(b), there is an upper limitation for stability condition, which is greater than K_{target_p}. The stability discussion in both scenarios explains an asymptotically stable approach for the optimal problem. According to [113], any system which is asymptotically stable is also convergent while the states are in their region of attractions which holds for the system discussed here.

To guarantee the stability, for the points that satisfy the criteria, we consider $W_c(P) = 1$, as the semi-curvature term is stable, otherwise $W_c(P) = 0$. Thus, the complete quadratic cost function can be considered as

$$\begin{aligned} E(X) &= \left\| \begin{bmatrix} \lambda M \otimes G \\ WD \\ \beta W_c A_c \end{bmatrix} X - \begin{bmatrix} 0 \\ WU \\ \beta W_c B_c \end{bmatrix} \right\|_F^2 \\ &= \| AX - B \|_F^2 \end{aligned} \tag{2.38}$$

Note that to integrate B_c in Eq. (7.28), we change the dimension of $\underline{C}^{-1}$ in Eq. (2.24) from $n \times 4n$ to $n \times 3$. Where $B_c = (\underline{B} + K_{target})\underline{C}^{-1}$, as the components are like matrix D, and accordingly $\underline{C}^{-1}$ is a sparse matrix.

Summarizing the process of establishing the correspondence and minimizing the cost function, that are both defined on the newly introduced semi-curvature, we propose Algorithm 1 to detail the process presented in Figure 3.4 for describing the proposed non-rigid ICP approach. In the algorithm, for each iteration with constant number of corresponding points and optimization parameters, the optimization process continues in an inner loop, until the changes in new X is less than a threshold of ϵ or the the number of iterations in the inner loop exceeds a defined number as $j^{\max}$.

2.4. EXPERIMENT SETUP

2.4.1. DATA-SET

In the experiment, besides simple geometrical shapes, the human foot, the lumbar vertebrae and the full human body are selected as typical geometric shapes for evaluating the proposed method. Data of those geometric shapes are collected from several data-sets. The first data-set being used is the SHREC'14 data-set [196], in which we selected the feet in data-set number 25 as the source surface. Before the experiment, the meshes of both feet were pre-processed to have a more uniform mesh using ACVD, a freely available software provided by Valette et al. [271]. Finally, two meshes, each with 5000 vertices, were used as the inputs of the experiment.

Algorithm 1 The proposed non-rigid ICP approach

```
Input: $X \in \mathscr{R}^{4n\times 3}$, $K \in \mathscr{R}^{n\times 1}$, and $K_{target} \in \mathscr{R}^{m\times 1}$
Output: $X_{optimal}$ to minimize $E(X)$ in Eq. (7.28)
Initialization
$k = 1$
while $k \le k^{\max}$ do
    $j = 1$
    while $\|X^{k,j} - X^{k,j-1}\| < \epsilon$ or $j < j^{\max}$ do
        $N(k)$ number of closest $u$ on $\mathscr{T}$ to $v_j$ on $\mathscr{S}(X^{k,j})$ ($\left[N(k) \in N_r n : N_{mean}\right]$)
        $N_{normal}(k)$ number of closest angle (among previous step), between normal vector of $u$ with $v_j$ ($N_{normal}(k) \in \left[min(M_{normal}, N(1)) : N_{mean}\right]$)
        $N_{mean}$ number of $u$ (among previous step), with lower $H$ in Eq. (3.22)
        $u_{corresponded}(j)$ = average of previous step
        $X^{k,j} = \frac{A^T B}{A^T A}$; ($A, B$ in Eq. (7.28))
        $j = j + 1$
    end while
    $X^{k}_{optimal} = X^{k,j}$
    $k = k + 1$
end while
```

The second data-set is part of the 3D DINED data-sets [270], consisting of 22 high-resolution 3D scans of foot (left and right) obtained from 11 people (4 females and 7 males). 3D scans were acquired by two Artec EVA scanners [18]. Before registration, we re-meshed the scans with the same method used for SHREC'14 and the number of vertices for each mesh is set as 10000.

To evaluate the proposed approach on more complicated freeform surfaces, two lumbar vertebrae were extracted from the data-set presented in [29], which includes a total of 86 models of lumbar vertebrae. In the data-set, lumbar vertebrae are mainly labelled by the prefix L and numbered 1 to 5 (e,g,. $L_4 - 20$ belongs to lumbar number 4 and case 20). We selected $L_4 - 20$ as the source surface and $L_1 - 17$, $L_1 - 18$, $L_1 - 19$ and $L_1 - 21$ as the target surfaces. The source is re-meshed to 20000 vertices through [271], and the targets are re-meshed to 8000 using the same method.

The full human body scans are selected from the Civilian American and European Surface Anthropometry Resource (CAESAR) data-set [210] as they are natural scans with holes, missed parts and natural noise which makes it suitable for assessment. In detail, the CAESAR data-set contains 3D human full body scans of the civilian populations of three North Atlantic Treaty Organization (NATO) countries; the United States of America (USA), The Netherlands, and Italy [210]. In this study, we selected the first 101 scans from the Dutch population as the target surface and evaluated through the predefined landmarks in each mesh. 74 landmarks are explained in [210] (from page 17 to page 30), while we used 73 of them (LM74 as butt block is neglected). For the source mesh, we used the full body template included in the Wrap 3 software [215].

Table 2.1: Parameters used in simulation configuration.

Parameter	Value	Description
$\lambda(k)$	1000:1	Stiffness trade-off
$\beta(k)$	1000:1	Semi-curvature trade-off
$\zeta(k)$	0:1	Distance gain in Eq.(3.22)
η	1	ζ power in Eq.(3.22)
N_{mean}	3	Target points averaging size
N_r	0.1	Initial ratio of n points on $\mathscr{S}$
$N(k)$	$N_r n$:N_{mean}	Target points pool size
M_{normal}	20	Initial number of the middle region members
$N_{normal}(k)$	$M_{normal} : N_{mean}$	Pool size for points with similar normal
ϵ	0.001	Convergence error threshold
$j^{\max}$	50	Convergence iteration threshold
$k^{\max}$	20	Number of iteration for the outer loop in Algorithm 1

2.4.2. PARAMETERS OF THE PROPOSED METHOD

Table 6.1 presents parameters used in the experiment. In establishing the correspondence, the algorithm is designed to emphasise on the semi-curvature at the beginning, and ends on the closest Euclidean distances to find the correspondence point. Thus, ζ changed from 0 to 1, and $1-\zeta^{\eta}$ changed from 1 to 0, regarding to Table 6.1.

During the minimization of the cost function, γ in G introduced in Section 2.3.4.1, was chosen to one. λ, named as gradual relaxation of the stiffness constraint, was employed for each method, decreasing from 1000 to 1. Regarding dependency of λ values to the dynamic of source surface, this value was manually defined so that only global deformations were considered in the beginning of registration. On the other hand, the lower limit of λ also depends on the data type [13]. Accordingly, a small λ may cause singularity of A in Eq. (7.28), which leads to instability of the solution. Therefore, our experiments started with a sufficiently high λ. A high value of λ was not problematic as λ had no effect on quality of the registration results, however more steps were expected, e.g. in Table 6.1, λ varied from 1000 to 1 in 20 iterations. As we want to have a fine match for the points with a larger weight on semi-curvature in the beginning but more emphasis on the distance at the end, we set the values of β using the similar strategy as in specifying λ, i.e. gradually reducing it from 1000 to 1 as shown in Table 6.1.

To have a smoother registration process, we were averaging a number of points from a set with size of 3 ($N_{mean}=3$). This N_{mean} number of points were offered by H in Eq. (3.22). Emphasizing on the semi-curvature at the beginning and the closest Euclidean distance at the end of simulation, N number of points on $\mathscr{T}$ in Algorithm 1, as a region to search for corresponding point to any vertices on $\mathscr{S}$, changed from 10 % of the total number of vertices in $\mathscr{T}$ to N_{mean}.

In the implementation of [42], the system of linear equations that arises in each step was solved with the help of the UMFPACK library [60].

2.4.3. OTHER METHODS FOR COMPARISON

In the experiment, we compared the performance of the proposed method with existing methods proposed by Amberg et al. [13], Lee et al. [136], Vestner et al. [274], Andriy et al. [180] and Hirose [102]. Briefly, Amberg's method accounts for an optimal step non-rigid ICP approach capable to employ different regularisations, while they are using a range of lowering stiffness parameter. However, the semi-curvature is not used. As the proposed method is based on Amberg's method, therefore in the implementation of Amberg's method, we removed the semi-curvature information from our method, i.e. the cost function of Eq. (2.10) and Eq. (3.3) are changed to:

$$E(X) = \left\| \begin{bmatrix} \gamma M \otimes G \\ WD \end{bmatrix} X - \begin{bmatrix} 0 \\ WU \end{bmatrix} \right\|_F^2 \tag{2.39}$$

Then Eq. (7.28) is employed for utilizing the semi-curvature in the non-rigid ICP registration.

Lee's method is able to establish correspondences between non-rigidly deformed shapes through mapping the shape to a unit Möbius sphere by centered conformal parameterization. Ultimately, they exert Fast Fourier Transformation (FFT) to detect the optimal rotational alignment between sphere meshes and perfect the registration process through optical flow. However, their method only works for Genus-zero shapes, which is a limitation. The executable source code used in this paper, is available in [137].

Vestner' method presents a methodology to specify the correspondences between two shapes which may be non-isometric shapes. The method uses kernel instead of distance as the descriptor, which makes the approach more sensitive to the quality of the mesh and the triangles size. The executable source code used in this paper, is available in [275].

Andriy introduced the CPD algorithm [180] employing Gaussian radial basis functions as a replacement for thin-plate splines which addresses a different version of regularizer. In the method, the rigid and non-rigid registrations are covered, while the approach is ill suited to account for a considerably large amount of outliers and in exploring among all the possible correspondences which is in general Non-deterministic Polynomial-time hard (NP-hard). The executable source code used in this paper, is available in [179].

Hirose proposed the BCPD method [102], using variational Bayesian inference theory to explain the coherent drift. The executable source code used in this paper, is available in [101]. All algorithms were implemented using Matlab®R2020a on a computing platform with an Intel® Core-i5™ 9600K 4.6 GHz processor.

2.5. EXPERIMENT RESULTS

2.5.1. EFFECT OF CURVATURE TERM DURING REGISTRATION

In this subsection, through a few experiments using some basic shapes, we investigate the effect of the semi-curvature term on the feature preservation after registration in terms of correspondences, and the mesh quality, especially for highly curved area.

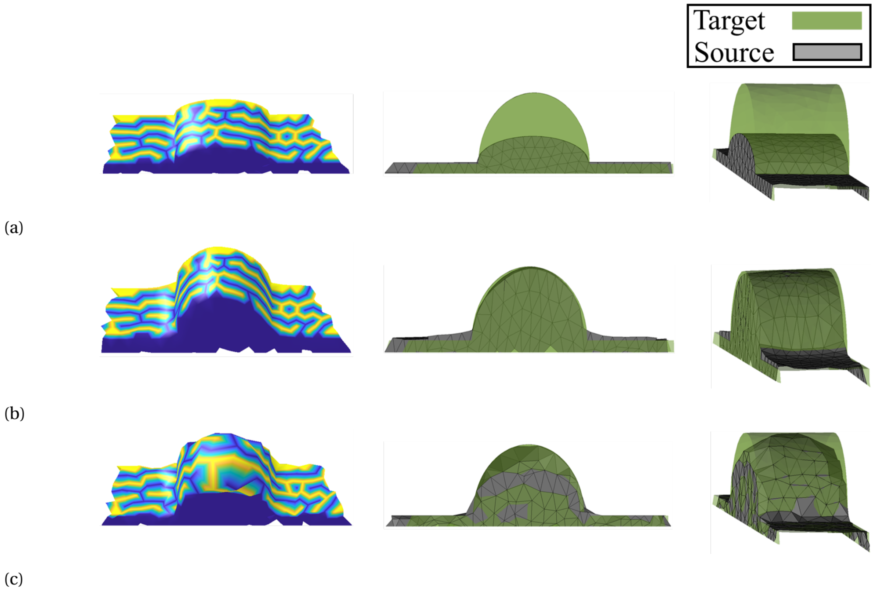


Figure 2.6: Curved shape preservation experiment (a) The initial setup (b) Registration using the proposed method (With the semi-curvature term). (c) Registration using Amberg's method (without the semi-curvature term).

FEATURE PRESERVATION

As we want to observe the effects of with and without the semi-curvature term, only Amberg's method and the proposed method were used. When the stiffness term is sufficiently low, in each iteration, the surface can be changed, and consequently the original curvature of the source surface will be changed. This causes an error in establishing the correspondences through Eq. (3.22). In this case the semi-curvature term contributes an additional logical connection between vertices and their neighbors, over the stiffness term to increased smoothness. Apparently, the improvements in high-curved areas are more visible than the flat areas due to the properties of the semi-curvature. To show the effect, a shape with ellipsoidal cross-section (gray) is used as $\mathscr{S}$, a shape with spherical cross-section (green) is employed as $\mathscr{T}$, and a colourful mesh of $\mathscr{S}$ is utilized shown in Figure 2.6. Figure 2.6(a) shows the initial condition, Figure 2.6(b) presents the registration through Eq. (7.28) and H in Eq. (3.22), and Figure 2.6(c) presents the registration through Eq. (7.28) without semi-curvature term and the H. A comparison of Figure 2.6(b) and (c) reveals that the ellipsoidal shape is preserved for all cross-sections when E_c is used due to both the additional meaningful connections of the neighbouring vertices and more the meaningful correspondences selection. The results are also more visible in the colorful graphs of Figure 2.6 left column, that the colors on the spherical cross-section without the semi-curvature are stretched and less preserved than the results with the semi-curvature.

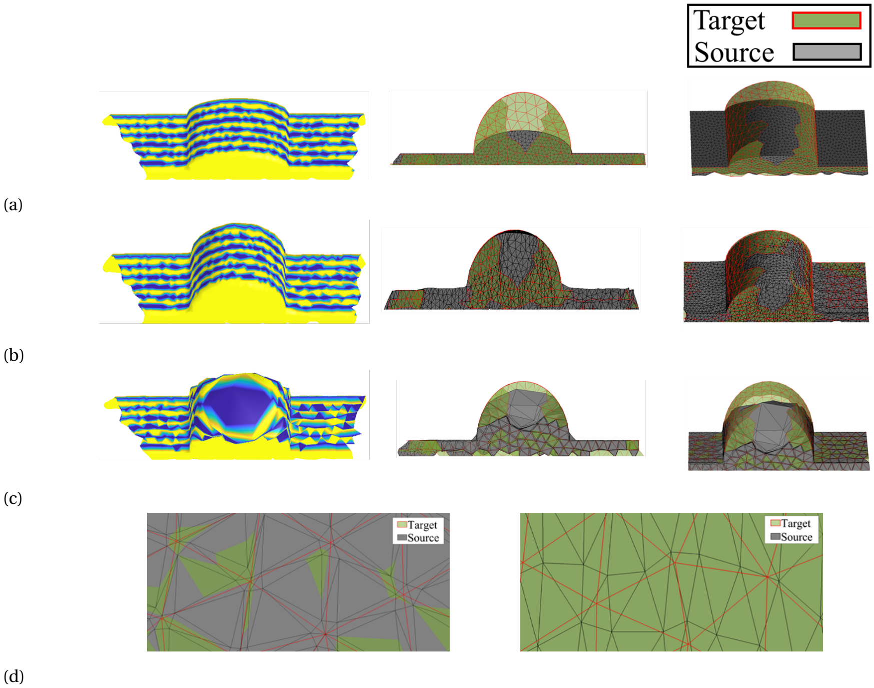
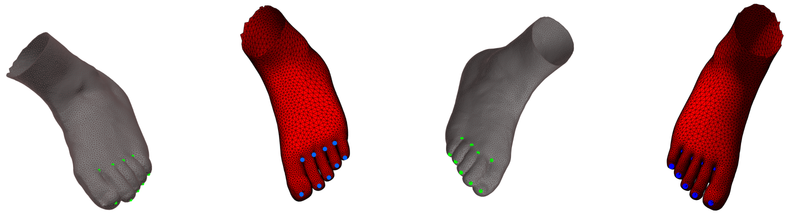


Figure 2.7: Missed part experiment (a) The initial setup (b) The proposed method (with the semi-curvature term). (c) Amberg's method (without the semi-curvature term) (d) Mesh quality: Amberg's method (left), The proposed method (with the semi-curvature term) (right).

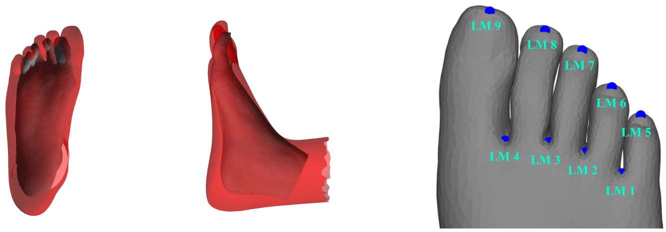
Missing parts and sparse / dense mesh

Using the proposed method, although we find more constraints for neighboring vertices of each point on source, extra unnecessary connections are also deleted based on H. After finding the correspondences, the number of connections to the corresponding points on the target is checked and if they are more than N_{mean}, we choose N_{mean} of them with minimum H. This correction avoids registering the points on source which are describing the missed part on the target. Only, they are wisely deformed based on stiffness term, and the semi-curvature term guarantees the uniform shape on the borders of the missed part. This is shown in Figure 2.7. In this experiment, Figure 2.7(a) shows the initial condition, Figure 2.7(b) defines the effects of curvature term on quality of registration for the missed part, and 2.7(c) depicts the registration without the semi-curvature term. Moreover, wisely deleting connections based on H preserves the original mesh detail (shape of the faces) and uniformity of the surfaces which is addressed by the semi-curvature term. Thus, the final registered mesh is very different from the target and the resolution is equal to the original source surface. The results can be explored through the shaded graphs showing that the semi-curvature term preserved the mesh structure for the missed part (Figure 2.7(b)(left)) as the faces with blue color is stretched around the missing part when the semi-curvature is not used (Figure 2.7(c)(left)). A comparison of



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

Figure 2.8: Meshes and LMs. (a) Target surface ($\mathscr{T}$) of left foot. (b) Source surface ($\mathscr{S}$) of left foot. (c) Target surface ($\mathscr{T}$) of right foot. (d) Source surface ($\mathscr{S}$) of right foot.



(a) (b) (c)

Figure 2.9: The initial setup: (a) Front view, (b) Side view. (c) LMs location on a sample foot.

Figure 2.7(d)(left) with Figure 2.7(d)(right) reveals that without the semi-curvature term in both H and E, the final mesh is going to be degenerate (i.e. the collapse into zero area faces) and some points have the same geometry. The degeneration happens when number of source vertices is higher than number of target vertices, thus many points on the source have the same corresponding points.

2.5.2. VALIDATION: THE FOOT SCANS

REGISTRATION PROCESS AND RESULTS

Figure 2.8 depicts the source ($\mathscr{S}$) and the target surfaces ($\mathscr{T}$) used in the registration. In the figure, location of manually defined LMs on the right/left foot are highlighted and those LMs are used to evaluate the registration results of using different methods. In all experiments, the initial conditions (Figure 2.9) are kept the same for a more accurate comparison.

The difference between landmarks in the registered $\mathscr{S}$ and the corresponding landmarks in the $\mathscr{T}$ represents the accuracy of the registration, which can be highly benefited from a non-rigid ICP method incorporated with a meaningful correspondence selection method and a robust convergence. For this, we defined a series of LM_j on all meshes as depicted in Figure 2.9(c), where $j = 1,\dots,9$. The registration results in different steps ($i = 1,6,13,$ and 20) of using the algorithms with/without semi-curvature (Amberg/our method) are presented in Figure 2.10. In the figure, it can be found that the proposed method is more robust against the interference of geometries in the toe area, which is clearly visible after $i = 13$. In Figure 2.11, the registered $\mathscr{T}$ and $\mathscr{S}$ using Am-

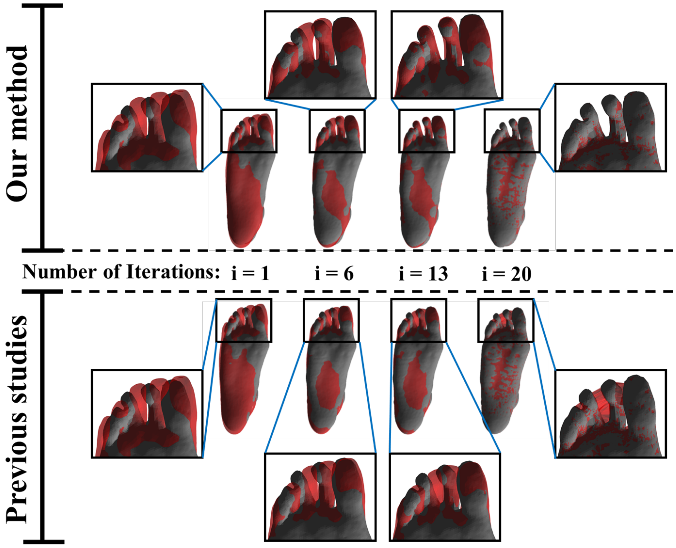
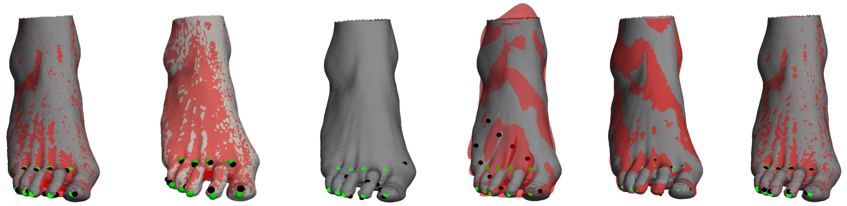


Figure 2.10: Registration results. Proposed method (top), Amberg's method (bottom).

berg's method, Lee's method, Vestner' method, the CPD method, the BCPD method and our approach are presented with the LMs, respectively. In the figure, green points represents LMs on $\mathscr{T}$ and black ones are LMS on $\mathscr{S}$ (only for Vestner' method, both the green and black points are on the target, as the approach only defines the correspondences and does not support any registration). It can be found that at the final stage there is a considerable distance between final black and green points in Figure 2.11(a)(b)(d)(e), especially for the tip of mid-toe compared to Figure 2.11(f), which utilizes our method.

Geometry-wise, there are many metrics used in evaluating the performance of the non-rigid ICP algorithm, such as Mean Absolute Error (MAE) [58], Root Mean Square Error (RMSE) [86], mean S value [46], average F1-measure (F1) [46] according to different applications. Among those, we investigate the RMSE metric for all the points on the registered $\mathscr{S}$, MAE for the landmarks and mesh quality values which allow us to evaluate the level of geometry preservation and accuracy. The foot registration results and the error of the LMs are reported in Table 2.2(a), where average Diagonal of the Bounding Box (DBB) for $\mathscr{S}$ for the source mesh is 362.63 mm. These results imply that not only the mean error of the proposed method is less, the standard deviations (SD) of errors are smaller as well. Accordingly, the semi-curvature term could offer more accurate corresponding points which means that the final transformed source points (black) are more close to the optimal positions (green). The major reason is that the proposed algorithm is more sensitive for points with higher curvature and usually flat areas (low curved points) are least likely to have LMs. Finally, the overall error per method in the table leads us to the conclusion that the proposed method is more accurate with less error and less SD.

Figure 3.10(a) depicts the percentage of correspondences including all 9 selected landmarks for all the 22 scans (y-axis) that have less distance error than a threshold (x-axis) [116]. Regarding the figure, our method (green line) detects the correspondences



(a) (b) (c) (d) (e) (f)

Figure 2.11: LM location comparison. (a) Amberg's method. (b) Lee's method. (c) Vestner's method. (d) CPD method. (e). BCPD method. (f) Our method.

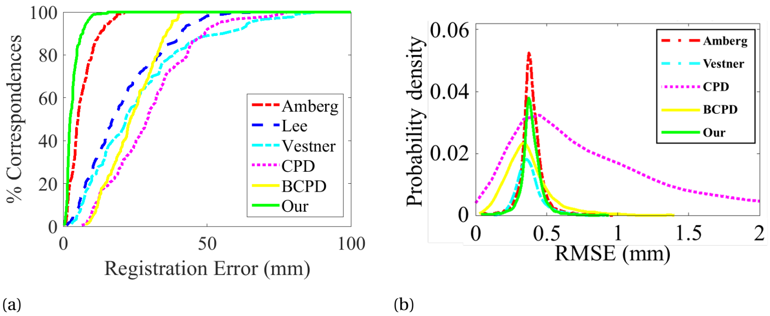
Table 2.2: Foot registration results.

a) Overall LM mean error ± Standard Deviation (SD) (mm) result.						
Mean DBB of $\mathscr{S}$ = 362.63 mm						
(lr)2-7	Amberg [13]	Lee [136]	Vestner [274]	CPD [180]	BCPD [102]	Our method
LM 1	5.9±4.4	20.3±14.5	29.4±23.3	41.4±18.1	24.1±19.3	**3.5±2.3**
LM 2	6.1±4.7	23.3±15.8	21.3±10.5	33.2±21.4	25.8±10.8	**2.9±3.1**
LM 3	5.1±4.1	26.4±18.2	20.3±10.6	28.5±22.5	25.6±11.9	**2.3±1.5**
LM 4	5.4±3.4	25.9±16.1	23.8±22.4	16.3±12.6	23.3±20.1	**3.0±1.9**
LM 5	6.7±5.2	18.9±12.1	39.1±31.9	54.7±27.9	14.1±9.9	**3.0±2.2**
LM 6	6.5±4.6	21.6±9.4	27.9±13.9	35.3±14.3	23.2±12.4	**3.3±2.1**
LM 7	7.6±5.3	21.5±8.1	25.9±9.7	22.8±16.6	21.1±8.5	**3.1±1.8**
LM 8	8.3±5.3	19.9±10.1	28.1±16.2	10.9±8.3	24.2±16.1	**3.7±3.6**
LM 9	6.0±2.9	7.3±4.1	21.3±22.8	8.9±6.2	12.9±7.7	**4.8±2.6**
Overall	6.4±4.5	20.6±13.6	26.3±19.2	30.5±15.3	21.6±8.9	**3.3±2.5**
b) Mesh quality mean value ± Standard Deviation (SD).						
Original mesh	Amberg [13]	Lee [136]	Vestner [274]	CPD [180]	BCPD [102]	Our method
0.92±0.07	0.51±0.48	0.65±0.33	NaN	0.71±0.25	0.78±0.16	**0.83±0.13**
c) The complete process computing time (s).						
	Amberg [13]	Lee [136]	Vestner [274]	CPD [180]	BCPD [102]	Our method
Time (s)	**9.3**	32.7	28.4	1266.5	680.3	17.1

earlier than the other five methods, and the errors are lower than 10 mm, while for Amberg (red line), Lee (blue line), Vestner (cyan line), CPD (purple line), and BCPD (yellow line), they are 25 mm, 68 mm, 88 mm, 75 mm, and 43 mm, respectively. The results confirm that the proposed methodology has better accuracy and robustness compared to other methods. Figure 3.10(b) shows the probability density based on the RMSE of each point on the registered $\mathscr{S}$ and the closest point on the $\mathscr{T}$ along the 22 scans. As can be seen Amberg (red) has the least spread and CPD has the largest sparsity in terms of closest point, this is because Amberg has no term to keep the original geometry of $\mathscr{S}$. Besides the error, our method has the least error of correspondences and acceptable level of the source mesh geometry preservation.

MESH QUALITY

Deformations of the computational mesh arising from optimization routines usually lead to a decrease of mesh quality or even destruction of the mesh [158]. Mesh quality is representative for various quality metrics for the shape of mesh elements [45], such as



(a) (b)

Figure 2.12: (a) Percentage correspondences according to registration error. (b) RMSE distribution.

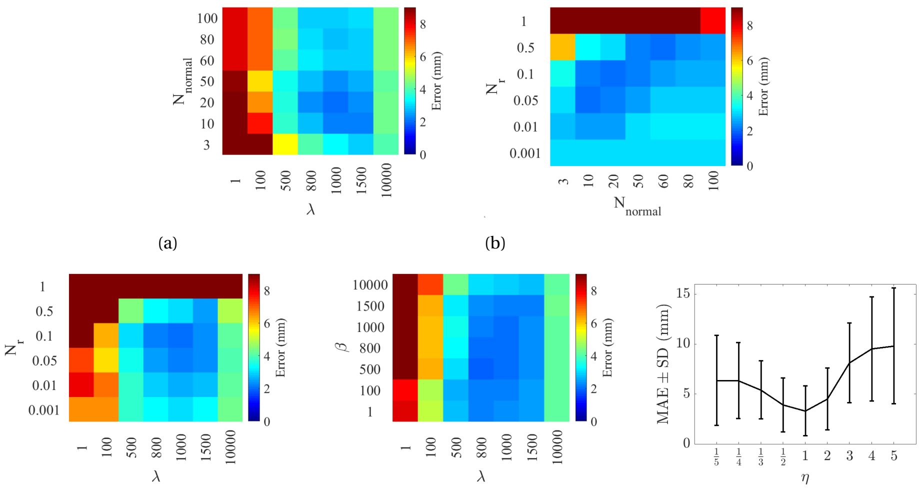
dihedral angles [81], the longest edge over the in-radius [88], and shape of the elements and smoothness [70]. Among various methods, we employ Liu and Joe's parameter approach [152] to compare the quality of meshes. The approach defines the quality per vertex (average of the mesh quality of the faces incident to the vertex) with a number between 0 and 1, in which 1 is the best quality and zero is the least quality. The mesh quality of a mesh is calculated as the mean of the quality of all vertices in the mesh. Accordingly, the mesh quality of registration results is reported in Table 2.2(b), where the mean and SD of the mesh quality of 22 original scans, and the mesh quality of the registration results using six methods are presented, respectively. The table defines 44%, 29%, 23% 15%, and 10% mesh quality loss for Amberg's method, Lee's method, the CPD method, the BCPD method and our approach respectively. As the output of Vestner' method is only the set of corresponded vertex ids, the mesh quality is not applicable.

RUN TIME DISCUSSION

The computing speed is used as the criterion to evaluate the efficiency of each mentioned method. Accordingly, here we report the average of computing time of the whole process for all the experiments in Table 2.2(c). The proposed method is slower than Amberg's methods for the complete process, which is mainly caused by the extra computing needed for the introduced semi-curvature term in Eq. (2.39). However, the needed computing times for minimizing the cost function are similar, 6.48 (s) for Amberg's method and 6.19 (s) for our method. It can be inferred that using the proposed method reduces 48% of the error at the cost of nearly same computing time for minimization. As Lee's method, Vestner' method, CPD, and BCPD take more computing time and the results are less accurate, it indicates that the proposed method is acceptable regarding the computing time, mainly due to the improved accuracy.

SENSITIVITY ANALYSIS

We initially perform a set of experiments using the 3D DINED data-sets of foot in [270] to investigate the sensitivity of the parameters λ, β, N_r and N_{normal}, and η for selecting the proper range of parameters [242]. As can be seen from Figure 3.5, which shows how the obtained LM error varies by altering, in pairs, the weights used in the cost function Eq. (7.28), the method is capable of obtaining a low the error value (blue areas) for a wide



(a) (b)

(c) (d) (e)

Figure 2.13: Sensitivity analysis showing error of the LMs for: (a) a domain of λ and N_{normal}, while $N_r = 0.1$; (b) a domain of N_{normal} and N_r, while $\lambda = 1000$; (c) a domain of λ and N_r, while $N_{normal} = 20$; (d) a domain of λ and β, while $N_{normal} = 20, N_r = 0.1$. (e) Sensitivity of error the LMs on η.

range of those parameters. This indicates the robustness level of the method, as it is less sensitive to the choice of the range of parameters. Theoretically, the robustness is expected as discussed in Section 2.3.5, the system is stable for the region of Eq. (2.36) and (2.37), which are infinite from one side. The robustness is desirable for practical applications, since it implies there is no need for fine tuning in order to achieve satisfactory results.

To investigate the effect of η, we show the average error of LMs for the all foot scans in Figure 3.5(e) based on MAE error of LMs and the SD. The error is minimal when η is equal to 1. If the value of η is in other regions, the errors might increase, however, it is still stable.

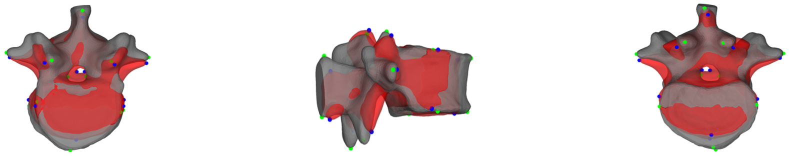
2.5.3. VALIDATION: THE LUMBAR VERTEBRA & THE FULL BODY

In line with Subsection 2.5.2, to evaluate the performance of the proposed method, we implement the method on the lumbar vertebrae data set and the full body explained in section 7.4.1.

THE LUMBAR VERTEBRA

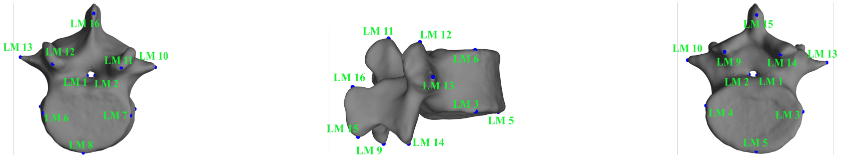
Using the same method in the previous section, we register the lumbar vertebra of L_4-20 in section 7.4.1 on the lumbar vertebrae of L_1-17, L_1-18, L_1-19 and L_1-21 using Amberg's method, the CPD method, the BCPD method and the proposed method. Lee's method is excluded since it is limited to genus 0 topology.

In all three experiments, the initial conditions for the four methods are the same. As an example, the initial condition of the L_4-20, and the L_1-17 is presented in Figure 2.14. For the assessment, we used 16 landmarks as shown in Figure 2.15. The reg-



(a) (b) (c)

Figure 2.14: Lumbar initial condition.



(a) (b) (c)

Figure 2.15: LM location in lumbar. (a) Top view. (b) 3D view (c) Bottom view.

istration results of using the four methods are compared in Figure 2.16, where in the target column of Figure 2.16, the quality of correspondences of LMs can be visualized with four highlighted areas. In the initial conditions, the foramen on the source surface is much smaller than the target and located on top of the target's foramen, and as Amberg's method uses closest distance to find the correspondences, the points may get stuck there and the hole disappears. However, the semi-curvature term can distinguish the differences as the curvature values are different. For the other two selected areas on the left-middle vertebra, the proposed method also outperforms Amberg's method in preserving regions with high curvature.

Errors between corresponding landmarks are reported in Table 2.3(a) where our method outperforms Amberg's method, CPD method, and BCPD method by 58% , 87%, and 81% respectively regarding the mean absolute error. Moreover, our approach, the CPD method, and the BCPD method preserved the same mesh quality with only 2% quality loss comparing to the original mesh reported in Table 2.3(b). However, Amberg's method lost 23% of the original mesh quality during the registration. Table 2.3(c) indicates that using the CPD method is very time consuming as the time duration for the experiment is about 8700%, 13400%, and 212% more than our method, Amberg's method, and the BCPD method.

THE FULL BODY

Similar to the lumbar vertebra registration, we repeated the experiments with full human body with more valid landmarks as validation. The initial condition and the landmarks of both the target and source meshes are depicted in Figure 3.9. The registration results are shown in Figure 3.12 and discussed in more detail in Table 3.4 for different numbers of points in the source mesh. The original source mesh contains 19882 vertices, and the registration results are reported employing the remeshed source mesh with 100%, 80%, 60%, 40%, and 20% of the total vertices of the original source mesh. According to Table 3.4(a), our method presents on average 62%, 86% and 66% less landmark errors

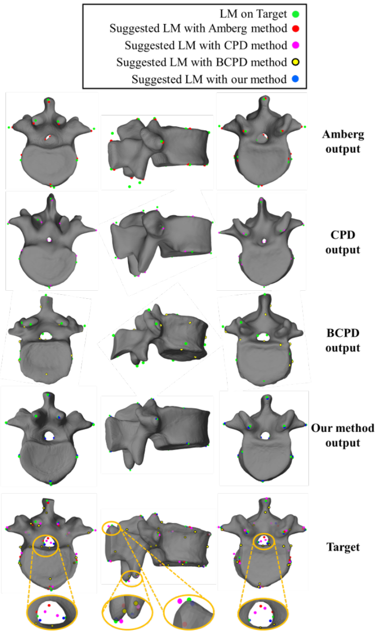


Figure 2.16: Lumbar registration.

Table 2.3: Lumbar vertebra registration results.

a) Overall LM mean error ±Standard Deviation (SD) (mm) result.				
	DBB of $\mathscr{S}$ = 90 mm			
Landmark ID	Amberg [13]	CPD [180]	BCPD [102]	Our method
LM 1	6.9±4.4	8.5±2.1	10.5±3.7	**3.2±0.5**
LM 2	3.8±2.7	8.3±3.9	10.2±4.4	**0.8±0.4**
LM 3	2.5±0.7	7.9±4.1	5.7±3.8	**1.2±0.4**
LM 4	2.3±1.4	10.3±1.1	5.9±2.1	**1.1±0.5**
LM 5	**2.0±1.5**	9.5±2.5	8.8±3.1	2.3±0.9
LM 6	3.2±1.3	8.2±2.3	7.3±1.9	**1.1±0.9**
LM 7	2.5±0.5	10.9±3.3	12.8±2.7	**1.1±0.9**
LM 8	1.9±1.2	8.8±3.6	5.1±1.5	**1.1±1.6**
LM 9	5.4±2.3	9.2±3.9	7.4±2.5	**0.9±0.4**
LM 10	3.9±1.6	11.1±1.5	5.9±2.2	**2.8±2.6**
LM 11	3.3±3.1	11.8±3.5	6.6±1.8	**1.4±0.3**
LM 12	3.9±3.5	8.8±3.4	8.2±4.8	**1.7±0.7**
LM 13	2.2±1.9	7.9±0.9	8.2±1.1	**0.8±0.2**
LM 14	4.1±2.9	7.9±0.7	6.5±1.1	**1.4±1.5**
LM 15	2.4±0.9	8.6±2.6	5.1±1.3	**0.9±1.1**
LM 16	4.8±2.4	10.2±4.9	9.3±3.6	**1.5±0.6**
Overall (mm)	3.4±2.5	9.2±4.3	7.7±2.9	**1.4±1.2**
b) Mesh quality mean value ±Standard Deviation (SD).				
Original mesh	Amberg [13]	CPD [180]	BCPD [102]	Our method
0.8 ±0.14	0.61±0.32	0.77±0.16	**0.78±0.17**	**0.78**±0.16
c) The complete process computing time (s).				
	Amberg [13]	CPD [180]	BCPD [102]	Our method
Time (s)	**14.1**	1895	869	21.6

Figure 2.17: Full body initial condition. (a) Front view. (b) Back view.

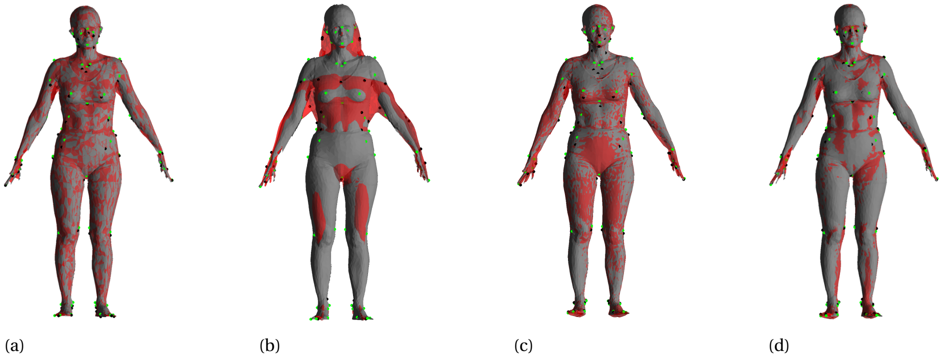
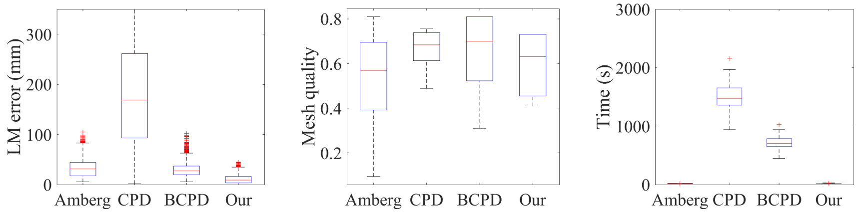


Figure 2.18: Full body registration. (a) Amberg. (b) CPD. (c) BCPD. (d) Our method.

comparing to Amberg's method, CPD method, and BCPD method respectively. Moreover, as in Table 3.4(b), the observed mesh quality loss comparing to the original source mesh is on average about 44% for using Amberg's method, 9% for using the CPD and the BCPD methods, and 15% for using our method. Again, Table 3.4(c) indicates that using the CPD method costs more computing time, accounting for 4820%, 6970% and 92% more of using our method, using Amberg's method and using the BCPD method, respectively. Also, the time for our approach and Amberg's approach is increasing linearly with the number of vertices, while the computing time of using the CPD and BCPD method is increasing exponentially, which often lead to large SD values. Figure 2.19 presents the boxplots of the registration results of using 100% the vertices of the source surface regarding the three criteria. In Figure 2.19(a), the landmark errors of 7373 samples (101 scans × 73 LMs) for each method are shown where our method has the smallest error. Figure 2.19(b) depicts the mesh quality with 101 samples per method. The results of using our method, the CPD, and the BCPD are in the same range. In terms of the computing time for registering the 101 scans (Figure 2.19(c)), registering scans by our and Amberg's methods are in the same range, both are much faster than using the CPD and the BCPD methods.

Table 2.4: Full human body registration results.

a) Overall LM mean error ± Standard Deviation (SD) (mm).					
		DBB of $\mathscr{S}$ = 2167.6 mm			
Total vertices %		Amberg [13]	CPD [180]	BCPD [102]	Our method
100		31.2±18.9	76.8±98.5	27.7±22.3	**10.6±7.6**
80		31.7±20.0	74.8±100.8	38.4±31.5	**12.4±8.6**
60		31.7±25.9	92.2±104	42.7±28.9	**13.1±12.3**
40		42.5±41.1	135.6±110.8	48.1±34.4	**16.7±12.2**
20		51.9±54.7	190.4±125.6	52.6±50.3	**17.5±20.0**
b) Mesh quality mean value ± Standard Deviation (SD).					
Total vertices %	Original mesh	Amberg [13]	CPD [180]	BCPD [102]	Our method
100	0.80±0.15	0.51±0.21	0.67±0.22	**0.69±0.25**	0.68±0.26
80	0.74±0.18	0.49±0.25	**0.74±0.16**	0.71±0.24	0.68±0.25
60	0.72±0.20	0.41±0.29	0.73±0.20	**0.73±0.24**	0.64±0.30
40	0.71±0.21	0.33±0.38	**0.67±0.22**	0.62±0.29	0.61±0.29
20	0.70±0.22	0.31±0.38	0.55±0.25	**0.59±0.31**	0.51±0.42
c) The complete process computing time ± Standard Deviation (SD) (s).					
Total vertices %		Amberg [13]	CPD [180]	BCPD [102]	Our method
100		**18.3±1.1**	1567±825	746±393	21.9±2.2
80		**16.7±1.7**	1221±1001	593±408	21.7±1.4
60		**12.5±0.9**	836±758	443±341	19.6±1.3
40		**10.6±1.1**	555±318	342±330	16.5±1.8
20		**7.9±0.4**	489±362	299±271	15.1±0.7



(a) (b) (c)

Figure 2.19: Boxplots of (a) Landmark error, (b) Mesh quality, and (c) Computing time regarding the full body registration with 100% of the source vertices.

2.6. CONCLUSION

This chapter presents a non-rigid ICP approach based on a newly defined semi-curvature. With similar properties to the Gaussian curvature, the semi-curvature has different mathematical and geometric characteristics. Based on these characteristics, we use the semi-curvature as part of the metric in establishing the correspondences and in the cost function, where the distance and the stiffness terms are embedded as well. Moreover, by increasing the logical dependency of vertices on their neighbors, the semi-curvature term preserves the features of a surface in a guaranteed stable region during the registration process, emphasising on mesh quality and convergence. Experimental results indicate that the proposed non-rigid ICP approach outperforms existing methods in the area where features are not prominent or there are interferences between/among features, as it is able to use the intrinsic properties of the complete surface during the registration. As the approach introduces a general methodology to integrate any number of linear or nonlinear terms (as long as they are linearisable) in the cost function with guaranteed stable region, we will study the extension of the cost function to other new terms in our future work. Moreover, establishing a logical balance between the deformation and the feature preservation via a time-varying adaptive stiffness term is another future direction. In addition, the semi-curvature is defined based on the 1-ring neighboring points, and the effect of including more rings and robustness to the noise on the target mesh will also be studied in the future research.

3

NON-RIGID REGISTRATION VIA INTELLIGENT ADAPTIVE FEEDBACK CONTROL

Preserving features or local shape characteristics of a mesh using conventional non-rigid registration methods is always difficult, as the preservation and deformation are competing with each other. The challenge is to find a balance between these two terms in the process of the registration, especially in presence of artefacts in the mesh. We present a non-rigid Iterative Closest Points (ICP) algorithm which addresses the challenge as a control problem. An adaptive feedback control scheme with global asymptotic stability is derived to control the stiffness ratio for maximum feature preservation and minimum mesh quality loss during the registration process. A cost function is formulated with the distance term and the stiffness term where the initial stiffness ratio value is defined by an Adaptive Neuro-Fuzzy Inference System (ANFIS)-based predictor regarding the source mesh and the target mesh topology, and the distance between the correspondences. During the registration process, the stiffness ratio of each vertex is continuously adjusted by the intrinsic information, represented by shape descriptors, of the surrounding surface as well as the steps in the registration process. Besides, the estimated process-dependent stiffness ratios are used as dynamic weights for establishing the correspondences in each step of the registration. Experiments on simple geometric shapes as well as 3D scanning datasets indicated that the proposed approach outperforms current methodologies, especially for the regions where features are not eminent and/or there exist interferences between/among features, due to its ability to embed the inherent properties of the surface in the process of the mesh registration.

Parts of this chapter have been published in:
Tajdari F, Huysmans T, Song Y. Non-rigid registration via intelligent adaptive feedback control. *IEEE Transactions on Visualization and Computer Graphics.* 2023 Jun 8;1-17 [243].

3.1. INTRODUCTION

IN the past decades, non-rigid registration is widely used in many applications like motion and shape analysis [96, 68], image registration for medical purposes [75, 258], etc. Although there are many forms of non-rigid mesh registration algorithms, the non-rigid Iterative Closest Point (ICP) registration drew more attention, as it is simple in the implementation with high efficiency and effectiveness regarding different types of applications [7]. Currently, non-rigid ICP was employed in a number of applications, e.g. statistical shape analysis [100], computer vision [138], multimedia applications [224], and 3D geometry of human body analysis [114, 250].

The non-rigid ICP registration is presented as a non-trivial and ill-defined problem, which contains a high number of Degrees-Of-Freedom (DOFs) where preserving features of the source surface during the registration process in the formulation of the algorithm [268] with respect to the topology of the target surface are always challenges. For example, creating meaningful robust correspondences in each step of the iteration [265], defining the unique time-varying stiffness level per vertex to maximize feature preservation and minimize mesh quality loss, and establishing the convergence regarding the suited minimum in the process of optimization [143]. Here, the features are defined as salient geometric characteristics [83], e.g. curvature.

Researchers utilized different approaches to tackle those challenges, i.e., utilizing landmarks (LMs) to enhance the corresponding selection [185, 109, 107, 217], embedding adaptive template [58], adding more terms representing shape characteristics in the cost functions [147], and using controlled point-based transformation, e.g. generalizing the affine model from linear to non-linear case [172] to have a more reasonable deformation. Although the aforementioned approaches could accurately determine the pattern of non-linear deformation among meshes to be registered, the challenges are not fully addressed, mainly due to the difficulty in balancing the desired non-rigid deformation and the preservation of salient features across the complete registration process.

To address the balance of several competing terms, researchers developed different methods capable of simultaneously controlling and identifying the unknown parameters of a system online (see, e.g., [229]), which could be integrated into the non-rigid registration methods as a controller scheme controls the feature preservation and mesh quality by identifying the level of "stiffness per vertex". "Stiffness per vertex" is the stiffness degree that a vertex has with its one-ring neighboring points. Therefore, the challenge can be formulated as a control problem in the scenario where the model structure is assumed known and parameters' values are unknown. A possible method to solve the problem is the Model Reference Adaptive Control (MRAC), which is designed to exploit conventional controllers while the controllers' parameters are updated based on model parameters identification. Though such methods have been widely used in controlling robotic systems [111, 241], online identification [85, 244], and noise filtering [281], etc. the use of MRAC in the area of non-rigid registration, e.g., mesh registration, is rare.

This chapter presents a novel globally robust approach for estimating the level of stiffness per vertex in each iteration of the non-rigid ICP via an intelligent MRAC framework to maximally preserve features with minimized mesh quality loss. In fact, here, we are addressing the establishment of a direct compromise between the level of deformation and preserving features per vertex; however, our previous work in [251] elaborates on first introducing a nonlinear shape descriptor to distinguish corresponding points, namely semi-curvature term, and then establishing a framework to solve a nonlinear problem with a linear optimizer, and was ill to address the optimum registration in vertex level, especially for the missed parts of the target surface. The main scientific contributions of this chapter are:

- We introduce an adaptive stiffness ratio estimator utilizing the intrinsic information of the surface to establish the correspondence as well as adjust the stiffness term in each step of the non-rigid ICP registration for better preservation of the shape properties;
- We prove that the estimator is globally asymptotically stable through a Lyapunov function;
- We integrate a data-driven ANFIS-based method to suggest the initial stiffness ratio values of the estimator to further improve the effectiveness and efficiency.

3.2. RELATED WORK

In non-rigid ICP registrations, the relationships between similar features on the source and the target surfaces are often used to establish meaningful dense correspondences. Besides manual labelling, pattern recognition techniques are also employed to automatically specify LMs, e.g. [185, 109, 107, 217]. However, those LMs are mostly positioned at the extrema of specific shape descriptors which are scarcely scattered across the surface and cannot describe the complete spectrum of the inherent properties of the surface. To embed extra features inside the registration algorithms, researchers investigated various inherent properties in the selection of correspondence. An iterative registration approach introduced in [58] combined ICP with Coherent Point Drift (CPD) to have a more robust correspondence selection. Recently, a Bayesian Coherent Point Drift (BCPD) method in [102] integrated the coherent drift in the variational Bayesian inference theory, while maintaining the rudimentary characteristics of the CPD method. In [313], a rigid ICP registration technique was introduced, employing curvature value resemblance to establish correspondences. However, the approach is sensitive to noise, and the computing time is exponentially growing regarding the number of vertices in the surfaces. Recently, a new semi-curvature term is introduced in [251] which considerably improved the corresponding selection accuracy. However, it was sensitive to the border shapes of an open mesh.

Shape characteristics may disappear when the output of non-rigid registration is very smooth, e.g. due to the inclusion of a smoothness functional in the cost function. For preserving the mesh quality of the surface, integrating regularization terms, e.g. ℓ_2-norm term[238], in the structure of the cost function helps. For example, Amberg et al. [12] presented an ℓ_2-norm cost function integrated with a stiffness term to maintain local mesh quality. Yang et al. [306] also used a sparse non-rigid registration scheme through an ℓ_1-norm cost function for feature preservation. However, the constraints of the position (e.g. adjacent piece-wise rigid deformation) were not adequately integrated with the model to incorporate the association with the piece-wise rigid deformation. Recently, Ayan [43] proposed a registration methodology utilizing an energy function incorporating the strength of local and global geometry through an middle level depiction of the point cloud. However, the approach is sensitive to the point density areas with detailed geometry features.

During an iterative non-rigid registration process, once the corresponding points are defined in each iteration, the path from the current condition of the points on the source mesh to the corresponded points on the target mesh is very challenging as the level of deformation to the corresponding point and level of the mesh quality are in contrast with each other. Thus, an optimal path should be capable to be adaptively updated in each iteration to allow for sufficient deformation per vertex while safeguarding the mesh quality of faces. As the solution is feedback-based, adaptive feedback control approaches might be a problem solver. Adaptive solutions have vast applications such as 3D meshing [110, 54], and mesh refinement [17]; however, there is no prior work about adaptive feedback solutions or adaptive control for 3D mesh registration. Instead, there are limited research works regarding adaptive image registration. In the area, non-rigid registration employing Radial Basis Functions (RBFs) [168, 299, 226] are interesting as they are meshless. In theory, knots can be located optimally and adaptively for a deformation area to reach the acceptable accuracy. Zhou et al. [62] proposed a RBF-based approach to address non-correspondent point-clouds registration. Recently, Zhang et al. [317] presented an implementation of RBF non-rigid registration with iterative knot-placement to adaptively decrease registration local error. However, quantitative evaluation of accuracy was lacking in RBF-based methods.

3.3. METHODOLOGY

3.3.1. PROBLEM FORMULATION

In the context of non-rigid registration, most of the available methods try to non-rigidly align a point cloud to the other point cloud in which this alignment accompanies deformation. However, in those methods, the meaning of the ideal deformation is vague and usually defined as reaching the least geometry error with the corresponding points; however, the deformation can destroy some features in the source mesh and/or reduce the quality of the source mesh. The biggest challenge here is that the methods use similar parameter values for all the points on the source mesh during registration, while each point in fact needs an advanced investigation based on its condition, to set those values. Thus, online parameter tuning for each vertex based on the course of registration for the vertex is a key to the challenge. To this end, a method that can optimize the degree of deformation in order to minimize the geometry error and maximize the degree of feature preservation is a breakthrough. In this paper, we introduce a novel method to adjust those parameters per-vertex online and implement it on the most simple registration method of ICP to study the impact of our method on the quality of registration. Instead of using the same values for the stiffness term value, we propose an adaptive estimator for generating proper vertex-wise stiffness values for the source surface. The procedure is depicted in Figure 3.1, and the used variables and parameters are reported in TABLE 3.1. According to the figure, in each iteration of the optimization process in the non-rigid ICP algorithm, the feedback-based adaptive estimator defines the proper stiff-

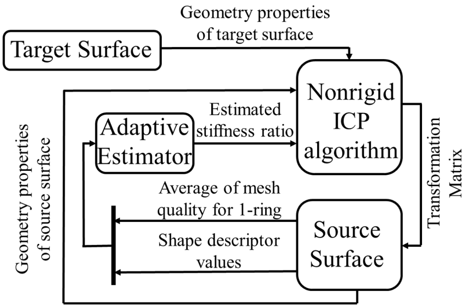


Figure 3.1: Non-rigid ICP and the adaptive estimator integration diagram. The feedback-based estimator adjusts the proper stiffness ratio per vertex, to minimize the mesh quality loss, and preserve the maximum possible feature of the source mesh.

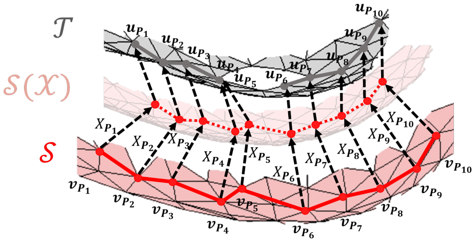


Figure 3.2: Non-rigidly registering the source surface to the target surface.

ness ratio per vertex using the geometry properties of the source surface, namely shape descriptor value and average of mesh quality, to maintain the highest possible deformation with the least source surface feature loss.

3.3.2. NON-RIGID ICP REGISTRATION

During the non-rigid ICP registration, $\mathscr{S} = (\mathscr{V}, \mathscr{E})$ as the source surface, includes n vertices in $\mathscr{V}$ and m edges in $\mathscr{E}$, is iteratively registered to the target surface $\mathscr{T}$. Figure 3.2 demonstrates one step of the registration process. Regarding the figure, the triangular meshes are used, and the vertices are indexed by numbers. In the step, first, the correspondences from vertices v_{P_i} in the source surface $\mathscr{S}$ (red) to vertices u_{P_i} in the target surface $\mathscr{T}$ (gray) are selected. Then v_{P_i} is transformed via locally affine transformation (X_{P_i}) towards the target surface $\mathscr{T}$. Here, the transformation matrix X_{P_i} for each vertex in the source mesh is a 3×4 transformation matrix which includes all possible translation and rotation transformations. Thus, the transformation matrix X of all vertices is defined in a $4n \times 3$ matrix as $X = \left[X_{P_1} \cdots X_{P_n}\right]^T$. The transformed source surface is $\mathscr{S}(X)$. This procedure iterates until an optimum state is achieved.

Here, according to a defined correspondence set, namely (v_{P_i}, u_{P_i}), a cost function based on Amberg [13] is determined and then the adaptive estimator is embedded accordingly. Amberg [13] proposed the non-rigid registration formulation as a combination of distance and stiffness terms summarised in the following formula

$$\bar{E}(X) = \| \begin{bmatrix} M \otimes G \\ WD \end{bmatrix} X - \begin{bmatrix} 0 \\ W\tilde{U} \end{bmatrix} \|_F^2 \tag{3.1}$$

where the sparse matrix D is formed to facilitate the transformation of the source points with the individual transformations in X through matrix multiplication and defined as $D = diag(v_{P_1}^T, v_{P_2}^T, \ldots, v_{P_n}^T)$. W explains a diagonal matrix composed of weights w_i. In addition, $\tilde{U} \in U$ includes the corresponding points from $\mathscr{S}$ to $\mathscr{T}$ which is a $n \times 3$ matrix, where $U = \left[u_{P_1}, u_{P_2}, \ldots, u_{P_n}\right]^T$. To regularise the deformation, an extra stiffness term is utilized. Employing the Frobenius norm $\|.\|_F$, the stiffness term minimizes changes in the transformations of adjacent vertices, via a weighting matrix $G = diag(1,1,1,\gamma)$. In the deformation process, γ contains a value to stress changes in the skew and rotational part against the translation part of the deformation. The value of γ can be determined according to data units and the types of deformation [12]. The node-arc incidence matrix M (e.g. Dekker [61]) of the source mesh topology is used to turn the stiffness term functional into a matrix form. As the matrix is unchangeable for directed graphs, the construction is one row for each edge of the mesh and one column per vertex. To define the node-arc incidence matrix of the source topology, the indices (i.e. the subscripts) of edges and vertices are addressed, for any edge of r which is linked to vertices (i, j), in r^{th} row of M, and the nonzero entries are $M_{ri} = -1$ and $M_{rj} = 1$.

3.3.3. ADAPTIVE ESTIMATOR DESIGN

A key objective in this research is to estimate and reconstruct a weighting matrix including ratios known as stiffness ratio ($\hat{g}_{P_i}$) per vertex P_i on $\mathscr{S}$, comprising G in Eq. (6.7). For this, we will define an Ordinary Differential Equation (ODE) based on the shape descriptor and the average of mesh quality where $\hat{g}_{P_i}$ will be its solution. To integrate the estimated $\hat{g}$ in the stiffness term in Eq. (6.7), we define the ratio in the form of a matrix as:

$$\hat{G}_{4n\times 4n} = \mathrm{diag}(\hat{g}_{P_i} I_{4\times 4}, \cdots, \hat{g}_{P_n} I_{4\times 4}) \tag{3.2}$$

which is embedded in the stiffness term

$$J_s = \|\big(M \otimes G\big)\,\hat{G}X\|_F^2. \tag{3.3}$$

Thus Eq. (6.7) is changed to:

$$J(X) = \| \begin{bmatrix} (M\otimes G)\hat{G} \\ WD \end{bmatrix} X - \begin{bmatrix} 0 \\ W\tilde{U} \end{bmatrix} \|_F^2 = \|AX - B\|_F^2, \tag{3.4}$$

The non-rigid registration can be achieved by optimizing the cost function in Eq. (3.4).

In this section, we first introduce the preliminary including the derivation of the ODE, the assumptions, the error system definition, and its stability and convergence in Section 3.3.3.1. In Section 3.3.3.2, we present the process of deriving the rules for estimating $\hat{g}_{P_i}$ using control theory that guarantees the asymptotic stability of the error system. In Section 3.3.3.3, we discuss the properties of estimation rules, and finally, we integrate the estimation rules into the non-rigid ICP scheme as Section 3.3.3.4.

PRELIMINARY

For formulating the estimator states per vertex, we use the integral error states [246, 242] for Eq. (3.5), as they are robust to uncertainty and are able to accommodate for zero error convergence [22]

$$e_s = \int_0^t k_s^t - k_s^0 \, d\tau, \;\; e_q = \int_0^t q^t - q^0 \, d\tau, \tag{3.5}$$

where k_s^t and q^t are the shape descriptor value and the average of mesh quality value of the faces incident an arbitrary vertex P_i on the source at time t, respectively. Accordingly, k_s^0 and q^0 are the values before registration, i.e. at time 0. Among different algorithms, Joe-Liu's parameter method [152] is used to juxtapose the meshes quality in this paper. The method determines the quality value per vertex may vary in a range among 0 and 1, where 1 defines the highest quality and zero explains the lowest quality. Following (3.5), the integral error system is considered as

$$\dot{e}_s = k_s^t - k_s^0, \;\; \dot{e}_q = q^t - q^0, \tag{3.6}$$

which is denoted in the matrix form as

$$\dot{x} = B_e u_e + r_e, \tag{3.7}$$

where

$$x = \begin{bmatrix} \int k_s^t \, dt \\ \int q^t \, dt \end{bmatrix}, u_e = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, B_e = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, r_e = \begin{bmatrix} k_s^0 \\ q^0 \end{bmatrix}. \tag{3.8}$$

u_1 and u_2 are shown to declare the dimension of the control input (u_e) which is a 2×1 matrix. We aim at controlling the state-space Eq. (7.12) through Model Reference Adaptive Control (MRAC) [229], which leads to penalisation of the tracking error. Accordingly, we assume a Proportional-action (P-action) feedback control law (see e.g., chapter 1 in[277]) that tries to penalise the error of tracking between the variable x, and the desired variables r_e with unknown gains as

$$\begin{aligned} u_e &= -\hat{g}_{P_i}(x - r_e) \\ &= -\hat{g}_{P_i} x + \hat{g}_{P_i} r_e, \end{aligned} \tag{3.9}$$

where $\hat{g}_{P_i}$ is an unknown scalar of the stiffness ratio for the point P_i that needs to be estimated. By assuming $\tilde{a} = \begin{bmatrix} -\hat{g}_{P_i} & \hat{g}_{P_i} \end{bmatrix}$, then,

$$u_e = \tilde{a} \begin{bmatrix} x \\ r_e \end{bmatrix}. \tag{3.10}$$

A model reference state is considered as

$$\dot{x}_M = -A_M x_M + B_M r_e, \tag{3.11}$$

where the optional predefined matrices A_M and B_M are used to establish a stable model reference in each iteration. Accordingly, the error is considered as the differences between the integral states (x), and the model reference states (x_M), where

$$\begin{aligned}
\dot{e} &= \dot{x} - \dot{x}_M \\
&= B_e(-\hat{g}_{P_i} x + \hat{g}_{P_i} r_e) + r_e + A_M x_M - B_M r_e + A_M x - A_M x \\
&= -A_M(x - x_M) + B_e\left(-\hat{g}_{P_i} + \frac{A_M}{B_e}\right)x + B_e\left(\hat{g}_{P_i} - \frac{B_M+I}{B_e}\right)r_e \\
&= -A_M e + B_e\left(-\hat{g}_{P_i} + \frac{A_M}{B_e}\right)x + B_e\left(\hat{g}_{P_i} - \frac{B_M+I}{B_e}\right)r_e.
\end{aligned} \tag{3.12}$$

This results in the following error dynamic system equation explained in Laplacian domain.

$$e = \frac{B_e}{sI + A_M}\begin{bmatrix} -\hat{g}_{P_i} + \frac{A_M}{B_e} & \hat{g}_{P_i} - \frac{B_M+I}{B_e} \end{bmatrix}\begin{bmatrix} x \\ r_e \end{bmatrix}, \tag{3.13}$$

where s is the Laplace variable and I is the identity matrix. As A_M can be selected to guarantee that the reference model becomes stable inherently i.e., A_M has negative eigenvalues therefore $(sI + A_M)$ is stable, the error dynamic of Eq. (7.18) is, then, stable in domain of time, if $\tilde{a}$ is bounded (or $\tilde{a}$ is converging).

Derivation of the estimation rule

Accordingly, to demonstrate that the $\tilde{a}$ is converging in the integral error dynamic in Eq. (7.18), a Lyapunov function for each of the vertex P_i is employed as follows:

$$\mathcal{V} = x\mathscr{P}x^T + \tilde{a}\Gamma_{P_i}^{-1}\tilde{a}^T, \tag{3.14}$$

where $\mathscr{P} \geq 0$ and $\Gamma_{P_i} > 0$ imply that $\mathcal{V} > 0$. Then, to prove the stability it would be sufficient to prove that $\dot{\mathcal{V}} \leq 0$ [113]. As

$$\frac{d\mathcal{V}}{dt} = \dot{x}^T\mathscr{P}x + x^T\mathscr{P}\dot{x} + \dot{\tilde{a}}^T\Gamma_{P_i}^{-1}\tilde{a} + \tilde{a}^T\Gamma_{P_i}^{-1}\dot{\tilde{a}}. \tag{3.15}$$

Based on Eq. (3.10) and considering $\theta_{P_i} = \begin{bmatrix} x \\ r_e \end{bmatrix}$, then $u_e^T = \theta_{P_i}^T\,\tilde{a}^T$, thus

$$\frac{d\mathcal{V}}{dt} = 2\mathscr{P}B_e x\theta_{P_i}^T\tilde{a}^T + 2\dot{\tilde{a}}\Gamma_{P_i}^{-1}\tilde{a}^T. \tag{3.16}$$

Following the assumption $\mathscr{P}B_e = c$, $e_{P_i} = cx$, and $\frac{d\mathcal{V}}{dt} = 0$, then from Eq. (7.22) we obtain

$$-2\,e_{P_i}\theta_{P_i}^T\tilde{a}^T = 2\dot{\tilde{a}}\Gamma_{P_i}^{-1}\tilde{a}^T; \tag{3.17}$$

Accordingly, to meet the stability criterion, the variation of $\tilde{a}$ by the time must follow the estimation rule as

$$\dot{\tilde{a}} = -\,e_{P_i}\;\theta_{P_i}^T\;\Gamma_{P_i}, \tag{3.18}$$

where Γ_{P_i} is the growth rate of the estimation rule. To define the estimation rule for $\hat{g}_{P_i}$, by multiplying $\bar{I}^T$ to both sides of Eq. (7.24), and knowing that $\bar{I}\bar{I}^T = 2I$ and $\tilde{a} = \hat{g}_{P_i}\bar{I}$, we will have

$$\dot{\hat{g}}_{P_i} I = -0.5 \; e_{P_{i_{2\times 1}}} \; \theta^T_{P_{i_{4\times 1}}} \; \Gamma_{P_{i\,4\times 4}} \; \bar{I}^T_{2\times 4} \tag{3.19}$$

To extend Eq. (3.19) to all vertices of the source mesh, we consider the matrix form as

$$\dot{\hat{\Pi}}_{2n\times 2n} = -0.5 \; E_{2n\times n} \; \Theta^T_{4n\times n} \; \Phi_{4n\times 4n} \; \Psi_{4n\times 2n}, \tag{3.20}$$

where Φ is a diagonal matrix composed of each Γ_{P_i} as $\Phi = \mathrm{diag}(\Gamma_{P_1}, \Gamma_{P_2}, \ldots, \Gamma_{P_n})$. Ψ is a diagonal matrix where each component on the diagonal of the matrix is $\bar{I}^T$. In addition, E is a diagonal matrix composed of each e_{P_i} as $E = \mathrm{diag}(e_{P_1}, e_{P_2}, \ldots, e_{P_n})$ and Θ is a diagonal matrix composed of each θ_{P_i} as $\Theta = \mathrm{diag}(\theta_{P_1}, \theta_{P_2}, \ldots, \theta_{P_n})$. As a result, $\dot{\hat{\Pi}}$ is a diagonal matrix composed of each $\dot{\hat{g}}_{P_i}$ as $\dot{\hat{\Pi}} = \mathrm{diag}(\dot{\hat{g}}_{P_1} I, \dot{\hat{g}}_{P_2} I, \ldots, \dot{\hat{g}}_{P_n} I)$.

THE GROWTH RATE (Γ_{P_i}) DESIGN

To simplify the Eq. (3.19), we design $\Gamma_{P_i} = \bar{\Gamma}_{P_i} I_{4\times 4}$, where $\bar{\Gamma}_{P_i}$ is a scalar. Empirical observations indicate that under condition of $|\dot{\hat{g}}_{P_i}| > 1$, high fluctuations in the estimation of $\hat{g}_{P_i}$ values are expected, which may lead to a decrease in the convergence level of the estimator. To prevent this, $\bar{\Gamma}_{P_i}$ is formulated in the form below to guarantee $\bar{\Gamma}_{P_i} \leq 1$ for a smoother converging towards the true values.

$$\bar{\Gamma}_{P_i} = \begin{cases} \frac{1}{|\dot{\hat{g}}_{P_i}|^{\zeta}}, & \text{if } |\dot{\hat{g}}_{P_i}| > 1, \\ |\dot{\hat{g}}_{P_i}|^{\nu}, & \text{if } |\dot{\hat{g}}_{P_i}| \leq 1. \end{cases} \tag{3.21}$$

Discussion I: In Eq. (3.21), ζ and ν are very effectual on the changes of $\bar{\Gamma}_{P_i}$ by the time. And large values of $\bar{\Gamma}_{P_i}$ can lead the system to an unstable condition as the system and control scheme are very dependent on the estimated stiffness ratio. Thus, $\bar{\Gamma}_{P_i}$ should be chosen so that the estimator is fast enough with the least overshoot. For the possible domain of ζ and ν, we list four conditions as follows.

1. $\zeta < 0$ and $\nu < 0$: This condition is unstable as $\bar{\Gamma}_{P_i}$ would be greater than 1 especially when $\dot{\hat{g}}_{P_i} \to 0$ results in $\bar{\Gamma}_{P_i} \to \infty$.

2. $\zeta < 0$ and $\nu > 0$: In this condition, for some cases where $|\dot{\hat{g}}_{P_i}| > 1$, $\bar{\Gamma}_{P_i}$ would be comparatively a large value, which increases the chance of having overshoots.

3. $\zeta \geq 0$ and $\nu \geq 0$: Only here, some stable points may be found via numerical trial and error methods as in this condition $\bar{\Gamma}_{P_i}$ is bounded ($\bar{\Gamma}_{P_i} \leq 1$).

4. $\zeta > 0$ and $\nu < 0$: Here and similar to condition 1, for $|\dot{\hat{g}}_{P_i}| < 1$, $\bar{\Gamma}_{P_i}$ has a large value which increases the probability of instability more than the condition 2.

In this research, condition 3 is selected for the possibility to guarantee the stability of the system. The ranges of ζ and ν will be studied in Section 3.4.4.

Table 3.1: Used variables and parameters.

Symbol	Description
The ICP method	
$\mathscr{S}$	Source surface
$\mathscr{T}$	Target surface
n	Number of vertices on $\mathscr{S}$
n_T	Number of vertices on $\mathscr{T}$
P_i	An arbitrary point on $\mathscr{S}$
v_{P_i}	Geometry of P_i on $\mathscr{S}$
u_{P_i}	Geometry of Corresponding point to v_{P_i} on $\mathscr{T}$
X_{P_i}	Transformation matrix for P_i
D	A matrix including geometry of all points on $\mathscr{S}$
U	A matrix including all the points on $\mathscr{T}$
$\tilde{U}$	A matrix including corresponding points from $\mathscr{S}$ to $\mathscr{T}$
G	Stiffness weighting matrix
The adaptive Estimator	
e_s	Integral error of shape descriptor for P_i
e_q	Integral error of average of mesh quality for P_i
k_s^t	Shape descriptor value at time t
k_s^0	Shape descriptor value before registration
q^t	Average of mesh quality value at time t
q^0	Average of mesh quality value before registration
t	Index of time (iterations)
$\tilde{a}$	Vector of unknown parameters
$\hat{g}_{P_i}$	Estimating unknown stiffness ratio at point P_i
θ_{P_i}	Estimator variables for point P_i
e_{P_i}	Error of estimator variables for point P_i
Γ_{P_i}	Growth rate of the estimator for P_i
$\hat{\Pi}$	Matrix of all estimating unknown stiffness ratio
Ψ	Constant matrix
Φ	Matrix of all growth rate values
E	Matrix of all error of estimator variables
Θ	Matrix of all estimator variables
$\hat{G}$	Matrix form of the estimated stiffness ratio
Correspondence selection and implementation via the estimator	
l	Index of updating the estimated stiffness ratio
$r_{\hat{g}}(P_i)$	A time-varying weight
h_{D_i}	A N by 1 matrix, including closest distance of points from $\mathscr{T}$ to P_i
h_{K_i}	A N by 1 matrix, including closest shape descriptor value of points from $\mathscr{T}$ to P_i
H_D	A matrix including h_{D_i}
H_K	A matrix including h_{K_i}
The simulation configuration	
$\hat{g}^{\max}$	Maximum stiffness
$\hat{g}^{\min}(k)$	Stiffness trade-off
N_{mean}	Target points averaging size
N_r	Initial ratio of n points on $\mathscr{S}$
$N(k)$	Target points pool size
ϵ	Convergence error threshold
$j^{\max}$	Convergence iteration threshold
$k^{\max}$	Number of iteration for the outer loop

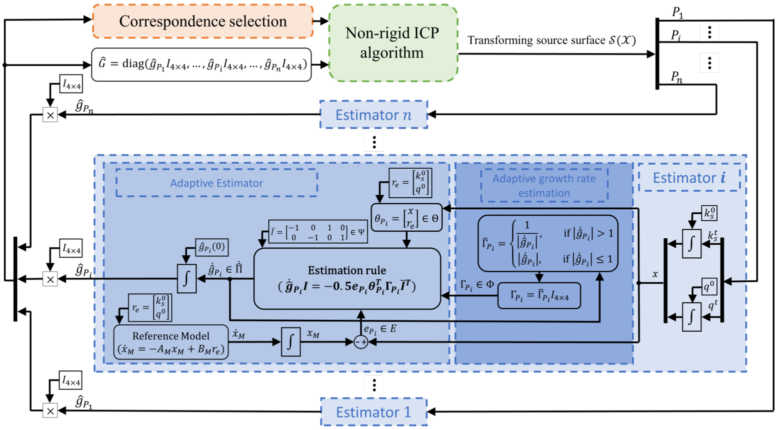


Figure 3.3: Closed-loop diagram of the proposed estimation and registration method integration framework.

THE ESTIMATOR FRAMEWORK

The estimator framework presented in this paper, shown in Figure 3.3, includes an estimator for each point P_i on $\mathscr{S}$ to predict the unknown stiffness ratio $\hat{g}_{P_i}$ during the registration process. It is expected that such a design can balance the two competing factors, shape deformation and feature preservation (k_s^t and q^t), in the non-rigid registration process.

The estimator framework is composed of two main parts: a) adaptive growth rate estimation (dark blue part), and b) the adaptive estimator (medium blue part). The dark blue part essentially defines a conservative time-varying growth rate ($\bar{\Gamma}_{P_i}$) value in Eq. (3.21) used in forming matrix Γ_{P_i} employed in the estimation rule Eq. (3.19). The medium blue part represents the estimation process of $\hat{g}_{P_i}$ which is the output of the estimator. The set of all $\hat{g}_{P_i}$ is then used in the correspondence selection (orange part) directly, and in the non-rigid ICP cost function (green part) via $\hat{G}$ in Eq. (3.4).

3.3.4. CORRESPONDENCE SELECTION AND IMPLEMENTATION VIA THE ESTIMATOR

The flowchart in Figure 3.4 presents more details of the presented process for obtaining the optimal X_{P_i} from v_{P_i} to u_{P_i} (where u_{P_i} is comparable with $X_{P_i} v_{P_i}$) in each iteration. In Figure 3.4, two nested loops are embedded: regarding the outer loop (where k is updated, counts the outer loop iterations), the weighting parameters related to the correspondence selection scheme are automatically updated; for the inner loop (where j is updated), the cost function for obtaining the optimal transformation matrix is incessantly being minimized according to the weights until the variation of the last two consequential Xs is less than a small value of ϵ or $j > j^{\max}$. As the $\hat{g}$ is never reset after each update in k, it is updated with index $l = (k-1)j^{\max} + j$. The nested loops in each of their iterations aim at referring various aspects of the measures, e.g. the Euclidean distance and the difference of curvature values of two corresponding points are affected by the estimated stiffness ratio $\hat{g}$.

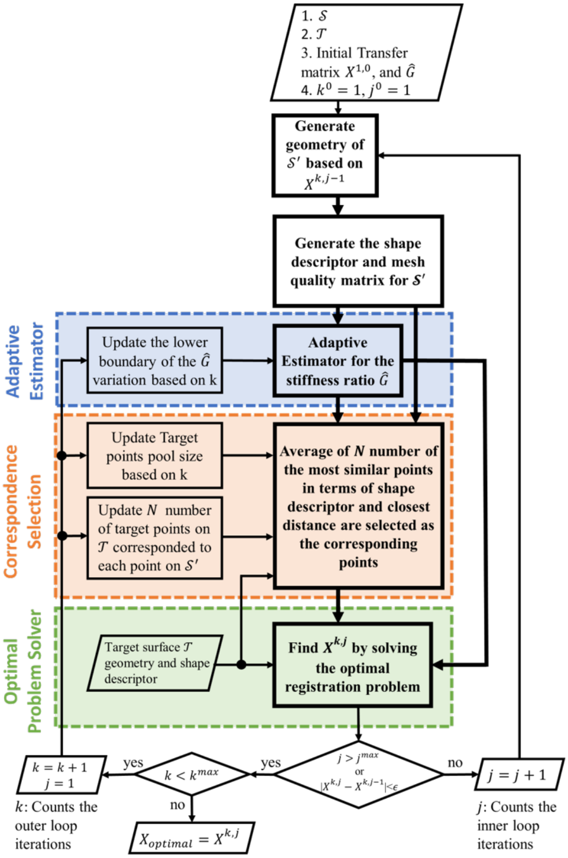


Figure 3.4: The proposed registration process consists of two nested loops. In the outer loop (where k is updated), the weighting parameters related to the correspondence selection and the adaptive estimator schemes are automatically updated; for the inner loop (where j is updated), the stiffness ratios are estimated and then the cost function for obtaining the optimal transformation matrix is incessantly being minimized to finally find the best solution for $X_{optimal}$.

ESTABLISH CORRESPONDENCES

In the context of establishing correspondences in an ordinary ICP method, the corresponding point for a vertex on $\mathscr{S}$ is the closest vertex on $\mathscr{T}$ if only the Euclidean distance is considered as the only criterion, and the inherent properties of the mesh are not considered. To embed the inherent properties, a novel formulation as Eq. (3.22) is proposed balances distance and shape similarity using the estimated $\hat{g}_{P_i}$. The formulation integrates both the Euclidean distance and the descriptor to determine the corresponding vertex on $\mathscr{T}$ as:

$$r_{\hat{g}}(P_i) = \frac{\hat{g}_{P_i} - \min(\hat{g})}{\max(\hat{g} - \min(\hat{g}))} \tag{3.22}$$

$$H = (1 - r_{\hat{g}})H_D + r_{\hat{g}} H_K. \tag{3.23}$$

Here, H_D belongs to the distance measures including h_{D_i} allied to P_i, and H_K belongs to the differential form of the shape descriptor term containing h_{K_i} allied to P_i. In addition, in Eq. (3.22), $r_{\hat{g}}(P_i)$ is a time-varying weight allied to P_i in the range of $[0,1]$ inspired by the corresponding estimated value of P_i as $\hat{g}_{P_i} \in \hat{g}$, e.g. it has a large value when the balance between the mesh quality and preserved shape descriptor value is low. Larger values of $\hat{g}$ define a higher stiffness requiring more convenient preservation of the features at the point location. The high values result in more emphasis on the shape descriptor term and vice versa. Practically, sparse similarities prohibit us to detect all points in $\mathscr{T}$ for establishing correspondences. Thus, only candidate correspondence points are used, considered in a pre-processed region containing N number of points in $\mathscr{T}$. Next, H_D and H_K in Eq. (3.23) can be elaborated as:

$$\left[h_{K_i}\right]_{N\times 1} = \frac{\frac{k_s(P_i)}{\max(K_s)} - \frac{[K_T]_{N\times 1}}{\max(K_T)}}{\max\left(\frac{k_s(P_i)}{\max(K_s)} - \frac{[K_T]_{N\times 1}}{\max(K_T)}\right)} \tag{3.24}$$

$$\left[h_{D_i}\right]_{N\times 1} = \frac{\operatorname{norm}\left(v^T_{P_i} - \left[\bar{U}\right]_{N\times 3}\right)}{\max\left(\operatorname{norm}\left(v^T_{P_i} - \left[\bar{U}\right]_{N\times 3}\right)\right)} \tag{3.25}$$

where K_T is composed of the shape descriptor values of the vertices on the $\mathscr{T}$, h_{D_i} is composed of a N by 1 matrix, including distances of the N vertices from $\mathscr{T}$ to P_i on $\mathscr{S}$ known as $\bar{U} \in U$. Likewise, h_{K_i} is composed of a N by 1 matrix, including divergence of the N vertices on $\mathscr{T}$ in terms of the shape descriptor value from P_i on $\mathscr{S}$. The preprocessing strategy to detect the area with N_{mean} vertices follows two procedures:

- N (decreasing value), initiates from $N_r n_T$ and ends with N_{mean} number of closest points on $\mathscr{T}$ to each point on $\mathscr{S}$ are selected, in which N_r represents a ratio with domain of [0 1] and n_T defines the number of points on $\mathscr{T}$.
- N_{mean} number of the candidate vertices with a decreasing value of H are selected from the previous step where the corresponding point is the average of these N_{mean} vertices' geometry.

After the correspondence points are established from the $\mathscr{T}$, we investigate the number of corresponding points from the $\mathscr{S}$ to each of the points of the $\mathscr{T}$. For the vertices corresponding to more than N_{mean} number of source points, we only pick the N_{mean} of them with minimum H.

DISCRETE TIME IMPLEMENTATION

In the implementation, we discretize the parameter estimation process by considering $\dot{\hat{\Pi}}(l) = \frac{\hat{\Pi}(l)-\hat{\Pi}(l-1)}{\Delta l}$; then, knowing that $\Delta l = 1$ as l is a sequentially increasing index, the estimation rule Eq. (3.20) turns into

$$\hat{\Pi}(l) = \hat{\Pi}(l-1) - 0.5\, E(l-1)\, \Theta(l-1)\, \Phi(l-1)\, \Psi. \tag{3.26}$$

Similarly, by considering $\dot{\hat{g}}_{P_i}(l) = \frac{\hat{g}_{P_i}(l)-\hat{g}_{P_i}(l-1)}{\Delta l}$, the discrete form of Eq. (3.21) will be

$$\bar{\Gamma}_{P_i}(l) = \begin{cases} \frac{1}{|\hat{g}_{P_i}(l)-\hat{g}_{P_i}(l-1)|^{\zeta}}, & \text{if } |\hat{g}_{P_i}(l)-\hat{g}_{P_i}(l-1)| > 1, \\ |\hat{g}_{P_i}(l)-\hat{g}_{P_i}(l-1)|^{\nu}, & \text{if } |\hat{g}_{P_i}(l)-\hat{g}_{P_i}(l-1)| \leq 1. \end{cases} \tag{3.27}$$

3.4. EXPERIMENT SETUP

3.4.1. DATASET

MAIN DATASET: FULL HUMAN BODY

In the experiment, realistic 3D scans of the full human body, including holes and spikes, are chosen as geometric shapes for assessing the presented method, as the holes are always challenges for mesh quality preservation. The full human body scans are chosen among the Civilian American and European Surface Anthropometry Resource (CAESAR) dataset [210]. Here, we selected the first 101 scans from the Dutch population as the target mesh and assessed via the pre-selected LMs in each mesh. 74 landmarks are defined in [210] (from page 17 to page 30), while we employed 73 of them (LM74 on the butt block is neglected). Regarding the source mesh, we recruited the full body mesh included in the Wrap 3 software [215]. To unify the inputs, we remeshed the source mesh in [215] to the same number of vertices as the original source mesh.

ASSESSMENT DATASET: THORACIC VERTEBRA

To evaluate the performance of the proposed approach on more complicated freeform surfaces, the twelve thoracic vertebrae in a complete spine in [257] which is extracted from [228] are employed. All the surface meshes are re-meshed to 8000 vertices [271], resulting in uniform triangle surface meshes, to have a ground truth for comparing the mesh quality variation between different methods. In the dataset, thoracic vertebrae are mainly labelled by the prefix T and numbered 1 to 12 (e,g., $T7$ belongs to thoracic number 7). We selected $T7$ as the source surface and the rest of the eleven vertebrae as the target surfaces. 25 Landmark Points (LMPs) for a thoracic vertebra are defined in [195], while we employed 12 of the LMPs with index 2, 3, 4, 5, 10, 19, 20, 21, 22, 23, 24, and 25.

3.4.2. PARAMETERS OF THE PROPOSED METHOD

TABLE 6.1 reports values of the parameters employed in the experiment which are explained in Table 3.1. During the minimization of the cost function, γ in G presented in Section 2.3.4, was selected as 1. To avoid large changes in the value, we defined a time-varying bound for the parameter with a gradual relaxation saturation operator as:

$$\text{sat}(\hat{g}_{P_i}(l)) = \begin{cases} \hat{g}^{\min}(k), & \text{if } \hat{g}_{P_i}(l) < \hat{g}^{\min}(k) \\ \hat{g}^{\max}, & \text{if } \hat{g}_{P_i}(l) > \hat{g}^{\max} \\ \hat{g}_{P_i}(l), & \text{otherwise,} \end{cases} \tag{3.28}$$

where $\hat{g}^{\min}(k)$ is the decreasing lower and $\hat{g}^{\max}$ is the constant upper bound for the input $\hat{g}_{P_i}(l)$. In the experiments $\hat{g}^{\max}$ is 1000 and $\hat{g}^{\min}$ is gradually decreased from 1000 to 1.

Table 3.2: Parameters used in simulation configuration.

Parameter	$\hat{g}^{\max}$	$\hat{g}^{\min}(k)$	N_{mean}	N_r
Value	1000	$\hat{g}^{\max}$:1	3	0.1
Parameter	$N(k)$	ϵ	$j^{\max}$	$k^{\max}$
Value	$N_r n_T$:N_{mean}	0.001	50	20

This indicates that in this case, $\hat{g}^{\min}$ and $\hat{g}^{\max}$ were specified independent of the source mesh defined and only global deformations were allowed at the beginning of registration. In addition, the lower limit of the boundaries relies on the type of data [13]. Thus, a small value of $\hat{g}^{\max}$ may result in a singularity of A in Eq. (3.4), which falls into instability condition during solving the optimal problem. Consequently, our experiments began with a considerably high $\hat{g}^{\max}$. To maintain a smoother registration process, we used the mean of a number of points from a set with size of 3 ($N_{mean} = 3$). This N_{mean} number of points regarding H in Eq. (3.22). N number of points on $\mathscr{T}$, as an area to explore for the corresponding point to any vertices on $\mathscr{S}$, varied from 10 % of the total number of vertices in $\mathscr{T}$ to N_{mean}. In the implementation of [42], the system of linear equations that arises in each step was solved with the help of the UMFPACK library [60].

3.4.3. REFERENCE MODEL DYNAMICS

Following the discussion presented in Section 3.3, a stable reference model (A_{ref}, B_{ref}) must be established, in which two of the states are considered as the integral of the other two states. Accordingly, a four-state dynamic system is utilized as the reference model which its states (x_{ref}) are proposed as below

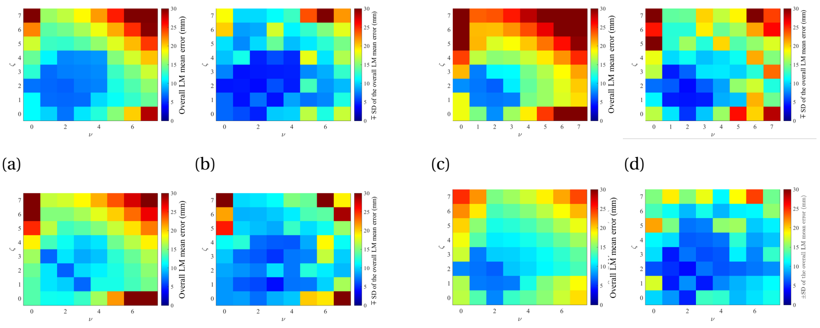
$$\dot{x}_{ref} = A_{ref} x_{ref} + B_{ref} r_e \tag{3.29}$$

$$A_{ref} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -10 & -10 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -10 & -10 \end{bmatrix}, B_{ref} = \begin{bmatrix} 0 & 0 \\ 10 & 0 \\ 0 & 0 \\ 0 & 10 \end{bmatrix}, \tag{3.30}$$

where $x_M = \left[x_{ref_{1,1}} x_{ref_{1,3}} \right]^T$ in Eq. (3.11). The system is globally stable to r_e as all the eigenvalues of A_{ref} are negative, and pair (A_{ref}, B_{ref}) is stabilisable (see Chapter 2 of [140]), which can be verified through, e.g, the Hautus-test [288, 245]. Please note that, in this chapter, for all the experiments that use the reference model we use the parameters in Eq. (3.30).

3.4.4. SENSITIVITY ANALYSIS REGARDING DIFFERENT SHAPE DESCRIPTORS

A set of experiments in the stable domain introduced in Discussion I are initially performed employing the scans from the Dutch population of the CAESAR dataset [210] to study the sensitivity of the parameters ν, and ζ per different shape descriptor aimed at selecting the proper range of the parameters and the descriptor. To this end, we select



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

(e) (f) (g) (h)

Figure 3.5: Sensitivity analysis showing for a domain of ν and ζ while $N_r = 0.1$, (left) Overall LM mean error; (right) SD of Overall LM mean error based on: (a, b) MC formulation; (c, d) GC formulation; (e, f) SI formulation; (g, h) CI formulation.

four shape descriptors (K_G) as Mean Curvature (MC), Gaussian Curvature (GC), Shape Index (SI), and Curvature Index (CI) from [262] with the following formulations:

$$\begin{cases} \mathrm{MC:} & \frac{\mu_1+\mu_2}{2}, \\ \mathrm{GC:} & \frac{\mu_1\,\mu_2}{2}, \\ \mathrm{SI:} & \frac{2}{\pi}\tan^{-1}\frac{\mu_2+\mu_1}{\mu_2-\mu_1}, \\ \mathrm{CI:} & \sqrt{\frac{\mu_1^2+\mu_2^2}{2}}, \end{cases} \tag{3.31}$$

where μ_1 and μ_2 are the principal curvatures at point P.

According to Figure 3.5, which demonstrates how the achieved error and Standard Deviation (SD) of error vary by altering, in pairs, the weights employed in the Γ function Eq. (3.21), the algorithm is able to achieve a low error value (blue areas) for a wide range of those parameters regarding each of the descriptor. Looking at the Overall LM mean error of the figure, all the descriptors have stable performance around $\zeta = 2$, and $\nu = 2$ which is visible with blue region; however, the area of the region explains the degree of robustness is changing. The robustness is theoretically expected because the system is globally asymptotically stable as investigated in Section 3.3, and here the expectation was proved as there was limited sensitivity regarding the choice of the range of parameters, and no fine-tuning is further needed to obtain the desirable results.

According to the figure, GC has the highest sensitivity as the color variation is the fastest and CI has the least sensitivity due to slow changes in the color. Meanwhile, MC has the biggest area with the least error (dark blue area), thus we choose MC formula for our calculation.

3.4.5. ANFIS INITIAL CONDITION PREDICTOR

Following Discussion II in Section 2.3.4, the main challenge of designing an ANFIS model is determining the proper inputs and output [255]. Firstly, the stiffness ratio is estimated through Mean curvature and mesh quality values (both explain the topology of

Table 3.3: Average RMSE and R^2 results of ANFIS performance for different MF types via the test data.

MF type	psigmf	dsigmf	pimf	gaussm2f
RMSE	0.71	0.70	1.87	0.73
R^2	0.77	0.78	0.71	0.76

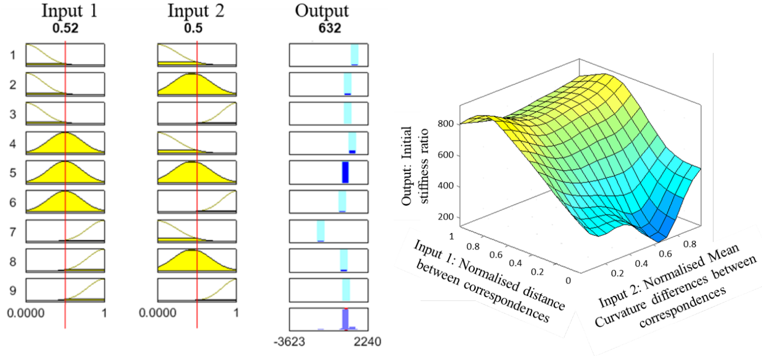
MF type	gaussmf	gbellmf	trapmf	trimf
RMSE	**0.69**	0.72	1.94	0.77
R^2	**0.89**	0.78	0.69	0.81

the mesh) explained in Section 3.3 and 3.4.4. Second, the distance between the point on $\mathscr{S}$ and its correspondence on $\mathscr{T}$ affects the level of stiffness i.e., when the distance is high, the stiffness should be high as well for being able to transfer the features and then register on the target. Thus, we consider the normalised distance (input 1) and the differences between the normalised the Mean Curvature (MC) values (input 2) of each point on the $\mathscr{S}$ and its correspondences on $\mathscr{T}$ before registering as the inputs of the ANFIS model. Consequently, we consider the initial stiffness ratio as the output of the model. We used the final results of the adaptive estimator in Figure 3.5 regarding the MC formulation in Eq. (3.31) as the label of the model, where $\zeta = 2$, and $\nu = 2$ consists of 101 registered meshes.

We considered the source mesh before registration as $\mathscr{S}$ and the registered source mesh via the adaptive estimator as $\mathscr{T}$ to generate the inputs of the training data of the model and the final estimated ratios as the ANFIS outputs. Therefore, the dataset includes 101 scans, and each scan contains 19882 vertices. 70 scans (equal to 70% of the dataset) are used as the train dataset and 31 (equal to approximately 30% of the dataset) are employed as test dataset [248, 252].

Eight different membership functions (MF) mentioned in Table 3.3 and explained in [165] were examined in the ANFIS model. The Root Mean Square Error (RMSE) and the correlation coefficient (R^2) between the anticipated values (initial stiffness ratio) and the true training values were employed to assess the performance of each MF, according to the lowest modeling RMSE and highest correlation coefficient. The results of averaging the RMSE and accuracy are reported in Table 3.3, where the experiments are based on ten runs over 500 training epochs with linear MF type for output, and three MFs for each input. Regarding the modeling results reported in Table 3.3, between the examined MFs, gaussmf performs to the highest accuracy and the smallest RMSE. Guassmf employs the general form of a gaussian function. The ANFIS rules were selected to predict the proper initial stiffness ratio due to the normalized topology of $\mathscr{S}$ and estimated normalized travel distance that resulted in 9 fuzzy rules. Figure 3.6(a) depicts these rules.

To investigate the sensitivity of the model's inputs variation to its output variation, the level of prediction is shown in Figure. 3.6(b). According to the figure, the level of the model's inputs through its output shows a smooth area around the employed range of the output, and inputs reached steady, which conveys the acceptable level of design relationships, sensitivity, and robustness of the predictor output to its inputs.



(a) (b)

Figure 3.6: ANFIS model. (a) Fuzzy rules. (b) Fuzzy surfaces for the intelligent estimator.

3.4.6. EMPLOYED METHODOLOGIES FOR COMPARISON

For the experiment, the performance of the proposed approach is compared to several existing non-rigid registration methods presented by Amberg et al. [12], Andriy et al. [180], and Hirose [102], where the executable source code for each method is available in [42], [179], and [101] respectively, and Tajdari et al. [251].

Briefly, Amberg's method introduced an optimal step non-rigid ICP methodology utilising a variety of regularisations, while employing a range of lowering stiffness parameter equals for all the vertices.

Andriy account for the CPD algorithm [180] utilising Gaussian radial basis functions as a substitution for thin-plate splines which investigates another type of regularizer. Through the approach, the rigid and non-rigid registrations are considered, although the method is not optimal in identifying a large number of outliers and in establishing all the possible correspondences, which is generally identified as Non-deterministic Polynomial-time hard (NP-hard). The executable source code employed in this paper is achievable in [179].

Hirose proposed the BCPD method [102], using variational Bayesian inference theory to convey the coherent drift. The executable source code employed in this paper is achievable in [101].

Tajdari introduced a novel geometry feature as semi-curvature for better classification and corresponding point selection performance. Also, he linearized the term and integrated to the Amberg's solver and showed the linear version is globally stable; however, the method was ill to preserve the mesh quality for the boundary of missed parts, e.g holes.

As both Amberg's method and Tajdari's method are using the same stiffness term integrated with the weighting matrix G same as Eq. (6.7), we are able to apply the estimator to both methods. We named these two variants as adaptive Amberg's method and adaptive Tajdari's method. Comparing our both proposed methods with the adaptive Amberg and the adaptive Tajdari may reveal the generalizability and extendability of our estimator for a variety of applications.

All methods were implemented by Matlab®R2020a on a computing platform with an Intel® Core-i5™ 9600K 4.6 GHZ processor.

3.5. EXPERIMENT RESULTS

3.5.1. EFFECT OF THE ADAPTIVE SCHEME DURING REGISTRATION

Here, we study the impact of the adaptive controller on the level of feature preservation after registration, via a few experiments employing some basic shapes, in terms of quality of mesh and the degree of followed topology, especially for highly curved areas, missed areas, disconnected areas, and misaligned areas.

FEATURE PRESERVATION

As we would like to investigate the impact of the estimator on the performance of the registration process only the methods with similar mechanisms are included.

To show the effect, a shape with three spherical cross-sections (gray) is used as $\mathscr{S}$ and a shape with three ellipsoidal cross-sections (green) is recruited as $\mathscr{T}$ depicted in Figure 3.7(a). Figure 3.7(b) shows the initial condition and the different regions for assessment as feature matching (part 1) to evaluate the accuracy of corresponding selection of points with similar normalized Mean curvature values, missed area (part 2), disconnected area (part 3), and misaligned part (part 4). We introduce a new measure named percent of mesh quality loss (Q^{loss}) for the evaluation as

$$Q^{loss} = 100 \times \left(1 - \frac{Q_S^{end}}{Q_S^0}\right) \tag{3.32}$$

where Q_S^0 and Q_S^{end} are the mesh quality of the source mesh for a region in the first and last intervals respectively. Figure 3.7(c) presents the performance of Amberg's method in the four regions where it loses 80%, 82%, 90%, and 85% of mesh quality for part 1, part 2, part 3, and part 4 respectively. Figure 3.7(d) figures the performance of Tajdari's method in the four regions with 15%, 18%, 20%, and 16% mesh quality loss for part 1, part 2, part 3, and part 4 respectively. Our adaptive method and our ANFIS adaptive method outputs are drawn in Figure 3.7(e) and (f) which report the same performance in terms of mesh quality loss of 7%, 8%, and 7% for the first three regions. In addition, our adaptive method has 2% less mesh quality loss than our ANFIS adaptive method for the part 4.

LEVEL OF CONVERGENCE TO THE TARGET MESH TOPOLOGY

Preserving the mesh quality and following the topology of the target mesh during registration are competing with each other. As the mesh quality preservation needs more rigidity and following the topology requires less rigidity. To investigate the level numerically we figured the histogram graph, mean value ($\bar{m}$), and standard deviation (σ) of the closest distance population from target mesh points to source mesh points in Figure 3.8 per part 1, part 2, part 3 and part 4, regarding each of the compared methods in Figure 3.7. According to Figure 3.8(a) where the part 1 is studied, Amberg's method has the highest σ, and our ANFIS adaptive method has the least σ. This shows our method could

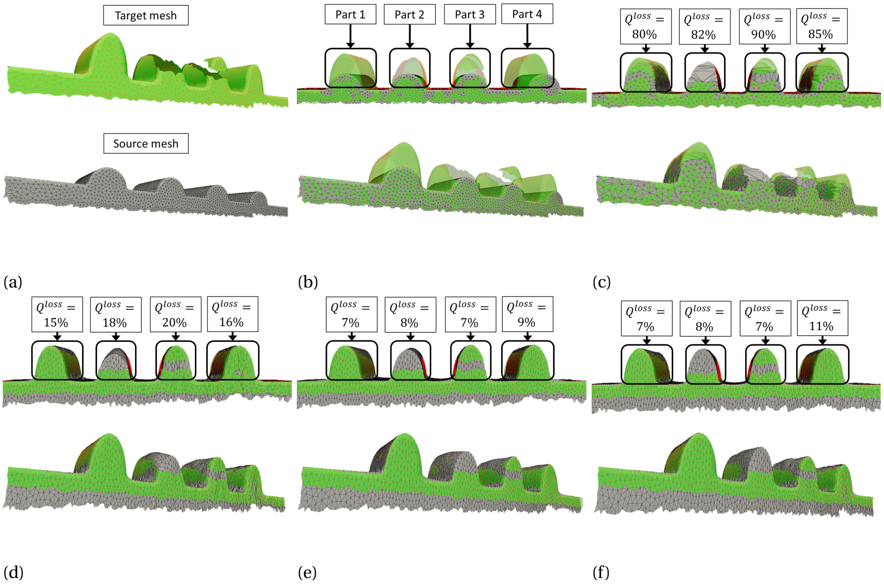


Figure 3.7: Curved shape preservation experiment: Front view on top, and the prospective view on bottom. (a) Target and source meshes (b) The initial condition (c) setup Registration using Amberg's method. (d) Registration using Tajdari's method method. (e) Our adaptive method. (f) Our ANFIS adaptive method.

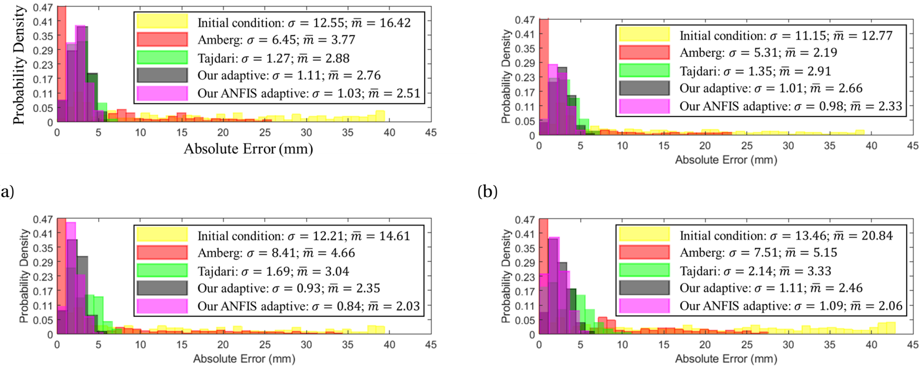
successfully reach the far points on the target mesh from the source mesh. However, as there is no missed area in the part 1, Tajdari's method has similar $\bar{m}$ and σ to our both methods.

Figure 3.8(b) investigates the performance of the method in presence of missed areas in the part 2. In this experiment, Amberg still has the highest $\bar{m}$ and σ values, while our both methods have better performance than Tajdari's method. Both of the proposed methods could follow the target mesh topology for about 28% better than Tajdari's method, which was expected due to better preservation of the source mesh feature close to the borders of missed area.

In addition, Figure 3.8(c) presents the results regarding the experiment of having disconnected areas as part 3. Although the conclusion for this part is similar to part 2, the advantage of our both methods is more visible here. Accordingly, our both methods outperformed for about 52% better than Tajdari's method in terms of $\bar{m}$, which shows the power of the adaptive estimator to face more borders of missed areas.

Furthermore, Figure 3.8(d) may highlight more clearly the importance of the ANFIS predictor than the other experiments. As the ANFIS adaptive method has 13% and 16% less $\bar{m}$ compared to the adaptive method for the part 3 and part 4 experiments respectively. This shows the ANFIS adaptive method has 23% improvement based on the adaptive method, which resulted from more chance of freedom by having logically lower stiffness values in the initial condition of the registration process for the ANFIS adaptive scheme.

Overall, comparing Figure 3.8(a), (b), (c), and (d) shows our ANFIS adaptive method



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

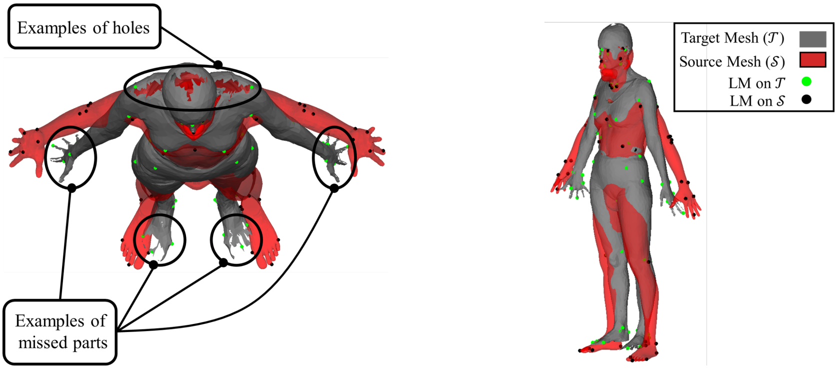
Figure 3.8: Histogram graph of closest distance from target mesh points to source mesh points where the Diameter of the Bounding Box (DBB) is 691 (mm). (a) Part 1: feature matching. (b) Part 2: missed part. (c) Part 3: disconnected part. (d) Part 4: Misaligned part.

has slightly better performance than our adaptive method in terms of $\bar{m}$ and σ values, this is because the ANFIS predictor is employed to avoid the stiffness ratio saturation meaning more flexibility for correct deformation of each point on the source mesh to follow the topology of the target mesh as rigid as possible. This is numerically investigatable by comparing Figure 3.8(a), (b), (c), and (d), where the ANFIS adaptive method has on average 6% and 13% less value of the σ and $\bar{m}$ than the adaptive method respectively.

Discussion III: Using gradually decreasing uniform stiffness level for all the points in each interval of the registration process used in Amberg et al. [12], and Tajdari et al. [251] may cause mesh quality loss by converging several points on $\mathscr{S}$ to one point on $\mathscr{T}$. The mesh quality loss especially happens in the final intervals of the registration process where the stiffness term is sufficiently low then the surface can be deformed unfavorably and the source face features can be lost. Although Tajdari's method could partly solve the challenge by improving the accuracy of corresponding selection and avoiding the confliction contrary to Amberg's method, Tajdari's method was ill to preserve the features and mesh quality of the source mesh on the boundaries of missed areas and disconnected areas.

3.5.2. FULL HUMAN BODY

In terms of validation, we use the CAESAR test dataset in Section 3.4.5 and the predefined landmarks introduced in Section 3.4.1. Visually, the initial condition and the selected landmarks are shown in Figure 3.9 for both the target and source meshes. The CEASAR dataset is suitable for assessment, as they are natural scans with holes, missed parts and natural noise e.g., a few examples are visually highlighted in Figure 3.9(a). In terms of validation, the registration results are reported employing the remeshed source mesh with 100%, 60%, and 20% of the total vertices of the original source mesh. The experiment results of the registration are reported in detail in Table 3.4, shown in Figure 3.12. Please note that for the experiments that the ANFIS predictor is employed, $\hat{g}^{\min}$ in Table 6.1



(a) (b)

Figure 3.9: Full body initial condition for one example among the test dataset. (a) Top-view. (b) Perspective view.

considered as constant and equal to 1.

LANDMARK ERROR

Based on the Table 3.4(a) where the length of the diagonal line of the Bounding Box (DBB) for $\mathscr{S}$ for the source mesh is 2167.6 mm, our both methods indicate on average about 78%, 91%, 75%, 30%, 80%, and 15% less landmark errors comparing to the Amberg's method, CPD method, BCPD, Tajdari's method, Adaptive Amberg's method, and adaptive Tajdari's method respectively. Figure 3.10(a) presents the percentage of correspondences for the case with 100% of the total vertices of the original source mesh including all 73 selected landmarks for all the 26 scans (y-axis) from the test data, that have less distance error than a threshold (x-axis) [116]. According to the figure, the proposed two methods (black and brown lines) find all the correspondences earlier than the other four methods, and the errors are less than 21 mm, while for Amberg (red line), CPD (purple line), and BCPD (yellow line), Tajdari (green line), adaptive Amberg (blue line), and adaptive Tajdari (cyan line), the errors are 91 mm, 371 mm, 98 mm, 40 mm, 66 mm, and 30 mm, respectively. The results prove that the presented two methods are more accurate and robust compared to other methods.

MESH QUALITY AND Q^{loss}

Based on Table 3.4(b), the observed mesh quality preservation improvement of the proposed two methods is on average about 48% better than Amberg's method, 12% than the CPD method, 9% than the BCPD method, 11% with Tajdari's method, 6% than the adaptive Amberg's method, and 2% with adaptive Tajdari's method.

Moreover, in terms of Q^{loss} (in Table 3.4(b)) for the proposed two methods, one can be observed that they have on average 3% the quality loss while it is 44% for Amberg's method, 12% for CPD method, 10% for BCPD method, 14% for Tajdari's method, 7% for adaptive Amberg's method, and 6% with adaptive Tajdari's method.

Figure 3.10(b) presents the sorted mean final value of $\hat{g}$ per vertex on $\mathscr{S}$, it can be inferred that the adaptive method tried to exert rigid registration on the Amberg's method

Table 3.4: Results of registration regarding the full human body dataset.

a) Overall LM mean error ± Standard Deviation (SD) (mm).								
Total vertices %	Amberg [12]	CPD [180]	BCPD [102]	Tajdari[251]	Adaptive Amberg	Adaptive Tajdari	Our adaptive	Our ANFIS adaptive
100	31.3±17.5	77.1±99.3	28.2±23.3	11.2±7.7	33.5±10.1	8.1±6.9	7.4±3.9	**7.2±4.2**
60	31.9±29.7	93.3±101	44.1±27.4	13.9±13.3	39.9±14.6	11±8.1	**10.2±5.6**	9.1±4.4
20	50.8±55.2	189.1±127.3	53.5±51.6	18.3±22.1	40.1±14.8	17.4±18.1	21.3±9.7	**16.8±7.5**

b) Mesh quality mean value ± Standard Deviation (SD); Percent of mesh quality loss mean value (Q^{loss}) % in Eq. (3.32).									
Total vertices %	Original mesh (Q_S^0)	Amberg [12]	CPD [180]	BCPD [102]	Tajdari[251]	Adaptive Amberg	Adaptive Tajdari	Our adaptive	Our ANFIS adaptive
100	0.80±0.15	0.53±0.24; 36	0.67±0.23; 16	0.69±0.27; 14	0.68±0.29; 15	0.71±0.3; 11	0.74±0.17; 7	**0.75±0.1; 6**	0.74±0.13; 7
60	0.72±0.20	0.40±0.27; 43	**0.71±0.22; 1**	0.71±0.20; 1	0.65±0.31; 11	0.68±0.11; 5	0.68±0.14; 5	0.70±0.16; 3	0.70±0.18; 3
20	0.70±0.22	0.30±0.40; 56	0.57±0.24; 21	0.58±0.33; 16	0.53±0.45; 27	0.66±0.17; 6	0.66±0.2; 6	**0.69±0.13; 1**	0.68±0.11; 3

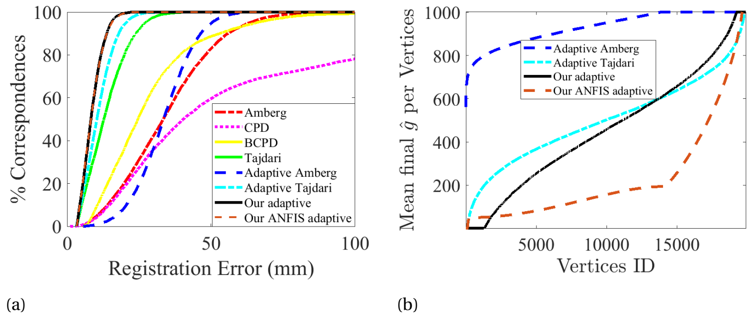
c) The run time mean value ± Standard Deviation (SD) (s).								
Total vertices %	Amberg [12]	CPD [180]	BCPD [102]	Tajdari[251]	Adaptive Amberg	Adaptive Tajdari	Our adaptive	Our ANFIS adaptive
100	**18.3±1.2**	1571±831	755±383	21.3±2.4	18.5±1.2	21.9±2.5	18.4±0.9	18.4±0.9
60	**12.5±0.9**	851±738	456±344	19.6±1.5	12.5±1.0	19.7±1.5	12.6±0.9	12.8±0.9
20	**8.0±0.4**	491±357	302±277	15.2±0.6	8.0±0.7	15.3±0.8	8.2±0.8	8.3±1.8

d) Mean distance between each point on $\mathscr{T}$ to the its closest point on registered S ± SD (mm).								
Total vertices %	Amberg [12]	CPD [180]	BCPD [102]	Tajdari[251]	Adaptive Amberg	Adaptive Tajdari	Our adaptive	Our ANFIS adaptive
100	**3.7**±5.5	17.7±12.5	6.1±3.2	13.1±3.7	16.1±11.9	18.7±5.4	9.3±3.7	6.2±**3.1**
60	**3.8**±5.5	19.2±13.8	6.8±3.9	16.3±5.2	16.3±11.6	20.3±7.8	10.1±3.6	7.1±**3.0**
20	**3.6**±5.9	20.1±13.5	7.4±4.1	19.1±7.1	16.4±11.7	24.7±9.9	10.6±3.7	7.5±**3.3**

as the correspondences are mainly selected incorrect, thus the mesh quality should be decreased (toward zero area faces). Accordingly, the estimated $\hat{g}$ is very high. Adaptive Tajdari, Our adaptive and ANFIS adaptive methods achieved way lower the estimated values than the Adaptive Amberg method, while still the ratio for our adaptive method stays less than adaptive Tajdari for more vertices (for about 15000 vertices over 19882 vertices), which means the correspondences are selected more accurate thus less deformation and less mesh quality loss is needed. Although our ANFIS adaptive method presents the same high quality registration as our adaptive method, the estimated stiffness ratios are considerably lower, which gives the opportunity to follow the topology of the target more accurately. This also can be observed visually by comparing the shown examples in Figure 3.11, and 3.12 for one example.

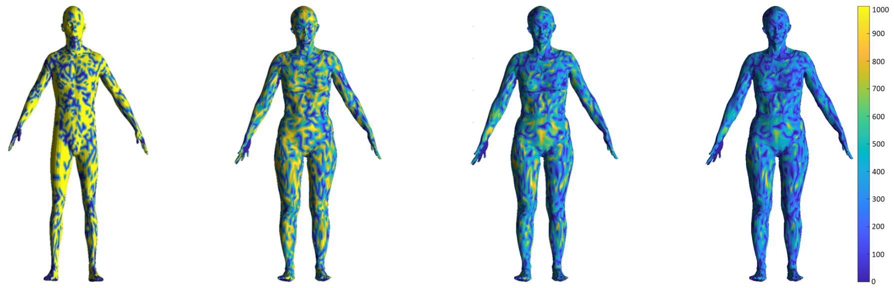
RUN TIME DISCUSSION

The computing speed is also a criterion to assess the performance of each mentioned method. Table 3.4(c) presents the computation time of all methods. Comparing the proposed two methods against other methods, the computing time reduction is about 0% less than Amberg's method, 99% less than CPD method, 97% less than BCPD method, 15% less than Tajdari's method, 0% less than adaptive Amberg's method, and 15% less than adaptive Tajdari's method. The results show our both methods have acceptable performance in terms of computing time regarding Amberg's method, however, both of them present relatively higher quality of registration regarding landmark error and mesh quality preservation.



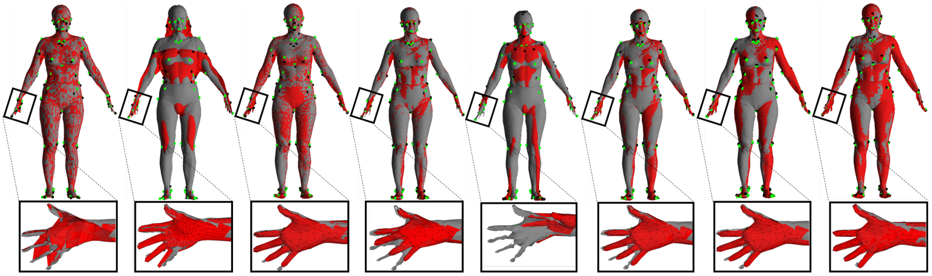
(a) (b)

Figure 3.10: Registration results via the test dataset for the case with 100% of the total vertices of the original source mesh. (a) Percentage correspondences according to registration error. (b) Sorted mean estimated stiffens ($\hat{g}$) per each vertex.



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

Figure 3.11: Estimated final stiffness ratio color map. (a) Adaptive Amberg. (b) Adaptive Tajdari. (c) Our adaptive method. (d) Our ANFIS adaptive method.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 3.12: Geometry of the full body registration results. (a) Amberg. (b) CPD. (c) BCPD. (d) Tajdari (e) Adaptive Amberg. (f) Adaptive Tajdari. (g) Our adaptive method. (h) Our ANFIS adaptive method.

Target-Source distance distribution

TABLE 3.4(d) explains the mean distance of each point on the $\mathscr{T}$ and the closest point on the registered $\mathscr{S}$ along the test data. According to the table, Amberg shows the least

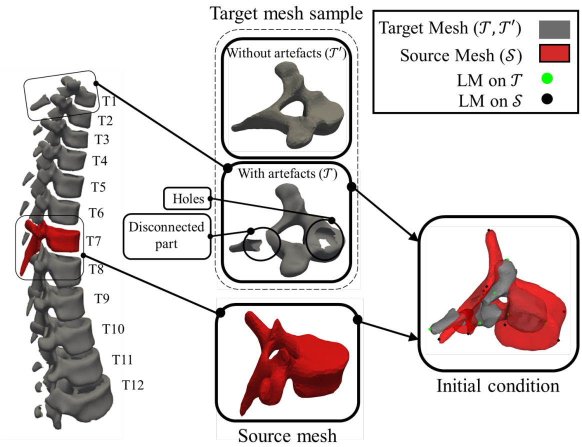


Figure 3.13: The used thoracic dataset for assessment, including T1 as the target mesh, T7 as the source mesh, and their initial condition before non-rigid registration.

sparsity and CPD shows the largest sparsity in terms of the closest point. This happens as Amberg did not embed sufficient terms to keep the original geometry of $\mathscr{S}$. In addition to the error comparison, our both methods perform with the least error of correspondences and highest level of the source mesh geometry preservation. While, our ANFIS adaptive method presents 34% less error than our adaptive method, which allows us to perform more logical registration resulting in better following the topology of the target with similar quality of LM errors and surface quality.

3.5.3. THORACIC VERTEBRA

Using the same method in the previous section, we register the thoracic vertebra of $T7$ with DBB of 42 (mm), from the dataset in Section 3.4.1.2 on the thoracic vertebrae of $T1,\ldots,T12$ excluding $T7$ (eleven vertebrae in total) using the methods in Section 3.4.6.

In all eleven experiments, the initial conditions for the eight methods are the same. As an example, the initial condition of the $T7$, and the $T1$ is presented in Figure 3.13. Please note that we manually made some artefacts in the target meshes shown in Figure 3.13 i.e., holes and disconnected parts, in order to investigate the challenges discussed in Section 3.1. In addition, the dataset is supposed to be able to highlight the advantages and disadvantages of each method as the variation of the topology of the vertebra from $T1$ to $T12$ is comparatively high [286]; however, all the vertebrae monopolize all the features belonging to a thoracic vertebra. For the used dataset in Section 3.4.1.2, the value of Cranial Endplate Depth (EPD), over spinous process length (SPL) introduced in [286] are varying from 0.68 to 2 which highlights the topology variation. For the assessment, we used the 12 landmarks explained in Section 3.4.1.2 and shown in Figure 3.13 with black dots (for $\mathscr{S}$) and green dots (for $\mathscr{T}$).

The registration results of using the eight methods are visually compared in Figure 3.14 through an example of registering the $T7$ on the $T1$, and numerically investi-

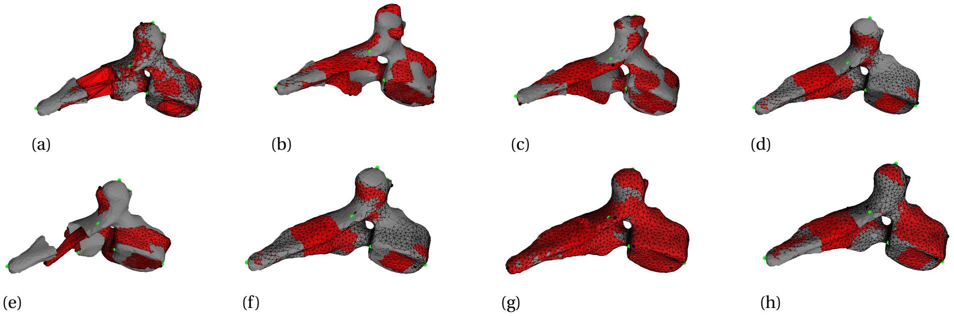


Figure 3.14: Geometry of the thoracic vertebra registration results. (a) Amberg. (b) CPD. (c) BCPD. (d) Tajdari (e) Adaptive Amberg. (f) Adaptive Tajdari. (g) Our adaptive method. (h) Our ANFIS adaptive method.

gated in Table 3.5 for all eleven registration results.

Errors between corresponding landmarks are reported in Table 3.5(a) where our methods outperform Amberg's, CPD, BCPD, Tajdari's, adaptive Amberg's, and adaptive Tajdari's methods by an average of 87%, 82%, 74%, 56%, 87%, and 53% respectively regarding the mean absolute error. In addition, our ANFIS adaptive method presents 12% less error than our adaptive method.

Regarding the mesh quality of the source mesh after registration, Table 3.5(b) shows that our methods present on average 7.5% the mesh quality loss, in which the value is about 48%, 25%, 14%, 18%, 2%, and 14% for Amberg's, CPD, BCPD, Tajdari's, adaptive Amberg's, and adaptive Tajdari's methods. According to the table, our methods have lower efficiency in terms of mesh quality only compared with adaptive Amberg's method, which comes from the fact that the adaptive solution found Amberg's algorithm path towards non-rigid registration, unfavorable and thus exerted as rigid as possible transformation. In other words, the adaptive Amberg's method presented high mesh quality as the registration was rigid mostly and thereby failed to match the target mesh. Furthermore, our ANFIS adaptive method presents 3% more mesh quality loss than our adaptive method.

In Table 3.5(c), it can be found that using the CPD method is very time-consuming as the time duration for the experiment is about 6810%, 87%, 4191%, 6480%, 4139%, 6015%, and 5755% more than Amberg's, BCPD, Tajdari's, adaptive Amberg's, adaptive Tajdari's, our adaptive, and our ANFIS adaptive methods.

In Table 3.5(d) contrary to the comparison using the full human body in Section 3.5.2.4, we compare the distance from the target meshes without the artefacts ($\mathscr{T}'$ in Figure 3.13) to the registered source mesh for the eight methods. Note that this comparison was not possible for the full human body experiment as the original data of the CAESAR had the holes, missed parts, and natural noise. According to Table 3.5(d), our methods outperform Amberg's, CPD, BCPD, Tajdari's, adaptive Amberg's, and adaptive Tajdari's methods by an average of 88%, 72%, 65%, 42%, 92%, and 52% respectively regarding the distance from the $\mathscr{T}'$ to the registered source mesh.

To have a holistic conclusion on the performance of the methods, we consider four matrices as $\mathscr{B}_i$ i.e., $i = 1, \ldots, 4$ where each of the matrices includes the mean values of

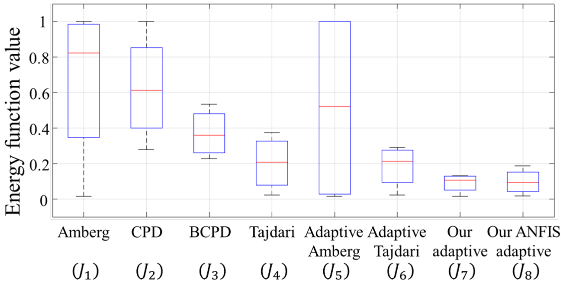


Figure 3.15: Energy values box-plot per each method.

Table 3.5: Results of registration regarding the thoracic vertebra dataset.

Original mesh (Q_S^0)	Amberg [12]	CPD [180]	BCPD [102]	Tajdari[251]	Adaptive Amberg	Adaptive Tajdari	Our adaptive	Our ANFIS adaptive
a) Overall LM mean error ± Standard Deviation (SD) (mm).								
–	5.9±4.1	4.3±3.6	2.9±3.2	1.7±1.1	6.1±4.9	1.6±1.1	0.8±0.9	**0.7±0.9**
b) Mesh quality mean value ± Standard Deviation (SD); Percent of mesh quality loss mean value (Q^{loss}) % in Eq. (3.32).								
0.84±0.11	0.44±0.29; 48	0.63±0.22; 25	0.68±0.19; 14	0.69±0.20; 18	**0.82±0.20; 2**	0.72±0.19; 14	0.79±0.12; 6	0.76±0.16; 9
c) The run time mean value ± Standard Deviation (SD) (s).								
–	**10.1±0.9**	691±328	369±252	16.1±1.7	10.5±1.1	16.3±2.2	11.3±1.0	11.8±1.1
d) Mean distance between each point on $\mathscr{T}'$ (without artifacts) to its closest point on registered S ± SD (mm).								
–	0.95±0.81	0.39±0.33	0.32±0.29	0.19±0.15	1.4±1.8	0.23±0.21	0.12±0.13	**0.10±0.09**

the four features used in Table 3.5. The mean value of each feature is named as $\mathscr{B}_i(z)$ i.e., $z = 1,\ldots,8$ where z is corresponding to the number of the compared method, in the Table 3.5 with the same order of the methods. For example, $\mathscr{B}_3$ includes the mean run time values for all methods, where $\mathscr{B}_3(4) = 16.1$ (corresponded to Tajdari's method in Table 3.5(c)). Note that from Table 3.5(b), we only considered the percent of mesh quality loss mean value e.g, $\mathscr{B}_2(4) = 18$. Then, an energy function of J_z for each of the methods can be calculated as:

$$J_z(i) = \frac{\mathscr{B}_i(z)}{max(\mathscr{B}_i)} \tag{3.33}$$

The values of each J_z are drawn in Figure 3.15, which shows that the overall performance by considering all matrices of comparison, our adaptive and ANFIS adaptive methods have the lowest value of the energy function which means the methods have the best performance. In addition, the median of the ANFIS adaptive method is slightly lower than our adaptive method, showing the positive impact of the ANFIS predictor on the overall performance of the method.

3.6. LIMITATIONS

The presented methods address the feature preservation during a non-rigid registration via controlling the stiffness per vertex term in each iteration. Though using both methods can achieve higher quality of mesh and less landmark errors, they have several limitations. The source and target mesh should have similar normalized curvature values for matching otherwise the adaptive estimator compels a rigid registration e.g., registering a human body on a dog body. This also prevents the method to tackle the challenges of high-frequency deforming objects e.g., a deforming cloth where the similar features of the target mesh to the source mesh are highly varying. Moreover, a weakness of using curvature values in establishing correspondences, is that the value is very sensitive to the noise, thus using mesh smoothing algorithms, e.g. Laplacian filter, is recommended prior to the registration. In addition, the global stability of (6.7) after integration of $\hat{G}$ in (3.4) is not investigated, as in Amberg [13], to prove the global stability of the solver.

3.7. CONCLUSION

In this chapter, we introduce a non-rigid ICP approach integrated with a novel adaptive feedback control scheme to estimate the stiffness ratio utilizing the gradient of the mesh quality and Mean curvature values per vertex. To facilitate the convergence procedure of estimation towards the true values, an ANFIS-adaptive-based predictor is integrated with the estimator. The ANFIS predictor proceeds based on the topology of both the source and the target meshes to anticipate the initial values for the adaptive estimator to facilitate the procedure of the estimation. Then, we embedded the estimated ratio as part of the metric in establishing the correspondences and in the cost function, where the distance and the stiffness terms are integrated as well. Furthermore, by adjusting the connectivity level of vertices on their neighbors (equal to stiffness), the ANFIS based adaptive estimator elaborates on preserving the features of a surface in a globally asymptotically stable region during the process of registration, stressing mesh quality and convergence. Experimental outcomes show that the presented non-rigid method outperforms the aforementioned approaches, especially in the highly curved areas, missed areas, disconnected areas, and misaligned areas. This highlights the ability of the proposed method to employ the inherent characteristics of the complete surface during the process of registration.

4

PROGNOSIS DEEP LEARNING FRAMEWORK FOR PEDIATRIC SPINAL DEFORMITIES

Predicting pediatric spinal deformity (PSD) from X-ray images collected on the patient's initial visit is a challenging task. This work builds on our previous method and provides a novel bio-informed framework based on a mechanistic machine learning technique with dynamic patient-specific parameters to predict PSD. We provide a geometry-based bone growth model that can be utilized in a range of applications to enhance the bio-informed mechanistic machine learning framework. The proposed technique is utilized to examine and predict spine curvature in PSD cases such as adolescent idiopathic scoliosis. The best fit of a segmented 3D volumetric geometry of the human spine acquired from 2D X-ray images is employed. Using an active contour model based on gradient vector flow snakes, the anteroposterior and lateral views of the X-ray images are segmented to derive the 2D contours surrounding each vertebra. Using minimal user input, the snake parameters are calibrated and automatically computed over the dataset, resulting in fast image segmentation and data collection. The 2D segmented outlines of each vertebra are transformed into a 3D image segmentation result. The Iterative Closest Point mesh registration technique is then used to establish a mesh morphing approach and creates a 3D atlas spine model. Using the comprehensive 3D volumetric model, one can automatically extract spinal geometry data as inputs to the mechanistic machine learning network.

Parts of this chapter have been published in:
Tajdari, M., **Tajdari, F.**, Shirzadian, P., Pawar, A., Wardak, M., Saha, S., Park, C., Huysmans, T., Song, Y., Zhang, Y.J. and Sarwark, J.F., 2022. Next-generation prognosis framework for pediatric spinal deformities using bio-informed deep learning networks. *Engineering with Computers*, 38(5), pp.4061-4084 [258].

4.1. INTRODUCTION

Deformities in the spine may cause pain, numbness, tingling sensation, loss of function, and even pulmonary and cardiac difficulties. If the spine loses its usual, graceful S-shape (when seen from the side), or if the spine loses its straightness, deformities may arise (viewed from front to rear). The term Pediatric Spinal Deformity (PSD) refers to such malformations of the spine in children, namely, scoliosis, kyphosis, and spondylolisthesis [285]. Scoliosis and kyphosis are conditions characterized by aberrant curvature of the spine, while spondylolisthesis is characterized by the displacement of vertebrae. PSD is caused by a variety of factors, the most important of which is bone growth, which controls the curvature and advancement of the deformity. There is still more to learn about the underlying etiology of PSD. Affecting roughly 7 million people in the United States, scoliosis is one of the most prevalent spinal abnormalities. Adolescent Idiopathic Scoliosis (AIS) is used to describe the condition in children and adolescents. AIS accounts for around 80% of all pediatric scoliosis cases, while affecting approximately 3% of adolescents under the age of 16 in the United States [123].

The treatment of scoliosis is mainly reliant on the shape and extent of spinal curvature, and specific treatment choices are usually determined by the surgeon's expertise. As a result, the development of a clinically validated, patient-specific model of the spine to assist surgeons in the prognosis of early-stage PSD would guide optimum surgical and non-surgical treatment options. For both screening and monitoring in present clinical practice, the lack of an appropriate safe, inexpensive, and accurate measuring technology is a major bottleneck. To track the evolution of deformities throughout adolescent growth, frequent imaging is essential [90] and the extracted features from medical images such as Cobb angle determine the severity of spinal deformity. The Cobb angle is the most often utilized measurement for determining the severity of spinal deformities.

It is common practice in spine surgery to use computer-aided procedures, such as determining the best path for the insertion of pedicle screws [120, 174, 91], as well as improved surgical navigation and a more comprehensive pre-operative surgical plan [177, 151]. One of the greatest barriers to incorporating these technologies into clinical practice is the time and effort necessary to generate patient-specific functional models from medical imaging. It entails a number of manual procedures and is time intensive, even for seasoned specialists [219, 40]. For example, image segmentation is both time-consuming and user-dependent [14, 145] because it requires to locate and segment vertebrae ahead of time [27]. Generating 3D detailed geometry of the spine from a 2D set of X-ray images is also a challenging task. Various manual processes are required to segment the obtained volumetric mesh in order to detect hard and soft tissue once the 3D model has been formed.

Machine learning (ML) approaches need a large amount of data to be trained and provide reliable results [255, 252]. The absence of reliable medical data for a given individual over time is one of the challenges in implementing ML for prognosis spinal deformity. Furthermore, these models are referred to as "data-hungry" approaches since they cannot forecast outside of the range of the training data [247]. Recent publications show that by including physics into the system, one may enhance the model's predictability range [206]. However, for spinal curve progression including bone formation, patient-specific governing physical equations with time-varying and geometric-based

Table 4.1: Contributions of proposed framework with respect to previous framework [257].

Steps of Framework		Limitations of Previous Framework	Contributions of Current Framework
1	2D image segmentation	Manual parameter adjustment	Calibrated parameters and semi-manual adjustment to expedite the process
2	3D reconstruction from bi-planar data and Volumetric mesh generation	Labor intensive	Fast and robust
3	Mechanistic deep learning method using bone growth model	Constant parameters which are independent of time and position	Patient-specific, position-sensitive and dynamic bio-informed mechanistic deep learning parameters

coefficients are unavailable.

Previously published research from our group has used X-ray imaging data to develop a prognostic framework for AIS [257]. This framework predicted spine morphology by combining clinical data acquired from X-ray images with mechanistic features such as stress distribution on the growing surface of the vertebrae extracted from a spine surrogate model along with the bone growth model. Although this unique framework used mechanistic data science for forecasting spine deformity, the previous method had a number of limitations that impede deploying this framework in real-world scenarios. One restriction is the manual parameter adjustment for image segmentation of each vertebra. The presented $3D$ reconstruction and geometry generation technique requires considerable labor since each tissue is generated separately and all tissues are assembled together. Moreover, the bone growth model has constant parameters throughout all patients and time steps. However, these assumptions are unrealistic since bone formation differs across ages and vertebrae.

The proposed framework (Figure 4.1) in this chapter builds on our previous work [257], with each step in the framework improved to meet real-world needs. Both the limitations of the previous framework and the contributions of the current framework are summarized in Table 4.1.

The organization of the chapter is as follows: After discussing related works in Section 4.2, we describe the framework to create the patient-specific 3D geometry from X-ray images in Section 3. In Section 4, we present our novel bio-informed mechanistic machine learning model for prognosis of pediatric spinal deformity. Lastly, we end with some conclusions and discuss future directions of this work.

4.2. Related works

4.2.1. Vertebrae image segmentation

Precisely segmenting the vertebrae is critical for subsequent analysis in an injury detection system. Statistical shape model (SSM)-based techniques have dominated previous work in vertebral segmentation [209, 10]. Based on a training set, these approaches cap-

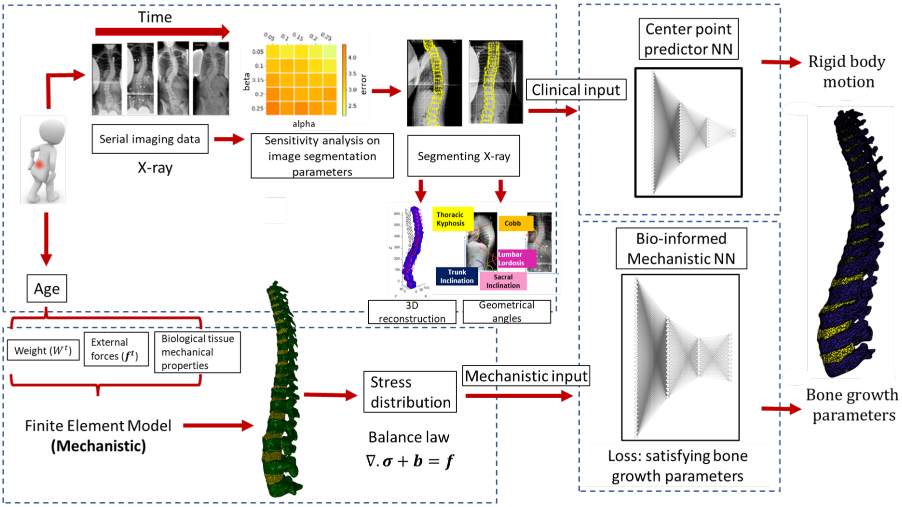


Figure 4.1: The overall workflow of the proposed framework. The geometric data of vertebrae is retrieved from 2D X-ray time-series images using an image segmentation algorithm whose parameters are calibrated through sensitivity analysis. The 2D data is then translated to 3D data where patient-specific features are extracted. The imaging data (clinical) and the mechanistic aspects such as the dynamic patient-specific bone growth model are passed through the bio-informed mechanistic neural network to predict spinal deformity.

ture statistical information on the shape and/or appearance of the vertebra. The mean shape is then manually or semi-automatically set close to the real vertebra, and a search process is used to converge the shape on the true vertebral boundaries. Latest evidence has used random forest-based machine learning (ML) models to achieve shape convergence [55, 10, 38, 208]. These approaches, however, are only efficient and accurate for the restricted data reported in the literature and cannot be applied to patient-specific datasets.

4.2.2. 3D MODEL DEVELOPMENT AND COMPUTATIONAL SIMULATION

Compared to MRI and computational tomography (CT) scanners, X-ray images are more commonly employed due to their accessibility, lower cost, shorter scanning time, and lower ionizing radiation levels particularly for adolescent patients. Statistical Shape Models (SSM) [56, 131, 80, 73] or Statistical Shape and Intensity Models (SSIM) were used to reconstruct bones from X-ray images. To prevent local maxima while optimizing the deformable model parameters, it is important to select a reasonable starting point [115, 207]. Recent years have seen the use of deep learning to recognize landmarks and triangulate them [115]. However, performing 3D reconstruction from two or more 2D photos using a deep learning technique remains a tough problem due to the complexity of describing a dimensional expansion in multi-view circumstances. The EOS imaging system (formerly, Biospace Med, Paris, France), the DIERS formetric scanner, and ultrasonography are examples of recent breakthroughs in diagnostic imaging for AIS [90]. The EOS imaging system is made up of two orthogonal pairs of X-ray tubes and detector units that allow for the simultaneous capture of anteroposterior (AP) and lateral (LAT) X-ray pictures while standing. EOS imaging can quickly scan the spine in 8–15 seconds, depending on the patient's height. From the EOS anteroposterior and lateral pictures, the software system sterEOS (EOS Imaging, Paris, France) can generate a highly accurate 3D model of the spine [63, 133, 69]. However, this technology is unable to divide distinct tissues in order to account for the many biological organs. Due to its expensive cost, it is not widely available in many medical centers.

To gain a better knowledge of load distribution and other mechanical features, the 3D generated patient-specific geometry may be utilized to construct a finite element model. There is no comprehensive automated workflow for anatomically correct FE simulations of the spine based on 2D X-ray data. A lot of work has been done on parametric FE models or a mix of statistical and FE models [25, 26]. However, those models either ignore essential patient-specific features or require a lot of manual labor, which necessitates a certain level of operator experience. Although efforts to automate the construction of FE models of the healthy spine have been performed [23, 40], the technique has never been integrated with deep learning-based segmentation algorithms or applied to diseased situations.

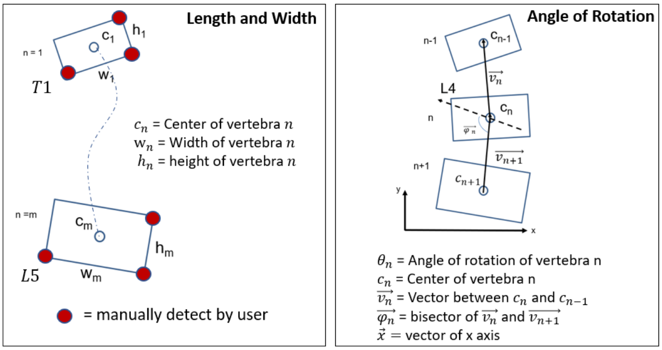
4.2.3. IMPLEMENTING ML FOR STUDYING SPINAL DEFORMITY

The application of ML in medical research has skyrocketed in recent decades. When it comes to applying ML for medical image analysis, there has always been the challenge on how to accurately integrate ML for disease diagnosis, prognosis, and therapy. A framework with such characteristics should always be able to capture the biological governing equation in order to offer extra information to the lacking training data. Recent studies have attempted to use AI to predict spinal deformities [190, 194, 148, 48]. These frameworks, however, cannot be applied to other disciplines. Recently, research has shown that by incorporating the system's underlying physical equations, the framework may forecast data outside of the projected range [257, 216]. Nevertheless, there are certain processes in between that need manual parameter adjustment and, as a result, cannot be implemented for real-time prognosis framework.

4.3. PATIENT-SPECIFIC IMAGE SEGMENTATION AND DATA GENERATION

4.3.1. IMAGE SEGMENTATION AND PARAMETER FITTING

Image segmentation of clinical X-ray images is carried out to extract features for the prognosis framework [319]. The corner points of each vertebra are identified and used as reference points to monitor variations in spine shape and bone formation over time. We implement semi-automated image segmentation using active contour, also known as, the snakes method [112]. A rectangular contour consisting of four corner points is initialized manually around each vertebra and evolved to capture the shape. We define the four corner points of each vertebra as landmarks. These are the key points which help in describing the spine geometry. The segmentation is carried out in 2D for both AP and LAT images. The evolution of the active contour is carried out iteratively through minimization of image energy, allowing it to converge at the edges of features. Snakes are considered energy-minimizing splines which are defined in a parametric form. The



(a) (b)

Figure 4.2: (a) Approximating length and width of vertebrae by picking reference corner points of T1 and L5; and (b) the approximation of angle of rotation for each vertebra.

total energy functional proposed in [112] considers both image and external constraint energy terms.

Active contour model is fast and is able to accurately segment each vertebra in the X-ray images. However, there are a few drawbacks. The accuracy of segmentation depends on how accurately the active contour is initialized. The energy functional which is minimized consists of many weighting parameters which are set empirically at the beginning of the segmentation. The optimized parameters of the snakes method which provides the contour around each vertebra are generally obtained through a trial and error process, which is time-consuming. We propose an improved technique to obtain optimum parameters for the accurate segmentation of clinical X-ray images. The parameters of the Snake algorithm are categorized into two groups: geometric parameters and fitting parameters.

Geometric Parameters: For each vertebra, we initialize the active contour as a rectangle and define geometric parameters such as the width (w_n), height (h_n) and rotation angle (θ_n) from the horizontal axis. n is the vertebra level number where $n = 1 \sim 12$ corresponds to the thoracic vertebra (T1, T2, $\cdots$, T12) and $n = 13 \sim 17$ represent lumbar vertebra (L1, L2, $\cdots$, L5). From the T1 to the L5 vertebra, we assume a linear relationship between each vertebra size and vertebra level, with L5 vertebra having the largest size. The length and width of T1 and L5 can be used to determine the corresponding length and width of other vertebrae. The user chooses three reference corner points (top right, bottom right and bottom left) for T1 (w_1 and h_1) and L5 (w_m and h_m) vertebrae as shown in Figure 4.2(a), and the parameters for the remaining vertebrae are computed using

$$\begin{aligned} h_n &= h_1 + (\frac{h_m - h_1}{N-1})\,(n-1), \\ w_n &= w_1 + (\frac{w_m - w_1}{N-1})\,(n-1). \end{aligned} \tag{4.1}$$

For each vertebra, the user manually selects the center point. The rotation angle can be obtained by

$$\theta_n = cos^{-1}(\frac{\vec{\phi}_n.\vec{x}}{|\vec{\phi}_n|.|\vec{x}|}) \tag{4.2}$$

where θ_n is the rotation angle of the vertebra n, $\vec{x}$ is the vector of horizontal axis and $\vec{\phi}_n$ is the bisector of the vector that connects the two adjacent vertebrae as shown in Figure 4.2(b). It is worth noting that the framework may be used for any number of vertebrae in any regions of interest, including cervical, thoracic, lumbar, or a combination of all three. By estimating geometric parameters for the lumbar spine, Figure 4.3 compares the ground truth and the modified snake method. The findings are in excellent accord with the ground truth, as shown. The ground truth segmentation result is evaluated using the snakes method in which the parameters are manually adjusted and the initialization of the contour is done manually for each vertebra.

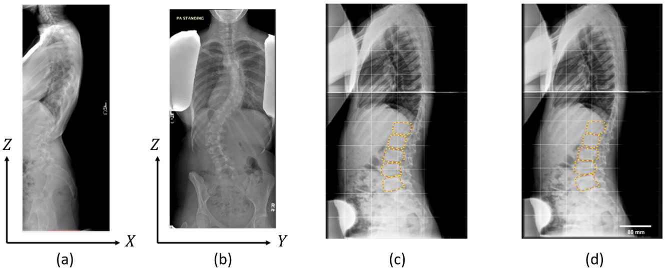


Figure 4.3: Segmentation of the X-ray images. (a) Lateral (LAT) and (b) anteroposterior (AP) view of the image obtained from X-ray data. Comparison between two segmented lumbar spine: (c) ground truth (the output of the snakes algorithm in which the parameters are manually adjusted) and (d) by estimating geometric parameters (angle of rotations, length and width of each vertebra).

Fitting Parameters: The weighting coefficients in the active contour model are set in order to move the contour around each vertebra. α and β are the weights associated with the first-order and second-order regularizing terms of the internal spline energy (Eq. (2) in [112]) that control the tension and rigidity of the snake, respectively. γ is the step size associated with the iterative update of the active contour and κ is the weighting coefficients associated with the derivatives of the external force terms (Eq. (17) in [112]). w_l, w_e and w_t are the weighting coefficients associated with the image, edge and terminal energy functionals, respectively [112]. By identifying the optimal sets of these parameters, one may precisely determine the curvature of the spine.

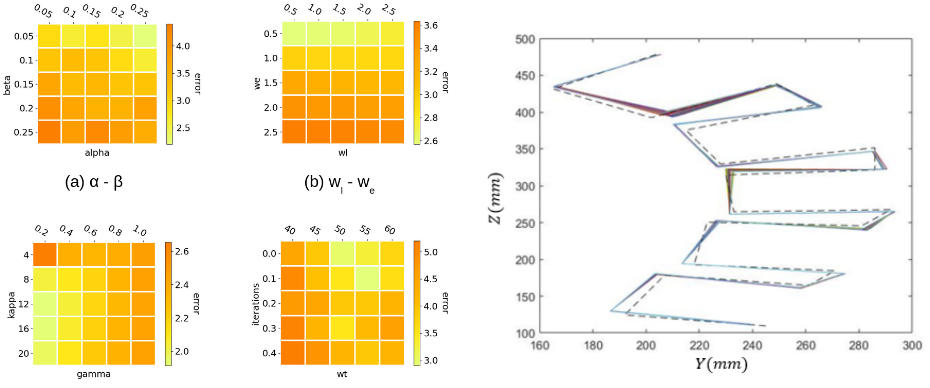
A sensitivity analysis [245, 242] is performed on each pair of fitting parameters, namely α-β,γ-κ, w_l-w_e and w_t-iterations, as illustrated in Figure 4.4 (a)-(d), to provide a better estimate on the optimized parameters. Here, for a particular pair of fitting parameters, the segmentation result is compared with the ground truth result and the average value of the normalized mean squared error is calculated for all the images in the dataset. The normalized mean squared error is given as

$$\frac{1}{\|N_s\|.\|N_v\|.\|N_\ell\|}\sum_{s=1}^{N_s}\sum_{v=1}^{N_v}\sum_{\ell=1}^{N_\ell}\left((x^s_{v,\ell}-\bar{x}^s_{v,\ell})^2+(y^s_{v,\ell}-\bar{y}^s_{v,\ell})^2\right) \tag{4.3}$$

where N_s, N_v and N_l are the number of images, number of vertebra and number of landmarks associated with each vertebrae. $x^s_{v,l}$ and $y^s_{v,l}$ are the positions of landmarks obtained by the current approach whereas $\bar{x}^s_{v,l}$ and $\bar{y}^s_{v,l}$ are the positions of the landmarks obtained from ground truth result.

To obtain the optimized parameters throughout the dataset, each pairwise parameter with the least normalized mean squared error is the initial value in the Simplex optimization. The final set of optimized parameters are used to segment the images, as shown in Figure 4.4 (e). We observe that the method is not very sensitive to the parameter values, which is desirable for practical applications, since there is no need for fine tuning of parameters in order to achieve satisfactory results.

Figure 4.5 shows the segmented spine in AP view, taking into account both the geometric and fitting parameters. First, the reference points of the T1 and L5 vertebrae will be identified and then the program will identify the contour around each vertebra.



(a) α - β (b) w_l - w_e

(c) γ - κ (d) w_t - iteration (e)

Figure 4.4: (a)-(d) Sensitivity analysis on the *fitting parameters* of the snakes algorithm to be provided as initial values for multi-variable optimization. (e) Optimized multi-variables of the snakes method for the lumbar spine. The dashed line denotes the ground truth manually picked four corner points of the vertebrae and solid lines are the optimized configuration. The normalized mean square error is 0.0187 corresponding to $\alpha = 0.25$, $\beta = 0.25$, $\gamma = 12$, $\kappa = 0.2$, $w_l = 0.5$, $w_e = 0.5$, $w_t = 0$, and $iteration = 50$.

4.3.2. 3D SHAPE RECONSTRUCTION FROM BI-PLANAR 2D DATA

Serial X-ray images from two orthogonal perspectives known as Anteroposterior (AP) and lateral (LAT) are used in this study to gather patient-specific data (Figure 4.3). Patient-specific characteristics are retrieved from X-ray images to assess spinal deformity prognosis. After performing segmentation and obtaining the contours from AP and LAT images, we perform 3D reconstruction from the 2D segmented contours in order to get the volumetric spine geometry. The different coordinate systems associated with the AP and LAT images present a challenge that needs to be addressed. Furthermore, it is challenging to keep track of the various coordinate systems of these images that have been longitudinally taken over time.

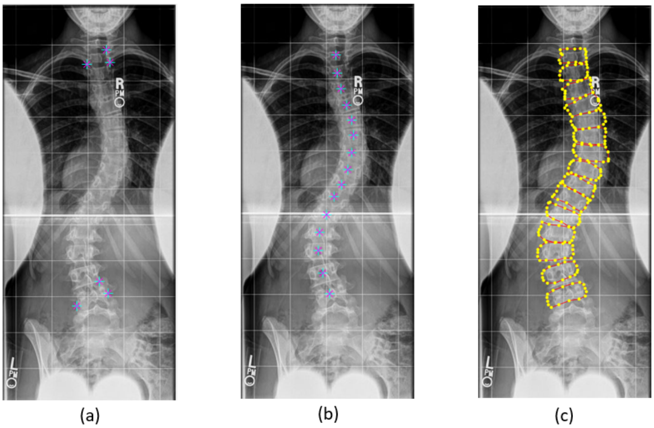


Figure 4.5: Three steps for image segmentation. (a) Picking the reference points of each vertebra, (b) picking the center points of all vertebrae and (c) segmented data by implementing optimized geometric and calibrating parameters using the snakes algorithm.

The initial step should be to calibrate the camera location in order to make all data consistent and comparable in scale. Here are the assumptions for the camera calibration step:

- *Reference of the coordinate system*: The coordinate system's reference point in AP and LAT perspectives is the center point of L5 (5^{th} lumbar vertebra).
- *Angle between AP and LAT views*: the angle between AP and LAT views is 90°.
- *Scaling criterion:* The images are taken with different scales. Images need to be scaled such that the heights of the spine in these two X-ray views are the same. It is assumed that AP view is fixed and the LAT view will be calibrated accordingly. The scaling factor between two pairs of images is calculated as

$$s = \frac{z_{max}^{AP} - z_{min}^{AP}}{z_{max}^{LAT} - z_{min}^{LAT}}, \tag{4.4}$$

where z_{max}^{AP} and z_{min}^{AP} are the maximum and minimum of z coordinates of landmarks in AP, respectively. Similarly, z_{max}^{LAT} and z_{min}^{LAT} are the maximum and minimum of z coordinates of landmarks in LAT, respectively.

As seen in Figure 4.6, the scaled data is subsequently fed into a 3D reconstruction procedure. The size of the bounding box for each vertebra is calculated by identifying the minimum and maximum in each direction, and the nodes in 3D space corresponding to that region are identified. The nodes on the boundary will also be detected using the MATLAB *convhull* function [166](Convex hull of the $3D$ reconstructed region). The $3D$ reconstruction geometry will be obtained as well as the boundary points for every vertebrae by performing the described technique and utilizing the scaled landmarks received from the X-ray images (Figure 4.7).

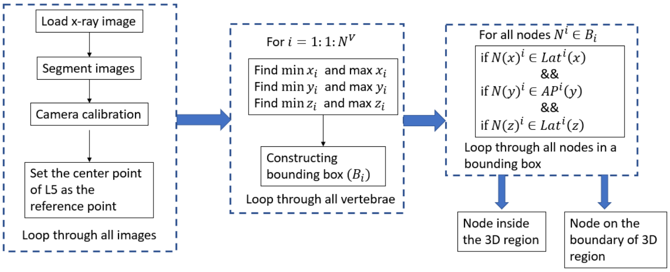


Figure 4.6: Flowchart of 3D reconstruction of vertebrae using 2D data.

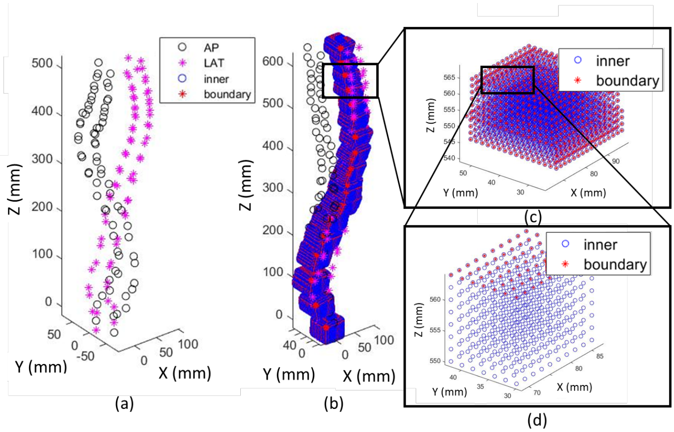
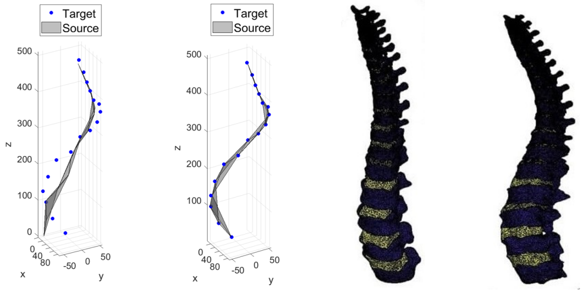


Figure 4.7: 3D reconstruction of vertebrae using 2D data. The camera parameters are calibrated before reconstruction based on the explanation in 4.3.2. (a) The segmented data of the AP and LAT views, (b) the 3D reconstructed geometry shaded in blue (obtained from the multiple inner volume blue points) using the bounding box algorithm; the magenta points denote the boundary (outer) surface and the blue points denote the inner volume, (c) and (d) zoomed-in views of the inner and boundary points of the third thoracic vertebra (T3).

4.3.3. VOLUMETRIC MESH GENERATION

The 3D reconstructed shape shown in Figure 4.7 cannot be directly used in to generate 3D volumetric mesh of an estimated spine from the 2D views, we use a twisting registration for the spine on the source mesh introduced in [256]. In our previous work [257], we used free-form deformation based on truncated hierarchical B-splines (THB-splines) for registration [191, 192, 193]. During registration, the detailed source geometry is taken from the atlas 3D geometry model which includes 128,205 vertices as explained in [257]. The 3D volumetric mesh composes of three components: intervertebral discs, lumbar vertebrae and thoracic vertebrae. The volumetric mesh of each vertebra is then segmented into growth plate, cortical bone and cancellous bone regions. Similarly, the volumetric mesh of each intervertebral disc is partitioned into Annulus Fibrosus and Nucleous Pulposus regions.

The employed registration method is based on the Iterative Closest Point (ICP) technique [13].In this method, we use the central points of each estimated vertebra and 8



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

Figure 4.8: Volumetric mesh registration on a point cloud including reconstructed landmarks from X-ray images. Before registration: (a) center line, (c) volumetric mesh; After registration: (b) center line, (d) volumetric mesh.

corner points of the corresponding bounding box as the target points. Then we define the point cloud as a surface and perform the nonrigid registration method with high stiffness ratio [251] , to register the mentioned corresponded points from source mesh (the Atlas 3D geometry in [257]) to the target mesh (generated in Section 4.3.2). The stiffness term prohibits the unfavourable rotation and causes in logical twisting instead. After finding the limited number of the transformation matrix as outcome of the registration method (17 (vertebra) × 8 (bounding box) + 17 (central of vertebra) = 153 transformation matrices in total), each point on the source mesh among the 128,205 vertices will be transformed based on the transfer matrix of each of the 153 points which is the closest to the vertex before registration. As shown in Figure 4.8 (a-b), the central points of the vertebra on the target spine and the source spine are depicted before and after registration respectively. Using the found transformation matrix of these 153 points, the complete volumetric spine is deformed from Figure 4.8 (c) to Figure 4.8 (d) which matches all the 8 landmarks of each vertebra.

Practically, the registration method allows us to generate a time-series mesh morphing of a patient's spine using the output of Section 4.3.2. Where, the X-ray images are serial images taken at different time steps from the spine. Through the approach introduced in Section 4.3.2, we have a low-quality time-series 3D point clouds of the spine. Here, we use each of the reconstructed spines in each time interval as a template for the mesh registration method explained in Chapter 2, which gives us a mesh morphing 4D (3D + time) [250, 253] data of the spine.

4.4. BIO-INFORMED MECHANISTIC MACHINE LEARNING

In this section, we propose our bio-informed mechanistic machine learning model that incorporates clinical data as well as mechanistic knowledge such as bone growth model and finite element results. We propose three machine learning models for prognosis of PSD: 3D-Clinical Neural Network; Bio-informed Clinical Neural Network; and Bio-informed Mechanistic Neural Network. The Bio-informed Clinical Neural Network is a stand-alone model that utilizes only the clinical data (i.e., X-ray images). Each of the latter two models is divided into two parts where both of them share the first part, called Center Point Predictor Neural Network. The Center Point Predictor Neural Network returns the center point of each vertebra and this output is fed into the second part of the respective model.

The major difference between the Bio-informed Clinical Neural Network and the Bio-informed Mechanistic Neural Network is the loss function of the second part. The former uses the bone growth parameters in the loss function whereas the latter takes both geometric features (position of landmarks and center points) and bone growth parameters into the loss function. This will be explained in detail in the following Section 4.3.

4.4.1. INTRODUCTION TO BONE GROWTH MODEL

One way to address the scarcity of patient-specific data is to incorporate the governing physical equation since it provides insights into the expected data. There are several aspects to consider when employing governing physical equations. The model must be validated over a large population, and some patient-specific time dependent constants must be calibrated for each individual patient at each time step if these models are used in computational simulations as well as ML frameworks.

The underlying physical equation that governs the progression of the spine is the bone growth model, because the AIS occurs during adolescence when bone growth is at its peak [47]. It is believed that bone formation is a consequence of the stress imposed on each vertebra's growth plate [234] located on the top and bottom of each vertebra. This assumption, however, may not be adequate to correctly depict bone growth, since other factors such as age, sex, and bone mineral density also affect bone formation. In the presented approach, we describe bone growth as follows:

$$\boldsymbol{G} = \boldsymbol{A} + \boldsymbol{B}\bar{\sigma} \tag{4.5}$$

where $\boldsymbol{G} = [G_X, G_Y, G_Z]$ are the growth rates along three normal directions and $\bar{\sigma}$ is the von Mises stress. There are two advantages of implementing von Mises stress. First, all stress components are taken into account and second, the definition of the stress is independent of the coordinate system. Vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ are patient-specific parameters that should be calibrated for each patient, where $\boldsymbol{A} = [A_X, A_Y, A_Z]$ are considered as growth rate for non-scoliotic spine and $\boldsymbol{B} = [B_X, B_Y, B_Z]$ are the regulating growth model parameters.

4.4.2. MODIFIED GROWTH MODEL

We aim to make the most use of all available data by exploiting medical features acquired from X-ray data, mechanistic features extracted from computational models, and addressing data scarcity while leveraging the underlying physical equation. The objective of the bio-informed mechanistic machine learning method is to reconstruct the curvature of the spine throughout time. However, the calibrated parameters generated by the system may be implemented into a computational model for monitoring and assessing bone formation on a local level.

FRAMEWORK EXPLANATION AND NOTATION

The parameters of the physical growth equation are *patient-specific, time-dependent, and position-dependent*. Thus, Eq. (4.5) for any arbitrary landmarks on the growth plates for time t and vertebra v can be rewritten as

$$\boldsymbol{G^{tv}} = \boldsymbol{A^{tv}} + \boldsymbol{B^{tv}}\bar{\sigma}^{tv} \tag{4.6}$$

where $\boldsymbol{A^{tv}} = [A_X^{tv}, A_Y^{tv}, A_Z^{tv}]$ and $\boldsymbol{B^{tv}} = [B_X^{tv}, B_Y^{tv}, B_Z^{tv}]$ are patient-specific time dependent, position dependent parameters that should be calibrated for each patient at a given time to calculate growth rates $\boldsymbol{G^{tv}} = [G_X^{tv}, G_Y^{tv}, G_Z^{tv}]$ using the von Mises stress at time t for vertebra v using the von Mises stress field from the surrogate finite element model explained in [257]. The loads, boundary conditions and material properties are the function of time t and is updated accordingly based on the age of the patient. The finite element model's output will contain the von Mises stress that corresponds to the growth landmarks' position. Indices X, Y and Z correspond to the global coordinate system shown in Figure 4.7. In plain terms, the modified growth equation describes the directional growth of a specific vertebra of a particular patient over a span of time. The spatiotemporal parameters of the equation track the co-ordinates of the vertebra over the time. One can imagine a bounding box around a vertebra that deforms in three dimensions. The growth equation tracks the three orthonormal components of the growth through the parameters. Hence, once solved, the modified growth equation can specify the later position of the bounding box.

Local and global coordinate systems: To obtain the growth model parameters that correspond to the growth deformation of the vertebrae without taking into account the rigid body motion caused by the movement of the bottom vertebrae, a local coordinate system is set up at the center point of each vertebra. The axes of the local coordinates have the same direction as the axes of the global coordinate system (as mentioned in Section 3, the center of global coordinate system is at the center point of $L5$ (the fifth lumbar vertebra)). Figure 4.9 (a) shows the global coordinate systems corresponding to the time t and the local coordinate system for the vertebra $L3$ at the same time (Figure 4.9 (b)). The landmarks that surround each vertebra are divided into two categories: those that are positioned on the growth plates (X_G) and those that dictate the vertebra's side shape (X_S) as shown in Figure 4.9 (c).

Growth parameters: The growth landmarks with comparable motion directions (due to growth) within the same vertebra should be analyzed together to represent the growth deformation of X_G landmarks. In the first step, nodes that behave similarly in each direction will be grouped together, as indicated in Table 4.2.

Table 4.2: Landmarks with the same growth behavior are clustered into the same group. The numbers are clarified in Figure 4.9 (d).

Local Growth Direction	Clustered Landmarks
x	$F = [1,2,3,4]$ $B = [5,6,7,8]$
y	$R = [1,3,5,7]$ $L = [4,2,6,8]$
z	$U = [1,2,5,6]$ $D = [3,4,7,8]$

In the y-axis, for example, landmarks 1, 3, 5 and 7 have the same growth order (positive local y), whereas landmarks 3, 4, 7 and 8 have the same growth direction in the z-axis

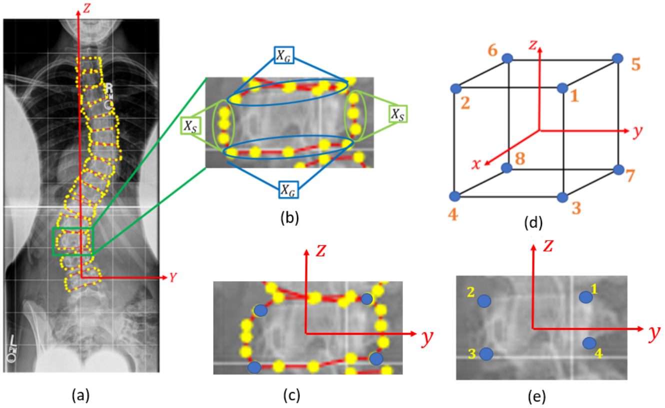


Figure 4.9: Description of the coordinate systems and landmarks. (a) Global coordinate system in AP view, (b) labeling landmarks to growth (X_G) and side (X_S), (c) local coordinate system of the vertebra $L3$ in AP view (the yellow dots are the 16 landmarks and the blue dots are the 4 corner landmarks), (d) labeling of the growth landmarks on vertebrae in local coordinate system in 3D view and (e) labeling of the growth landmarks on the $L3$ in local coordinate system in 2D view using the 4 corner points.

(negative local z) as visualized in Figure 4.9. The growth equation of each landmark (j) in local coordinates (x, y, z) for time (t) and vertebra (v) can be written as:

$$\boldsymbol{G}_{\boldsymbol{j}}^{\boldsymbol{tv}} = \boldsymbol{A}_{\boldsymbol{j}}^{\boldsymbol{tv}} + \boldsymbol{B}_{\boldsymbol{j}}^{\boldsymbol{tv}} \bar{\sigma}_j^{tv} \tag{4.7}$$

where $\boldsymbol{A}_{\boldsymbol{j}}^{\boldsymbol{tv}} = [A_{xj}^{tv}, A_{yj}^{tv}, A_{zj}^{tv}]$ and $\boldsymbol{B}_{\boldsymbol{j}}^{\boldsymbol{tv}} = [B_{xj}^{tv}, B_{yj}^{tv}, B_{zj}^{tv}]$ are patient-specific time dependent, position dependent parameters for landmark (j) that should be calibrated for each patient at a given time for every vertebra. Indices x, y and z represent the three directions of local coordinate system shown in Figure 4.9. It should be noted that $\bar{\sigma}_j^{tv}$ is the von-Mises stress on landmark j which is independent of the coordinate system.

Calculating growth parameters

Nodes with comparable growth behavior have the same growth parameters, as listed in Table 4.2. These parameters are considered to be the same for each vertebra. The first step in determining growth parameters for each vertebra v is to create growth equations for all nodes as presented in Eq. (4.7).

The equations are rewritten to represent the clustered label in each direction, using the same growth parameter assumption as mentioned before as

$$G_{xj}^{tv} = \begin{cases} A_{xF}^{tv} + B_{xF}^{tv}\bar{\sigma}_j^{tv} & x_j^{tv} < 0, \\ A_{xB}^{tv} + B_{xB}^{tv}\bar{\sigma}_j^{tv} & x_j^{tv} \geq 0, \end{cases} \tag{4.8}$$

$$G_{yj}^{tv} = \begin{cases} A_{yR}^{tv} + B_{yR}^{tv}\bar{\sigma}_j^{tv} & y_j^{tv} < 0, \\ A_{yL}^{tv} + B_{yL}^{tv}\bar{\sigma}_j^{tv} & y_j^{tv} \geq 0, \end{cases} \tag{4.9}$$

$$G_{zj}^{tv} = \begin{cases} A_{zU}^{tv} + B_{zU}^{tv}\bar{\sigma}_j^{tv} & z_j^{tv} < 0, \\ A_{zD}^{tv} + B_{zD}^{tv}\bar{\sigma}_j^{tv} & z_j^{tv} \geq 0 \end{cases} \tag{4.10}$$

where x_j^{tv}, y_j^{tv} and z_j^{tv} are the coordinates in the local system of landmark j for the vertebra v at time t. The vectors $\mathscr{A}^{tv}$ and $\mathscr{B}^{tv}$ for a vertebra v at time t will be defined as

$$\begin{aligned} \mathscr{A}^{tv} &= [A_F^{tv}, A_B^{tv}, A_R^{tv}, A_L^{tv}, A_U^{tv}, A_D^{tv}], \\ \mathscr{B}^{tv} &= [B_F^{tv}, B_B^{tv}, B_R^{tv}, B_L^{tv}, B_U^{tv}, B_D^{tv}], \end{aligned} \tag{4.11}$$

to define the patient-specific, time-dependent, and position-dependent scalar value of growth parameters for each vertebra. To discretize the growth parameters in each direction, the vectors

$$\begin{cases} \mathscr{A}_x^{tv} = [A_F^{tv}, A_F^{tv}, A_F^{tv}, A_F^{tv}, A_B^{tv}, A_B^{tv}, A_B^{tv}, A_B^{tv}], \\ \mathscr{A}_y^{tv} = [A_R^{tv}, A_L^{tv}, A_R^{tv}, A_L^{tv}, A_R^{tv}, A_L^{tv}, A_R^{tv}, A_L^{tv}], \\ \mathscr{A}_z^{tv} = [A_U^{tv}, A_U^{tv}, A_D^{tv}, A_D^{tv}, A_U^{tv}, A_U^{tv}, A_D^{tv}, A_D^{tv}], \end{cases}$$

$$\begin{cases} \mathscr{B}_x^{tv} = [B_F^{tv}, B_F^{tv}, B_F^{tv}, B_F^{tv}, B_B^{tv}, B_B^{tv}, B_B^{tv}, B_B^{tv}], \\ \mathscr{B}_y^{tv} = [B_R^{tv}, B_L^{tv}, B_R^{tv}, B_L^{tv}, B_R^{tv}, B_L^{tv}, B_R^{tv}, B_L^{tv}], \\ \mathscr{B}_z^{tv} = [B_U^{tv}, B_U^{tv}, B_D^{tv}, B_D^{tv}, B_U^{tv}, B_U^{tv}, B_D^{tv}, B_D^{tv}], \end{cases} \tag{4.12}$$

are defined corresponding to the growth parameters in each direction for all the growth landmarks. In the presented frame work, the number of growth landmarks is $X_G = 8$, corresponding to 8 corner points shown in Figure 4.9. The post-processed data of serial X-ray images (described in Section 4.3.2) are used to directly assess the growth parameters by monitoring the locations of the growth landmarks in the two consecutive images. To represent the matrix form of growth parameters for each vertebra, matrices $\mathbb{A}$ and $\mathbb{B}$ are developed as

$$\mathbb{A}^{tv} = [[\mathscr{A}_x^{tv}]^T \ [\mathscr{A}_y^{tv}]^T \ [\mathscr{A}_z^{tv}]^T], \quad \text{and} \quad \mathbb{B}^{tv} = [[\mathscr{B}_x^{tv}]^T \ [\mathscr{B}_y^{tv}]^T \ [\mathscr{B}_z^{tv}]^T]. \tag{4.13}$$

The dimension of $\mathbb{A}^{tv}$ and $\mathbb{B}^{tv}$ are 8×3 corresponding to 8 growth landmarks and 3 local coordinate directions (x, y and z).

4.4.3. SPINAL DEFORMITY PROGNOSIS FRAMEWORK

A bio-informed machine learning framework is introduced by incorporating our knowledge from clinical data (X-ray images) and the modified bone growth computational model explained in section 4.4.2. The landmark position X is related to the features I

at a given time by an unknown bio-physics equation which can be generically defined as a mapping. Note that I is referring to general features for now. It will be specified in the following sections depending on the choice of the model. The overall structure of a neural network can also be described as a mapping, i.e.:

$$\begin{cases} Unknown\ bio-physics\ equation:\ \boldsymbol{X} = F_{Unknown-BioPhy}(\boldsymbol{I}) \\ Neural\ network\ mapping:\ \boldsymbol{X} = F_{FFNN}(\boldsymbol{I}) \end{cases} \tag{4.14}$$

where F_{FFNN} is the feed forward neural network (FFNN) that uses clinical and mechanistic features I as input, and generates landmark coordinates X as the output. The notation used throughout this section is defined in Table 4.3. To predict the position of a landmark, the input neurons would be clinical and mechanistic features, the hidden neurons would act as a multiplicative, functional decomposition of the unknown bio-physics equation that estimates the unknown function required to map input features to coordinates in the output neurons. In this project, MATLAB is used to build the FFNN and to train the neural network parameters [167].

Each neuron within every layer of a generic FFNN receives the output value from each neuron in the previous layer as input and produces a single output. This procedure is carried out for each layer. For an arbitrary number of layers and neurons per layer, the value of the j^{th} neuron in layer l for the s^{th} sample (either a training sample or prediction) may be written as:

$$a_{j,s}^{l} = \begin{cases} I_{j,s}, & \text{if } l = 1 \text{ (input layer)} \\ A(b_j^l + \sum_{i=1}^{N_N(l-1)} W_{ji}^{l-1} a_{i,s}^{l-1}), & \text{if } 1 < l < N_L \text{ (hidden layers)} \\ b_j^l + \sum_{i=1}^{N_N(l-1)} W_{ji}^{l-1} a_{i,s}^{l-1}, & \text{if } l = N_L \text{ (output layer)} \end{cases} \tag{4.15}$$

where A is an activation function. In the training part, this research uses *ReLU* (Rectified Linear Unit) function defined as the positive part of its argument, [4]: $f(x) = x^+ = max(0, x)$, and each neuron is computed using a different weight $W_{ij}^{l=1}$ and bias $b_j^{l=2}$, where i is the neuron in the previous layer. Finally, the overall response – the predicted landmark coordinates – is given by:

$$X_{j,s}^{\text{predicted}} = a_{j,s}^{N_L}. \tag{4.16}$$

The FFNN can learn the unknown governing bio-physics equation based on the loss function and input and output features. The loss function may be constructed to train the FFNN to offer accurate estimates while also supplying patient-specific parameters that can be considered as the network's hyperparameter. Multiple FFNN approaches are presented in the next section, each with a distinct loss function and input-output features.

MACHINE LEARNING FRAMEWORK EXPLANATION

3D-Clinical Neural Network ($FFNN_{CL}$): Given the coordinates of a vertebra's landmarks at time t, the clinical neural network attempts to predict the vertebra's landmark coordinates at time $t + \Delta t$; see Figure 4.10. The input vectors are $\boldsymbol{I}_s^t = [\bar{\boldsymbol{X}}_s^t, \boldsymbol{\alpha}_s^t, t, \Delta t]$, where

Table 4.3: Notation table of variables used in the feed forward neural network.

t	Counting index for number of time steps
$\boldsymbol{I_s^t}$	Vector of input , $s = 1, ..., N_T$
$I_{j,s}$	jth entry of I vector of input , $s = 1, ..., N_T$
$\boldsymbol{O_s^t}$	Vector of output , $s = 1, ..., N_T$
s	Counting index for number of samples (training or validation, depending on context)
m	Counting index for number of landmarks on each vertebra
N_m	Number of landmarks on each vertebrae
l	Counting index for number of layers
i	Counting index for neurons in a given layer
j	Counting index for neurons in another layer
N_T	Number of training samples
N_L	Number of layers in the neural network
$N_N(l)$	Number of neurons in layer l
$a_{j,s}^l$	Neuron value for j^{th} neuron in l^{th} layer and for s^{th} sample
W_{ji}^l	Weight connecting the i^{th} neuron in layer l to the j^{th} neuron in layer $l+1$
b_j^l	Bias of the j^{th} neuron in layer l
A	Activation function
F_{FFNN}	Feedforward neural network function
$\mathscr{H}$	The height of the spine

$\bar{\boldsymbol{X}}_s^t$ indicates the landmark coordinates of sample s at time t and $\boldsymbol{\alpha}_s^t$ denotes the global angles (described in [257]) characterizing the 3D shapes of the spine at that time. The outputs are vectors $\boldsymbol{O}_s^t = [\boldsymbol{X}_s^{t+\Delta t}]$, where $\boldsymbol{X}_s^{t+\Delta t}$ denotes the expected 3D coordinates of the vertebra's corner points at time $t + \Delta t$.

The relative approximation error for this model is calculated as

$$\delta_{CL} = \frac{1}{\mathscr{H}.N_T.N_m} \sum_{s=1}^{N_T} \sum_{m=1}^{N_m} \left\| \boldsymbol{X}_{sm}^{t+\Delta t} - \bar{\boldsymbol{X}}_{sm}^{t+\Delta t} \right\|^2, \tag{4.18}$$

where δ_{CL} denotes the relative error of the clinical neural network, $\boldsymbol{X}_{sm}^{t+\Delta t}$ is the coordinates of the landmark m of vertebra s predicted by the neural network, and $\bar{\boldsymbol{X}}_{sm}^{t+\Delta t}$ is from clinical dataset.

3D-Clinical neural network (FFNN$_{\text{CL}}$)**:**

Input: $\boldsymbol{I}_s^t = [\bar{\boldsymbol{X}}_s^t, \boldsymbol{\alpha}_s^t, t_0, t]$.

Output: $\boldsymbol{O}_s^t = [\boldsymbol{X}_s^{t+\Delta t}]$.

Find $W_{ij}^l (l \in \{1,2,3\}), b_j^l (l \in \{2,3,4\})$ for each patient to minimize

$$loss_{CL}^t = \frac{1}{N_T} \frac{1}{N_m} \sum_{s=1}^{N_T} \sum_{m=1}^{N_m} \left\| \boldsymbol{X}_s^{t+\Delta t} - \bar{\boldsymbol{X}}_s^{t+\Delta t} \right\|^2, \tag{4.17}$$

$$\text{where} \quad \boldsymbol{O}_s^t = [\boldsymbol{X}_s^{t+\Delta t}] = \left[b_q^{l=4} + \textstyle\sum_{k=1}^{N_N[l=3]} W_{qk}^{l=3} A\left(a_{k,s}^{l=3}\right) \right]_{q=1}^{N_N[l=4]},$$

$$a_{k,s}^{l=3} = b_k^{l=3} + \textstyle\sum_{j=1}^{N_N[l=2]} W_{kj}^{l=2} A\left(a_{j,s}^{l=2}\right),$$

$$a_{j,s}^{l=2} = b_j^{l=2} + \textstyle\sum_{i=1}^{N_N[l=1]} W_{ji}^{l=1} I_{i,s}^t,$$

$\bar{\boldsymbol{X}}_s^{t+\Delta t}$ is from clinical dataset.

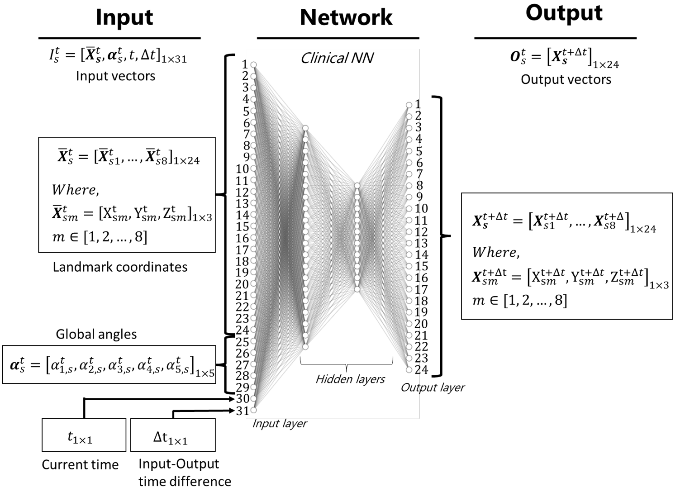


Figure 4.10: The structure of the 3D-clinical feed-forward neural network $FFNN_{CL}$ developed to predict the coordinates of the landmarks at time $t+\Delta t$.

Center Point Predictor Neural Network ($FFNN_{CR}$)**:** Given the set of landmarks' coordinates $\bar{\boldsymbol{X}}_s^t$, global angles $\boldsymbol{\alpha}_s^t$, and von Mises stress $\bar{\boldsymbol{\sigma}}_s^t$ at landmarks at time t, $FFNN_{CR}$ predicts the coordinates of the center of the sample at time $t+\Delta t$; see Figure 4.11. More formally, the input to $FFNN_{CR}$ is $\boldsymbol{I}_s^t = [\bar{\boldsymbol{X}}_s^t, \boldsymbol{\alpha}_s^t, t, \Delta t, \bar{\boldsymbol{\sigma}}_s^t]$ and its output vector is $\boldsymbol{O}_{CR,s}^t = [\boldsymbol{C}_s^{t+\Delta t}]$, where $\boldsymbol{C}_s^{t+\Delta t}$ denotes the center coordinates of the sample s at time $t+\Delta t$. The results of this network are used in the following neural networks.

Bio-informed Clinical Neural Network ($FFNN_{BC}$)**:** Using the clinical data to predict bone growth parameters, $FFNN_{BC}$ predicts the parameters of the physical growth equation given the set of landmark coordinates, global angles, and von Mises stress at the landmarks; see Figure 4.12. Input vectors are $\boldsymbol{I}_s^t = [\bar{\boldsymbol{X}}_s^t, \boldsymbol{\alpha}_s^t, t, \Delta t, \bar{\boldsymbol{\sigma}}_s^t]$, where $\bar{\boldsymbol{X}}_s^t$ is the landmarks' coordinates at time t, $\boldsymbol{\alpha}_s^t$ is the set of global angles, and $\bar{\boldsymbol{\sigma}}_s^t$ is the set of von Mises stress. The outputs of the network are vectors $\boldsymbol{O}_s^t = [\mathscr{A}_s^{t+\Delta t}, \mathscr{B}_s^{t+\Delta t}]$ predicting the growth model parameters in $3D$ for each landmark. Combining the results of $FFNN_{BC}$ with the center points resulted from $FFNN_{CR}$, we are able to predict the coordinates of the landmarks of the sample at time $t+\Delta t$.

For any sample s and any landmark m on s, define the landmark $\boldsymbol{X}_{BC}^{t+\Delta t}$ predicted by $FFNN_{BC}$ as

$$\boldsymbol{X}_{BC}^{t+\Delta t} = \bar{\boldsymbol{X}}_{BC}^{t} + \Delta \boldsymbol{C}_s^t + (\mathbb{A}_{BC}^{t+\Delta t} + \mathbb{B}_{BC}^{t+\Delta t} \sigma_{BC}^t)\Delta t, \tag{4.21}$$

where $\mathbb{A}_{BC}^{t+\Delta t}$ and $\mathbb{B}_{BC}^{t+\Delta t}$ are obtained from $\mathscr{A}_s^{t+\Delta t}$ and $\mathscr{B}_s^{t+\Delta t}$, which are the outputs of the second model, $\bar{\boldsymbol{X}}_{BC}^t$ is the coordinates of landmark m of sample s at time t obtained from clinical dataset, and $\Delta \boldsymbol{C}_s^t$ is the change in the center of sample s from time t to time $t+\Delta t$, obtained from the results of the first neural network. By this definition, this

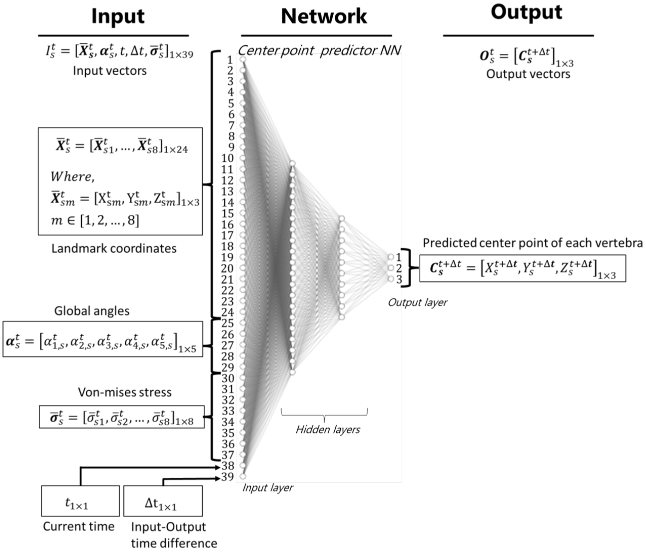


Figure 4.11: The structure of the center point prediction feed-forward neural network $FFNN_{CR}$ developed to predict the location of the center point of the vertebra at time $t + \Delta t$.

model's relative approximation error, δ_{BC}, is calculated as

$$\delta_{BC} = \frac{1}{\mathscr{H}.N_T.N_m} \sum_{s=1}^{N_T} \sum_{m=1}^{N_m} \left\| \boldsymbol{X}_{sm}^{t+\Delta t} - \bar{\boldsymbol{X}}_{sm}^{t+\Delta t} \right\|. \tag{4.22}$$

Bio-informed Mechanistic Neural Network ($FFNN_{BM}$): Similar to $FFNN_{BC}$, in the first step of this model, $FFNN_{CR}$ predicts the coordinates of the center point of the vertebra at time $t + \Delta t$, followed by computing the transition vector from the center point of the sample from time t to $t + \Delta t$. The neural network $FFNN_{BM}$ then predicts the set of growth model parameters as an output using a mechanistic loss function; see Figure 4.12. The input vectors are $\boldsymbol{I}_s^t = [\bar{\boldsymbol{X}}_s^t, \boldsymbol{\alpha}_s^t, t, \Delta t, \bar{\boldsymbol{\sigma}}_s^t]$ with the same definitions for $\bar{\boldsymbol{X}}_s^t$, $\boldsymbol{\alpha}_s^t$, and $\bar{\boldsymbol{\sigma}}_s^t$. The output of $FFNN_{CR}$ is $\boldsymbol{O}_{CR,s}^t = [\boldsymbol{C}_s^{t+\Delta t}]$ which remains the same and the output vectors to the second step is $\boldsymbol{O}_s^t = [\mathscr{A}_s^{t+\Delta t}, \mathscr{B}_s^{t+\Delta t}]$, with $\mathscr{A}_s^{t+\Delta t}$ and $\mathscr{B}_s^{t+\Delta t}$ representing the parameters of the physical growth equation.

Similar to the $FFNN_{BC}$, we define the predicted landmark m of sample s, for each m and s, as

$$\boldsymbol{X}_{BM}^{t+\Delta t} = \bar{\boldsymbol{X}}_{BM}^t + \Delta \boldsymbol{C}_s^t + (\mathbb{A}_{BM}^{t+\Delta t} + \mathbb{B}_{BM}^{t+\Delta t} \sigma_{BM}^t) \Delta t. \tag{4.25}$$

The definitions are identical to those stated in the $FFNN_{BC}$ formulation. The relative approximation error for $FFNN_{BM}$ model, represented by δ_{BM}, is calculated using

$$\delta_{BM} = \frac{1}{\mathscr{H}.N_T.N_m} \sum_{s=1}^{N_T} \sum_{m=1}^{N_m} \| \boldsymbol{X}_{BM}^{t+\Delta t} - \bar{\boldsymbol{X}}_{BM}^{t+\Delta t} \|^2. \tag{4.26}$$

Bio-informed-Clinical neural network:

Step 1 (FFNN$_{\mathrm{CR}}$)*:*

Input: $\boldsymbol{I}_s^t = [\bar{\boldsymbol{X}}_s^t, \boldsymbol{\alpha}_s^t, t, \Delta t, \bar{\boldsymbol{\sigma}}_s^t]$.

Output: $\boldsymbol{O}_{CR,s}^t = [\boldsymbol{C}_s^{t+\Delta t}]$.

Find $W_{ij}^l (l \in \{1,2,3\}), b_j^l (l \in \{2,3,4\})$ for each patient to minimize

$$loss_{CR}^t = \frac{1}{N_T} \sum_{s=1}^{N_T} \left\| \boldsymbol{C}_s^{t+\Delta t} - \bar{\boldsymbol{C}}_s^{t+\Delta t} \right\|^2, \tag{4.19}$$

where $\boldsymbol{O}_{CR,s}^t = [\boldsymbol{C}_s^{t+\Delta t}] = \left[b_q^{l=4} + \sum_{k=1}^{N_N[l=3]} W_{qk}^{l=3} A\left(a_{k,s}^{l=3}\right)\right]_{q=1}^{N_N[l=4]}$,

$a_{k,s}^{l=3} = b_k^{l=3} + \sum_{j=1}^{N_N[l=2]} W_{kj}^{l=2} A\left(a_{j,s}^{l=2}\right)$,

$a_{j,s}^{l=2} = b_j^{l=2} + \sum_{i=1}^{N_N[l=1]} W_{ji}^{l=1} I_{i,s}^t$,

$\bar{\boldsymbol{C}}_s^{t+\Delta t}$ is from clinical dataset.

Step 2 (FFNN$_{\mathrm{BC}}$)*:*

Input: $\boldsymbol{I}_s^t = [\bar{\boldsymbol{X}}_s^t, \boldsymbol{\alpha}_s^t, t, \Delta t, \bar{\boldsymbol{\sigma}}_s^t]$.

Output: $\boldsymbol{O}_s^t = [\mathscr{A}_s^{t+\Delta t}, \mathscr{B}_s^{t+\Delta t}]$

Find $W_{ij}^l (l \in \{1,2,3\}), b_j^l (l \in \{2,3,4\})$ for each patient to minimize

$$loss_{BC}^t = \frac{1}{N_T} \sum_{s=1}^{N_T} \left\| \mathscr{A}_s^{t+\Delta t} - \bar{\mathscr{A}}_s^{t+\Delta t} \right\|^2 + \left\| \mathscr{B}_s^{t+\Delta t} - \bar{\mathscr{B}}_s^{t+\Delta t} \right\|^2, \tag{4.20}$$

where $\boldsymbol{O}_s^t = [\mathscr{A}_s^{t+\Delta t}, \mathscr{B}_s^{t+\Delta t}] = \left[b_q^{l=4} + \sum_{k=1}^{N_N[l=3]} W_{qk}^{l=3} A\left(a_{k,s}^{l=3}\right)\right]_{q=1}^{N_N[l=4]}$,

$a_{k,s}^{l=3} = b_k^{l=3} + \sum_{j=1}^{N_N[l=2]} W_{kj}^{l=2} A\left(a_{j,s}^{l=2}\right)$,

$a_{j,s}^{l=2} = b_j^{l=2} + \sum_{i=1}^{N_N[l=1]} W_{ji}^{l=1} I_{i,s}^t$,

$\bar{\mathscr{A}}_s^{t+\Delta t}$ and $\bar{\mathscr{A}}_s^{t+\Delta t}$ are from clinical dataset.

IMPLEMENTING BIO-INFORMED MECHANISTIC MACHINE LEARNING FRAMEWORKS

Data collection and feature selection: The database is comprised of clinical X-ray images. The snakes algorithm is capable of generating an infinite number of landmarks around each vertebra using X-ray images. Landmarks are classified into two types: those located on growth plates (X_G) and those that determine the vertebral sides (X_S), as illustrated in Figure 4.9. The number of X_G landmarks on each 2D plane in this application is four, signifying the corner points that are expected to be on the growth plates. As a result, in the 3D framework presented ($X_G = 8$ and $X_S = 8$) as shown in Figure 4.13. At $t_0 = 124$ months, the patient's first X-ray image is obtained. The next four X-ray images are obtained at $t = 139, 149, 156, 168$ months and used to train the neural network. The remaining three X-ray images, collected at $t = 160, 179, 187$ months, are utilized to compare with the neural network findings as shown in Table 4.5. Each X-ray image depicts the form of the patient's spine at a specific age. Based on the explained machine learning framework explained in section 4.4.3.1 $FFNN_{CL}$, $FFNN_{CR-BC}$ and $FFNN_{CR-BM}$ are setup as explained in Table 4.4. The number of training samples, 68, corresponds to the 4 training datasets (Table 4.5), with each dataset containing 17 vertebrae. Depending on the NN methodology, each sample has a different size. The three test datasets listed in Table 4.5 are represented by the number of test samples ($17 \times 3 = 51$) in Table

Bio-informed-Mechanistic neural network:

Step 1 (FFNN$_{\mathrm{CR}}$)*:*

Input: $\boldsymbol{I}_s^t = [\bar{\boldsymbol{X}}_s^t, \boldsymbol{\alpha}_s^t, t, \Delta t, \bar{\boldsymbol{\sigma}}_s^t]$.

Output: $\boldsymbol{O}_{CR,s}^t = [\boldsymbol{C}_s^{t+\Delta t}]$.

Find $W_{ij}^l (l \in \{1,2,3\}), b_j^l (l \in \{2,3,4\})$ for each patient to minimize

$$loss_{CR}^t = \frac{1}{N_T} \sum_{s=1}^{N_T} \left\| \boldsymbol{C}_s^{t+\Delta t} - \bar{\boldsymbol{C}}_s^{t+\Delta t} \right\|^2, \tag{4.23}$$

where $\boldsymbol{O}_{CR,s}^t = [\boldsymbol{C}_s^{t+\Delta t}] = \left[b_q^{l=4} + \sum_{k=1}^{N_N[l=3]} W_{qk}^{l=3} A\left(a_{k,s}^{l=3}\right) \right]_{q=1}^{N_N[l=4]}$,

$a_{k,s}^{l=3} = b_k^{l=3} + \sum_{j=1}^{N_N[l=2]} W_{kj}^{l=2} A\left(a_{j,s}^{l=2}\right)$,

$a_{j,s}^{l=2} = b_j^{l=2} + \sum_{i=1}^{N_N[l=1]} W_{ji}^{l=1} I_{i,s}^t$,

$\bar{\boldsymbol{C}}_s^{t+\Delta t}$ is from clinical dataset.

Step 2 (FFNN$_{\mathrm{BM}}$)*:*

Input: $\boldsymbol{I}_s^t = [\bar{\boldsymbol{X}}_s^t, \boldsymbol{\alpha}_s^t, t, \Delta t, \bar{\boldsymbol{\sigma}}_s^t]$.

Output: $\boldsymbol{O}_s^t = [\mathscr{A}_s^{t+\Delta t}, \mathscr{B}_s^{t+\Delta t}]$.

Find $W_{ij}^l (l \in \{1,2,3\}), b_j^l (l \in \{2,3,4\})$ for each patient to minimize

$$loss_{BM}^t = \frac{1}{N_T} \frac{1}{N_m} \sum_{s=1}^{N_T} \sum_{m=1}^{N_m} \| (\bar{\boldsymbol{X}}_{sm}^t + \Delta \boldsymbol{C}_s^t + (\mathbb{A}_{sm}^{t+\Delta t} + \mathbb{B}_{sm}^{t+\Delta t} \bar{\sigma}_{sm}^t) \Delta t) - \bar{\boldsymbol{X}}_{sm}^{t+\Delta t} \|^2, \tag{4.24}$$

where $\boldsymbol{O}_s^t = [\mathscr{A}_s^{t+\Delta t}, \mathscr{B}_s^{t+\Delta t}] = \left[b_q^{l=4} + \sum_{k=1}^{N_N[l=3]} W_{qk}^{l=3} A\left(a_{k,s}^{l=3}\right) \right]_{q=1}^{N_N[l=4]}$,

$a_{k,s}^{l=3} = b_k^{l=3} + \sum_{j=1}^{N_N[l=2]} W_{kj}^{l=2} A\left(a_{j,s}^{l=2}\right)$,

$a_{j,s}^{l=2} = b_j^{l=2} + \sum_{i=1}^{N_N[l=1]} W_{ji}^{l=1} I_{i,s}^t$,

$\mathbb{A}_{sm}^{t+\Delta t}$ and $\mathbb{B}_{sm}^{t+\Delta t}$ are obtained from $\mathscr{A}_s^{t+\Delta t}$ and $\mathscr{B}_s^{t+\Delta t}$, and

$\bar{\boldsymbol{X}}_{sm}^t, \bar{\boldsymbol{X}}_{sm}^{t+\Delta t}$ are from clinical dataset and $\bar{\sigma}_{sm}^t$ is from surrogate model,

$\Delta \boldsymbol{C}_s^t = F_{\mathrm{FFNN_{CR}}}(\boldsymbol{I}_s^t, W_{ij}^l, b_j^l)$.

Table 4.4: Neural network setup for neural networks.

NN Components	3D-Clinical	Bio-inf. Clinical		Bio-inf. Mechanistic	
	FFNN$_{\mathrm{CL}}$	FFNN$_{\mathrm{CR}}$	FFNN$_{\mathrm{BC}}$	FFNN$_{\mathrm{CR}}$	FFNN$_{\mathrm{BM}}$
# training samples	68	68		68	
Size of training samples	31	39		39	
# test samples	51	51		51	
Size of test samples	24	3	12	3	12
# Hidden layers	2	2	2	2	2
Neurons in layer 1	20	20	20	20	20
Neurons in layer 2	10	10	10	10	10
Activation function	*ReLU* function				

4.4. It should be noted that the growth landmarks are tested for each framework, and the relative approximation error indicated in Table 4.6 is determined for all landmarks, including growth and side landmarks. The data is normalized before being fed to the

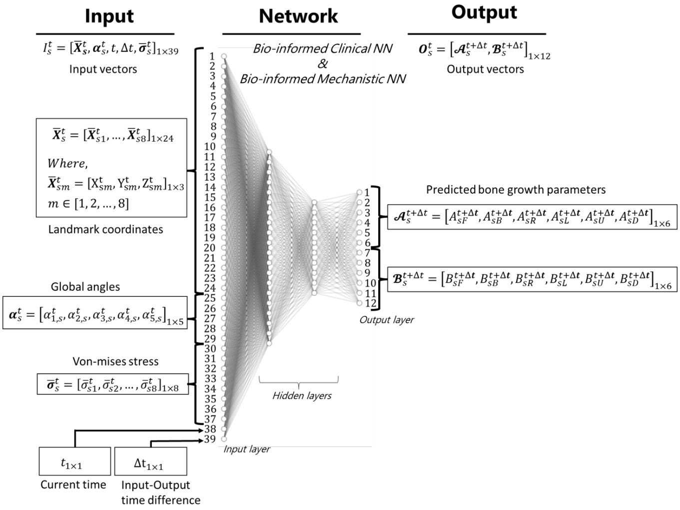


Figure 4.12: The architecture of a neural network for predicting the physical growth equation parameters. This structure is used in both $FFNN_{BC}$ and $FFNN_{BM}$ by defining two separate loss functions (one mechanistic and one non-mechanistic).

Table 4.5: Data collection for $FFNN_{CL}$, $FFNN_{CR-BC}$ and $FFNN_{CR-BM}$ for a single patient whose X-rays images were collected serially over time.

Identification of X-ray Images	Age of Patient (Months)
Initial X-ray Image	124
Output training X-ray images	{139, 149, 156, 168}
Output testing X-ray images	{160, 179, 187}

model since the range and units of the input matrix are not the same.

4.4.4. RESULTS AND CROSS VALIDATION ON THE DATA

The findings obtained from the patient described in Table 4.5 are presented here to illustrate the accuracy of the bio-informed mechanistic neural network prediction. The NN reconstruction process is the same as that explained in section 4.4.3.1. A breakdown of the data used for training and testing can be seen in Table 4.5. Figure 4.14 shows a comparison of the ground truth (actual data acquired from X-ray scans) and $FFNN_{CL}$ for the age of 160 months which is inside of the range of the training data and 179 and 187 for the outside of the range of the training data. The results are visualized for the eight corner points signifying the growth landmarks. Due to the 3D-Clinical NN's nature, the framework fails to forecast the ground truth. This framework is built on interpolation, and when it comes to the patient-specific prediction, where the available data is insufficient, the model cannot learn from the available data, hence it fails to predict the outcomes. Figure 4.15 shows the results for the $FFNN_{CR-BC}$ framework applied on

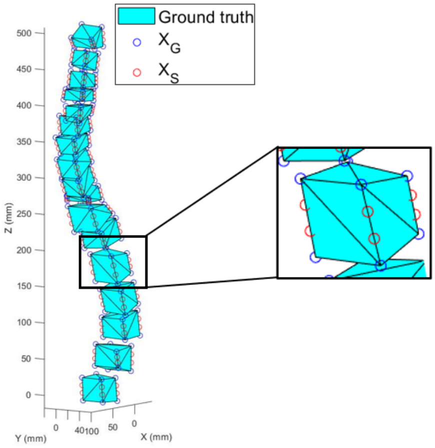
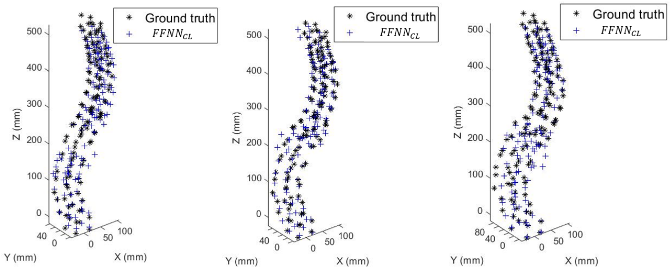


Figure 4.13: Illustration of the growth landmarks and side landmarks. There are 8 growth landmarks ($X_G = 8$) and 8 side landmarks ($X_S = 8$).

the same dataset. The results are $3D$ reconstructed based on the growth parameters obtained by $FFNN_{CR-BC}$ (Figs. 4.15 a, c and e) and the corresponding AP views are visualized in Figs. 4.15 b, d and f respectively. Because the findings are recreated using the bone growth model, this framework delivers a more accurate prediction than $FFNN_{CL}$. The data are next examined for the $FFNN_{CR-BM}$, as illustrated in Figure 4.16. Due to the implementation of the mechanistic loss function, the results are the best when compared to the $FFNN_{CL}$ and $FFNN_{CR-BC}$ and are compatible with the ground truth obtained from X-ray images.

A cross validation study is carried out to see how the input and output data impact the framework error, as shown in Table 4.6. The term “cross validation” refers to the process of determining how well a prediction model will work in practice [236]. We preserve the past two ages as test cases and rotate the test case inside the prior age span since we are more interested in prediction outside of the range. We can see from Table 4.6 that the prediction error of $FFNN_{CR-BM}$ is always lower than that of $FFNN_{CL}$ and $FFNN_{CR-BC}$, and that it can predict future spine curvature with a small error. The relative errors are calculated based on the Eqs. (4.18), (4.22) and (4.26) for $FFNN_{CL}$, $FFNN_{CR-BC}$ and $FFNN_{CR-BM}$ respectively. This observation can be explained by the fact that $FFNN_{CL}$ operates as an interpolation function and cannot provide accurate prediction when there is insufficient training data. Since $FFNN_{CR-BC}$ employs the bone growth equation to reconstruct the expected geometry, as explained in Eq. 4.20, it performs better than $FFNN_{CL}$. Finally, $FFNN_{CR-BM}$ performs the best since it modifies the loss function as mentioned in Eqs. 4.23 and 4.24 while also rebuilding geometry using the biological bone growth model. Table 4.7 refers to the cross validation on the $2D$ data using the existing Mechanistic framework [247] and $FFNN_{CR-BM}$. As it is shown, for each testing case, the present structure is more efficient as the parameters for bone growth are calibrated taking into account the effects of time and position. The suggested



(a) (b) (c)

Figure 4.14: Differences between pure data science prediction ($FFNN_{CL}$) and ground truth (the results obtained by X-ray images) at age of (a) 160 Months (inside of the range of the trained data), (b) 179 Months (outside of the range of the trained data) and (c) 187 Months (outside of the range of the trained data). The landmarks are eight corner points of each vertebra. It is obvious that $FFNN_{CL}$ cannot predict the ground truth.

Table 4.6: Cross validation study for three different neural networks. The table lists the relative approximation error of the predicted landmarks (X_G and X_S as shown in Figure 4.13) of each vertebra using the different neural networks. Datasets are categorized by age (in months) of the tracked patient. For each trial case, the testing data is shown in blue cells and the training data is shown in white cells.

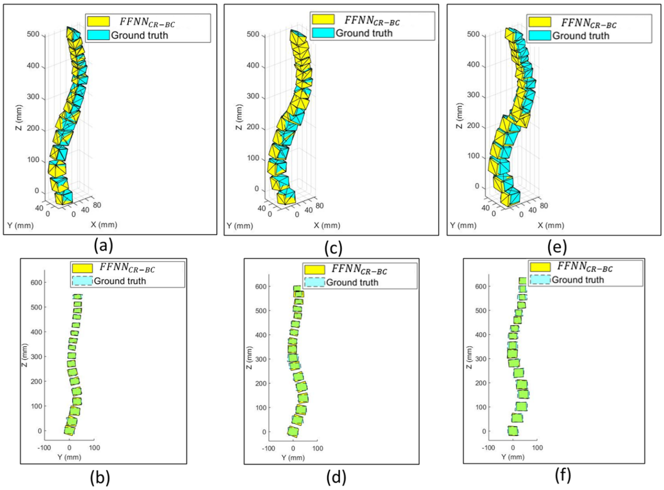
Age / NN	Prediction Error						
	139	149	156	160	168	179*	187*
$FFNN_{CL}$	**74.33**					**52.57**	**50.31**
$FFNN_{CR-BC}$	**1.09**					**0.10**	**0.18**
$FFNN_{CR-BM}$	**0.68**					**0.02**	**0.04**
$FFNN_{CL}$		**66.49**				**52.34**	**50.10**
$FFNN_{CR-BC}$		**0.19**				**0.05**	**0.13**
$FFNN_{CR-BM}$		**0.10**				**0.02**	**0.03**
$FFNN_{CL}$			**59.92**			**52.65**	**50.38**
$FFNN_{CR-BC}$			**0.090**			**0.091**	**0.16**
$FFNN_{CR-BM}$			**0.08**			**0.02**	**0.032**
$FFNN_{CL}$				**57.64**		**52.51**	**50.26**
$FFNN_{CR-BC}$				**0.03**		**0.08**	**0.18**
$FFNN_{CR-BM}$				**0.01**		**0.02**	**0.03**
$FFNN_{CL}$					**57.35**	**52.50**	**50.25**
$FFNN_{CR-BC}$					**0.03**	**0.09**	**0.13**
$FFNN_{CR-BM}$					**0.015**	**0.018**	**0.02**

* Prediction at age outside of the range of trained data.

$FFNN_{CR-BM}$ enhances prediction accuracy by 40% for the inside of the range and 84.3% for the outside of the range.

4.5. DISCUSSION AND CONCLUSION

In this study, we trained a bio-informed mechanistic deep learning model for prognosis of pediatric spinal deformity. The training data consists of clinical data and mechanistic features. The clinical data are extracted from anteroposterior and lateral views of the X-ray images using the active contour image segmentation model whose parameters are calibrated through sensitivity analysis. Next, the personalized 3D spine models are established through the mesh morphing technique which consists of registering a volumetric atlas spine model on each of the extracted clinical data. Thus, any geometry-dependant feature within the patient-specific models is detectable and trackable to calibrate the bone growth model parameters. The mechanistic features are obtained from



(a) (c) (e)

(b) (d) (f)

Figure 4.15: Differences between Bio-informed Clinical prediction ($FFNN_{CR-BC}$) and ground truth (the results obtained by X-ray images) at age of 160 Months (inside of the range of the trained data, a-b), 179 Months (outside of the range of the trained data, c-d), and 187 Months (outside of the range of the trained data, e-f). Subpanels (a, c, e) show the 3D view and subpanels (b, d, f) show the 2D view on AP plane. The landmarks are eight corner points of each vertebra. It is obvious that $FFNN_{CR-BC}$ can predict the results close to ground truth.

the bone growth model which takes the stress distributions of the spine as inputs. A dynamic patient-specific bone growth model is proposed to enhance the accuracy of the model. To measure the stress distribution, a patient-specific 3D finite element model is generated based on the 2D clinical data and the 3D reconstruction algorithm. By merging medical data with a mechanistic model, the presented framework can address the problem of limited data for the patient-specific study. The proposed model is capable of predicting the spinal curve of a single patient, either inside or outside the training range. This study is unique in that it provides patient-specific, time-dependent, and position-dependent parameters that can be calibrated throughout the dataset. In addition, the proposed bio-informed deep learning network with the modified bone growth model was shown to achieve competitive or even superior performance against other state-of-the-art learning-based methods.

In conclusion, the workflow described in this article can be a useful and innovative guide for the early detection and treatment planning of spinal illnesses such as scoliosis, lordosis, and kyphosis. Furthermore, the framework may be used for dynamic finite element analysis of various tissues at a smaller scale. It can also be potentially extended to other image-based studies, including tumor progression and cardiovascular applications. Future research will look at the same framework over a larger dataset to construct a real-time, patient-specific, optimal treatment plan based on the first patient visit. Moreover, a virtual reality application will be developed to visualize the spine in 3D and guide surgical training in orthopaedic surgery.

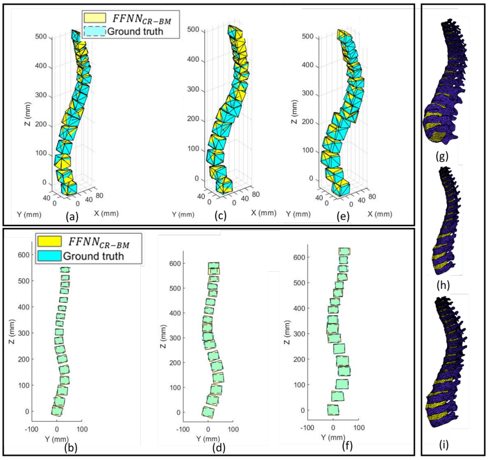


Figure 4.16: Differences between Bio-informed Mechanistic predictions ($FFNN_{CR-BM}$) and ground truth (results obtained by X-ray images) at age of 160 Months (inside of the range of the trained data, a, b and g), 179 Months (outside of the range of the trained data, c, d and h) and 187 Months (outside of the range of the trained data, e, f and i). Subpanels (a, c, e) show the 3D view, subpanels (b, d, f) show the 2D view on AP plane and subpanels (g, h, i) show the 3D reconstructed detailed geometry. The landmarks are eight corner points of each vertebra. $FFNN_{CR-BM}$ can clearly predict outputs that are close to ground truth.

Table 4.7: Cross validation study for two different neural networks on 2D data for AP view. The relative approximation error of the predicted landmarks (X_G and X_S as shown in Figure 4.13) of each vertebra was calculated using the different neural networks. Datasets are categorized by age (in months) of the tracked patient. For each trial case, the testing data is shown in blue cells and the training data is shown in white cells. The mechanistic framework $FFNN_{ME}$ is borrowed from [248]. For each trial case, the Bio-informed Mechanistic approach ($FFNN_{CR-BM}$) had better performance.

Age / NN	Prediction Error						
	139	149	156	160	168	179*	187*
$FFNN_{CR-BM}$	**0.14**					**0.08**	**0.03**
$FFNN_{ME}$	**0.30**					**0.12**	**0.67**
$FFNN_{CR-BM}$		**0.05**				**0.07**	**0.02**
$FFNN_{ME}$		**0.07**				**0.3**	**0.53**
$FFNN_{CR-BM}$			**0.03**			**0.06**	**0.03**
$FFNN_{ME}$			**0.04**			**0.17**	**0.20**
$FFNN_{CR-BM}$				**0.04**		**0.08**	**0.03**
$FFNN_{ME}$				**0.08**		**0.24**	**0.32**
$FFNN_{CR-BM}$					**0.04**	**0.06**	**0.03**
$FFNN_{ME}$					**0.03**	**0.24**	**0.44**

* Prediction at age outside of the range of trained data.

III

4D HUMAN BODY SHAPE SCANNING

5

OPTIMAL POSITION OF CAMERAS DESIGN IN A 4D FOOT SCANNER

Optical motion capturing explains the three-Dimensional (3D) position estimation of points through triangulation employing several depth cameras. Prosperous performance relies on level of visibility of points from different cameras and the overlap of captured meshes in-between. Generally, the accuracy of the estimation is practically based on the camera parameters e.g., location and orientations. Accordingly, the camera network configurations play a key role in the quality of the estimated mesh. This chapter proposes an optimal approach for camera placement based on characteristics of a depth camera D435i - Intel RealSense. The optimal problem includes a cost function that contains several minimisation and maximisation terms. The minimisation terms are distance of the cameras to the center of the scanning object, resolution error, and sparsity. And the maximisation terms are distance between each two pair of cameras, percent of captured point from an object, and the level of overlap between cameras. The object is designed based on practical experiments of human walking and is a bounding box around one step of dynamic foot work-space from heel strike posture to toe-off posture. The accuracy and robustness of the algorithms are assessed via experiment measurement, and sensitivity to the number of cameras is investigated. Accordingly, the experiment results determined that the scanning accuracy can be as high as 2.5 % based on a reference scan with a high-end scanner (Artec Eva).

Parts of this chapter have been published in:
Tajdari, F., Eijck, C., Kwa, F., Versteegh, C., Huysmans, T. and Song, Y., 2022, August. Optimal position of cameras design in a 4D foot scanner. *In International design engineering technical conferences and computers and information in engineering conference* (Vol. 86212, p. V002T02A044). American Society of Mechanical Engineers. [253].

5.1. INTRODUCTION

Foot shape is recognized as a very diverse character among the population, containing sex [298, 126, 127], and age [267]. This diversity may not be considered in footwear sizing, as available footwear fitting standards require only length and width of foot, and arch length to adjust to standardized shoe sizes [36]. Such diversity challenges the users to define a proper fit, leading to non-suited footwear in terms of comfortable and fitting increase the chance of occurring injury during ambulation [65]. Comfort-based footwear fit has determined direct positive impact on enhancing biomechanical performance [200], decreasing the chance of movement-related injury [178], and practical-wise it is the most important factor for users to choose footwear [163]. Thus, Footwear should be fully/semi personalised fit to satisfy any customer in terms of comfort, safety and fitting [173].

Available approaches of designing footwear established on employing static lasts, which supposes the foot includes rigid segments. Due to the fact that the foot is a deformable part of the body [246, 242], the assumption usually leads to failure investigating the dynamic movements in the foot morphology, specifically for a loaded foot in estimation of ankle joint mechanics [315, 276]. Proofs have been studied on articular motion through the loaded foot [159, 290], which impacts linear foot measurements, especially during transitioning between postures i.e., from sitting to standing [302] or during the stance phase of gait [94]. The measurements of the dynamic foot recommend morphological variations happening, which are not reachable via static linear and circumferential measurements. Accordingly, characteriing the diverse population of foot shapes would be challenging within individuals in presence of different loading scenarios e.g., gait. One solution to the challenge may be recruiting motion tracking techniques for capturing the foot deformation for shoe penalisation.

Due to the inherent connection between virtual reality systems and naturalistic applications, motion tracking techniques as the key component of the systems should be sufficiently robust and accurate [132]. Partly, motion tracking has been employed to estimate the position and direction of the viewing gate to define the eyepoint to be able to render the most accurate perspective of images. Furthermore, body motion tracking [249, 248] is usually deployed to establish avatars which virtually models the human to the viewer. Meanwhile, optical motion tracking, as a well-known technique, may cover the eyepoint estimation and the dynamic avatar generation. To do so, arrangement of several cameras is an undeniable part of any motion tracking technique. The arrangement should guarantee the free movement of the users to be able to track the targeted features. The feature points (generally detectable via passive or active landmarks) are identified from camera outputs which give the 3D position of the landmarks through triangulation between the multiple camera outputs [74, 161, 251]. Apparently, the accuracy and the quality of the 3D reconstructed point-clouds [92, 257] rely on the number of the employed cameras and their configuration. However, the number is competing with the cost of the final product and partly the speed of the process. Thus, this chapter focuses on an automatic method to optimally define the position and orientation of a set of cameras in a foot scanner.

Recent research investigating the cameras' configuration have mainly addressed two main challenges that may result in inaccurate reconstruction: (1) Locating the finite number of cameras to maximally penalize the accuracy error of the final triangulation performance and (2) Locating the cameras to establish the optimum views with existence of occlusion.

The accumulative error was analyzed by Sanders-Reed [218] in 3D position estimation via two-camera triangulation. According to the analysis, the configuration of cameras plays a key role on the level of the accumulative error which defined a 90 degree angle between the sensors as the optimal configuration. The accumulative error is maximised when the sensors have the same orientation (either 0 or 180 degree). Olague and Mohr [186] studied the properties of the accumulative error in the process of the triangulation and validated the outputs practically. They proposed a metric of the accumulative error via extracting the component with maximum value through the diagonal of covariance matrix of the 3D point reconstruction. The maximum value indicates the highest accumulative error regarding a certain configuration (position and orientation). To achieve a global optimal solution to the optimal problem, a genetic algorithm is applied. To simplify the problem, Olague et al. considered a constraint that cameras must be located on a plane with a constant distance from the target point. Cowan and Kovesi [57] specified a group of constraints impacting the quality of the cameras' network to circumvent the time-consuming trial-and-error method for designing the configuration of the cameras' network (locations). They mapped the captured object into its surfaces. Then, the solution regarding each surface was extracted individually and independently according to a group of predefined constraints and next, each of the local solutions were integrated. The integration presented a global solution to the optimal problem. For each constraint e.g., resolution loss, sparsity, and occlusion, they determined a valid 3D volume that guarantees any point in that area fulfil the related constraint. Then, they considered the overlap area of those individual 3D volumes as checking area(s) where all constraints are satisfied. Mason [164] continued the idea of Cowan and Kovesi [57] with extended constraints and offered a novel model to determine the optimal placement. The model introduced by Cowan and Kovesi only covers sedentary occlusion known as self-occlusion. Mason's heuristic approach defines a solution for a specific target, however, the result can be a non-optimal output regarding too much simplifications of the target. Moreover, moving targets are big restrictions for Mason's approach. To avoid facing the object surfaces, Wu, Sharma, and Huang [294] contemplate the overlap of the mapped pyramids instead of dealing with a computationally expensive numerical measurement to tackle with the unknown bounds. Chen and Davis [44] assumed a probabilistic model for the occlusion, where a vertical plane was supposed as the occluder next to the target point. The model anticipates the visibility of a target point by at least two cameras considering all valid orientations of the occluder around the vertical axis. Though the Chen and Davis' model perfectly tackles the dynamic occlusion regarding the error metric, the convergence angles of the cameras are not addressed. As a result, the method proceeds to define configurations leading to comparatively large errors in target points alignment outperformed by poor triangulation regarding locating nearly parallel cameras.

The limited space, the number of cameras and the intrinsic and extrinsic properties of the cameras restrict the positioning of the tracking cameras especially for foot scanning as the bottom of the foot is very complicated to be scanned. Generally, the cameras configuration considering the space constraint and having maximum overlap between cameras and maximum visibility of the foot to the cameras is very challenging. This chapter proposes an optimal approach to locate a group of depth cameras D435i - Intel RealSense for a pre-defined box-shape work-space to maximize the percentage of observed points of a target object with minimum accuracy error of the 3D reconstruction in the existence of dynamic occlusion. Generally, work-space location of the target objects is known in many augmented and virtual reality systems. For instance, Rahimian et al. requested the participants to walk across a single-lane road [205], and after some iterations the valid work-space would be identified for a single step of walking. The approach we present estimates the optimal positions and orientations of a group of cameras for a predefined set of target points. Where the best set considers, first, all possible locations of occluders and specifies the cameras' locations in order to have the maximum visibility of the target points between at least two triangulable views i.e., the views' that have the maximum distance with each other and the maximum overlap of capturing the target points. Secondly, the placements should consider that the targets are visible from the widest range of viewpoints to have an acceptable level of overlap for alignment.

We propose a method to define the optimal positions for 2 to 7 cameras regarding the motion tracking, where always one of the cameras is located on the bottom of the work-space.

5.2. METHODOLOGY

In order to derive the equations of modeling a depth camera, a number of simplified assumptions were considered and the rest of the parameters are optimally designed based on the desired work-space and related constraints.

5.2.1. ASSUMPTIONS

To model the camera, we convey the camera parameters based on the distance to a detected point. The parameters that are considered in this chapter are the spatial error, the resolution, and the acceptable viewing angle for Depth Camera D435i - Intel RealSense. The error of a 3D point location is the distance between the real measurement and the location estimation by the camera. Practically, the fixed resolution of the camera might lead to sparsity when the distance between the object and the camera becomes large and the error also increases by the distance. According to [9], the error has a parabolic relation with the distance as shown in Figure 5.1(a), where the error has the following equation.

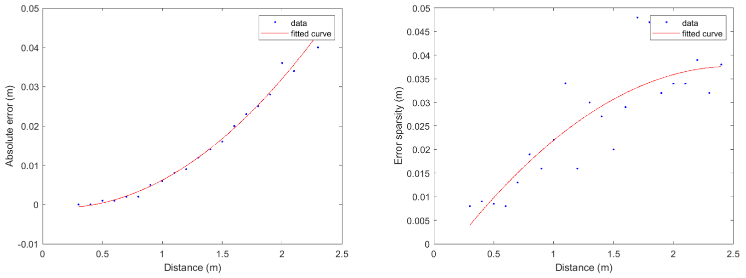
$$Error_{resolution}(dis) = p_1\ dis^2 + p_2\ dis + p_3; \tag{5.1}$$

where dis unit is in meter, and p_1, p_2, and p_3 are the parameters of the fitted curve as follows.

$$p_1 = 0.009458;\ \ p_2 = -0.002578;\ \ p_3 = -0.000655; \tag{5.2}$$

Moreover, the viewing angle of the camera is assumed as a constant and equal to 60° based on [9].

With the same approach for the resolution error and from [9], the sparsity variation has also a parabolic relation with the distance as shown in Figure 5.1(b), where the func-



(a) (b)

Figure 5.1: Camera parameter.

tion has the following equation.

$$Error_{sparsity}(dis) = p_4 \ dis^2 + p_5 \ dis + p_6; \tag{5.3}$$

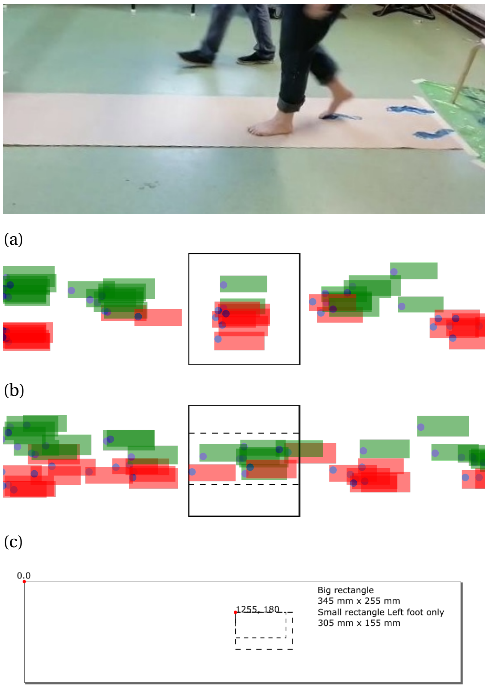
where p_4, p_5, and p_6 are the parameters of the fitted curve as follows.

$$p_4 = -0.006931; \ \ p_5 = 0.03473; \ \ p_6 = -0.005851; \tag{5.4}$$

5.2.2. WORK-SPACE DEFINITION

Currently the assumption is made that the participants place their feet in the middle of the walkway. However, in practice, a slight variance of location will happen due to the fact each person is unique. To capture the location of the heel strike of each participant, we invited 8 participants (3 females and 5 males), 6 within the age range of 20 and 30, 2 around 60 years, with a varying height between 1.6m and 2.1m and Body Mass Index (BMI) between 21 and 28.7. Paint was applied to their feet and they were instructed to walk across in several manners shown in Figure 5.2(a) to see if it is possible to reduce the variance of the location of the heel strike while ascertaining a natural walk. The resulting paintings where digitally overlayed to show where the average foot location was of each participant. Several runs were conducted with different commands. During one walk the participants were instructed to place their foot in the middle of a drawn box. The next walk the participants were instructed to place their feet in a certain position from the edge and were given the command to just walk to the other side.

Both results of the test were digitized and overlaid to show where the participants placed their feet. These images can be seen in Figure 5.2(b) and (c). It was found that the users, with the proper instruction and foot placement, were able to land their feet naturally in the same area. The area is shown as a bonding box in Figure 5.2(d). Considering the right foot, the left foot and both feet these three scenarios of scanning, we extracted three bounding boxes reported in Table 5.1 where we considered the biggest box for our calculation regarding finding the optimal camera positioning. It worth mentioning that the measured maximum height of the foot/ankle during walking belongs to toe-off posture. See the reported height in the table.



(a)

(b)

(c)

(d)

Figure 5.2: (a) Participants were instructed to try to land their feet in the middle without further instructions. (b) Participants were instructed to where to place their feet and to walk to the other side. (c) Average foot position participants. (d) the practical test-bed.

5.2.3. OPTIMAL DESIGN OF THE FOOT SCANNER

As the future application of the scanner is to scan feet during walking, the work-space of any dynamic foot is considered as predefined as discussed in the previous section and equal to a rectangular box. According to Figure 7.2, r_i is the distance between i^{th} camera to the center of the box, d_{ij} belongs to the distance between i^{th} camera and j^{th} camera, and G is the center of the object. In addition, the (x_i, y_i, z_i) corresponded to the location of i^{th} camera are optimized through the following cost function.

$$J_{tot} = \begin{cases} Minimizing: & J_r, J_e, J_s \\ Maximizing: & J_d, J_v, J_o \end{cases} \tag{5.5}$$

where J_{tot} is the total cost function including both minimization and maximization terms. Regarding the minimization, J_r explains a term regarding the distance of each camera to the center of the work-space, denoted as r_i in Figure 7.2, and is considered as follows where n is the number of cameras.

$$J_r = \sum_{i=1}^{n} |r_i|^2 \tag{5.6}$$

Table 5.1: The choice of options.

Parameter	Cases		
	Left foot	Right foot	Both feet
Width (cm)	15.5	13.2	25.5
Height (cm)	17.5	17.5	17.5
Length (cm)	30.5	29.8	34.5

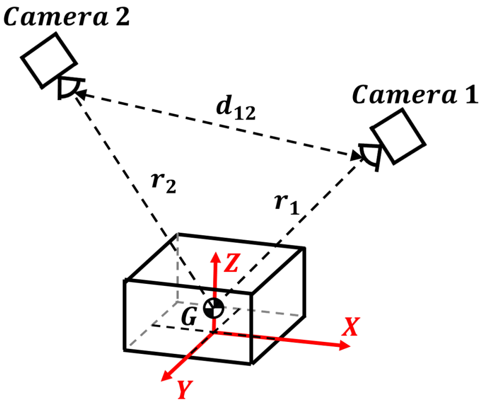


Figure 5.3: Camera parameter.

J_e is explaining the error of resolution calculated from (5.1) where we considered r_i as the dis in (5.1) for each camera to simplify the calculations. Accordingly, J_e would be extracted as follows

$$J_e = \sum_{i=1}^{n} |Error_{resolution}(r_i)|^2 \tag{5.7}$$

The last minimization term as J_s defines the sparsity of the points in (5.3) and is denoted as follows.

$$J_s = \sum_{i=1}^{n} |Error_{sparsity}(r_i)|^2 \tag{5.8}$$

Regarding the maximization, we have three terms of J_d, J_v and J_o which refer to the distance between each two cameras (e.g., 10 distances for 5 cameras), the number of the captured points on the box through all the cameras, and the overlap level between cameras, respectively. These teams should be maximised to have the maximum visible area of the object with the cameras and avoid having too close cameras to each other. The J_d is calculated as follows.

$$J_d = \sum_{i=1}^{n} \sum_{j=i}^{n} |d_{ij}|^2 \tag{5.9}$$

where d_{ij} is the distance between i^{th} camera and j^{th} camera. Accordingly, J_v is calculated as

$$J_v = \left(\frac{N_v}{N_{tot}}\right)^2 \tag{5.10}$$

where N_v is the number of points captured with at least one camera and N_{tot} denotes the total number of the points on the rectangular shape box. Moreover, J_o is generated as

$$J_o = \frac{1}{N_{tot}^2 n^2} \sum_{i=1}^{N_{tot}} n_{o_i}^2 \tag{5.11}$$

where n_{o_i} defines the number of cameras that could capture the point i^{th}. Thus, the optimal problem is finding a set of optimal design camera positions in X as follows

$$X = \begin{bmatrix} x_1 & y_1 & z_1 & \dots & x_{n-1} & y_{n-1} & z_{n-1} & z_n \end{bmatrix} \tag{5.12}$$

such that

$$\min J_{tot} = \alpha_1 J_r + \alpha_2 J_e + \alpha_3 J_s - \alpha_4 J_d - \alpha_5 J_v - \alpha_6 J_o \tag{5.13}$$

subject to

$$\begin{aligned}
\text{Camera classification}: &\begin{cases} -z_1 \leq 0 \\ \vdots \\ -z_{n-1} \leq 0 \\ z_n \leq 0 \end{cases} \\
\text{Walking way}: &\begin{cases} -|y_1| + 0.3 \leq 0 \\ \vdots \\ -|y_{n-1}| + 0.3 \leq 0 \end{cases} \\
\text{Allowed work-space of cameras}: &\begin{cases} |y_1| - 0.5 \leq 0 \\ \vdots \\ |y_{n-1}| - 0.5 \leq 0 \end{cases} \\
\text{Bottom camera}: &\begin{cases} y_n = 0 \\ x_n = 0 \end{cases}
\end{aligned} \tag{5.14}$$

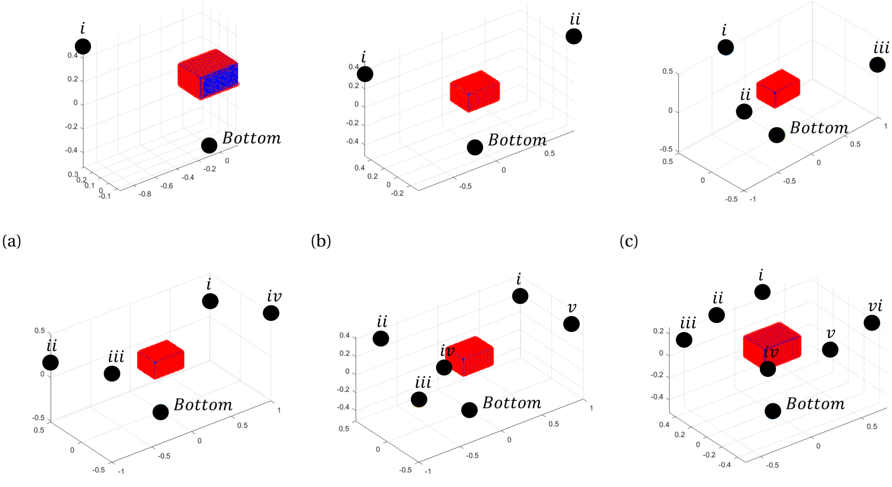
By applying the algorithm in [130], the optimal parameters are designed and for each of the cases the values related to the both feet will be obtained.

5.3. RESULTS

Here, we perform several experiments with differing number of cameras to find the optimum position of the cameras. Also we investigate the effects of the number of cameras in terms of the overlap percentage of pairs of cameras for the minimum number of cameras as well as the maximum overlap to maximise the performance of any mesh registrations methods.

Table 5.2: The optimal positions as [x,y,z] in cm and percent of captured area of the box ($\sqrt{J_v}$).

Camera ID	Number of cameras on top					
	1	2	3	4	5	6
i	[92,31,49]	[-91,50 49]	[-29,50,50]	[100,50,17]	[100,50,17]	[28,48,26]
ii		[92,-32,49]	[-100,-50,50]	[-100,50,17]	[-69,50,29]	[-22,43,25]
iii			[100,-50,17]	[-100,-50,50]	[-100,-50,17]	[-65,30,23]
iv				[100,-49,49]	[-68,-50,43]	[-55,-50,26]
v					[100,-30,23]	[25,-43,17]
vi						[63,-30,17]
Bottom	[0,0,-51]	[0,0,-51]	[0,0,-50]	[0,0,-48]	[0,0,-47]	[0,0,-45]
$\sqrt{J_v}$	0.54	0.99	1	1	1	1



(a) (b) (c)

(d) (e) (f)

Figure 5.4: The optimal camera positions. (a) One camera on top. (b) Two cameras on top. (c) Three cameras on top. (d) Four cameras on top. (e) Five cameras on top. (f) Six cameras on top.

5.3.1. SOLVING THE OPTIMAL PROBLEM

In all the experiments one camera is on the bottom of the box (negative values for z axis) and the rest of the cameras have positive value along the z axis as defined in (5.14). The box includes 3708 points as $N_{tot} = 3708$. In addition the optimization parameters in (5.13) are considered as $\alpha_1 = 200$, $\alpha_2 = 1.2$, $\alpha_3 = 1.2$, $\alpha_4 = 1$, $\alpha_5 = 15$, and $\alpha_6 = 1$. The values are defined via trial and error method emphasising on having as small as possible scanner dimensions (high $\alpha_1 = 200$). The results of solving the optimal problem are numerically reported in Table 5.2. According to the table, except the cases with having one or two cameras on top, the other cases have J_v equal to one meaning all the points on the box were visible to at least one camera. The position of the cameras are also visually shown in Figure 5.4, where the points with red color on the box are the captured points with at least one camera and the blue ones are not captured with any cameras.

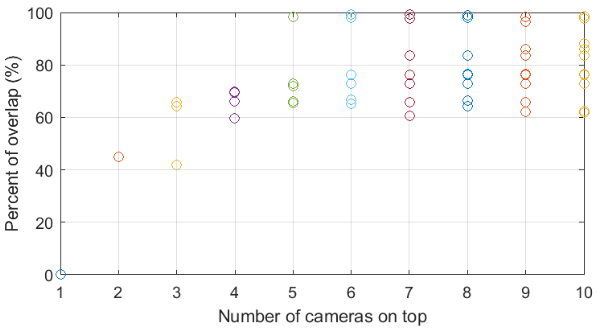


Figure 5.5: Percent of overlap variation based on the number of cameras on top.

5.3.2. PERCENT OF OVERLAP BETWEEN PAIR OF CAMERAS

A crucial step to reconstruct a scanned object through multiple depth cameras is using surface registration methods. Surface registration is capable to assemble multiple 3D point clouds in a common coordinate system via aligning the overlapping parts of the point clouds [311, 309]. The point clouds may contain characteristics of a single shape as a mesh structure to explain the surface of the 3D objects or scenes. To form a complete 3D shape, different 3D captured datasets of an object from different viewpoints are needed which should cover all the targeted areas of the object with acceptable level of overlap between each pair of the views. This facilitates a well-performed registration to recombine all the datasets and reconstruct as accurate as possible the 3D mesh of the original scanned object or scenes [139]. In the registration process of a multiple-camera scanner, these datasets must be registered on each other and the quality of the registration is very depending on the level of overlap between each two pair of cameras from the object. Thus, here we investigate the degree of overlap based on the number of used cameras with optimal positions.

Overlap percent (P_o^i) from camera i^{th} to camera j^{th} is calculated as follows.

$$P_o^i = \frac{N_o^{i,j}}{N_o^i} \times 100 \tag{5.15}$$

where, $N_o^{i,j}$ is the number of points on the scanned object captured with both camera i^{th} and camera j^{th}, and N_o^i is the total number of points captured with camera i^{th}. To study the effect, we designed the optimal camera locations through minimizing the cost function in (5.13) for different number of cameras on top and reported the overlap percent for each pair of cameras in Figure 5.5. According to the figure, the overlap range increases by increasing the number of top cameras up to 5. After the 5 cameras on top the range is more or less constant. Thus, 5 cameras on top should be enough to have acceptable registration. However, as we would like to have a symmetric design we prefer 6 cameras on top to have three cameras on each sides.

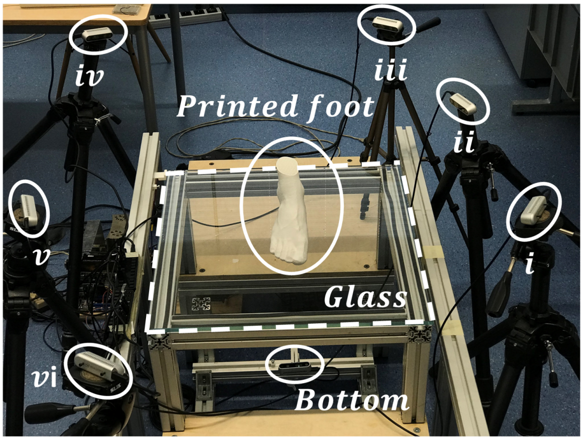


Figure 5.6: Scanner.

Table 5.3: Components of the final product in Figure 7.1.

Component	Description	Application
i-vi and Bottom	RealSense D435i	Capturing 3D point cloud
Glass	ClearVision low-iron glass plate	For capturing sole
Printed foot	Printed out via Ultimaker S5	For accuracy evaluation

5.3.3. SCANNER PROTOTYPE

The scanner used in this study is the first prototype of a 4D foot scanner at TUDelft [128]. As shown in Figure 7.1, the scanner utilizes seven RealSense D435i depth cameras to capture a 4D foot scan. In order for the user to start and end a scanning process, two AdaFruit VL53L0X ToF distance sensors (30 to 1000 mm range) are integrated, which are connected through an Adafruit TCA9548a Multiplexer (MUX). A rectangular-shaped glass platform is able to support users of up to 200 kg in the context of walking. To this end, the final scanner is fitted with a tempered ClearVision low-iron glass plate from AGC's Planibel Extra Clear glass collection. This type of glass is known for its high light transmission rate (92% for 6 mm). The glass plate has a nominal thickness of 12 ± 0.3 mm and a width and height of 600 ± 2 mm and 600 ± 3 mm, respectively. These dimensions match the scanning plate support platform of Vidmar's (2020) 4D foot scanner setup. The glass was chosen over polycarbonate (PC) or polymethyl methacrylate (PMMA) because of their sensitivity to scratching. It is desired to have a durable scanning solution, which also includes the quality of the scanning surface. Since glass has excellent scratch-resistant and light transmission properties, and there is a constant interaction between the walking users and the scanning plate, glass is chosen over other materials.


(a) (b) (c)

Figure 5.7: No-load foot. (a) The source mesh. (b) Scanned through Artec Eva scanner. (c) Printed Artec Eva output via Ultimaker S5 printer.

5.3.4. ACCURACY

Through the explained 3D registration method in [250] which uses a probability function [252, 247, 255, 254], we use seven cleaned frames captured by each camera to reconstruct the foot. According to the method each cleaned data-set is rigidly registered on a reference model to build a rough point-cloud of the foot, then the reference model is non-rigidly registered on the point-cloud to have a meaningful mesh. Due to the larger overlap between each data-set captured by each camera in the proposed method than the study in [250], their registration method shows better performance though our camera arrangement comparing to their non-optimal camera arrangement. Accordingly, the used reference model (source mesh) is shown in Figure 6.7(a).

To evaluate the accuracy of the results with the introduced algorithm, we scanned a foot with Artec Eva scanner in a no-load condition as shown in Figure 6.7(b), then the scanned foot was printed out via Ulimaker S5 scanner shown in Figure 6.7(c). The printed foot is scanned with our scanner in static mode which is comparable with the scan data of the Artec Eva. Next, the output of the Artec Eva scanner dimensions are compared with the corresponded dimensions of our method output data depicted in Figure 6.8. We first find the scaling ratio which maps the length of the reconstructed foot to the length of the foot collected with the Artec Eva scanner and then calculate the errors for the other dimensions as width, ball width, and ball angle according to [263]. Based on Figure 6.8, L_r, W_r, BW_r, and α_r explain the length, width, ball width, and ball angle of the real scanned foot in the no-load case with the Artec Eva scanner respectively. And, L_e, W_e, BW_e, and α_e define the length and width of the estimated foot through our method respectively. The scaling factor of r is considered as follows

$$r = \frac{L_r}{L_e} \tag{5.16}$$

where in our experiment, $L_r = 222.2$ mm, $L_e = 0.3701$ mm which results in $r = 600.4$. Thus any dimension extracted from our method is multiplied to the r and compared with the corresponded values on the Artec Eva scanner output summarised in Table 6.3. According to the table, the errors for the width is 1.8 mm, for the ball width is 2.5 mm, and for the ball angle is 0.3°, which shows the average percentage error based on the Artec Eva scanner is about 2.5%.

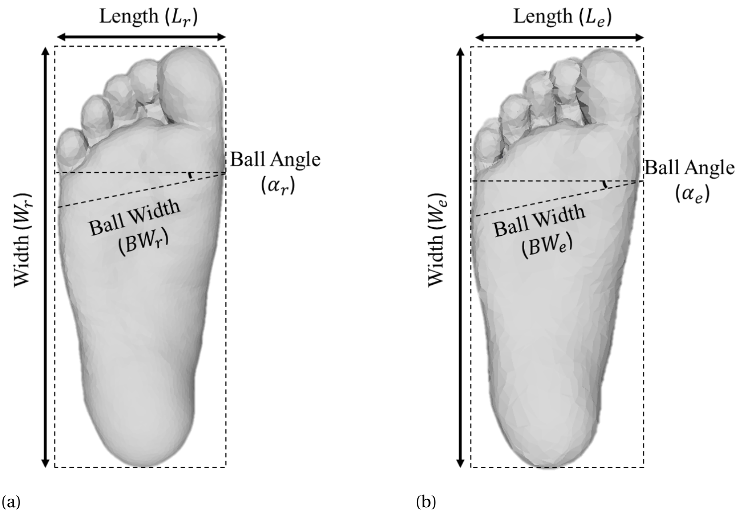


Figure 5.8: Dimension definition: (a) Artec Eva. (b) Our method.

Table 5.4: Error results of the parameters introduced in Figure 6.8 with $r = 600.4$.

Parameter	Artec Eva	Our approach	MAE
Width (mm)	$W_r = 87.3$	$r \times W_e = 89.0$	1.7
Ball Width (mm)	$BW_r = 92.6$	$r \times BW_e = 95.1$	2.5
Ball Angle	$\alpha_r = 10.8°$	$\alpha_e = 11.1°$	0.3°

5.4. CONCLUSION

This chapter proposes an approach to find the optimal number, position and orientation of depth cameras to scan a foot for fast 3D mesh reconstruction. The optimisation problem aims to have minimum camera scanning error, sparsity, and the scanner dimension, and to have maximal overlap between scans captured by an adjacent pair of cameras. The results showed the scanner with seven cameras (one on the bottom and six on top) have the most optimum performance. The future works lie on optimal design of a scanner equipped by Azure Kinect cameras to improve the accuracy of individual scans. In addition, the optimal problem will be integrated with a term regarding the lighting, as it is one of the most important factors on the quality of the captured data. Moreover, a human factor term will be integrated in the cost function to assure the participants reproduce their everyday-walking step in a comfortable manner, which is essential for ergonomics study and product design, e.g. shoe design.

6

DYNAMIC MESH RECONSTRUCTION BASED ON 3D NONRIGID REGISTRATION

Fitting apparel and apparel in performing different activities is essential for the functional yet comfortable experience of the user. 4D scans, i.e. 3D scans in continuous timestamps, of the body (part) in performing those activities are the basis for the design of garments/apparel in 4D. In this chapter, we proposed a semi-automatic workflow for constructing 4D scans of the body parts with the emphasis on registering noisy scans at a given timestamp. Continuous 3D scans regarding the moving body parts are captured first from different depth cameras from different view angles. In a given timestamp, the collected 3D scans are roughly aligned to a template using the rigid Iterative Closest Points (ICP) algorithm. Then these scans are further registered using a newly proposed non-rigid Iterative Closest-Farthest Points (ICFP) algorithm, in which correspondences between the source and the target are established by either closest or farthest points based on the newly defined logical distance concept and the probability theory. Experimental results indicated that the ICFP method is robust against noise and the scanning accuracy can be as high as 3.4 %. It also reveals that, for the human foot, the differences of ball width and ball angles between the loaded and the unloaded situation can be as large as 8 mm and 2 degrees, respectively. This highlights the importance of using 4D scan in designing garments and apparel.

Parts of this chapter have been published in:
Tajdari, F., Kwa, F., Versteegh, C., Huysmans, T. and Song, Y., 2022, August. Dynamic 3d mesh reconstruction based on nonrigid iterative closest-farthest points registration. *In International design engineering technical conferences and computers and information in engineering conference* (Vol. 86212, p. V002T02A051). American Society of Mechanical Engineers. [250].

6.1. INTRODUCTION

Due to comparatively huge changes in shape during deforming a dynamic object, dynamic anthropometry and 4D scanning have attracted a lot of attention in clothing technology recently [35, 39]. The 4D scanning gives valuable information regarding human body deformation during moving, conducting an activity and dynamic workload, establishes the fundamentals of comprehensive ergonomic fit design e.g., personal protective equipment, workwear, sportswear, and other practical garments [34, 173].

The concept of body measurement started with using measurement tape, which was time-consuming and required considerable effort [220]. Advancement in technology introduces the 3D scanning tool-kits [312, 74, 249, 248, 246, 242], which improves the effectiveness and efficiency of measuring the body shapes [51, 253]. However, the 3D body measuring techniques [92, 257] and scanners only allowed the collecting data in static postures, e.g., hand posture in [311, 309]. Thus, the scan captures the muscular system aiming to hold the static position and neglects the interaction between the body and environment during movement. As a result, garments are still designed based on the anthropometric standard position specified in standards and sizing framework as standing upright, legs hip widespread, arms slightly abducted [20, 176]. However, body proportion diverges from static to dynamic status, especially during intense activities such as sports. Therefore, it is important to investigate the body during motion, to extract meaningful features alterations i.e., dimensions, and transfer the the features into the garment enhancement process. To this end, the technological development from 3D to 4D scanning approaches establishes scanning in motion [266, 3].

Kirk et al. [118] introduced a method to evaluate the raising and lowering of the body surface with respect to connection with position changes via measurement lines on the knee. Schmid et al. [221] proposed seam damages were usually affected by changes of body geometry. The research group of Ashdown studied dimensional variation on the basis of 3D point cloud for the upper and the lower body half [50, 135, 49]. Based on the state-of-the-art methodology the deformation of the body surface was studied through discreet static postures. However, there is a paucity of literature regarding measurement alteration during movement and only a few researchers address this challenge [20, 176].

The concept of dynamic anthropometry was prosperously implemented mainly for the enhancement of high-performance sportswear. The challenges were decreasing muscle fatigue, increasing comfort, and resistance diminution and performance development [314, 119]. Firstly, Morlock et al. [157] studied deviation in body measurements from work to sports corresponded postures of men and women that the population sample was very large [176]. Next, body motion is mainly observed through Motion Capture Technology. The method captures the deformation by tracking a few pre-selected landmarks on the body mostly for articulation reconstruction. The approach has broad applications such as video game design, computer-animated movies, or biomechanical analyses in the fields of medicine and sports. The acquisition of 3D surfaces with motion scanning has been studied by many researchers such as [189, 6, 291, 78]. A finite number of landmarks highlighted with trackable markers are mounted on the human body and the position of the landmarks was estimated by the systems.

A very recent topic of capturing the dynamic body surface deformation over the moving articulation attracted a lot of attention, as it is often challenged by capture speed, computational speed, and accuracy of the measurements [266]. Most approaches exploit equipment founded on the concept of light and depth sensor technology. Employing techniques and measurement principles i.e., triangulation or light section algorithms, the surfaces are estimated. However, in each time frame, the methods fail to capture the 4D pattern (correspondence) but instead deliver an independant 3D mesh describing shape of the scanned object, which needs higher knowledge to extract the 4D features.

In this chapter, we present a workflow of automatically extracting dynamic features from the foot during walking. In detail, the method first aligns 3D point clouds extracted from seven depth cameras around the walking foot through the well-known rigid Iterative Closest Points (ICP) registration method. Then a reference foot model is nonrigidly registered on the aligned 3D raw data to define the patterns and features in each time interval. In the proposed nonrigid ICP algorithm, a new corresponding point selection approach is proposed and implemented for a more effective and effcient registration.

The rest of the chapter is presented as follows: first, the methodology is introduced in Section 7.3 where the techniques of the rigid alignment and nonrigid registration are discussed. In Section 6.3.1, the setup of the experiments is introduced. In Section 7.5, experimental results on the comparison of the proposed approach output and the reference data provided by Artec Eva scanner are presented. Finally, a short conclusion is drawn and future research directions are highlighted as well.

6.2. METHODOLOGY

In this section, we introduce a methodology that reconstructs a temporally corresponded 3D mesh out of different 3D pieces of point cloud through a source mesh, i.e., making a complete 3D shape using a few different depth cameras around the object. For proceeding with the methodology, we use a rigid ICP registration technique to register each piece containing the captured point clouds by each camera, correctly on the source mesh, and then the source mesh will be registered on the combined pieces (target mesh) through a non-rigid ICP registration method. Due to a high degree of noise, the corresponding points selection is modified from only considering the closest point to logical distance selection. The logical distance selection may choose the closest or farthest point on the target to a point on the source as the corresponding point, based on the topology of the reconstructed target. By iteratively reconstructing 3D shape out of the extracted 3D data from each camera in a time frame, we are able to make time series of 3D scans known as 4D scanning.

6.2.1. RIGID REGISTRATION

In the proposed research, we focus on the geometry of the foot. Firstly, we estimate the points visible from each camera on the source mesh used in [251] which is scaled to have approximately the same foot length as the scanned data with one of the side cameras (e.g., the red points on a side camera in Figure 6.1 are used to scale the gray source mesh in order to have the same length as the red point-cloud). These estimated points from each camera are used as the target for any extracted data from the relevant camera. Then using the rigid ICP method we are able to find the extrinsic transformation matrix that is

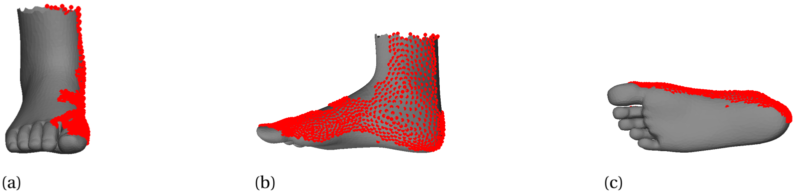


Figure 6.1: Pre-defined points on the source mesh for a side camera (the camera 3 in Figure 7.1): (a) Front view. (b) Side view. (c) Bottom view.

able to map each piece of data to the correct location on the source mesh. As an example of showing the predefined points on the source mesh, we depicted the predefined points for a side camera in Figure 6.1(e.g., camera 3 described in Section 7.3.1).

6.2.2. FILTERING

Usually, the output of the depth cameras is very noisy, and also there may be some point clouds that are not presenting any targeted features, e.g. from the background. Thus, first we filter the 3D point cloud of the target mesh ($\mathbb{T}$).In this process, the closest point for each point on $\mathbb{T}$ is found on the source mesh ($\mathbb{S}$), and the distance between each point on $\mathbb{T}$ and its corresponding point on $\mathbb{S}$ are compared with a distance limit (d_l) as follows.

$$d_l = m + 2\sigma \tag{6.1}$$

where, m is the average of all the closest distance from $\mathbb{T}$ to $\mathbb{S}$, and σ is the standard deviation of the closest distance population. The points that have a distance greater than the d_l will be rejected as follows.

$$\begin{cases} keep, & \text{if } d_p \leq d_l \\ reject, & \text{if } d_p > d_l \end{cases} \tag{6.2}$$

where d_p is the distance from a point p on $\mathbb{T}$ to its corresponding point on $\mathbb{S}$. It is worth mentioning that we consider the distance from $\mathbb{S}$ to $\mathbb{T}$ only, many useful points on $\mathbb{T}$ might be excluded due to the imperfect alignment of the pieces during establishing the $\mathbb{T}$.

6.2.3. NONRIGID CORRESPONDING SELECTION

In the use of the conventional ICP method, given a point on the $\mathbb{S}$, the closest point on the $\mathbb{T}$ is considered as its corresponding point. As only the Euclidean distance is used in establishing the correspondences, in terms of a high level of noise or disturbances, the topography of the original $\mathbb{S}$ may be lost. Thus, we use a logical distance instead of the closest point approach to establish the correspondences. Accordingly, firstly the closest point from each point on $\mathbb{T}$ to $\mathbb{S}$ is established. Regarding the miss alignment and disturbances for some points, the found distance is comparatively high, which may never be considered as a corresponding point regarding the conventional ICP corresponding point selection algorithm. However, not all the far points are spare as there may be con-

Algorithm 2 The proposed corresponding selection approach

```
Input: Target point clouds ($\mathbb{T}$) and Source point clouds ($\mathbb{S}$)
Output: Corresponding points
Initialization
SourceID: Find the closest point from each point on $\mathbb{T}$ to a point on $\mathbb{S}$
SourceDis: The distance between the two corresponding points regarding SourceID
IDclass = unique(SourceID): Defines all the involved points from source as correspondences
Bound = mean(SourceDis) + std(SourceDis)
$k = 1$
while $k \leq$ length(IDclass) do
    ResembleID = find(SourceID == IDclass(k))
    DistanceMatrix = SourceDis(SourceID(ResembleID))
    if isempty(find(SourceDis(ResembleID)<Bound,1)) then
        The corresponding point is the farthest
    else
        The corresponding point is the closest
    end if
    $k = k + 1$
end while
```

siderable error with rigid registration then some useful points can be located in far distance from the source mesh. In addition if the deformation is very high e.g., closed finger hand as a target versus opened finger hand as a source mesh highlights the disability of using only closest distance to find the correspondences. Thus, we define a varying distance boundary by iteration that increases the probability of finding the useful point although they are far. In the process of finding correspondences from each point on $\mathbb{T}$ to $\mathbb{S}$, each point on $\mathbb{S}$ may be selected for more than one point on the $\mathbb{T}$. In this case, we logically select either the closest or farthest point. In each iteration of the registration process, we define a boundary distance as follows for the corresponding distance matrix from $\mathbb{T}$ to $\mathbb{S}$ as follows.

$$Bound = m + \zeta\sigma \tag{6.3}$$

where m and σ are the mean and standard deviation of the counted distance for the correspondences from $\mathbb{T}$ to $\mathbb{S}$, accordingly. If for a point on $\mathbb{S}$ a population of points $\mathbb{T}$ are selected which includes more than one point, then we consider the point with the highest distance among the population, if all the distance for the population is greater than the bound. Otherwise, we select the closest point as the corresponding point to the point on $\mathbb{S}$. The process of the corresponding point selection is explained in Algorithm 2.

Discussion: To study the probability of success with the considered value in (7.5), statistical solutions are employed. First, we assume that the closest distance from points on $\mathbb{T}$ to points on $\mathbb{S}$ establishes a normally distributed population with an average of m and standard deviation of σ. Accordingly and based on Miller's studies of Freund's statistics [82] used in [252, 247, 255], the probability indicator z is defined as follows.

$$z = \frac{Bound - m}{\sigma} \tag{6.4}$$

$$= \frac{m + \zeta\sigma - m}{\sigma} = \zeta \tag{6.5}$$

Using Table III of the book [82], while $z = \zeta$, the probability (P) of having a distance more than the bound is

$$P(Distance > Bound) = 0.5 - P(z) = 0.5 - P(\zeta) \tag{6.6}$$

It should be noted that in the Table III of the book, z is bounded where $z \in [0, 3.09]$ which $0 \leq P(z) \leq 0.499$. Thus, $P(Distance > Bound)$ is always equal or less then 0.5. As the methodology is iterative with a maximum iteration of $j^{\max}$, the probability of missing a useful point would be $P(Distance > Bound)^{j^{\max}}$ which would be a very small value. In fact, greater values of the $Bound$ resulted from greater values of ζ, means the closest point which is often used in the conventional ICP algorithm is addressed in the selection. While, smaller values of the $Bound$ (smaller value of ζ) emphasises more on selecting the farthest point which may map all the source mesh vertices to the farthest point and results in the source mesh topology loss. There must be an optimal value of ζ that balances the closest and farthest selection process as discussed in Section 6.3.2.

6.2.4. NONRIGID REGISTRATION

In this section, based on the established correspondences, a cost function based on Amberg [13] is defined. Amberg [13] proposed the non-rigid registration formulation as a combination of distance and stiffness terms summarised in the following formula

$$E(X) = \left\| \begin{bmatrix} \alpha M \otimes G \\ WD \end{bmatrix} X - \begin{bmatrix} 0 \\ WU \end{bmatrix} \right\|_F^2$$

$$= \| AX - B \|_F^2 \tag{6.7}$$

where, The sparse matrix D is formed to facilitate the transformation of the source vertices with the individual transformations contained in X via matrix multiplication, and denoted as $D = diag(v_1^T, v_2^T, \ldots, v_n^T)$, where $v_i \in \mathbb{S}$ and $i = 1, \ldots, n$, and n is the number of vertices on the $\mathbb{S}$. W is a diagonal matrix consisting of weights w_i. α is the stiffness constraint. To regularise the deformation, an additional stiffness term is introduced. Using the Frobenius norm $\|.\|_F$, the stiffness term penalizes the difference of the transformations of neighboring vertices, through a weighting matrix $G = diag(1, 1, 1, \gamma)$. During the deformation, γ is a parameter to stress differences in the skew and rotational part against the translational part of the deformation. The value of γ can be specified based on data units and the types of deformation [13]. The node-arc incidence matrix M (e.g. Dekker [61]) of the template mesh topology is employed to convert the stiffness term functional into a matrix form. As the matrix is fixed for directed graphs, the construction is one row for each edge of the mesh and one column per vertex. To establish the node-arc incidence matrix of the source topology, the indices (i.e. the subscripts) of edges and vertices are addressed, for any edge of r which is connected to vertices (i, j) , in r^{th} row of M, and the nonzero entries are $M_{ri} = -1$ and $M_{rj} = 1$.

Table 6.1: Parameters used in simulation configuration.

Parameter	Value	Description
α	1000:1	stiffness ratio
N_{mean}	3	Target points averaging size
ϵ	0.001	Convergence error threshold
$j^{\max}$	50	Convergence iteration threshold
$k^{\max}$	20	Number of iteration for the outer loop

6.3. EXPERIMENT SETUP

6.3.1. PARAMETERS OF THE PROPOSED METHOD

Table 6.1 presents parameters used in the experiment. During the minimization of the cost function, γ in G introduced in (6.7), was chosen to one. The stiffness constraint, α, is the lowering scalar. In the experiments α is decreasing from 1000 to 1. Regarding dependency of α values to the dynamic of source surface, the bounds of these values were manually defined so that only global deformations were considered in the beginning of registration. On the other hand, the lower limit of the α also depend on the data type [13]. Accordingly, a small α may cause singularity of A in Eq. (6.7), which leads to instability of the solution. Therefore, our experiments started with a sufficiently high α.

To have a smoother registration process, we averaged a number of points from a set with size of 3 ($N_{mean} = 3$). In the implementation of [42], the system of linear equations that arises in each step was solved with the help of the UMFPACK library [60].

6.3.2. TUNING OF ζ

In this section we numerically investigate the percentage of mean mesh quality loss and percentage of the target vertices involved in the nonrigid registration for a domain of ζ. There is an optimum point which maximise the involved vertices and minimises the mesh quality loss. Accordingly, we consider a cost function as follows.

$$\min J = \sum \frac{|\bar{Q}^{final} - \bar{Q}^{0}|}{\bar{Q}^{0}} + \frac{N^{T}_{in}}{N^{T}_{tot}} \tag{6.8}$$

where $\bar{Q}^{final}$ and $\bar{Q}^{0}$ are the average of mesh quality for all vertices on the source mesh before and after registration respectively. Also, N^{T}_{in} is the number of vertices from target employed as corresponding points during the nonrigid registration process, and N^{T}_{tot} is the total number of vertices on the target mesh. As stated in Section 6.2.3, ζ equals to z and has an acceptable range in $[0,\ 3.09]$. The results reported in Figure 6.2 shows best performance where $\zeta = 1.7$.

6.3.3. SCANNER INTRODUCTION

The scanner used in this study is the first prototype of a 4D foot scanner at TUDelft [128]. As shown in Figure 7.1 and reported in Table 6.2, the scanner utilizes seven RealSense D435i depth cameras to capture a 4D foot scan.

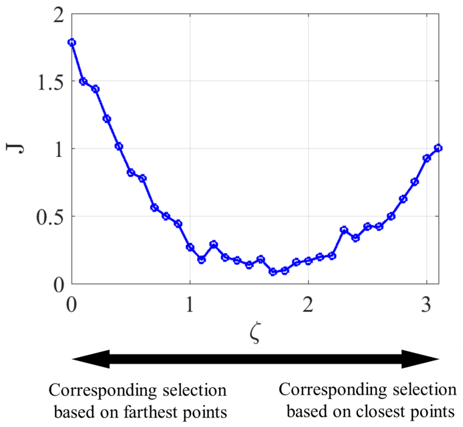


Figure 6.2: Variation of J for a domain of ζ.

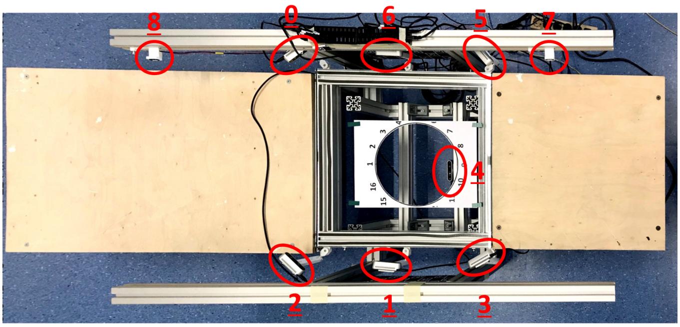


Figure 6.3: Scanner.

6.4. RESULTS

6.4.1. RAW DATA

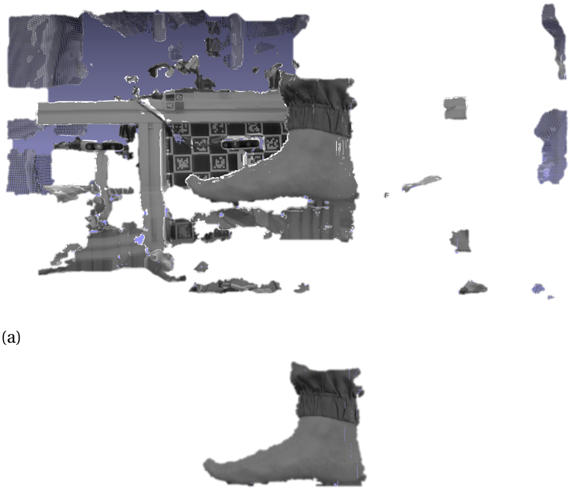
Using the scanner, we were able to capture the geometric shape of the foot as shown in Figure 6.4(a), which is very noisy and the foot part is not the majority of the extracted data. Thus, we manually deleted the spare parts through MeshLab software. In which the result is depicted in Figure 6.4(b).

6.4.2. FRAME SELECTION

During scanning a foot many frames are captured by each camera while the captured frames are not synchronised with each other. In this chapter, we manually looked at the frame by frame extracted with each camera and picked the ones that corresponded to the same posture. We considered five main postures as shown in Figure 6.5 as: 1. Heel strike, 2. Foot-flat, 3. Midstance, 4. Heel-off, and 5. Toe-off. Accordingly, the selected frames from an experiment is shown in Figure 6.6.

Table 6.2: Components of the final product in Figure 7.1.

Part Number	Description	Application
0-6	RealSense D435i	capturing 3D point cloud
7-8	AdaFruit VL53L0X ToF stance sensors	Define start and end of scanning



(a)

(b)

Figure 6.4: Raw data from one of the RealSense depth camera: (a) Uncleaned. (b) Cleaned.

6.4.3. ACCURACY

Employing the discussed technique in Section 7.3, we assess the accuracy of the scanner through two static 3D objects of human foot and human hand as follows.

FOOT MODEL:

The used source mesh is shown in Figure 6.7(a). To evaluate the effectiveness of the introduced algorithm, we scanned a foot with an Artec Eva scanner in a no-load case as shown in Figure 6.7(b), then the scanned foot was printed out via Ulimaker S5 scanner shown in Figure 6.7(c). The printed foot is scanned with our scanner in static mode which is compared with the scan data of using the Artec Eva. Next, some dimensions shown in Figure 6.8 are compared between the two scanners' output. We firstly find the scaling ratio which maps the length of the reconstructed foot to the length of the foot collected with the Artec Eva scanner and then calculate the errors for the other dimensions as width, ball width, and ball angle according to [263]. Based on Figure 6.8, L_r, W_r, BW_r, and α_r explain the length, width, ball width, and ball angle of the scanned foot with the Artec Eva scanner respectively. And, L_e, W_e, BW_e, and α_e define the length and width of the estimated foot through our method respectively. The scaling factor of r is

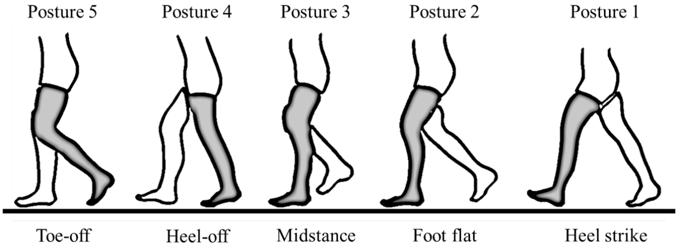


Figure 6.5: Main postures during walking.

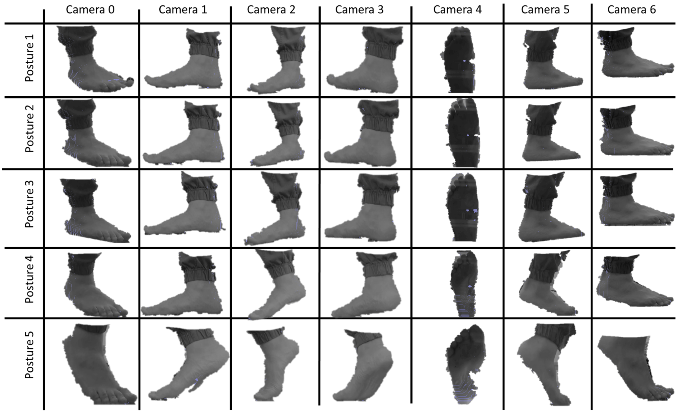


Figure 6.6: Selected frames.

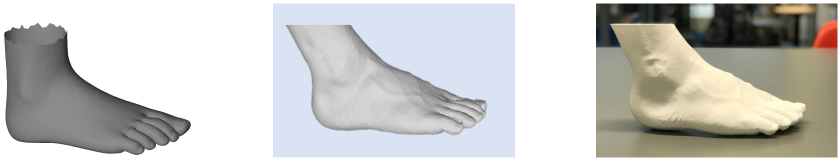
considered as follows,

$$r = \frac{L_r}{L_e} \tag{6.9}$$

where in our experiment, $L_r = 222.2$ mm, $L_e = 0.3561$ mm which results in $r = 623.9$. Thus any dimension extracted from our method is multiplied to the r and compared with the corresponded values on the Artec Eva scanner output summarised in Table 6.3. According to the table, the errors for the width is 3.2 mm, for the ball width is 3.4 mm, and for the ball angle is 0.9°, which shows the average percentage error is about 3.4%.

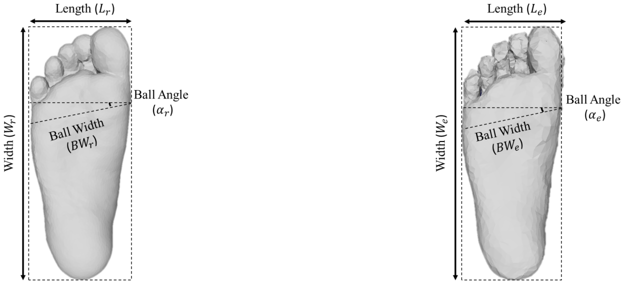
Table 6.3: Error results of the parameters introduced in Figure 6.8 while the scaling ratio $r = 623.9$.

Parameter	Artec Eva	Our approach	MAE
Width (mm)	$W_r = 87.3$	$r \times W_e = 84.1$	3.2
Ball Width (mm)	$BW_r = 92.6$	$r \times BW_e = 89.2$	3.4
Ball Angle	$\alpha_r = 10.8°$	$\alpha_e = 11.7°$	0.9°



(a) (b) (c)

Figure 6.7: No-load foot. (a) The source mesh. (b) Scanned through Artec Eva scanner. (c) Printed Artec Eva output via Ultimaker S5 printer.



(a) (b)

Figure 6.8: Dimension definition: (a) Artec Eva. (b) Our method.

HAND MODEL:

As an extra validation, we used a hand model in [312] as shown in Figure 6.9. According to the figure, the 3D printed hand, the Artec Eva scan of the hand, the used source mesh, and the output of our scanner are depicted in Figure 6.9(a), (b), (c), and (d) respectively. Furthermore, we rigidly aligned the output of our scanner with the Artec Eva scan as shown in Figure 6.9(e) to visually compare the accuracy of the output. To numerically investigate the deviation of our scanner output with the Artec eva scan we figured the histogram of the distance between a point on our scan and its closest point on the Artec eva scan. Accordingly, the average error is about 2 mm which is acceptable in ergonomics study.

6.4.4. POSTURE RECONSTRUCTION

Through the explained method in this chapter, we use seven cleaned frames for each posture as the input of the proposed method. Accordingly, the output is depicted in Figure 6.11 for the bottom view and the side view of the reconstructed scans of one foot shown in Figure 6.6 over three feet we reconstructed in this chapter. The deformation of the source mesh is logically related to the corresponded posture, which shows the acceptable visual performance of the approach. In addition, the deformation of each the foot width is compareable between adjacent frames, where the posture with more load on front of foot has bigger width shown in Figure 6.11. To study the deformation numerically, we reported the introduced dimensions in Section "Accuracy", for different postures per case in Figure 6.12 which shows the deformation direction of all the feet are

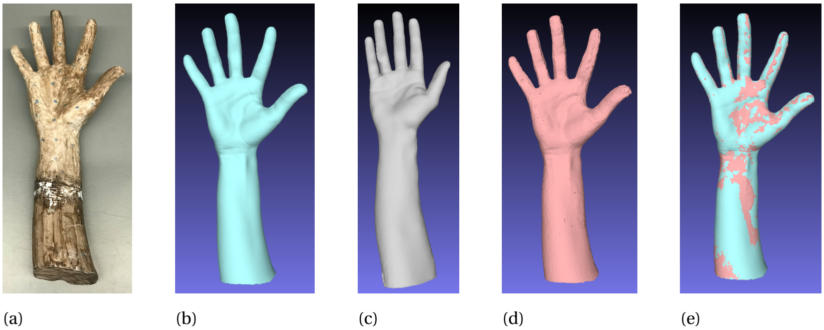


Figure 6.9: Hand evaluation: (a) Real model. (b) Artec Eva scan. (c) Source mesh. (d) Our method scan. (e) Aligned our output to Artec Eva scan.

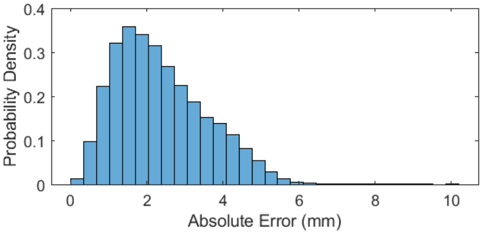


Figure 6.10: Error sparsity histogram.

the same however the amplitude is different. The feet belong to three cases as follows:

- **Case 1:** Male, from Middle-east, 30 years old, with Body Mass Index (BMI) equal to 27.8.
- **Case 2:** Male, from Asia, 25 years old, with BMI equal to 26.8.
- **Case 3:** Male, from Europe, 24 years old, with BMI equal to 21.9.

The figure shows the variation of the foot per case in different postures and please note that in the posture 5, the length of the foot is not the real length and is the distance between tip of the big toe to the heel of the foot, and the bending of the foot is not considered. This resulted in a sudden decrease of the foot length in the figure.

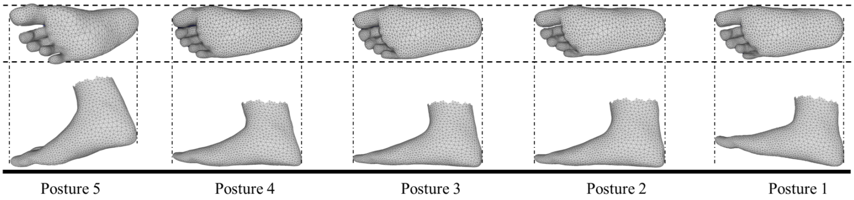


Figure 6.11: Results of posture reconstruction for the case 1.

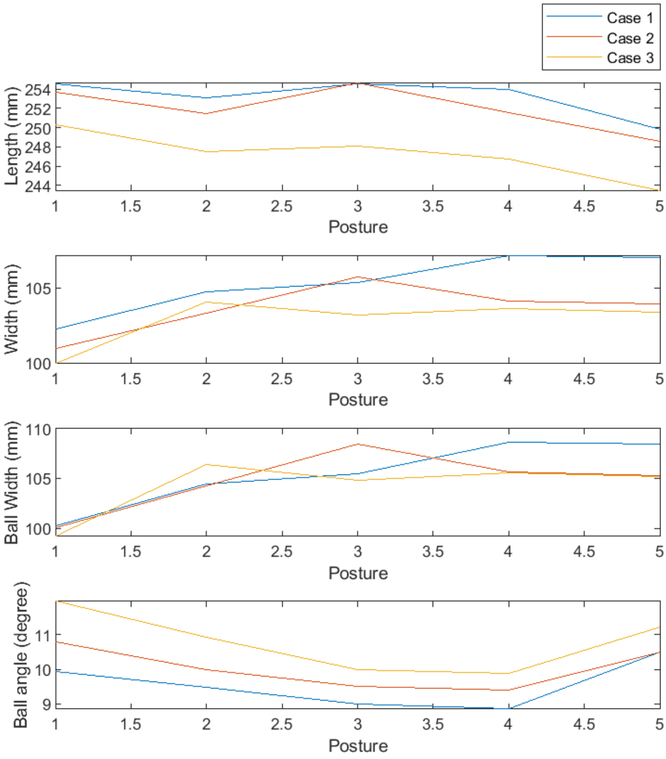


Figure 6.12: Results per different cases.

6.5. CONCLUSION

The method presented in this chapter demonstrates a workflow of extracting any time-varying features from dynamic scans of a deformable object. The semi-automated procedure includes pre-defining points which may be seen by each depth camera, manually cleaning the raw data from each camera, rigid registration of the extracted data from each camera to the corresponded predefined points, and finally non-rigidly registering a source mesh on the point cloud reconstructed through the rigid registration. Reconstructing the final mesh through the non-rigid registration for several frames in a time series provides a 4D feature outcome (for specific points) or 4D trackable scanning (for all the points) which is unique. Comparison of the output of our algorithm with the dimensions extracted from the scanning data of the same object using an Artec Eva scanner shows the estimated dimensions have acceptable error. Further developments include deriving a full automated approach for the filtering, the predefined point selection, and corresponded frame selection, a higher frequency scanner, and integrating temporal super-resolution repetitive motion techniques to have higher resolution of the 4D scanning.

7

4D FEET: WALKING FOOT SHAPES RECONSTRUCTION VIA ADGC-LSTM NETWORK

4D scans of dynamic deformable human body parts help researchers have a better understanding of spatiotemporal features. However, reconstructing 4D scans based on multiple asynchronous cameras encounters two main challenges: 1) finding the dynamic correspondences among different frames captured by each camera at the timestamps of the camera in terms of dynamic feature recognition, and 2) reconstructing 3D shapes from the combined point clouds captured by different cameras at asynchronous timestamps in terms of multi-view fusion. In this chapter, we introduce a generic framework that is able to 1) find and align dynamic features in the 3D scans captured by each camera using the nonrigid iterative closest-farthest points algorithm; 2) synchronize scans captured by asynchronous cameras through a novel ADGC-LSTM-based network, which is capable of aligning 3D scans captured by different cameras to the timeline of a specific camera; and 3) register a high-quality template to synchronized scans at each timestamp to form a high-quality 3D mesh model using a non-rigid registration method. With a newly developed 4D foot scanner, we validate the framework and create the first open-access data-set, namely the 4D feet. It includes 4D shapes (15 fps) of the right and left feet of 58 participants (116 feet in total, including 5147 3D frames), covering significant phases of the gait cycle. The results demonstrate the effectiveness of the proposed framework, especially in synchronizing asynchronous 4D scans using the proposed ADGC-LSTM network.

Parts of this chapter are available in arxive:
Tajdari F, Huysmans T, Yao X, Xu J, Song Y. 4D Feet: Registering Walking Foot Shapes Using Attention Enhanced Dynamic-Synchronised Graph Convolutional LSTM Network. *available in arXiv preprint arXiv:2307.12377* [259]*, and submitted to IEEE Open Journal of the Computer Society.*

7.1. INTRODUCTION

Human movements often lead to large shape deformation of different body parts. 4D scanning, which is able to capture 3D geometric shapes over time, attracted a lot of attention for a better understanding of the dynamic anthropometry in different applications [35, 39], as 4D scans help researchers establish the fundamental understanding of human movements and give valuable information regarding human body deformation while performing different types of activities. Outcomes of research on 4D scans can be applied in many areas, e.g. building virtual avatars, performing (virtual) ergonomics evaluations, developing computer games, designing personal protective equipment, workwear, sportswear, and other practical garments [34].

To acquire 4D scans, multiple (depth) cameras are often used. Those cameras can be synchronized for capturing continuous images at a given moment. However, it is difficult to balance the needed resolutions of the images, the needed time duration, the buffer of the depth cameras, the data transfer rate, the computing power, and the storage [33]. For instance, to capture 640×576 depth images by 6 cameras at 30 frames per second (fps), the needed bandwidth is about 2Gb/s. This poses challenges in the design of a 4D scanning system, especially for a low-cost system. "Dropped frames" are frequently observed in the captured data, mainly due to that the huge amount of to-be-transferred data leads to a nonlinear accumulative delay regarding each camera, even if all cameras are hardware-based synchronized [171]. In a practical case of using 6 Azure Kinect DK cameras for 15 fps 4D scanning, as in Figure 7.1, even when all cameras are hardware synchronized, we found that there are on average 2 milliseconds delays for each frame acquired by those cameras in a 3-seconds scanning session. Note that the delay is accumulative, i.e. at the beginning of the scanning all the cameras' outputs are well-aligned based on their clocks; however, the longer the duration of the scanning is, the more the delay accrues, resulting in a divergence of the geometry in each frame regarding the timestamps.

Researchers developed different methods, on both hardware and software, to tackle those challenges. In the scanning of a human body (parts), a possible approach is to use the prior knowledge of human actions and the associated dynamic features to synchronize the captured frames. In the past decades, recognizing human dynamics features has attracted a lot of attention in the field of computer vision. The developed 3D human action recognition methods can be roughly classified as the RGB video-based approaches [278, 280], skeleton-based methods [223, 225], depth image-based methods [307, 308, 300] and the point cloud-based method [282]. Although the existing methods are proven to be effective in many applications, e.g., video surveillance, human-computer interaction, sports analysis [239, 199], most of them are limited to employ (depth) images as the input, and the recognized 3D actions as the output. Extracting point-to-point correspondences among sequential point clouds from multiple views e.g., cameras, is rarely investigated [282, 203, 31, 153]. On this topic, there are two fundamental challenges: 1) establishing the dynamic connectivity among asynchronous images (scans) captured by different cameras in terms of dynamic feature synchronization i.e. temporal correspondence, and 2) extracting meaningful dynamic features from the combined camera views for accurate analysis of deformation, i.e. multi-view fusion.

In this chapter, using a newly developed low-cost 4D foot scanner based on 6 Microsoft Azure Kinect DK depth cameras, we developed a framework to synchronize and register the captured asynchronous images on significant phases of the gait cycle, resulting in a new open-access 4D Feet data-set of 58 subjects (116 feet). Our main contributions are:

- Establishing and implementing a simple and effective framework to synchronized spatiotemporal asynchronous scans captured from multiple cameras and to track a point's correspondences in all the frames to extract dynamic features of each vertex e.g, velocity;
- Developing a novel Attention Enhanced Dynamic-Synchronised Graph Convolutional (ADGC)-LSTM network to synchronize the dynamic features extracted from different cameras besides existing algorithms;
- Presenting the first 4D mesh-morphed walking foot open-access data-set (4D Feet), as a validation of the proposed framework.

7.2. RELATED WORK

7.2.1. SKELETON-AND-DEPTH-BASED ACTION RECOGNITION

The skeleton-based approach and the depth-based approach are often used in recognizing dynamic features of human actions based on prior knowledge [153]. Regarding skeleton-based 3D action recognition, sequence-based approaches, and graph-based approaches are often used. Via describing the skeleton as a sequence of joints, the sequence-based approaches [223, 225] employed the RNN (Recurrent Neural Network) based methods to extract temporal connectivity among those featured points. The graph-based approaches [144, 318] often utilized GCN (Graph Convolution Network) to exploit spatiotemporal connectivity by considering the skeletal structure as a graph, where the featured points are considered as the points of the graph. Regarding depth-based 3D action recognition, the available methods [187, 307, 184, 279] mainly use the visualization features through 2.5D depth maps. Although both approaches are able to give a reasonably good estimation of the 3D actions that the target subject performed, it is difficult to form a generic framework to fully extract dynamic features based on a few featured points of the moving object, which may cause a reduction in the performance of 3D action recognition.

7.2.2. 3D POINT CLOUDS ACTION RECOGNITION

Deep learning tools play a key role in extracting human actions via 3D point clouds, which is widely employed in recent studies [203, 156, 142, 237, 154, 95, 155, 28]. Among them, PointNet [202] as one of the most recent methods in the area, employs a set of MLPs on each of the individual vertices to identify the unique features. Next, it utilizes a max-pooling layer to generate the global identifier for each point cloud which does not use any geometry-based connectivity of the local neighboring structure. Contrary to these single-frame-based point cloud analysis approaches, in this chapter, we present a simple and effective framework for time series 3D shape reconstruction and action recognition, in which we explicitly use temporal information in the motion stream to identify dynamic features.

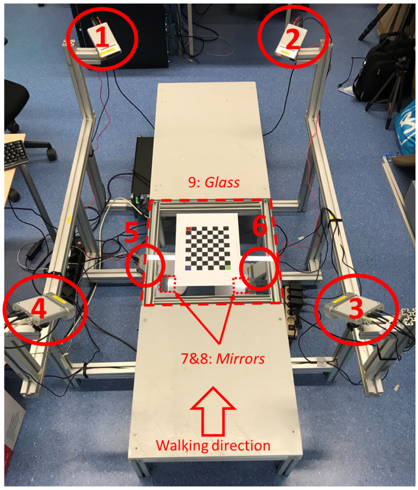


Figure 7.1: The TU Delft 4D foot Scanner.

7.3. METHODOLOGY

In this section, we first present the newly developed 4D foot scanner which is used to capture 4D data. Then, details of the proposed novel ADGC-LSTM network are presented. Based on the hardware and the novel ADGC-LSTM network, we introduce the framework for synchronizing asynchronous 4D scans from different cameras. Finally, we introduce the mesh registration method to along synchronized scans of different cameras at each timestamp.

7.3.1. THE 4D FOOT SCANNER

A 4D foot scanner was developed at TU Delft [128] for acquiring dynamic foot shape data. Figure 7.1 presents the next generation of the 4D foot scanner which utilizes six Microsoft Azure Kinect DK cameras to capture the 4D foot shapes, where four cameras are installed on the top (id 1, 2, 3, and 4) and two cameras are at the bottom (id 5, and 6). To adapt to the minimal focal distance($\sim$ 50 cm) of the cameras at the bottom, two first-surface mirrors (id 7 and 8) were placed on the floor to "fold" the optical path for lowering the height of the scanner for a better user experience. A 9 mm thickness plexglass (id 9) was installed on the footpath to enable capturing the shape of the bottom of the foot while a subject is walking.

The spatial positions and orientations of all cameras were optimized to maximize the resolutions of the captured scans and the intersections of effective view volumes of 6 cameras [128]. To transform the captured data to a global coordinate system, we used a two-sided checkerboard shown in Figure 7.1, and the code in [87] is utilized.

7.3.2. ATTENTION ENHANCED DYNAMIC-SYNCHRONISED GRAPH CONVOLUTIONAL LSTM NETWORK

In the analysis of sequential geometric shapes, many studies suggested that the LSTM, as a transformation of RNN, has a strong capability to understand long-term time dependency of the phenomenons e.g., understanding temporal dynamics of limited points-network (skeleton) sequences. However, using LSTM alone is difficult for incorporating spatial relations in the limited points-network-based action recognition. To this end, AGC-LSTM [225], as an extension of LSTM, was developed to incorporate not only unique features of spatial configuration and temporal dynamics but also the coincident relationships between the spatial domain and temporal domain.

In the process of capturing moving objects (4D scanning), the requirements of the needed movement ranges and the limited views of the cameras are always contradictory factors. This often results in a compromise in the design of the 4D scanner, either with a very small working envelope with limited movements or with sparse points in some of the captured point clouds. Commercial systems may employ more camera modules in 4D scanning; however, at the cost of investment and increased complexity. As the human movements are part of nature and cannot be constrained in a limited range, we target at building correspondences between/among sparse point clouds. Therefore the principles of the graph convolution model which has been broadly employed in sequential data with limited points-network nodes were adopted. Establishing the graph model plays a fundamental role in the graph convolution algorithm. Available graph convolution models e.g, AGC-LSTM, have several limitations for example using single graph structures, illed-correspondences among points, and inadequate discrimination of dissimilar actions. Here we develop a graphic model according to the Dynamic-Synchronised Graph based on the dynamic points, aiming at generating more sparse dynamic features to enhance the capability of the AGC-LSTM model in classifying spatiotemporal features and improve the precision of the action recognitions. The proposed novel method presented here is named Attention Enhanced Dynamic-Synchronized Graph Convolutional LSTM Network (ADGC-LSTM). The details of the algorithm are presented below.

Following the structure of LSTM, the ADGC-LSTM includes three gates: the input gate i_t, forgetting gate f_t, and output gate o_t. The input X_t, hidden state H_t, and cell memory C_t are graph structure data, and the graph structure is generated by the ICFP (Nonrigid Iterative Closest-Farthest Points) algorithm explained in Section 7.3.3.1. The graph convolution operator in the ADGC-LSTM, cell memory C_t, and hidden state H_t can be used to extract temporal dynamics, and include spatial structure information. Figure 7.2(a) describes the structure of an ADGC-LSTM layer. Figure 7.2(b) describes the structure of the ADGC-LSTM unit. Equation (7.1) describes the functions of the ADGC-

LSTM unit.

$$\begin{aligned}
i_t &= \sigma\left(W_{xi} \star gX_t + W_{hi} \star gH_{t-1} + b_i\right)\\
f_t &= \sigma\left(W_{xf} \star gX_t + W_{hf} \star gH_{t-1} + b_f\right)\\
o_t &= \sigma\left(W_{xo} \star gX_t + W_{ho} \star gH_{t-1} + b_o\right)\\
u_t &= \tanh\left(W_{xc} \star gX_t + W_{hc} \star gH_{t-1} + b_c\right)\\
C_t &= f_t \odot C_{t-1} + i_t \odot u_t\\
\hat{H}_t &= o_t \odot \tanh(C_t)\\
H_t &= f_{att}(\hat{H}_t) + \hat{H}_t
\end{aligned} \tag{7.1}$$

where $\star g$ defines the graph convolution operator and $\odot$ defines the Hadamard product. $\sigma(.)$ denotes the sigmoid activation function. u_t denotes the modulated input. $\hat{H}_t$ explains an intermediate hidden state. $W_{xi} \star gX_t$ defines a graph convolution of X_t with W_{xi}. The used graph convolution is the same as the graph convolution employed for the Graph Convolutional Neural (GCN) network in [305] with K number of labels.$f_{att}(.)$ is an attention network that can select the diverse information of key nodes. The output $H_t$reinforces the information of key nodes, without neglecting the information of non-focus nodes, aiming at better integrity of spatial information.

The ADGC-LSTM network logically insists on key nodes by using a soft attention mechanism that automatically quantifies the emphasis level of the key nodes. The importance of the spatial attention network is depicted in Figure 8.3. The intermediate hidden state ($\hat{H}_t$) of ADGC-LSTM contains persistent spatial structure information and temporal dynamics. The state practically improves the selection of the key nodes procedure. In order to guarantee that independent degree weights are established and reinforce the significance of dissimilar nodes for dissimilar types of actions, we employed a query feature as:

$$q_t = \mathrm{relu}\left(\sum_{i=1}^{N} W\hat{H}_{t_i}\right) \tag{7.2}$$

where W defines the trainable parameter matrix, and N is the number of nodes in the graph. Thus the attention scores of all nodes would be specified as:

$$\alpha_t = \mathrm{sigmoid}\left(U_s \tanh\left(W_h\hat{H}_t + W_q q_t + b_s\right) + b_u\right) \tag{7.3}$$

where $\alpha_t = (\alpha_{t_1}, \alpha_{t_2}, \ldots, \alpha_{t_N})$, and U_s, W_h, and W_q are the trainable matrices. b_s and b_u are the bias. A non-linear function $sigmoid$ is employed regarding the probability of selected key joints. The hidden state H_{t_i} of node v_{t_i} is considered as $(1+\alpha_{t_i})\hat{H}_{t_i}$. The attention enhanced hidden state H_t is considered as an input for the next ADGC-LSTM layer. In the final layer of the ADGC-LSTM network, the accumulation of all node features is classified as a global feature F_t^g, and the weighted sum of focused nodes is classified as a local feature F_t^l:

$$F_t^g = \sum_{i=1}^{N} H_{t_i}; \;\; F_t^l = \sum_{i=1}^{N} \alpha_{t_i}\hat{H}_{t_i}. \tag{7.4}$$

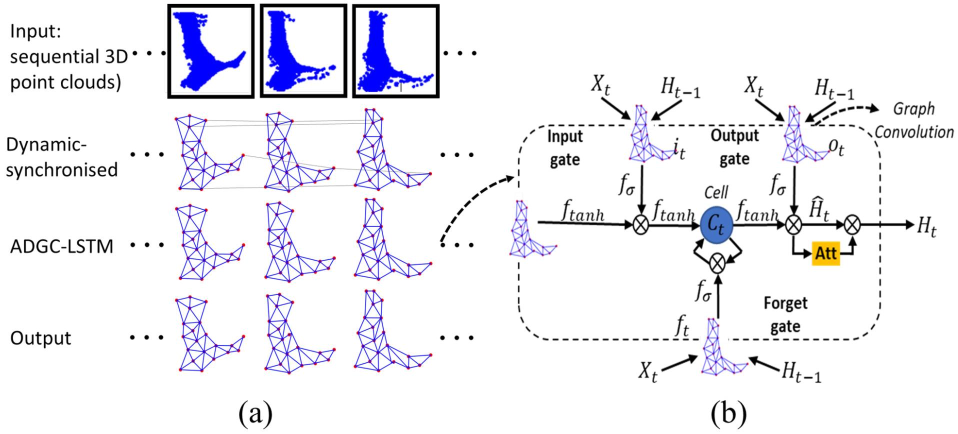


Figure 7.2: The structure. a) One ADGC-LSTM layer; b) One ADGC-LSTM unit adapted from [225].

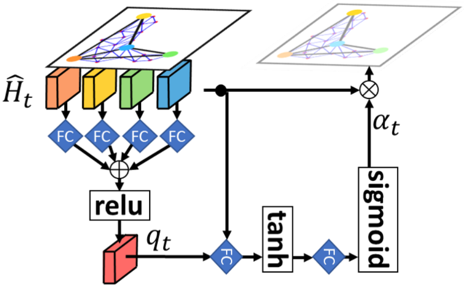


Figure 7.3: Illustration of spatial attention mechanism principal adapted from [225].

7.3.3. TIME SYNCHRONIZATION

In Figure 8.6, the workflow of the proposed time synchronization method is presented based on the data captured by the 4D scanner and the proposed novel ADGC-LSTM network. In the following, Section 7.3.3.1 focuses on explaining the process of synchronizing data for each camera as the first step in the workflow. Then in Section 7.3.3.2, the aligned data from all cameras is synchronized using the proposed ADGC-LSTM network. Section 7.3.3.3 describes details in the implementation of ADGC-LSTM regarding scans from different cameras.

CONSTRUCTION OF THE DYNAMIC-SYNCHRONISED GRAPH BASED ON THE DYNAMIC POINTS

After scanning, each camera gives a set of time series 3D point clouds, and there are not logical correspondences among them. This prevents us to explain any dynamic features between the frames as the correspondence of points from one frame to the other frame, known as dynamic points, does not exist. Figure 7.6 (the first row of each sub-figure) presents this "lack of correspondence", where we selected two points (highlighted with red and green colors) in the first frame of each camera, and tracked these points in the

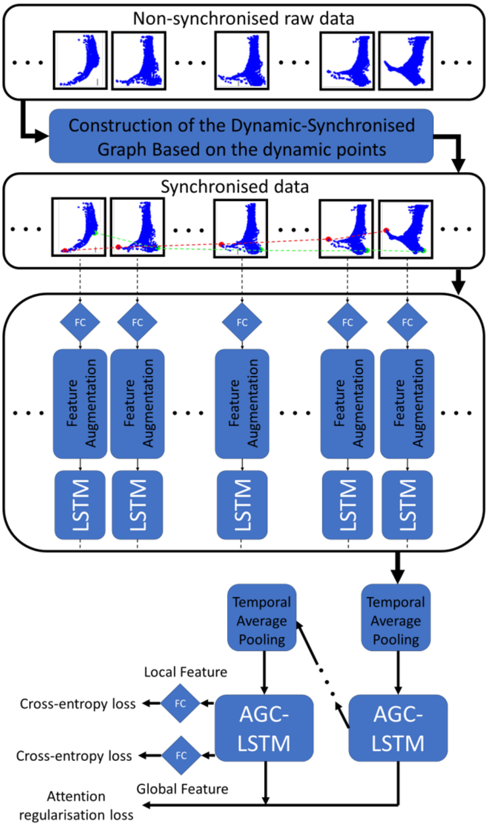


Figure 7.4: Time synchronization procedure.

rest frames using point ids in the acquired point clouds. To be able to have meaningful dynamic features between frames of a camera needed as the key nodes used for the ADGC-LSTM network in Section 7.3.2, we established the correspondences of points using a novel extended version of the Nonrigid Iterative Closest-Farthest Points (ICFP) scheme [250] which guarantees to find proper corresponded points in a limited number of iterations from a Source mesh ($\mathbb{S}$) to a Target mesh ($\mathbb{T}$). In the process of finding correspondences from each point on $\mathbb{T}$ to $\mathbb{S}$, initially, each point on $\mathbb{S}$ may have multiple corresponding points on the $\mathbb{T}$. In this case, we logically select either the closest or the farthest point. In each iteration of the registration process, a boundary distance (l) in (7.5) is defined as the corresponding distance matrix from $\mathbb{T}$ to $\mathbb{S}$.

$$l = m + \zeta\sigma \tag{7.5}$$

where m and σ are the mean and standard deviation of the counted distances for the correspondences from $\mathbb{T}$ to $\mathbb{S}$, respectively. ζ is the probability indicator in [250] regulates the closest-farthest point selection and has an acceptable range in [0, 3.09] refer to Table III of the book in [82]. In [250], the ζ is a predefined constant variable that is nu-

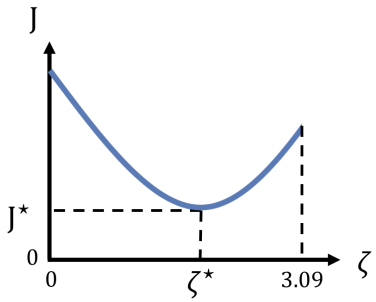


Figure 7.5: Assumed $J-\zeta$ shape.

merically defined based on registering a source foot on only one target foot to minimise a concave parabolic cost function (J) including two terms of percentage of mean mesh quality loss and percentage of the target vertices involved in the nonrigid registration. However, there is no guarantee that the selected ζ results in the minimization of the cost function for any other target foot. Thus, here we extend the corresponding selection criterion by designing an adaptive ζ finds the minimum cost function by iterations and can be implemented on any other registering shapes. Assuming J from [250]

$$\min J = \sum \frac{|\bar{Q}^{final} - \bar{Q}^{0}|}{\bar{Q}^{0}} + \frac{N_{in}^{T}}{N_{tot}^{T}} \tag{7.6}$$

where $\bar{Q}^{final}$ and $\bar{Q}^{0}$ are the average of mesh quality for all vertices on the source mesh before and after registration respectively. Also, N_{in}^{T} is the number of vertices from target employed as corresponding points during the nonrigid registration process, and N_{tot}^{T} is the total number of vertices on the target mesh.

7.3.3.1.1 Assumption: For the design of the estimator, we formulate a parabolic $J-\zeta$ relationship. In particular, we employ the following function describing the $J-\zeta$ relationship, also depicted in Figure 7.5,

$$J = a\zeta^{2} + b\zeta + c, \tag{7.7}$$

where $a \in \mathbb{R}_{>0}$ and $b \in \mathbb{R}_{>0}$ are unknown parameters, and C is equal to $J(0)$ defined as initial condition of J which is assumed known; function (7.7) has a minimum point $(\zeta^{\star}, J^{\star})$ as

$$\zeta^{\star} = \frac{-b}{2a}; \qquad J^{\star} = \frac{-b^{2}}{4a} + J(0). \tag{7.8}$$

7.3.3.1.2 Adaptive ζ design: By replacing the nominal values of $J^{\star}$ and $\zeta^{\star}$ in (7.7), the error of J from $J^{\star}$ is

$$J - J^{\star} = a(\zeta^{2} - \zeta^{\star 2}) + b(\zeta - \zeta^{\star}). \tag{7.9}$$

Now, we introduce the integral error states

$$E_q = \int \left(J - J^{\star}\right) dt; \quad E_{\zeta} = \int \left(\zeta - \zeta^{\star}\right) dt, \tag{7.10}$$

allowing to define the integral error system

$$\dot{E}_J = q - q^{\star}; \quad \dot{E}_{\zeta} = \zeta - \zeta^{\star}, \tag{7.11}$$

that can be reformulated as

$$\dot{x} = B_e u_e + r_e, \tag{7.12}$$

where

$$x = \begin{bmatrix} \int J\, dt \\ \int \zeta\, dt \end{bmatrix}, u_e = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \zeta^2 - \zeta^{\star 2} \\ \zeta - \zeta^{\star} \end{bmatrix} \tag{7.13}$$

$$B_e = \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}, r_e = \begin{bmatrix} J^{\star} \\ \zeta^{\star} \end{bmatrix}. \tag{7.14}$$

We propose controlling system (7.12) using MRAC [229], which let us simultaneously identify the unknown parameters a and b (both appearing in B_e) and minimise the tracking error. In order to proceed, we introduce the feedback control law (see, e.g., Chapter 1 in [277])

$$u_e = -\hat{\Pi}(x - r_e), \tag{7.15}$$

where $\hat{\Pi}$ is an unknown matrix that needs to be estimated. We then introduce a model reference

$$\dot{x}_M = -A_M x_M + B_M r_e, \tag{7.16}$$

where A_M is designed as a positive definite matrix and B_M is an arbitrarily defined matrix in which they guarantee stable model reference dynamics. Let us define the error between the integral states and the model reference $e = x - x_M$, whose dynamics are defined as (see [229])

$$\begin{aligned}
\dot{e} &= \dot{x} - \dot{x}_M \\
&= B_e(-\hat{\Pi}x + \hat{\Pi}r_e) + r_e + A_M x_M - B_M r_e + A_M x - A_M x \\
&= -A_M(x - x_M) + B_e(-\hat{\Pi} + A_M)x + B_e(\hat{\Pi} - \frac{B_M - I}{B_e})r_e \\
&= -A_M e + B_e(-\hat{\Pi} + A_M)x + B_e(\hat{\Pi} - \frac{B_M - I}{B_e})r_e
\end{aligned} \tag{7.17}$$

Knowing $v = \begin{bmatrix} x \\ r_e \end{bmatrix}$, the error dynamic system would be

$$\begin{aligned}
e &= -A_M e + \begin{bmatrix} -B_e\hat{\Pi} + B_e A_M & B_e\hat{\Pi} - B_M + I \end{bmatrix} v \\
&= -A_M e + \tilde{\phi} v
\end{aligned} \tag{7.18}$$

where $\tilde{\phi} = \begin{bmatrix} -B_e\hat{\Pi} + B_e A_M & B_e\hat{\Pi} - B_M + I \end{bmatrix}$. We assume that $\tilde{\phi} = \hat{\phi} - \phi$, where $\hat{\phi}$ is an estimating matrix, namely $\hat{\phi} = \begin{bmatrix} -B_e\hat{\Pi} & B_e\hat{\Pi} \end{bmatrix}$ and ϕ is the unknown constant matrix, namely $\phi = \begin{bmatrix} -B_e A_M & B_M - I \end{bmatrix}$ defines $\dot{\phi} = 0$. Thus,

$$\dot{\tilde{\phi}} = \dot{\hat{\phi}}. \tag{7.19}$$

We can observe that the error dynamic of (7.18) is bounded over time if $\tilde{\phi}$ is bounded, and the error is asymptotically stable if $\tilde{\phi}$ converges to zero, considering that $-A_M$ is selected as a stable matrix with negative eigenvalues, while B_M and B_e are constant matrices.

In order to study the convergence of $\tilde{\phi}$ to zero, a Lyapunov function is employed as follows:

$$\mathcal{V} = e\mathscr{P}e^\top + \tilde{\phi}\Gamma^{-1}\tilde{\phi}^\top, \tag{7.20}$$

where $\mathscr{P} \geq 0$ and $\Gamma > 0$ imply that $\mathcal{V} > 0$. In order to guarantee stability, it is sufficient if $\dot{\mathcal{V}} \leq 0$, then

$$\frac{d\mathcal{V}}{dt} = \dot{e}\mathscr{P}e^\top + e\mathscr{P}\dot{e}^\top + \dot{\tilde{\phi}}\Gamma^{-1}\tilde{\phi}^\top + \tilde{\phi}\Gamma^{-1}\dot{\tilde{\phi}}^\top. \tag{7.21}$$

By replacing $\dot{e}$ from (7.18), and considering (7.19), we obtain

$$\frac{d\mathcal{V}}{dt} = -A_M e\mathscr{P}e^\top - e\mathscr{P}e^\top A_M^\top + 2e\mathscr{P}v^\top\tilde{\phi}^\top + 2\dot{\hat{\phi}}\Gamma^{-1}\tilde{\phi}^\top. \tag{7.22}$$

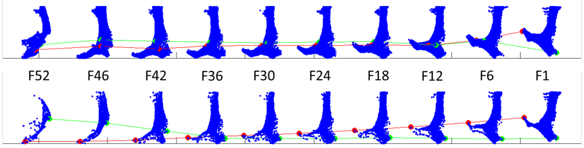
As A_M is positive definite, $-A_M e\mathscr{P}e^\top - e\mathscr{P}e^\top A_M^\top$ is negative semi-definite matrix, thus $\frac{d\mathcal{V}}{dt} \leq 0$ if and only if

$$2e\mathscr{P}v^\top\tilde{\phi} + 2\dot{\hat{\phi}}\Gamma^{-1}\tilde{\phi}^\top = 0, \tag{7.23}$$

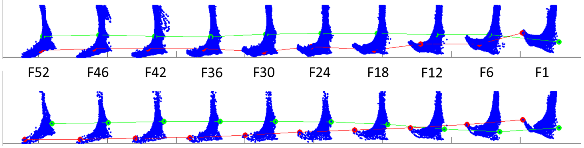
which is a sufficient condition for stability where the changes in the estimating unknown matrix $\hat{\phi}$ is

$$\dot{\hat{\phi}} = -e\mathscr{P}v^\top\Gamma, \tag{7.24}$$

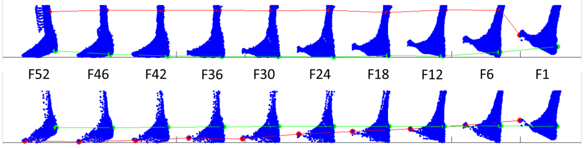
where Γ is known as the *growth rate* of the estimation law. Using (7.24), the instruction of ϕ and $\hat{\phi}$, and $\tilde{\phi} \to 0$, we may conclude that $\hat{\Pi} \to A_M$, and $B_e \to (B_M - I)\hat{\Pi}^{-1}$, which results in $\zeta \to \zeta^\star = \frac{-B_{e_{1,2}}}{2B_{e_{1,1}}}$ from (7.8), and (7.14). Thus, having $\zeta \to \zeta^\star$ and for a point on $\mathbb{S}$, if a number of points on $\mathbb{T}$ are selected, we consider the point with the largest distance among the selection, if all the distance for the population is greater than the l. Otherwise, we select the closest point as the corresponding point to the point on $\mathbb{S}$. The ICFP scheme was used to find the available correspondences for all frames captured by a single camera e.g, top-rows of the sub-figures in Figure 7.6. In the implementation, we use the ICFP to match each consequent pair of frames (e.g. i^{th} frame and $(i+1)^{th}$ frame), starting from the first frame to the last frame, e.g. for 100 captured frames, 99 pairs were used to generate the correspondences matrix. Apparently, not all points in a frame have correspondences in the neighboring frames, as a new frame may not be able to capture all the points captured in the previous frame e.g, from comparing Figure 7.6(d)-top with Figure 7.6(d)-bottom after the 42^{th} frame (F42) the density of Dynamic-Synchronised Graph is reduced. Thus, some frames with very low-density point-clouds are skipped due to the lack of correspondences.



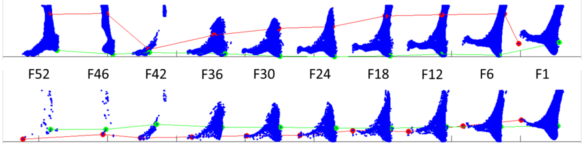
(a) Camera 1



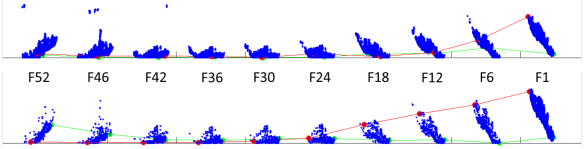
(b) Camera 2



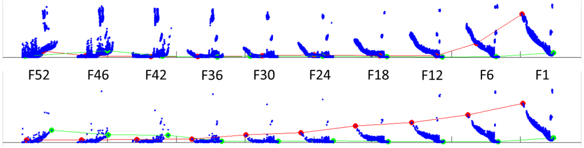
(c) Camera 3



(d) Camera 4



(e) Camera 5



(f) Camera 6

Figure 7.6: Top: Raw data; Bottom: Dynamic-Synchronised Graph as the key nodes.

ADGC-LSTM NETWORK

In this section, an end-to-end attention enhancement Dynamic-Synchronised Graph Convolution LSTM network (ADGC-LSTM) for points-network-based action behavior recognition, i.e. human walking, is explained. The captured data of the moving object from each camera can be described as a unique class that has considerable overlap(s) with other classes with respect to the configuration of the cameras. All the ADGC-LSTM networks designed for each Dynamic-Synchronised Graph in this chapter also have similar structural characteristics. First, the linear layer and the shared LSTM layer were employed to collect the feature information for each of the graph. Then, the feature information of each graph was fed into the proposed three layers of the ADGC-LSTM as node depiction to consider the spatiotemporal features in the model.

7.3.3.2.1 Dynamic-synchronised Graph Model Based on Human feet dynamic points: Firstly a linear layer and LSTM layer were employed to convert the 3D coordinate of each key node into a high-dimensional feature space regarding the key node-network sequence. The preliminary linear layer maps the 3D coordinates onto a 256-dimensional vector, as the geometric features P_t, i.e., P_{ti} defines the geometry feature of key node i. As it includes only geometry information, P_{ti} is effective to proceed with the learning process regarding spatial structure features in graph models. The differential feature V_{ti} between two sequential frames, facilitates the dynamic feature understanding used to train the ADGC-LSTM. The sequential group of features is able to explain a more sparse domain of feature information better, while the differential of the features is more sensitive to the changes of the feature vectors. Thus, the LSTM layer was utilized to avoid having unnecessary sensitivity between the sequential feature groups. Equation (7.25) presents this proposition.

$$\begin{aligned} E_{ti} &= f_{lstm}\Big(concat\big(P_{ti}, V_{ti}\big)\Big) \\ &= f_{lstm}\Big(concat\big(P_{ti}, P_{(t-1)i}\big)\Big) \end{aligned} \tag{7.25}$$

where E_{ti} is the augmented featured of key node i at time t.

7.3.3.2.2 Learning of the ADGC-LSTM: Finally, the global feature F_t^g and local features F_t^l at each timestamp were converted to scores o_t^g and o_t^l of each class. According to (7.1), the predicted probability of the i^{th} class can be obtained as:

$$\hat{y}_{ti} = \frac{e^{o_{ti}}}{\sum_{j=1}^{C} e^{o_{ti}}}, i = 1, \cdots, C \tag{7.26}$$

In the training process, taking into account the hidden state of each time interval, the ADGC-LSTM includes short-term dynamics and the loss function with the structure in (7.27), extracted to the train model as:

$$\begin{aligned} L = &-\sum_{t=1}^{T_3}\sum_{i=1}^{C} y_i \log \hat{y}_{ti}^{g} - \sum_{t=1}^{T_3}\sum_{i=1}^{C} y_i \log \hat{y}_{ti}^{l} \\ &+ \bar{\lambda}\sum_{j=1}^{3}\sum_{n=1}^{N}\left(1 - \frac{\sum_{t=1}^{T_j}\alpha_{tnj}}{T_j}\right)^2 + \bar{\beta}\sum_{j=1}^{3}\frac{1}{T_j}\sum_{t=1}^{T_j}\left(\sum_{n=1}^{N}\alpha_{tnj}\right)^2 \end{aligned} \tag{7.27}$$

where $y = (y_1, \cdots, y_c)$ is the ground-truth label. T_j denotes the number of time intervals on the j^{th} ADGC-LSTM layer. The third term is considered to emphasize equally to variation of featured points. The final term is to restrict the number of interested nodes. $\bar{\lambda}$ and $\bar{\beta}$ are weight decaying coefficients.

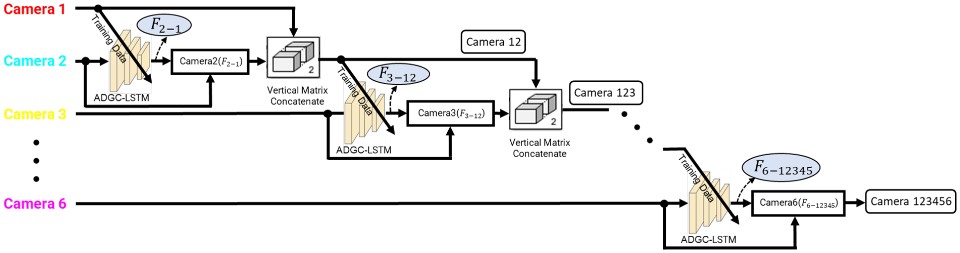


Figure 7.7: Hierarchical learning-synchronization process.

IMPLEMENTATION OF ADGC-LSTM

To synchronize the captured frames between each pair of cameras, we use the generated Dynamic-Synchronised Graph of each camera. We use one camera's Dynamic-Synchronised Graph as the supervisor to train our ADGC-LSTM Network and the other one for validation. In this case, we use a hierarchical learning process to have the maximum overlap between cameras with the shown framework in Figure 7.7. In the figure, firstly we synchronize Camera 2 with Camera 1 (where the corresponding frames of Camera 2 to Camera 1 is F_{2-1}) and name the overall point cloud as Camera 12 (Camera 1 with synchronized Camera 2). Then synchronize Camera 3 with Camera 12 (with frame set of F_{3-12}) and name it as Camera 123. Then we continue with Camera 4, Camera 5, and finally Camera 6, to have all the cameras synchronized based on Camera 1.

7.3.4. MESH REGISTRATION

Based on the established correspondences, a cost function based on Tajdari et. al [251] is defined for registering meshed at each time step. Tajdari et. al [251] proposed the non-rigid registration formulation as a combination of distance (W, D, U), stiffness (M, G), and semi-curvature (W_c, A_c, B_c) terms summarised in the following formula

$$\begin{aligned} E(X) &= \| \begin{bmatrix} \alpha M \otimes G \\ WD \\ \beta W_c A_c \end{bmatrix} X - \begin{bmatrix} 0 \\ WU \\ \beta W_c B_c \end{bmatrix} \|_F^2 \\ &= \| AX - B \|_F^2 \end{aligned} \tag{7.28}$$

where, The sparse matrix D is formed to facilitate the transformation of the source vertices with the individual transformations contained in X via matrix multiplication, and denoted as $D = diag(v_1^T, v_2^T, \ldots, v_n^T)$, where $v_i \in \mathbb{S}$ and $i = 1, \ldots, n$, and n is the number of vertices on the $\mathbb{S}$. W is a diagonal matrix consisting of weights w_i. α is the stiffness constraint. To regularise the deformation, an additional stiffness term is introduced. Using the Frobenius norm $\|.\|_F$, the stiffness term penalizes the difference of the transformations of neighboring vertices, through a weighting matrix $G = diag(1, 1, 1, \gamma)$. During the deformation, γ is a parameter to stress differences in the skew and rotational part against the translation part of the deformation. The value of γ can be specified based on data units and the types of deformation [13]. The node-arc incidence matrix M (e.g. Dekker [61]) of the template mesh topology is employed to convert the stiffness term into the

Table 7.1: The anthropometric data.

Sex	Age	Shoe size	Height	Weight	BMI
♀	24.0±5.1	37.5±1.5	161.1±9.0	55.9±9.4	21.5±2.6
♂	26.2±6.4	42.9±1.6	178.8±8.8	73.3±11.7	22.9±3.0

matrix form. As the matrix is fixed for directed graphs, the construction is one row for each edge of the mesh and one column per vertex. To establish the node-arc incidence matrix of the source topology, the indices (i.e. the subscripts) of edges and vertices are addressed, for any edge of r which is connected to vertices (i, j), in r^{th} row of M, and the nonzero entries are $M_{ri} = -1$ and $M_{rj} = 1$.

7.4. EXPERIMENT SETUP

7.4.1. DATA-SET

OUR DATA-SET

Using the proposed 4D scanner and the novel framework, we tried to build an open-access data-set of 4D feet data regarding significant phases of the gait cycle such as initial contact, foot flat, midstance, heel lift, and toe-off. An experiment was designed and approved by the local human research ethical committee. In the experiment, after a brief explanation, participants first read and signed the consent forms. Subjects under 18 had their consent forms signed by their parents/legal guardians. Then each subject was guided to walk through the glass bridge with his/her bare feet twice regarding the left and the right feet, respectively. Both feet of 59 subjects (26 females (♀) and 33 males (♂)) were scanned while the data of participant 53 was not saved and was excluded from the data-set, resulting in a data-set with 58 subjects. Among them, 55 subjects are right-handed and the rest are left-handed. The age of the population ranges from 6 to 50 years old where the mean age is 24 for females (♀) and 26.2 for males (♂). Their normal shoe sizes range from 32 to 46 (European sizes, 20-29.3 CM). To be more inclusive and address the diversity of the population, we invited subjects from different countries such as The Netherlands, Belgium, Italy, Spain, Latvia, Slovenia, Swaziland, Turkey, Iran, India, Thai, China, Japan, Costa Rica, Mexico, Cameroon, Nigeria. The anthropometric data of the population can be found in Table 7.1.

DATA-SET FOR REGISTRATION

In the experiment, both the right and the left feet shapes in data-set number 25 in the SHREC'14 data-set [196] were selected as the source surface. Before the experiment, the meshes of both feet were pre-processed for a more uniform mesh using ACVD, a freely available software provided by Valette et al. [271]. The acquired two meshes, each has 5000 vertices, were used as the inputs of the experiment as the source meshes for the nonrigid registration regarding the left and the right foot, respectively.

7.4.2. METHODS FOR COMPARISON

We compare the proposed methods in the framework with the following methods with similar state-of-the-art [98]:

- ARIMA [287]: Auto-Regressive Integrated Moving Average method is one of the well-known methods to anticipate the future values in a time sequential data-set.
- VAR [320]: Vector Auto-Regressive finds the pairwise connectivity between time-sequential data-sets.
- LSTM [103]: Long-Short Term Memory network, is a variant of RNN network.
- GRU [53]: Gated Recurrent Unit network, is a specific RNN network.
- STGCN [305]: A Spatial-Temporal Graph Convolution model is developed based on automatic learning of both the spatial and temporal patterns.
- GeoMAN [149]: A multi-level attention-based RNN model aimed for the geo-sensory time sequential anticipation problem.

Root mean square error (RMSE) of the geometry based on closest points is used as the metric.

7.4.3. ADGC-LSTM PARAMETERS' CONFIGURATION

In the experiments, a fixed length of $T = 40$ is used in (7.27) from each graph sequence as the input. Regarding the ADGC-LSTM, we assumed the neighbor set of each node includes only nodes directly connected with itself. Regarding a fair comparison with ST-GCN [305], the graph labeling function in ADGC-LSTM divides the neighbor set into $K = 3$ subsets according to [305]. In the training process, the Adam optimizer [117] is employed to optimize the network. The parameters of $\bar{\lambda}$ and $\bar{\beta}$ are set to 0.01 and 0.001, respectively. We set the initial learning rate to 0.0005 which is reduced in every 15 epochs by multiplying 0.1 to the learning rate. In addition, we discretize the parameter estimation formula in (7.24) by considering $\dot{\hat{\phi}} = \frac{\hat{\phi}(k+1)-\hat{\phi}(k)}{\Delta k}$, $k \in \mathbb{Z}_{\geq 0}$; then, knowing that $\Delta k = 1$ as k is a sequentially increasing index (the index of intervals in the registration process), the estimation rule (7.24) turns into

$$\hat{\phi}(k+1) = \hat{\phi}(k) - e(k)\mathscr{P} v^{\top}(k)\Gamma, \tag{7.29}$$

where we assumed $\Gamma = 0.8 I_{4\times 4}$, and $\mathscr{P} = 1$. Regarding the used mesh registration method in Section 7.3.4, we use the same parameter values in [251] regarding (7.28).

7.5. RESULTS

7.5.1. MOTION SYNCHRONISATION

To evaluate the effectiveness of the synchronization, we developed a K-fold-like scheme where: 1) we used the mean Closest Points Geometry Distance (CPGD) [13] values between adjacent point clouds as the metric and 2) for each camera, we compared its synchronized scans to the merged results of other 5 cameras at each timestamp. That is, in the i^{th} frame and after synchronization with each of the aforementioned methods in Section 7.4.2, we exclude the j^{th} camera points from the complete foot and calculate the CPGD of the camera j^{th} points with the remaining points. We repeat this process for all other cameras and the average values of errors are presented in Table 7.2. According to

Table 7.2: RMSE results of the comparison based on Closest Points Geometry Distance (CPGD), and Percent of Improvement (PI) comparing to the raw data, for the left and right foot.

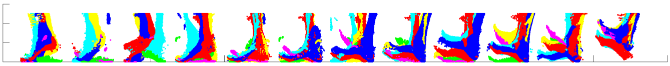
Method	Left foot		Right foot	
	CPGD (cm)	PI(%)	CPGD (cm)	PI(%)
raw data	7.01	–	7.35	–
ARIMA	6.81	2.8	5.95	19.1
VAR	3.24	53.8	4.22	42.6
LSTM	1.62	76.9	1.58	78.5
GRU	1.71	75.6	1.59	78.3
STGCN	1.21	82.7	1.14	84.4
GeoMAN	1.13	83.8	1.03	85.9
Our	**0.64**	**90.8**	**0.72**	**90.2**

the table, our proposed method outperforms all the other methods for the both the left and right feet in the data-sets.

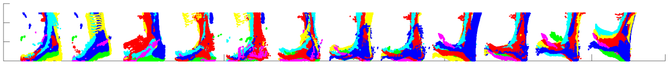
According to Table 7.2 and Figure 7.8, one can be seen is that generally the output of the non-deep learning methods e.g., ARIMA and VAR, demonstrate a higher error than the deep learning methods e.g., LSTM, GRU, STGCN, GeoMAN. This is investigated numerically and the results are presented in Table 7.2, which shows that the deep-learning methods could averagely improve the performance for about 80% in terms of PI, revealing the limited abilities of the non-deep-learning methods to tackle non-linearity and complexity in time series analysis. Among the deep-learning methods, the models that simultaneously consider temporal and spatial correlations, e.g. STGCN, GeoMAN, and the proposed method, outperform other deep-learning-based methods including LSTM and GRU for about 11% in terms of PI. Where, GeoMAN slightly outperforms STGCN in terms of PI, defining that the multi-level attention mechanisms employed in GeoMAN enhance finding the correlation among dynamic features of the feet. Our ADGC-LSTM network, achieved better results than other included state-of-the-art methods, confirming the performance of the proposed method in describing spatial-temporal features of the walking foot.

7.5.2. ESTABLISHING REGISTRATION-BASED DATA-SET

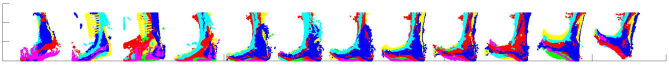
Through the explained method in Section 7.3.4, we register the synchronised frames in Section 7.4.1.2 at each time step to establish a mesh morphed 3D geometry as Figure 7.9. In this regard, we can track not only the geometry of any point but also the dynamic features of the point such as velocity, and acceleration. In addition, we can compare the geometry or dynamic features of any points among all captured feet shapes. To this end, we numerically investigated the deformation variation of a few well-known foot dimensions in Table 7.3. The dimensions are length (L_f), width (W_f), and ball width (BW_f) according to [263] used in [250], and their variations are ΔL_f, ΔW_f, and ΔBW_f. Where the operator Δ defines the differences between the maximum and minimum value of the dimension for a participant during walking. By calculating the average foot length (L_{ave}) of all the feet in our data-set (both left and right feet) as 24.3 cm, we can see from



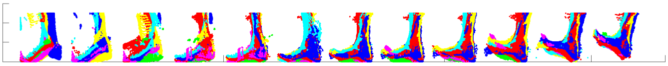
(a) ARIMA



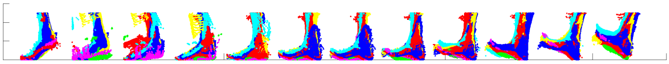
(b) VAR



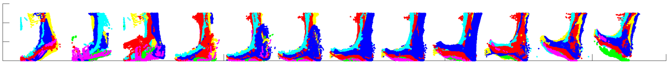
(c) LSTM



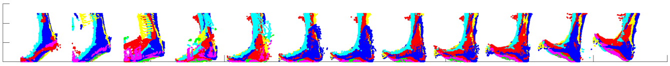
(d) GRU



(e) STGCN



(f) GeoMAN



(g) Our

Figure 7.8: Results of time synchronisation with different methods.

Table 7.3 that the variation of L_f is about 3% of L_{ave}, and W_f and BW_f are about 5% of L_{ave}, which are a considerable variation and highlights the importance of 4D scanning, and 4D studying of human actions.

7.5.3. DATA-SET COMPARISON

To the best of our knowledge and referring to Sections 7.1 and 7.2, there are few articles that developed a 4D data-set based on a software-based frame time-synchroniser, while we recognised a few works with similar state-of-the-art results summarised in Table 7.4 as Walking Foot [250], 4DComplete [146], Dynamic foot [36], and SURP [188]. According to the table, we compare the presented data-sets in the works with the results of our work in this paper using several matrices: Number of objects, Number of cameras, speed, total frames, time-delay synchronisation, and accuracy.

Briefly, the work in [250] presents a step-by-step semi-automated framework to reconstruct a full walking foot using 7 RealSense cameras, 4DComplete [146] includes animation sequences of animals and humans body, the work in [36] introduced a human feet data-set based on a parametric statistical shape model, and SURP [188] is a data-set including different human body part and full human body shapes.

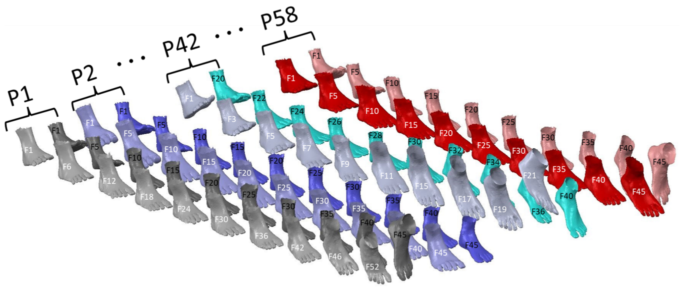


Figure 7.9: 4D Feet. We present a new 4D data-set of 58 Participants (P1,···,P58 in the figure), including 5147 frames of 3D scans. The raw 3D scans (meshes) were collected at 15 fps through a novel 4D foot scanner including 6 Azure Kinect DK cameras. Then we showed how to synchronize the cameras through a novel deep-learning-based framework, and establish a mesh-morphed data-set.

Table 7.3: Foot dimensions variation results.

Parameter	Left foot	Right foot
ΔL_f (cm)	0.71±0.60	0.69±0.62
ΔW_f (cm)	1.1±0.9	1.0±0.9
ΔBW_f (cm)	1.2±0.9	1.1±1.0

According to Table 7.4, the work in [250] presents more accuracy than our work (due to manual filtering and time-delay synchronisation), Dynamic foot [36] shows higher speed than our work, and SURP [188] has more total frames than our work; however, the framework we introduced is fully automated, especially for time-delay synchronisation which is completely novel, and the data-set we presented is comparatively including more than four times objects than the other works. In addition, excluding the semi-automated work in [250], our work outperforms the other compared works in accuracy for an average of 45% which is a considerable achievement.

7.6. CONCLUSION

In this chapter, we proposed a generic framework to synchronize and register asynchronously captured point clouds of a moving and deforming object, namely the hu-

Table 7.4: Comparison of available data-sets.

Work in	Number of objects	Number of cameras	Speed (fps)	Total frames	Time-delay synchronisation	Accuracy (cm)
Walking Foot [250]	3	7 (RealSense)	5	15	Manually	0.2
4DComplete [146]	31	1 (unknown)	unknown	1.972 k	not applicable	3.74
Dynamic foot [36]	30	6 (RealSense)	90	1.771 k	Manually	1
SURP [188]	Not available	8 (3DMD)	10	1200 k	Manually	0.9
Ours	116	6 (Azure Kinect)	15	5.147 k	Automatically	0.68

man foot, through a novel ADGC-LSTM-based network and a non-rigid registration algorithm. We implemented the framework on the data captured from a novel 4D foot scanner to acquire the first 4D open-access feet data-set with the focuses on 1) finding the dynamic connectivity among 3D scans captured at different timestamps of each camera in terms of dynamic feature synchronization and 2) extracting meaningful dynamic features from the combined views of multiple cameras for estimating the amplitude of the deformation. Experiment results show that our method improved the synchronization process on average by about 30% compared to other state-of-the-art methods. Meanwhile, the quality of the acquired 4D scan was comparatively high regarding the deformation of each part of the foot, and such information can be useful in different applications, e.g. footwear design. Further developments include establishing a 4D Statistical Shape Model (SSM) of human foot as a tool to study the gait and foot deformations. Also, due to inconsistency in capturing speeds of different cameras, there are differences in the resolutions of the frames, which might be improved by using temporal super-resolution repetitive motion methods.

IV

DESIGN AUTOMATION

8

PERSONALIZED PRODUCT DESIGN THROUGH DIGITAL FABRICATION

Personalized designs bring added value to the products and the users. Meanwhile, they also pose challenges to the product design process as each product differs. In this chapter, with the focus on personalized fit, we present an overview as well as details of the personalized design process based on design practice. The general workflow of personalized product design is introduced first. Then different steps in the workflow such as human data/parameters acquisition, computational design, design for digital fabrication, and product evaluation are presented. Tools and methods that are often used in different steps in the process are also outlined where in human data acquisition, 3D scanning, and digital human models are addressed. For computational design, the use of computational thinking tools such as abstraction, decomposition, pattern recognition and algorithms are discussed. In design for digital fabrication, additive manufacturing methods (e.g. FDM), and their requirements on the design are highlighted. For product evaluation, both functional evaluation and usability evaluation are considered and the evaluation results can be the starting point of the next design iteration. Finally, several case studies are presented for a better understanding of the workflow, the importance of different steps in the workflow and the deviations in the approach regarding different contexts. In conclusion, we intend to provide designers a holistic view of the design process in designing personalized products as well as help practitioners trigger innovations regarding each step of the process.

Parts of this chapter have been published in:
Minnoye, A.S.L., **Tajdari, F.**, Doubrovski, E.Z.L., Wu, J., Kwa, F., Elkhuizen, W.S., Huysmans, T. and Song, Y., 2022, August. Personalized product design through digital fabrication. *In International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (Vol. 86212, p. V002T02A054). American Society of Mechanical Engineers [173].

8.1. INTRODUCTION

PErsonalized products are products that are designed and manufactured to satisfy the needs of individual customers, ranging from functional requirements to aesthetics [160]. Enabled by Industry 4.0 and especially the advanced manufacturing techniques [232], the production goal of personalized products is shifting towards the added value for consumers, who are also engaged in specifying the requirements of, or even designing, their own products. Personalized products can be grouped into three categories [204], i.e.,

- Personalization in Identity: This category focuses on the perception of the product; The unique form, texture, color, print, smell, taste, sound, feel, etc. provide added value for the customers; e.g. Apple® offers customer a service of printing his/her names on the AirPod®;
- Personalization in Capabilities: In this category, the design focuses on the personalized functions of the product. The unique performance of the products that is enabled by extra ingredients (electrical, mechanical, fluidic, and thermal components) demonstrates the added values of the product, e.g. adding electric roof in a car configuration;
- Personalization in Fit: It addresses the presence of the personalized product regarding the interactions between the product and the consumer, the environment and/ or other products that are used by the consumer. Physical characteristics of the product, such as shape, size, mass, area, quantity, color palette, etc. and the personalized interactions (e.g. comfort), present the added values of the personalized products in this category, e.g. custom fit shoes/chair/glasses.

Personalized products may help businesses in different ways, e.g. generate more sales; increase the profit margin; stand out from the competition; lower the inventory costs; have a deep insight of the needs of customers; increase customer loyalty and power the online business [212]. Meanwhile, personalized products also pose challenges for businesses as the products require flexibility and fast responses from the business [211]. In particular, a seamless information flow needs to be established across customer relation management (CRM), supply chain management (SCM), enterprise resource planning (ERP), the manufacturing process and the logistics [8]. It requires a complete transformation of business. Additionally, due to the nature of personalized products, they are hard to be reused. That is, an ill fitted/used product cannot be repaired or reused. Therefore, sustainability in the design of products, e.g. using more recyclable materials, design for recycling, should be addressed across the complete life cycle of the products.

In the area of design and manufacturing, traditionally, designers often design a series of products while taking into account the variations in consumers' wishes, sizes and other requirements. When designing personalized products, the designer is responsible for making a "modifiable" template. This template, often a script or computer program, processes user data (dimensions, text) and outputs the finalized design geometry and specifications [261]. Regarding manufacturing, agile manufacturing enabled by Industry 4.0 is able to quickly respond to the customer's demands [8]. As an important enabler of agile manufacturing, additive manufacturing methods (i.e. 3D printing) has been attracting attention due to its high flexibility and cost-effectiveness. Currently, additive manufacturing methods are widely used in manufacturing personalized products such as medical implants [260]. In this chapter, with the focus on “personalized fit” products, we summarize the design process of personalized products and address the key steps based on design practice. The rest of the chapter is organized as follows: first the general workflow of designing personalized products is discussed. Then the role of human data acquisition in the personalized fit is presented in Section 3. In Section 4, computational design is introduced as a bridge between designers and the data. Furthermore, design for digital fabrication is studied in Section 5, and methods of evaluations are explained in Section 6. Several cases studies are presented in Section 7 to highlight different steps in the design practice and finally, a short conclusion is drawn and possible improvements for the future works are highlighted as well.

8.2. The workflow

Based on the design requirements and using the computational design approach, the general workflow (Figure 8.1) of designing personalized products can be divided into the following iterative steps: 1) Human data/parameters acquisition; 2) Generate design using computational design tools; 3) Design for digital fabrication; 4) Product evaluation. Besides, human models and 3D scanning techniques are often used for data acquisition, and to generate design templates for computational design. Human data/parameter acquisition In this step, data and parameters of individual body shapes are collected/generated. Besides parameters regarding the context of the design, two methods are often used in acquiring human body shape data:

- Direct data collection from 3D scanning, CT and/or MRI.
- Data augmentation from collected data, i.e., using a digital human model to generate (part of) the 3D human body shape.

8.2.1. Generate design using computational design tools

Prior to the personalized product design process, a set of design templates are often created. Instances of the template are controlled by the data (e.g. dummy human body shape), and/or parameters (e.g. the required length of a hand splint and the thickness of the splint). With the newly acquired data/parameters, a personalized design can be generated automatically based on those templates.

8.2.2. Design for digital fabrication

In this step, materials and manufacturing requirements are used as the inputs, and the personalized design is further tuned for the selected manufacturing methods, selected materials and optimized for the specific manufacturing process.

8.2.3. Product evaluation

A product needs to be evaluated regarding its functionality and usability. For personalized products, extensive evaluations should be conducted on different boundary condi-

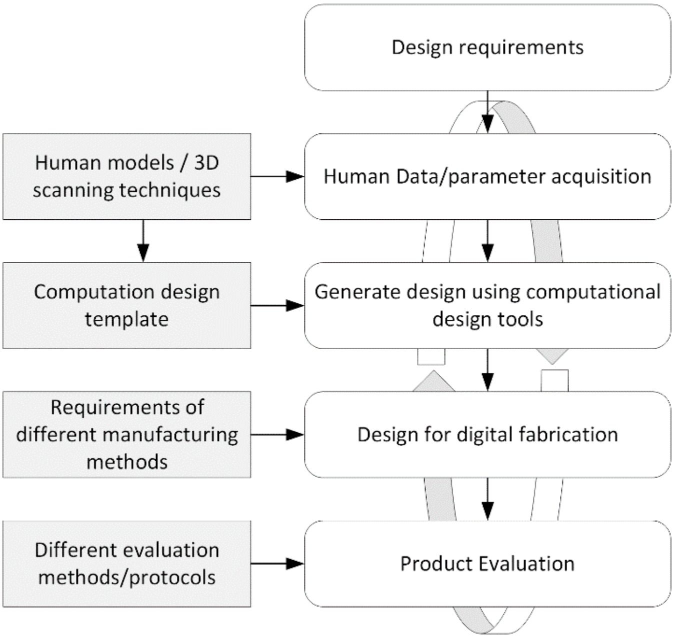
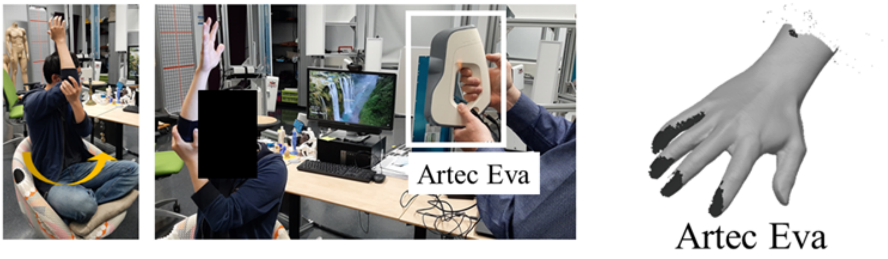


Figure 8.1: The workflow.

tions (with a safety margin) regarding engineering, manufacturing and ergonomics factors. Evaluation results shall be feedback to the designer for the next iteration of the design.

In personalized product design for digital fabrication, a series of software tools can be used. For human data acquisition based on 3D scanning methods, tools such as Geomagic®, Agisoft Metashape® are often used. In the category of computational design, Rhino® (and Grasshopper®) is frequently mentioned. For computer aided design (CAD), Solidworks®, CATIA®, NX®, Autodesk Fusion® etc. are often used and in the area of computer aided engineering (CAE) analysis, Ansys®, Abaqus®, CATIA®, NX®, Adams®, SimScale®, etc. are tools that are often applied. For computer aided-manufacturing (CAM) tools, besides machine specific tools, 3D expert®, Cura®, Slic3r®, Materialise 3-matic®, etc. are frequently applied in simulation and toolpath generation. It is worth mentioning that the uses of those tools are not exclusive but complementary. Designers often select tools based on the desired functionality, usability, familiarity, and availability.

The aforementioned workflow and the detail steps are more indications rather than guidelines. In practice, designers often take short iterations to improve the design, or sometime they even swap the steps following the needs of a particular context. For instance, in the process of design for digital fabrication, designers may modify the computational design templates to fit the manufacturing requirements; or they will simply print a (part of the) prototype to verify the set parameters. Another example is that for the known scenarios, CAE simulations are always performed prior to the manufacturing



(a) (b)

Figure 8.2: 3D scanning of human hands using Artec® Eva® scanner, courtesy of [312].

process to tune the design. However, CAE simulations are also conducted in the evaluation for a comparison of the experiment and the simulation results to find protentional problems in setting up the simulation, e.g. boundary conditions, material properties. The new scenario discovered during evaluation are often simulated as well for the next design iteration.

8.3. HUMAN DATA/PARAMETER ACQUISITION

The shape of humans differs by nature. “Personalized Fit” requires not only the desired parameters of the product, but also data regarding the 3D body shape of the individual. For acquiring human 3D shapes, two methods are often used: 1) 3D scanning - to acquire the exact body shape of individuals; and 2) Data augmentation using digital human models.

8.3.1. 3D SCANNING OF HUMAN BODY

Many techniques, e.g., structured lights, time-of-flight scanning, laser scanning, computed tomography and photogrammetry, have been developed and used in digitizing the 3D shape of the human body. In real life scenarios, the challenges in digitizing human-body parts are mainly the complex geometry and the potential movement of the human body during the scanning process. Figure 8.2(a) presents the scanning of the hand where an Artec Eva® was held by a researcher. A participant sat on a rotary chair and raised his hand over his head. While the subject was asked to keep the body and the hand as stable as possible, a second researcher (not shown in the picture) slowly rotated the chair with instructions from the first researcher holding the scanner. After a few tries, it was possible to acquire the scan of the hand within 40-50 seconds. However, the quality of the scan (Figure 8.2(b)) strongly depends on the cooperation of the subject and researchers [312]. To solve this problem, researchers/practitioners sometime use another strategy which is to scan the negative shape, e.g. podiatrists use the scan of the foot imprint on the foot impression foam to design customized orthotics [284]. However, this process can only get the shape information of one posture.

In the area of high-speed digitization of human body shapes, 3D optical scanning and photogrammetry are two techniques that are often used. 3D optical scanning meth-

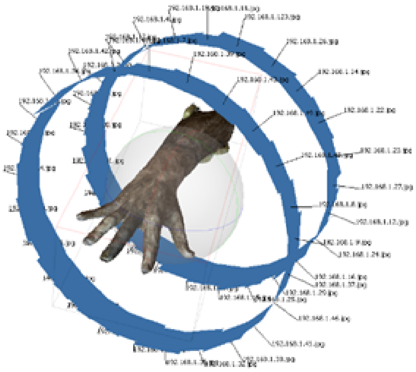


Figure 8.3: 3D hand reconstruction using photogrammetry (blue planes are identified camera image planes), courtesy of [312].

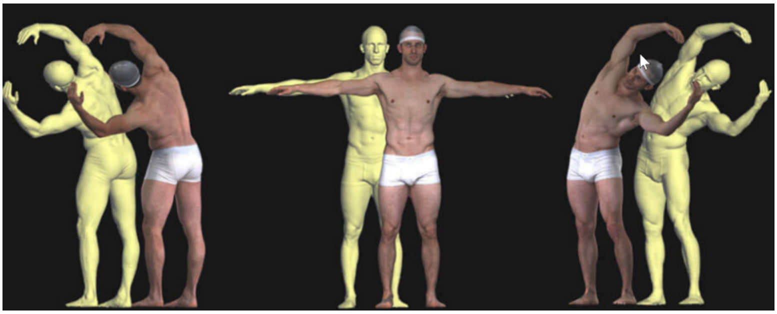


Figure 8.4: 3dMD® scanning system (courtesy of [2]).

ods project either a laser or a structured light pattern onto a region of interest (ROI). The images of the ROI are then captured by cameras. Based on image processing and the use of the triangulation method, the external shape of the ROI can be acquired. Examples of those scanners are the Artec Eva® [18] and Capture 3D® [41] . Photogrammetry is the process of creating a 3D scan of a human/object using multiple images of the object taken at different angles. The principle of photogrammetry is to match the same featured points (e.g. pixels or regions) in overlapped regions of different images and then compute the coordinates of those points using the triangulation method. The Scale-Invariant Feature Transform (SIFT) [59] is a typical algorithm used to identify those featured points. Examples of software tools that use photogrammetry technique to construct 3D models are Agisoft Metashape® [5], and Meshroom® [11], a free and open-source 3D reconstruction software. Figure 8.3 shows the process of reconstructing a human hand 3D model from 50 images using Agisoft Photoscan®. Many commercial scanners often combine photogrammetry with another technique. A typical example is the 3dMD® scanning system [2], which is built on multiple Modular Camera Units (MCUs). MCUs utilize a hybrid of stereo photogrammetry and structured light tech-

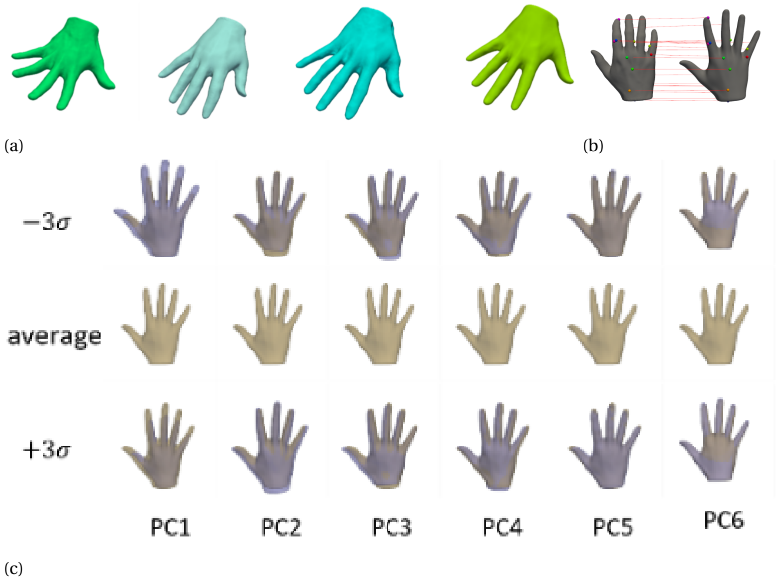


Figure 8.5: Building a hand SSM model. (a) 3D scans of the hand. (b) Establishing correspondences between scans. (c) Change the coefficient of the principal components (PCs) of the hand SSM model.

nology. Depending on the needs of the applications, the system can be configured to capture the shape of the whole body as well as different parts.

8.3.2. DIGITAL HUMAN MODEL

A digital human model is a parametric virtual representation of the variation of a human characteristic based on large sample database. Based on a few anthropometric measurements, it is possible to use a 3D digital human model to approximate shapes of the human body with reasonable accuracy. A digital human shape model can be constructed based on different requirements using different tools. A statistical shape model (SSM) of the human body, which can represent the 3D shape with a limited number of parameters with a certain accuracy, can be utilized as a high-fidelity digital representation of the body in many applications. In general, a SSM can be created based on digitization, establishing correspondence and modelling these three steps.

- Digitization: In the digitization step, many 3D scans representing the whole population are collected. For example, Figure 8.5(a), we show some scans of the human hands. All scans are triangle meshes, i.e., the surface is represented by triangles that are connected by their shared edges or vertices.

- Establishing correspondence: In this step, the correspondences between vertices/triangles of a triangular mesh and a reference model (template) are established as Figure 8.5(b). Many algorithms were developed for this purpose, for instance, the non-rigid iterative closest points methods (non-rigid ICP, e.g. [251]).

	Measure	Persona 1	Group 1
X	Stature 2	1800 mm	1500 – 1840 mm
Y	Face length 85	129 mm	98 – 114 mm

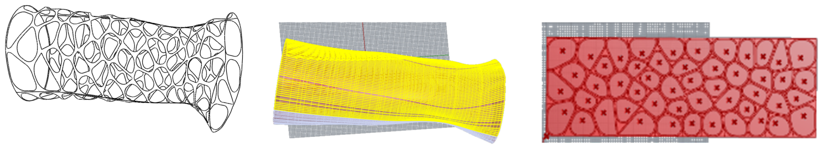
Figure 8.6: The Mannequin tool from the DINED platform [64].

- Modelling: In this step, all meshes are aligned together (brought together as close as possible). Possible posture variations can be corrected using different methods, such as using embedded skeletons [311]. A statistical shape model can be built based on the mean of corresponding vertices and the variances of each vertex regarding the corresponding vertex in the mean model. The variances (Figure 8.5(c)) can be simplified using dimension reduction methods, e.g. the principle component analysis method, which is a dimension-reduction tool that can be used to reduce a large set of variables to a small set that still contains most of the information present in the large set.

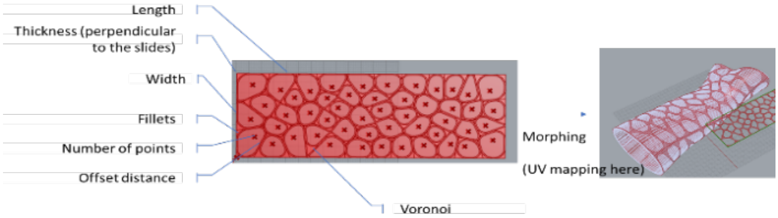
An example of online digital human models is the DINED database [106], developed in TU Delft as Figure 8.6. Since 2000, the focus has been moving to the application of 3D scanning in anthropometry. This has resulted in various 3D data collection as well as research into the analysis and presentation of 3D anthropometric data and its applications in design. Designers can easily acquire a human shape model based on a few parameters, e.g. the statue, BMI. It is worth mentioning that the accuracy of the model strongly depends on that if the target user is in the population that the datasets were that the model built on, the inputs of the users as the default parameters of the model is often the mean values of the population.

8.4. COMPUTATIONAL DESIGN

Computational design has the possibility to help designers explore the solution space in a systematic and comprehensive manner by utilizing the computing power of computers and the intelligence of embedded algorithms. Using computational design, the role of designers is evolving from designing an explicit shape, to programming instructions for a computer to generate up with (unique instances of) a design, automating steps of the design process. The basis of the computational design is computational thinking [289].



(a) The concept (b) Abstraction and decomposition



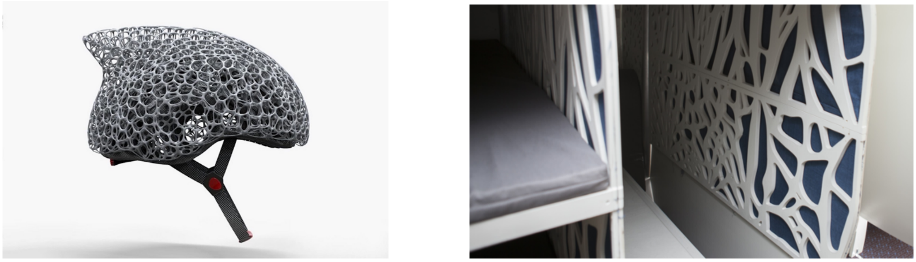
(c) Pattern recognition & Algorithms

Figure 8.7: An example of using computational thinking in design.

Computational thinking is a systematic approach to tackle ambiguities of design, sophistications and open-ended optimal problems through exploiting fundamental computer science principles and practices. Thus, computational thinking contains a large variety of computational-related components, while we list the most design-oriented components as follows:

- Abstraction: which encourages designers to focus on the core idea of the design instead of being lost in the complexity and details;
- Decomposition: which allows designers to logically divide a new challenge in the field into several related problems and increase the manageability of the workflow, e.g., using the divide-and-conquer strategy;
- Pattern recognition: which help designers find the “rhythms” in the design to simplify and accelerate the design process;
- Algorithm: which translates the designers’ idea to a set of ordered instructions that utilize the computing power to automate the design process and optimize the design.

Figure 8.7 shows an example of using computational thinking in design. The task in this case is to design a personalized splint for a patient with bone fractures on the forearm as Figure 8.7(a). In the abstraction of the design, we notice the similarity of human forearms shapes, and it is possible to design a "reference design" of the splint as a template on a dummy shape of the forearm. Further thinking indicates that the splint can be built on the forearm by mapping a planar pattern as Figure 8.7(b). In the construction of the model, the holes on the planar shape can be automatically generated by different algorithms, e.g. the Voronoi algorithm. And the planar shape can be morphed to the forearm using UV mapping (Figure 8.7(c)). For a geometric model generated by computational



(a) (b)

Figure 8.8: (a) The Voronoi Bicycle Helmet designed by Yuefeng Zhou, Zhecheng Xu and Haiwei Wang. (b) Panels of aircraft interiors of the flying-v project [269].

design tools, it often utilizes a set of parameters. Parametric design/modelling is the "*creation of a 3D geometric model using a series of pre-programmed rules or algorithms based on data and design parameters*" [32]. Using parametric design, the 3D model can be generated and updated automatically based on the data and parameters specified by the designer(s). Design parameters of personalized products can be intrinsic and extrinsic parameters of 3D shapes:

- Intrinsic parameters of a 3D shape can be interpreted as a human description and interpretation of a shape or object [231]. In the case of the Voronoi Bicycle Helmet (Figure 8.8(a), an intrinsic parameter of the helmet can be the width of the helmet;
- Extrinsic parameters can be considered as parameters of geometrical entities which actually define the shape or object. Computers use extrinsic parameters to generate and update a 3D model, e.g. the position of each vertex in the helmet.

With the human centered design principle, it is important that the parameters of a parametric design should be interpretable for personalized products, preferably by the customers to drive a design towards their needs and wishes.

8.4.1. GENERATIVE DESIGN & TOPOLOGY OPTIMIZATION

The idea of generative design is to explore the design space in a systematic and automatic manner, and thus to generate the optimal design under prescribed design specifications. It is an algorithm-driven design process. Example approaches include shape and topology optimization, shape grammar-based design, machine learning based design methods, among others [292]. For instance, Wu et. al. utilized Generative Adversarial Network (3D-GAN) to generate 3D objects from a probabilistic space by leveraging recent advances in volumetric convolutional networks and generative adversarial nets [293].

One widely recognized generative design technique is topology optimization [295], which has been increasingly used in personalized product design, either for reducing the weight without compromised to the stiffness of the design, and/or for a better design aesthetic. In topology optimization, the problem of structure design is reformulated as finding the optimal distribution of material in a discretized design domain. The optimized layout is not restricted to its initial topology, opening up possibilities for superior structural performance over manual designs based on engineers' intuition and experience. Topology optimization aims to find the optimal structural layout for certain applications such as finding a structure that is optimized for maximum stiffness under a given load or torque. The optimization problem is often solved using an iterative approach. Figure 8.8(b) presents panels of aircraft interiors for which an innovative topology optimization method [296] was applied in the design process.

8.5. DESIGN FOR ADDITIVE MANUFACTURING

Additive Manufacturing (AM), or 3D Printing is a Digital Manufacturing technology that is increasingly being used in the architecture, engineering, and construction sector, and could also be considered a prerequisite for fabricating many personalized fit designs [182].

Depending on the requirements of the design, an AM process and material needs to be selected from the large variety of available AM processes and materials. Fused Deposition Modelling (FDM), Stereolithography (SLA) and/or Selective Laser Sintering (SLS) are often used in printing plastic products, and for metal parts/products, Powder Bed Fusion, Direct Energy Deposition and Binder Jetting are often applied. Recent developments on printing soft materials, conductive materials [105, 303] and multi-material printing [213] offer new opportunities for personalized products. For instance, the Stratasys J735 multi-color/material 3D printer uses the polyjet technique and is able to print with multiple materials, ranging from rigid opaque materials, to transparent and rubber-like soft materials. Figure 8.9(a) presents a personalized dress that was printed using a multiple materials [173]. For optimal performance of the printed product, it is essential to consider different parameters of the chosen AM process. AM-related design considerations need to be taken into account from the beginning of the design process. Each AM system has a set of design rules. For example, design rules for wall thickness, tolerances, and possible overhangs directly influence the possible structures that can be fabricated. Also, the engineering properties, e.g. thermal or mechanical properties, of 3D printed materials differ from the original materials, mainly due to the non-uniformness introduced by the manufacturing process.

Next to design rules, there are also manufacturing-related decisions that need to be taken and which have an influence on part properties. A well-known example in the use of Fused Deposit Modelling (FDM) is the building orientation as shown in Figure 8.9(b). Different building orientations will result in different support structures, different stiffness of the prototype/products, different surface finishing. Details of those constrains are available in different knowledge bases, e.g. [1]. Referring to the 3D printing practices, e.g. FDM, many practical issues need to be considered in setting the printing parameters for a specific printing technique, or even for a particular machine, such as build plate adhesion type and strategies to avoid warping.

As many personalized products come in contact with the user, it is important to choose materials that can be in contact with the human body without any adverse effects, i.e. biocompatibility is recommended for such products. The FDA (U.S. Food & Drug Administration) approved many Class I and Class II materials that can be used to construct personalized products, such as Polylactic acid (PLA) filaments or PETG Filament, which can be used in Class II products [76]. Regarding metals, Titanium Alloy

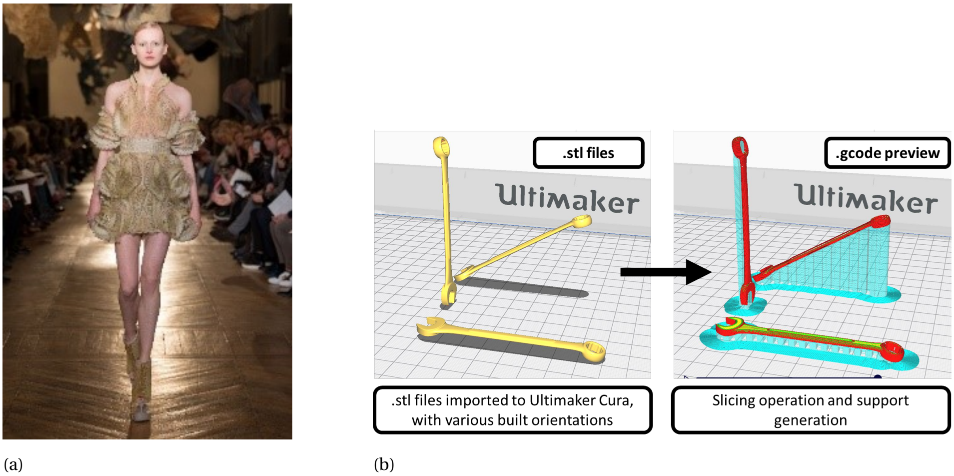


Figure 8.9: (a) A 3D printed dress, design by Iris van Herpen, presented at Galerie de Minéralogie et de Géologie in Paris, Photographed by Yannis Vlamos and courtesy of [173]. (b) An example of adjusting printing parameters according to needs, i.e. preview of a part with different building directions using FDM.

Figure 8.10: A design of hand splints, but sizes may vary according to different hand shapes [310].

(Ti-6Al-4V) is often used in medical implants.

8.6. DESIGN EVALUATION

Personalized products can share the same design template, but the model has to be adjusted based on different body shapes. For example, the thickness of the hand splint might need to be increased for larger hands as Figure 8.10. Due to size variations, it can be difficult to set one evaluation standard. An evaluation strategy needs to be set regarding different boundary conditions of the design based on the functionality and usability of a product. Scenario creation Usage scenarios with the user(s), the product(s), the environment and possible interactions may help the designer to have a better understanding of the boundary conditions for evaluating the products. A possible procedure in defining the scenarios can be:

1. Define the persona: It is worth mentioning that personalized products differ due to their nature. Multiple personas are often needed regarding personalized fit, e.g., P5 and P95 of the population, or randomly generated examples for verification;

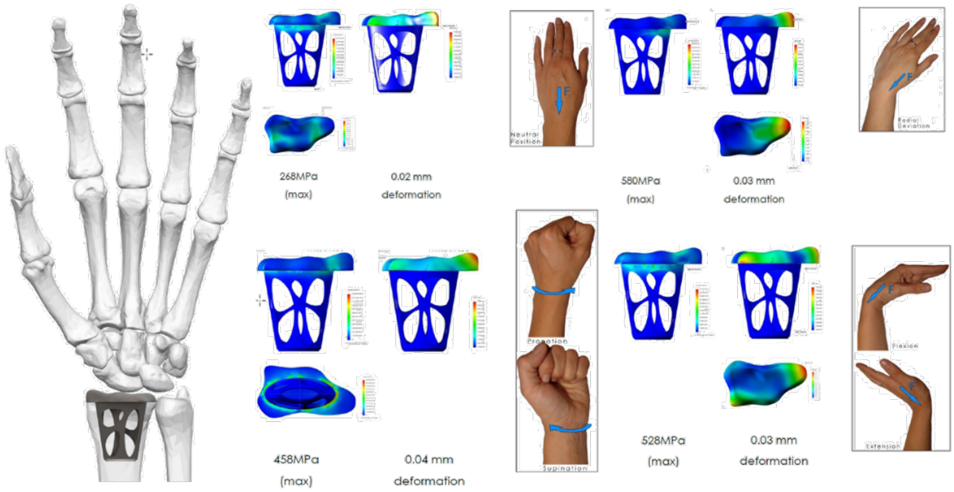


Figure 8.11: FEM simulation of a new patient-specific prosthesis for resurfacing of the distal radius based on different scenarios (CAE tool: Solidworks®), Courtesy of [72].

2. Define the starting point and the tasks of the scenario(s); Multiple scenarios are often needed in the evaluation as well;
3. Explore the stakeholders and the product use environments;
4. Write different stories regarding the user activities in different scenarios;
5. Explore the extrema (boundary conditions) regarding the use of the product in those stories. For the functionalities of the product, the boundary conditions can be identified through different usage scenarios and manufacturing variations. Mechanical, electrical, fluidic, thermal properties are often evaluated against those boundary conditions.

8.6.1. Functional evaluation

As indicated in Section 2, CAE simulations can be use prior to the manufacturing and/or in the evaluation of the product, depending on the requirements and the case. Through CAE simulations, it is possible to predict/analyze the performances of a prototype in a cost-effective manner. Figure 8.11 presents the Finite Element Method (FEM) simulation of a patient specific implant in different scenarios (with different loading conditions).

8.6.2. Usability evaluation

Subjective evaluation is often used in the usability evaluation. Questionnaires are an important tool(s) in subjective evaluation. There are many types of questionnaires for evaluating different types subjective feelings, and using verified questionnaires may accelerate the evaluation process. Figure 8.12 presents a simple comfort/discomfort questionnaire, which is often used to evaluate the level of comfort/discomfort that the user

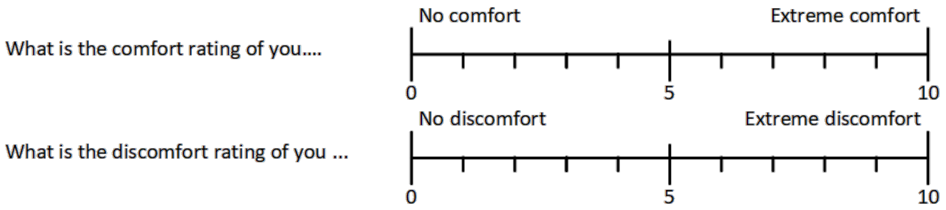


Figure 8.12: Comfort and discomfort questionnaire.

feels before, during and after using the product. "..." in the figure can be changed by different contexts. A list of questionnaires for comfort evaluation of different products can be found in [16]. Objective measures regarding the user and the environment are also often used in the evaluation [230]. Parameters of the environment include vibration, light, noise, smell, etc. Physical parameters of the user(s) include anthropometry, contact interface pressure, movements, etc. Physiological parameters of the user(s) include heart rate, heart rate variability (HRV), blood pressure, etc. Literature indicates that there are many relations between/among objective measures and subjective feelings, e.g., features in HRV are associated with the stress level [15]. With the digital twin technology [99], and especially the embedded real-time sensors in the products, the future personalized product might be able to predict the feeling of the user or even change its form for a better performance.

8.6.3. Experiments design

Experiments or trials are the “gold” standard in product evaluation. In the design of the experiment to evaluate the design, the following aspects might be considered:

- The hypothesis(s) to be testified by the experiment;
- The setup and the location of the experiment;
- The metrics (function & usability) will be used in the evaluation;
- The measurement methods and devices that may generate data to support the metrics;
- The data management plan (DMP) and the post-processing methods of the collected data;
- The target users of the products, i.e. experiment participants should be representative for the target population;
- The safety of users in the use of the product; a risk management matrix can be a good addition;
- In any experiments where the user(s) is involved, we need to consider the ethics and apply permissions from the ethical committee (incl. informed consent);

Figure 8.13: Personalized products fabricated by 3D printing.

8.6.4. RISK MANAGEMENT

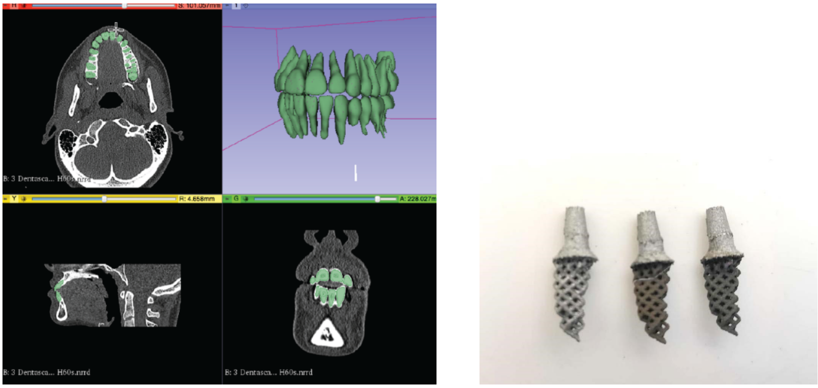
Risk management is an important requirement in the development of personalized products especially many personalized products, such as a hand splint for managing fractured bones, can be categorized as Class I or II medical devices [201]. For Medical devices - Application of risk management to medical devices (ISO 14971:2019 [108]) specifies the terminology, principles and a process for risk management of medical devices. Following the scenarios defined before, according to the functional and usability evaluation results, the designer can often use a risk assessment matrix [77] to explore the ways of mitigating the potential risks, especially regarding the severe and catastrophic risks.

8.7. CASE STUDIES

In the past decade, “personalized product design though digital fabrication” is a focus of the Faculty of Industrial Design Engineering (IDE) at Delft University of Technology (TU Delft), from both research and education perspectives. For instance, in Figure 8.13, students’ design of personalized products fabricated by 3D printing are presented. The authors have explored and developed the workflow, advanced algorithms. Meanwhile, they also explored various novel applications in the area of personalized designs. In this chapter, based on those design practices, we summarized our experiences and present the outcomes as the “best practice” for personalized product design though digital fabrication. In the rest of this section, we present some case studies as typical examples. It is worth mentioning that in different cases studies, the focuses on different steps of the workflow might be different, depending on the context of the design.

8.7.1. CASE: PERSONALIZED DENTAL

Implant – Courtesy of [169] This project uses the computational design method, to design the lattice structure for a personalized dental implant. Currently, the process of placing dental implants is a long process with a 3-5 % failure rate, mainly caused by the lack of osseointegration (the integration of living bone and an artificial implant) and infection. This patient-specific implant has a porous structure to promote osseointegra-



(a)

(b)

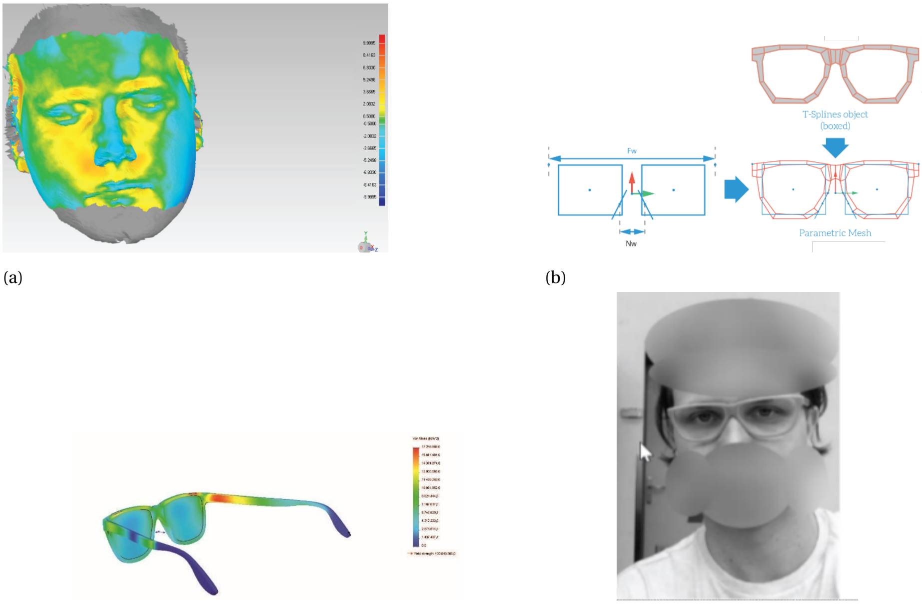
Figure 8.14: Design of patient specific tooth implants, courtesy of [169].

tion. Another advantage of using this product is that bone healing after extraction (of the tooth) is not required, which saves 3-6 months of the procedure.

In the personalized design, the shape of the root of the tooth was acquired by a cone-beam computed tomography system (CBCT). Using image segmentation software, the shape of the root can be retrieved (Figure 8.14(a)). A lattice structure was designed following the shape of the root using Rhino® and Grasshopper®. Such porous structure was created to promote osseointegration, where bone can grow into the pores of the structure for a better fixation. The prototypes were made of Ti6Al4V alloy manufactured by selective laser melting (SLM). The surface of the implant was treated by plasma-electrolytic oxidation (PEO) with silver nanoparticles to create an antimicrobial surface. Figure 8.14(b) shows the printed implant (left), surface treated by PEO (middle), surface treated by PEO and silver plating (right). The design was validated with oral surgeons and all of them see added values in the patient-specific design as compared to current solutions, which only have a set of “standard shapes”.

8.7.2. Case: Personalized sunglasses

It is difficult for the customers to find sunglasses that they like and fit to their face. Personalized sunglasses with adjustable aesthetic style as well as ergonomics fit might meet the needs of consumers. The aim of this project is to develop customized sunglasses that fit individual users. Photogrammetry technique was selected for acquiring human face models. The user was asked to make a video of their face while turning his/her head from right to left in 30 seconds. A 3D model can be constructed based on photos extracted from the video with an accuracy of 1 mm regarding the critical areas of the face for wearing glasses, such as nose bridge, ears, and eye positioning. (Figure 8.15(a)). Several templates of sunglasses were designed in different styles. For each template, parameters were introduced to adjust the template for the best fit regarding a given 3D scan. An example is presented in Figure 8.15(b) where parameter Fw is the width of the face and Nw (a crucial parameter for the user’s comfort) is the width of the nose bridge. The Selective laser sintering (SLS) printing method was selected as the manufacturing



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

Figure 8.15: Personalized sunglasses, courtesy of [272].

method where PA12 was selected as the material. FEM simulations were conducted on the small and the large designs in different scenarios to verify the stiffness of the personalized sunglasses structure (Figure 8.15(c)). Several prototypes were also produced and tested by the users and Figure 8.15(d) shows one of the prototypes worn by the user.

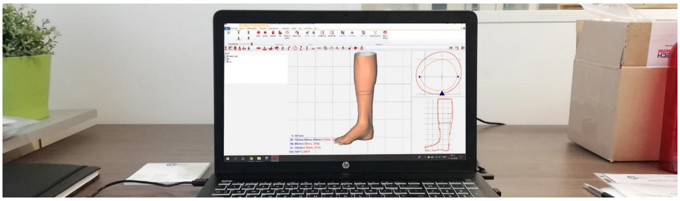
8.7.3. CASE STUDY: ANKLE FOOT ORTHOSIS

Patients diagnosed with drop foot syndrome often encounter difficulties in walking. An Ankle Foot Orthosis (AFO) is an orthopedic aid that limits the plantar flexion of the foot to ensure a safe walking gait for the patient. Currently, these AFOs are often acuum formed over a machined foam following the shape of the patient's leg. In this case study, we present the design of the personalized AFO using the 3D scanning and the additive manufacturing method. The 3D shape of the lower leg of the patient was scanned first where some extra measurements were collected manually (Figure 8.16(a)), mainly to make up the missing part(s) in the 3D scans. The 3D scanning data was post processed (Figure 8.16(b)) and then the AFO was designed using computational design tools, in this case Rhino® and Grasshopper® (Figure 8.16(c)). The AFO was printed and tested regarding engineering parameters (Figure 8.16(d)) as well as usability (Figure 8.16(e)).

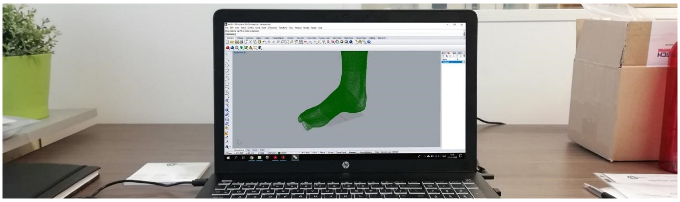
8.8. CONCLUSION

Personalized product design brings added value to the product(s), meanwhile it also poses challenges to the design process. In this chapter, we give an overview of differ-

(a) Human data acquisition



(b) Human modelling



(c) Computational Design

(d) Functional evaluation

(e) User evaluation

Figure 8.16: Design a patient specific AFO, courtesy of [273], in (e) the blue/grey block is the pressure gauage).

ent steps of the personalized product design process based on the best practice. Within the proposed iterative workflow, the body shape of the user(s) is collected first, either by 3D scanning or based on digital human models. Computational design tools are used to fit an existing design template to the acquired body shape. With design for digital fabrication tools, the design is further optimized for a better manufacturability. Both functional and user evaluation methods are introduced for evaluating the design. We want to address that this chapter is based on the design practice, and we expect that it will give designers a holistic view of the design process in designing personalized products as well as help practitioners trigger innovations regarding different steps in the process.

Meanwhile, our design practice also indicated that there are many aspects can be improved in the workflow. In the area of human data acquisition, shapes of human body (parts) used in the design is in mainly 3D, which often resemble a static pose. Research on acquiring 4D shapes and using those dynamic body postures in design is undergoing. For computational design, more advanced algorithms in the area of generative design and topology optimization are under development. In design for digital manufacturing, different types of digital materials will be introduced and new additive manufacturing methods, e.g. cold spray, are under investigation. For product evaluation, engineering properties of materials manufacturing by additive manufacturing methods, especially about their anisotropic properties, are being embedded in the material database for a better prediction of the behaviors of the product in use. Among all potential improvements, perhaps the most important is to develop a software platform to support personalized product design through digital fabrication. Currently, many software tools are used in the design process, and designers have to shift among different tools for the desired functions. This is expensive regarding both the cost and the needed skills. A platform which is able to synthesize the needed functions may help designers avoid possible errors in swapping tools and accelerate the design process, therefore improve the effectiveness and the efficiency in designing.

V

DISCUSSION AND RECOMMENDATION

9

SUMMARY AND FUTURE WORK

In this chapter, we first summarize the contributions of this Ph.D. research and provide answers to the three research questions presented in Chapter 1. The implications of this Ph.D. project are highlighted from the perspectives of design engineering and sustainability. Finally, discussions on the limitations of this research and possible future work are presented.

THIS Ph.D. project provides effective solutions for automating the processing of 3D/4D geometry. Firstly, new algorithms were developed and implemented to address accurate 3D point cloud non-rigid registration, both as surface mesh and volumetric mesh. The main focus of these methods was to ensure robustness, convergence, and speed. Additionally, new hardware, specifically a 4D scanner, was designed and manufactured to enable 4D scanning of moving objects, achieving a frame rate of 30 frames per second. To facilitate the synchronization of time-wise aligned frames, a novel time-synchronization software was implemented on the scanner. By employing the 3D point cloud registration methods, a comprehensive mesh-morphing 4D scan product was created. Finally, the integration of all the developed software and hardware for design purposes is discussed. The subsequent section presents the answers to the research questions posed in Chapter 1, highlighting the main contributions of this Ph.D. research.

9.1. ANSWERS TO THE RESEARCH QUESTIONS

- **RQ1:** *What type of registration methods can contribute to establish 3D human shape model and their use as prior knowledge in fitting new subject scans?*

The answer to **RQ1** is presented in Part II, which includes Chapter 2, Chapter 3, and Chapter 4, with a primary focus on 3D static point clouds. Through our observations, we found that during the registration process, a source mesh, whether it's a surface mesh or a volumetric mesh, may lose some of its features as deformation and feature preservation compete with each other. In Part II, we explored the idea of balancing these competing factors either globally across all registering points or individually for each registering point within an existing surface registration method, namely the non-rigid ICP algorithm.

  - **RQ1.1:** *How to integrate a shape descriptor to the nonrigid ICP algorithm to consider the shape characteristics?*

The answer to **RQ1.1** is discussed in Chapter 2, which aims to present a non-rigid ICP approach based on a newly defined concept called semi-curvature. The semi-curvature exhibits similar properties to Gaussian curvature but possesses distinct mathematical and geometric characteristics. Leveraging these unique properties, we incorporate the semi-curvature into the metric for establishing correspondences and the cost function within the registration method. Additionally, the distance and stiffness terms are embedded in the cost function. By enhancing the logical dependency of vertices on their neighbors, the semi-curvature term preserves surface features within a stable region throughout the registration process, prioritizing mesh quality and convergence.

The effectiveness of the registration methods is evaluated using various human body parts, including 22 right and left feet, 101 full human bodies from the CAESAR dataset, and 4 lumbar vertebrae of the human spine. Through extensive testing, our method demonstrates the ability to achieve a logical balance between the competing terms. This is evaluated based on three matrices: the quality of correspondences selection, the quality of the surface after registration, and the time required for the convergence process. In comparison to the several approaches with the same state-of-the-art, the method presented in Chapter 2 outperformed them in terms of the quality of correspondences selection (75%), the quality of the surface after registration (8%), and the time spent on the convergence process (82%).

- **RQ1.2:** *How to derive a nonrigid ICP-based algorithm that is robust to artefacts in a 3D shape?*

The answer to **RQ1.2** is explored in Chapter 3, which introduces a non-rigid ICP approach combined with an innovative adaptive feedback control scheme to estimate the stiffness ratio. This estimation is achieved by utilizing the gradient of the mesh quality and mean curvature values per vertex. To enhance the convergence procedure towards accurate estimation, an ANFIS-based predictor is integrated with the estimator. The ANFIS predictor utilizes the topology of both the source and target meshes to anticipate initial values for the adaptive estimator, facilitating the estimation procedure. Subsequently, the estimated stiffness ratio is incorporated into the metric for establishing correspondences and the cost function, alongside the distance and stiffness terms. By adjusting the connectivity level of vertices based on their neighbors (equivalent to stiffness), the ANFIS-based adaptive estimator effectively preserves surface features within a globally asymptotically stable region during the registration process, placing emphasis on mesh quality and convergence.

The effectiveness of the registration method is validated using various human body parts, including 101 full human bodies from the CAESAR dataset and 11 thoracic vertebrae of the human spine. Through comprehensive testing, our method demonstrates the ability to achieve a logical balance between the competing terms. This evaluation is based on four criteria: the quality of correspondences selection, the quality of the surface after registration, the time required for the convergence process, and the level of topology alignment between the source and target meshes after registration. In comparison to the method presented in Chapter 2, the method described in Chapter 3 outperformed it with respect to the quality of correspondences selection (50%), the quality of the surface after registration (60%), and the topology alignment level of the source mesh and target mesh (20%).

- **RQ1.3:** *How to implement the nonrigid-ICP registration method for registering volumetric meshes on each other?*

The response to **RQ1.3** is discussed in detail in Chapter 4 and is illustrated through a case study involving the development of a bio-informed mechanistic deep learning model for the prognosis of pediatric spinal deformity. The training data utilized in this study consists of both clinical data and mechanistic features. The clinical data is extracted from anteroposterior and lateral views of X-ray images using an active contour image segmentation model. To ensure optimal performance, the parameters of the active contour model are calibrated through sensitivity analysis. Subsequently, personalized 3D spine models are established using the mesh morphing technique. This involves registering a volumetric atlas spine model onto each of the extracted clinical data, enabling the detection and tracking of any geometry-dependent features within the patient-specific models. The bone growth model parameters are calibrated based on these features. The mechanistic features are obtained from the bone growth model, which takes the stress distributions of the spine as inputs. To measure the stress distribution, a patient-specific 3D finite element model is generated using a 2D clinical data and a 3D reconstruction algorithm. By integrating medical data with a mechanistic model, the proposed framework addresses the challenge of limited data for patient-specific studies. The presented model has the capability to predict the spinal curve of an individual patient, even beyond the training range. This means it can make accurate predictions for patients who fall outside the range of the training data, enhancing its applicability and generalization capabilities.

The effectiveness of the model is verified using the human body parts such as 5 full human volumetric spines. Finally, in Chapter 4, the suggested FFNNCR-BM enhanced geometry prediction accuracy by 40% for the inside of the range spine data and 84.3% for the outside of the range.

- **RQ2:** *What type of tools and the registration methods can contribute to establish 4D human shape model and their use as prior knowledge in fitting new subject scans?*

This research question is addressed in Part III, which comprises Chapter 5, Chapter 6, and Chapter 7. These chapters focus on the collection and processing of time-series 3D point cloud data, commonly referred to as 4D scanning. Starting with the 4D scanner, it is crucial to achieving an optimal design that satisfies various criteria, such as employing the minimum number of cameras, minimizing the scanner's dimensions, maximizing resolution and accuracy, and enabling comprehensive observation of dynamic objects. With this in mind, we formulated and investigated a novel optimization problem that encompasses these criteria and facilitates optimal design considerations. On the other hand, capturing a moving deformable object using asynchronous cameras introduces significant nonlinear accumulative delays, often caused by data transfer congestion. Consequently, the outputs of the cameras are not aligned in terms of time and geometry. To tackle this issue, we explored effective solutions for time synchronization and dynamic geometry registration of the camera outputs.

  - **RQ2.1:** *How to optimally define the position and orientation of a set of cameras in a foot scanner?*

To address **RQ2.1**, Chapter 5 explored a methodology for determining the optimal number, position, and orientation of depth cameras to achieve fast 3D mesh reconstruction of a foot. The optimization problem focused on minimizing camera scanning error, sparsity, and scanner dimensions, while maximizing the overlap between scans captured by adjacent camera pairs. The findings demonstrated that a scanner configuration comprising seven cameras (one positioned at the bottom and six positioned on top) yielded the best performance.

The experimental results revealed that the scanning accuracy, when compared to a reference scan obtained using a high-end scanner (Artec Eva), could reach up to 2.5% error against the output of Artec Eva scanner.

  - **RQ2.2:** *How to establish meaningful dynamic features between frames of a camera which guarantees to find the proper corresponded points in a limited number of iterations from a Source mesh to a Target mesh?*

The response to **RQ2.2** is detailed in Chapter 6, which presents a comprehensive workflow for extracting time-varying features from dynamic scans of deformable objects. This semi-automated procedure involves several steps. Firstly, points that can be observed by each depth camera are pre-defined. Secondly, the raw data from each camera is manually cleaned. Subsequently, the extracted data from each camera is rigidly registered to the corresponding pre-defined points. Finally, a source mesh is non-rigidly registered onto the point cloud obtained through the rigid registration process. By reconstructing the final mesh using non-rigid registration for multiple frames in a time series, a unique 4D feature outcome (for specific points) or 4D trackable scanning (for all points) is achieved. This workflow demonstrates the capability to extract dynamic information from the scans and provides a comprehensive understanding of the deformable object's behavior over time.

Comparison of the output of our algorithm with the dimensions extracted from the scanning data of the same object using an Artec Eva scanner shows the method is robust against noise and the scanning accuracy can be as high as 3.4 % compared to the Artec Eva scanner.

  - **RQ2.3:** *How to design a higher level of software-based synchronisation method between cameras in a 4D foot scanner?*

The answer to **RQ2.3** is investigated in Chapter 7, which introduces a generic framework for synchronizing and registering asynchronously captured point clouds of moving and deforming objects. This framework incorporates a novel ADGC-LSTM based network and a non-rigid registration algorithm. To demonstrate the effectiveness of the proposed framework, it was implemented using data captured from a novel 4D foot scanner, resulting in the creation of the first 4D open-access feet dataset. Within this framework, a method was developed to determine the dynamic connectivity between different time intervals in each camera, focusing on dynamic feature synchronization. Additionally, meaningful dynamic features were extracted from the combined views of multiple cameras to estimate the magnitude of deformation.

The experimental results demonstrate that our method significantly improved the synchronization process by an average of approximately 30% when compared to other state-of-the-art methods. Additionally, the quality of the acquired 4D scan was notably high, particularly in terms of capturing the deformation of each part of the foot. Such detailed information can prove highly valuable in various applications, such as footwear design and analysis.

- **RQ3:** *How can we use the 3D/4D human data in the design automation for UPPS ?*

The investigation of potential solutions to address **RQ3** is discussed in Chapter 8, in which the treatment of this chapter is more general and that this treatment gives a better idea of how methods discussed in Chapters 2–7 fit within the whole UPPS workflow.

To the best of our knowledge, the availability of a workflow that systematically guides designers in utilizing new methods and tools, such as the findings in Chapters 2–7, is relatively rare. To address this gap, we provide an overview of the various steps involved in the personalized product design process, based on best practices.

In the proposed iterative workflow, the first step involves collecting the body shape of the user(s) through either 3D scanning or digital human models. Computational design tools are then utilized to fit an existing design template to the obtained body shape. The design is further optimized for improved manufacturability using design for digital fabrication tools. Both functional and user evaluation methods are introduced to assess the design. Chapter 8 is grounded in practical design experience, and it aims to provide designers with a comprehensive understanding of the design process for personalized products. Additionally, we anticipate that it will inspire practitioners to innovate at various stages of the process.

9.2. LIMITATIONS AND FUTURE WORKS

In broad terms, this dissertation examines the influence of geometry processing on design automation, resulting in the development of multiple methods and tools. Being a part of the Next UPPS project, which is discussed in Section 1.2, the future work aims to integrate all the developed methods and tools into a cohesive and user-friendly toolkit for designers working in the field of personalized fit. However, there are certain limitations associated with the presented methods and tools. In this section, we will discuss these limitations and outline potential avenues for future research in the three main areas discussed in Section 1.4: 3D human body shape processing, 4D human body shape scanning, and design automation.

9.2.1. 3D HUMAN BODY SHAPE PROCESSING

This section pertains to Part II, where several methods are discussed for registering 3D point clouds with each other. However, it should be noted that these methods are limited in their ability to handle point clouds that are not monotonically related. For instance, when attempting to register healthy with pathological data such as amputee or deformity, these methods are expected to perform poorly. Additionally, these methods primarily rely on searching for local features for matching and may struggle to address significant differences between the point cloud being registered and the registering point cloud. For example, they may encounter difficulties in identifying features between a sitting person and a standing person. Overcoming this limitation is considered as part of future work in this area.

The methods presented in this dissertation offer a general framework for integrating a variety of linear or nonlinear terms (as long as they are linearizable) into the cost function while ensuring a stable region. A potential future direction is to investigate the expansion of the cost function to incorporate additional new terms in our ongoing research. Furthermore, in Chapter 2, the concept of semi-curvature is defined based on the 1-ring neighboring points. It would be valuable to explore the effects of including more rings and evaluate the robustness of the method against noise in the target mesh. These aspects can be considered as potential areas for future research.

9.2.2. 4D HUMAN BODY SHAPE SCANNING

In Part III, as well as in this section, we made contributions to the design of a 4D foot scanner and developed 4D scanning techniques inspired by the methods employed to address **RQ1** in Part II. Consequently, the limitations discussed in Section 9.2.1 remain applicable in this area. Furthermore, in the step involving the construction of the Dynamic-Synchronized Graph based on dynamic points, the performance of the ICFP method is highly sensitive to missed frames or very low-density point clouds. When the density is reduced, many local features become obscured, making it challenging to reconstruct these features from a specific frame. As a result, we chose to skip missed or low-density frames captured by the cameras. Additionally, it is important to note that the entire process is conducted offline due to the time required for frame synchronization and computational calculations. Also, the presented data-set considered normal people and abnormal, or disabled people are not collected. The approaches to these limitations are also part of future works.

Furthermore, future work in this area will focus on the optimal design of a scanner that is equipped with Azure Kinect cameras, using a dynamic template foot. This approach will allow for the consideration of dynamic features when solving the optimization problem. Additionally, the optimization problem will be expanded to include a term related to lighting, as it is a critical factor in the quality of the captured data. Moreover, a human factor term will be integrated into the cost function to ensure that participants can comfortably reproduce their everyday walking steps. This consideration is vital for ergonomics studies and product design, such as shoe design, where the comfort and usability of the product for the user are paramount.

9.2.3. DESIGN AUTOMATION

Based on the design practice presented in this thesis, there are several aspects within the workflow that offer room for improvement. In the realm of human data acquisition, the current design focuses on specific body parts such as the foot, full body, and spine in the 3D/4D space. However, exploring the implementation of the introduced tools and methods for other body parts, such as the heart [37], would be an interesting avenue for future research.

In computational design, ongoing development of more advanced algorithms in the areas of generative design and topology optimization holds promise, as these advancements may also address the 4D behavior of the human body.

Among the various potential improvements, the development of a software platform dedicated to supporting personalized product design through digital fabrication is perhaps the most significant. Currently, designers have to rely on multiple software tools throughout the design process, which can be costly in terms of both expenses and required skill sets. A unified platform that synthesizes the necessary functions would help designers avoid errors that may arise from switching between tools, while also accelerating the design process and improving its effectiveness and efficiency.

BIBLIOGRAPHY

[1] 3D Hubs. “Knowledge base: Quality articles for engineers and designers to learn about Digital Manufacturing”. In: *Hubs Protolabs company* (2022). URL: `https://www.hubs.com/knowledge-base/`.

[2] 3DMD. “3D scanning systems”. In: *3DMD* (2022). URL: `https://3dmd.com/`.

[3] “4D Dynamic Scanner”. In: *Max Planck Institute for Intelligent Systems* (2020). URL: `https://ps.is.mpg.de/pages/4d-capture`.

[4] Abien Fred Agarap. “Deep learning using rectified linear units (ReLU)”. In: *arXiv preprint arXiv:1803.08375* (2018).

[5] Agisoft. “Metashape”. In: *Agisoft* (2020). URL: `https://www.agisoft.com/`.

[6] E Aguiar. “Performance Capture Methods”. In: *Proceedings of the European Conference on Computer Vision (ECCV), Zürich, Swiss* (2014).

[7] Wilbert G Aguilar et al. “RRT* GL based optimal path planning for real-time navigation of UAVs”. In: *International Work-Conference on Artificial Neural Networks* (2017), pp. 585–595.

[8] Shohin Aheleroff et al. “Mass Personalisation as a Service in Industry 4.0: A Resilient Response Case Study”. In: *Elsevier Ltd* 50 (Oct. 2021).

[9] Min Sung Ahn et al. “Analysis and noise modeling of the Intel RealSense D435 for mobile robots”. In: *2019 16th International Conference on Ubiquitous Robots (UR)* (2019), pp. 707–711.

[10] SM Masudur Rahman Al Arif et al. “Improving an active shape model with random classification forest for segmentation of cervical vertebrae”. In: *International Workshop on Computational Methods and Clinical Applications for Spine Imaging* (2016), pp. 3–15.

[11] Alice Vision. “Meshroom ”. In: *Meshroom* (2022). URL: `https://alicevision.org/`.

[12] Brian Amberg, Sami Romdhani, and Thomas Vetter. “Optimal step nonrigid ICP algorithms for surface registration”. In: *2007 IEEE Conference on Computer Vision and Pattern Recognition* (2007), pp. 1–8.

[13] Brian Amberg, Sami Romdhani, and Thomas Vetter. “Optimal step nonrigid icp algorithms for surface registration”. In: *2007 IEEE Conference on Computer Vision and Pattern Recognition* (2007), pp. 1–8.

[14] J Andrew et al. “Spine Magnetic Resonance Image Segmentation Using Deep Learning Techniques”. In: *6th International Conference on Advanced Computing and Communication Systems (ICACCS)* (2020), pp. 945–950.

[15] Shabila Anjani. “Aircraft interiors, effects on the human body and experienced comfort”. In: *Delft University of Technology* (2021).

[16] Shabila Anjani et al. “PCQ: Preferred Comfort Questionnaires for product design”. In: *IOS Press* 68.s1 (2021), S19–S28.

[17] Oscar Antepara, Néstor Balcázar, and Assensi Oliva. “Tetrahedral adaptive mesh refinement for two-phase flows using conservative level-set method”. In: *Wiley Online Library* 93.2 (2021), pp. 481–503.

[18] “Artec Eva. Fast 3D scanner for professionals”. In: *Artec3D* (2018). URL: `https://www.artec3d.com/portable-3d-scanners/artec-eva`.

[19] Doris Aschenbrenner et al. “Draft: using digital twins for Ultra-Personalized Product and Services”. In: *Report of UPPS project for ASME* (2019).

[20] S P Ashdown. “Improving body movement comfort in apparel”. In: *Elsevier* (2011), pp. 278–302.

[21] Susan P Ashdown and Lucy Dunne. “A study of automated custom fit: Readiness of the technology for the apparel industry”. In: *Sage Publications Sage CA: Thousand Oaks, CA* 24.2 (2006), pp. 121–136.

[22] Karl Johan Åström and Tore Hägglund. “PID controllers: Theory”. In: *Instrument Society of America Kesoarch Triangle Park, NC* 2 (1995), pp. 59–120.

[23] Mamadou T Bah, Prasanth B Nair, and Martin Browne. “Mesh morphing for finite element analysis of implant positioning in cementless total hip replacements”. In: *Elsevier* 31.10 (2009), pp. 1235–1243.

[24] Haris Balta et al. “3D registration and integrated segmentation framework for heterogeneous unmanned robotic systems”. In: *Multidisciplinary Digital Publishing Institute* 12.10 (2020), p. 1608.

[25] Maxim Bashkuev, Sandra Reitmaier, and Hendrik Schmidt. “Effect of disc degeneration on the mechanical behavior of the human lumbar spine: A probabilistic finite element study”. In: *Elsevier* 18.10 (2018), pp. 1910–1920.

[26] Maxim Bashkuev, Sandra Reitmaier, and Hendrik Schmidt. “Relationship between intervertebral disc and facet joint degeneration: A probabilistic finite element model study”. In: *Elsevier* 102 (2020), p. 109518.

[27] Thomas Baum et al. “Automatic detection of osteoporotic vertebral fractures in routine thoracic and abdominal MDCT”. In: *European Radiology* 24.4 (2014), pp. 872–880.

[28] Jens Behley et al. “Semantickitti: A dataset for semantic scene understanding of lidar sequences”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2019), pp. 9297–9307.

[29] H Bennani, B McCane, and J Cornwall. “Three dimensional (3D) lumbar vertebrae data set”. In: *Ubiquity Press* 15 (2016).

[30] Uttaran Bhattacharya and Venu Madhav Govindu. “Efficient and robust registration on the 3d special euclidean group”. In: *Proceedings of the IEEE International Conference on Computer Vision* (2019), pp. 5885–5894.

[31] Hakan Bilen et al. “Action recognition with dynamic image networks”. In: *IEEE* 40.12 (2017), pp. 2799–2813.

[32] BIM WIKI. “Parametric modelling”. In: *CIOB* (2021). URL: `www.designingbuildings.co.uk`.

[33] Christopher John Blackwell, Javid Khan, and Xianzhong Chen. “54-6: Holographic 3D Telepresence System with Light Field 3D Displays and Depth Cameras over a LAN”. In: *SID Symposium Digest of Technical Papers* 52.1 (2021), pp. 761–763.

[34] Lynn M Boorady. “Functional clothing—Principles of fit”. In: *NISCAIR-CSIR, India* (2011).

[35] Lynn M Boorady et al. “Protective clothing for pesticide applicators: A multi-method needs assessment”. In: *Journal of Textile and Apparel, Technology and Management* 6.2 (2009).

[36] Abhishektha Boppana and Allison P Anderson. “Dynamic foot morphology explained through 4D scanning and shape modeling”. In: *Elsevier* 122 (2021), p. 110465.

[37] Bart Bosmans et al. “Aortic root sizing for transcatheter aortic valve implantation using a shape model parameterisation”. In: *Springer* 57 (2019), pp. 2081–2092.

[38] PA Bromiley, JE Adams, and TF Cootes. “Localisation of vertebrae on DXA images using constrained local models with random forest regression voting”. In: *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging* (2015), pp. 159–171.

[39] Elizabeth Bye, Karen L Labat, and Marilyn R Delong. “Analysis of body measurement systems for apparel”. In: *Sage Publications Sage CA: Thousand Oaks, CA* 24.2 (2006), pp. 66–79.

[40] Julius Quinn Campbell and Anthony J Petrella. “An automated method for landmark identification and finite-element modeling of the lumbar spine”. In: *IEEE* 62.11 (2015), pp. 2709–2716.

[41] CAPTURE 3D. “Industrial Metrology 3D Solutions”. In: *CAPTURE 3D* (2022). URL: `https://www.capture3d.com/`.

[42] Nash Charlie. “nricp - Non-rigid iterative closest point”. In: *MATLAB* (2020). URL: `https://www.mathworks.com/matlabcentral/fileexchange/54077-optimal-step-nonrigid-icp`.

[43] Ayan Chaudhury. “Multilevel Optimization for Registration of Deformable Point Clouds”. In: *IEEE* 29 (2020), pp. 8735–8746.

[44] Xing Chen and James Davis. “An occlusion metric for selecting robust camera configurations”. In: *Springer* 19.4 (2008), pp. 217–222.

[45] Xinhai Chen et al. “Developing a new mesh quality evaluation method based on convolutional neural network”. In: *Taylor & Francis* 14.1 (2020), pp. 391–400.

[46] Shiyang Cheng et al. “Statistical non-rigid ICP algorithm and its application to 3D face alignment”. In: *Elsevier* 58 (2017), pp. 3–12.

[47] Jason Pui Yin Cheung et al. “Curve progression in adolescent idiopathic scoliosis does not match skeletal growth”. In: *Association of Bone and Joint Surgeons* 476.2 (2018), p. 429.

[48] Jae-sung Cho et al. “Scoliosis screening through a machine learning based gait analysis test”. In: *Springer* 19.12 (2018), pp. 1861–1872.

[49] Jiyoung Choi and Kyunghi Hong. “3D skin length deformation of lower body during knee joint flexion for the practical application of functional sportswear”. In: *Elsevier* 48 (2015), pp. 186–201.

[50] Sun-Yoon Choi and Susan P Ashdown. “Application of lower body girth change analysis using 3D body scanning to pants patterns”. In: *The Korean Society of Clothing and Textiles* 34.6 (2010), pp. 955–968.

[51] Sunyoon Choi and Susan P Ashdown. “3D body scan analysis of dimensional change in lower body measurements for active body positions”. In: *SAGE Publications Sage UK: London, England* 81.1 (2011), pp. 81–93.

[52] Haili Chui and Anand Rangarajan. “A new point matching algorithm for non-rigid registration”. In: *Elsevier* 89.2-3 (2003), pp. 114–141.

[53] Junyoung Chung et al. “Empirical evaluation of gated recurrent neural networks on sequence modeling”. In: *arXiv preprint arXiv:1412.3555* (2014).

[54] Luca Cirrottola et al. “Adaptive deformation of 3D unstructured meshes with curved body fitted boundaries with application to unsteady compressible flows”. In: *Elsevier* 433 (2021), p. 110177.

[55] Timothy F Cootes. “Fully automatic localisation of vertebrae in CT images using random forest regression voting”. In: *Computational Methods and Clinical Applications for Spine Imaging: 4th International Workshop and Challenge, CSI 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, October 17, 2016, Revised Selected Papers* 10182 (2017), p. 51.

[56] Timothy F Cootes et al. “Active shape models-their training and application”. In: *Elsevier* 61.1 (1995), pp. 38–59.

[57] Cregg K Cowan and Peter D Kovesi. “Automatic sensor placement from vision task requirements”. In: *IEEE* 10.3 (1988), pp. 407–416.

[58] Hang Dai, Nick Pears, and William Smith. “Non-rigid 3d shape registration using an adaptive template”. In: *Proceedings of the European Conference on Computer Vision (ECCV) Workshops* (2018), pp. 0–0.

[59] Zhu Daixian. “SIFT algorithm analysis and optimization”. In: *IASP 10 - 2010 International Conference on Image Analysis and Signal Processing* (2010), pp. 415–419.

[60] Timothy A Davis. “Algorithm 832: UMFPACK V4. 3—an unsymmetric-pattern multifrontal method”. In: *ACM* 30.2 (2004), pp. 196–199.

[61] M Dekker. “Mathematical programming”. In: *CRC, May* 4 (1986).

[62] Qingqiong Deng, Mingquan Zhou, and Zhongke Wu. “An automatic non-rigid registration method for dense surface models”. In: *2010 IEEE International Conference on Intelligent Computing and Intelligent Systems* 1 (2010), pp. 888–892.

[63] Sylvain Deschênes et al. “Diagnostic imaging of spinal deformities: Reducing patients radiation dose with a new slot-scanning X-ray imager”. In: *LWW* 35.9 (2010), pp. 989–994.

[64] DINED. “Anthropometric database TU Delft Ergonomics”. In: *Dined.Nl, 2011. http://dined.io.tudelft.nl/ergonomic/* (2020).

[65] Jessica A Dobson et al. “The three-dimensional shapes of underground coal miners’ feet do not match the internal dimensions of their work boots”. In: *Taylor & Francis* 61.4 (2018), pp. 588–602.

[66] Wesley Doorsamy, B Paul, and Tshilidzi Marwala. “The disruptive fourth industrial revolution”. In: *Springer* (2020).

[67] Zjenja Doubrovski, Jouke C Verlinden, and Jo MP Geraedts. “Optimal design for additive manufacturing: opportunities and challenges”. In: *International design engineering technical conferences and computers and information in engineering conference* 54860 (2011), pp. 635–646.

[68] Ian L Dryden and Kanti V Mardia. “Statistical shape analysis: with applications in R”. In: *John Wiley & Sons* 995 (2016).

[69] Jean Dubousset et al. “A new 2D and 3D imaging approach to musculoskeletal physiology and pathology with low-dose radiation and the standing position: the EOS system”. In: *Bulletin de l’Academie Nationale de Medecine* 189.2 (2005), pp. 287–97.

[70] Raul Durand, BG Pantoja-Rosero, and Vicente Oliveira. “A general mesh smoothing method for finite elements”. In: *Elsevier* 158 (2019), pp. 17–30.

[71] Roberto M Dyke et al. “Non-rigid registration under anisotropic deformations”. In: *Elsevier* 71 (2019), pp. 142–156.

[72] M Eekhout. “Design of a new patient-specific prosthesis for resurfacing of the distal radius”. In: *Delft University of Technology* (2014).

[73] Moritz Ehlke et al. “Fast generation of virtual X-ray images for reconstruction of 3D anatomy”. In: *IEEE* 19.12 (2013), pp. 2673–2682.

[74] Willemijn Elkhuizen et al. “Gloss, color, and topography scanning for reproducing a Painting’s appearance using 3D printing”. In: *ACM New York, NY, USA* 12.4 (2019), pp. 1–22.

[75] Jingfan Fan et al. “BIRNet: Brain image registration using dual-supervised fully convolutional networks”. In: *Elsevier* 54 (2019), pp. 193–206.

[76] FDA. “3D Printing of Medical Devices”. In: *FDA* (2020). URL: `https://www.fda.gov/medical-devices/products-and-%20medical-procedures/3d-printing-medical-devices`.

[77] Fda. “Methodological Approach to Developing a RiskRanking Model for Food Tracing FSMA Section 204 (21 U.S. Code § 2223)”. In: *FDA* (2020). URL: https://www.fda.gov/media/142247/download.

[78] *Fit to Deformation.* URL: http://www.drematrix.de/projects/fit-to-deformation/.

[79] Flavio et al. Fogliatto. “The mass customization decade: An updated review of the literature”. In: *Elsevier* 138.1 (2012), pp. 14–25.

[80] Ted Julien Tchinde Fotsin et al. “Shape, pose and density statistical model for 3D reconstruction of articulated structures from X-ray images”. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2019), pp. 2748–2751.

[81] Lori A Freitag. “On combining Laplacian and optimization-based mesh smoothing techniques”. In: *Argonne National Lab., IL (United States)* (1997).

[82] John E Freund, Irwin Miller, and Marylees Miller. “John E. Freund’s Mathematical Statistics: With Applications”. In: *Pearson Education India* (2004).

[83] Ran Gal and Daniel Cohen-Or. “Salient geometric features for partial shape matching and similarity”. In: *ACM New York, NY, USA* 25.1 (2006), pp. 130–150.

[84] Madeline Gannon, Tovi Grossman, and George Fitzmaurice. “Tactum: a skin-centric approach to digital design and fabrication”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (2015), pp. 1779–1788.

[85] Joseph E Gaudio et al. “Parameter estimation in adaptive control of time-varying systems under a range of excitation conditions”. In: *IEEE* (2021).

[86] Xuming Ge. “Non-rigid registration of 3D point clouds under isometric deformation”. In: *Elsevier* 121 (2016), pp. 192–202.

[87] Andreas Geiger et al. “Automatic camera and range sensor calibration using a single shot”. In: *2012 IEEE international conference on robotics and automation* (2012), pp. 3936–3943.

[88] Paul Louis George. “Tet meshing: construction, optimization and adaptation”. In: *8th international meshing roundtable* (1999), pp. 133–141.

[89] Syed Zulqarnain Gilani et al. “Geometric facial gender scoring: Objectivity of perception”. In: *Public Library of Science* 9.6 (2014), e99483.

[90] Steven Girdler et al. “Emerging techniques in diagnostic imaging for idiopathic scoliosis in children and adolescents: A review of the literature”. In: *Elsevier* 136 (2020), pp. 128–135.

[91] J Goerres et al. “Spinal pedicle screw planning using deformable atlas registration”. In: *IOP Publishing* 62.7 (2017), p. 2871.

[92] Lyè Goto et al. “Analysis of a 3D anthropometric data set of children for design applications”. In: *Proceedings 19th Triennial Congress of the IEA* 9 (2015), p. 14.

[93] Venu Madhav Govindu. “Lie-algebraic averaging for globally consistent motion estimation”. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.* 1 (2004), pp. I–I.

[94] Stefan Grau and Bettina Barisch-Fritz. “Improvement of safety shoe fit-evaluation of dynamic foot structure”. In: *Taylor & Francis* 10.3 (2018), pp. 179–187.

[95] Xiuye Gu et al. “Hplflownet: Hierarchical permutohedral lattice flownet for scene flow estimation on large-scale point clouds”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2019), pp. 3254–3263.

[96] Kaiwen Guo et al. “Real-time geometry, albedo, and motion reconstruction using a single rgb-d camera”. In: *ACM New York, NY, USA* 36.4 (2017), p. 1.

[97] Kaiwen Guo et al. “Robust non-rigid motion tracking and surface reconstruction using l0 regularization”. In: *Proceedings of the IEEE International Conference on Computer Vision* (2015), pp. 3083–3091.

[98] Shengnan Guo et al. “Attention based spatial-temporal graph convolutional networks for traffic flow forecasting”. In: *Proceedings of the AAAI conference on artificial intelligence* 33.01 (2019), pp. 922–929.

[99] Bin He, Yu Song, and Yan Wang. “Digital twin-driven design and manufacturing”. In: *American Society of Mechanical Engineers Digital Collection* 21.3 (2021).

[100] Tobias Heimann and Hans-Peter Meinzer. “Statistical shape models for 3D medical image segmentation: a review”. In: *Elsevier* 13.4 (2009), pp. 543–563.

[101] Osamu Hirose. “Executable source code for Bayesian Coherent Point Drift”. In: *github* (2020). URL: https://github.com/ohirose/bcpd.

[102] Osamu and Hirose. “A Bayesian Formulation of Coherent Point Drift”. In: *IEEE* (2020).

[103] Sepp Hochreiter and Jürgen Schmidhuber. “Long short-term memory”. In: *MIT Press* 9.8 (1997), pp. 1735–1780.

[104] Hidekata Hontani, Takamiti Matsuno, and Yoshihide Sawada. “Robust nonrigid ICP using outlier-sparsity regularization”. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition* (2012), pp. 174–181.

[105] Tao Hou et al. “Design of 3D Wireless Power Transfer System Based on 3D Printed Electronics”. In: *Institute of Electrical and Electronics Engineers Inc. IEEE Access* 7 (2019), pp. 94793–94805.

[106] Toon Huysmans et al. “Tijdschrift voor Human Factors 4 Tijdschrift voor Human Factors-jaargang 45-nr. 1-april 2020 DINED Mannequin A new form of anthropometry”. In: *Delft University of Technology* (2020). URL: http://dined.io.tudelft.nl.

[107] Bulat Ibragimov and Tomaž Vrtovec. “Landmark-Based Statistical Shape Representations”. In: *Elsevier* (2017), pp. 89–113.

[108] ISO. In: *ISO 14971:2019 Medical devices — Application of risk management to medical devices* (2019).

[109] Tao Jiang et al. “Consistent as-similar-as-possible non-isometric surface registration”. In: *Springer* 33.6-8 (2017), pp. 891–901.

[110] Can Jin, Shouguo Li, and Xu Yang. “Adaptive three-dimensional aggregate shape fitting and mesh optimization for finite-element modeling”. In: *American Society of Civil Engineers* 34.4 (2020), p. 04020020.

[111] Hongzhe Jin et al. “A dynamic parameter identification method for flexible joints based on adaptive control”. In: *IEEE* 23.6 (2018), pp. 2896–2908.

[112] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. “Snakes: Active contour models”. In: *Springer* 1.4 (1988), pp. 321–331.

[113] Hassan K Khalil and Jessy W Grizzle. “Nonlinear systems”. In: *Prentice hall Upper Saddle River, NJ* 3 (2002).

[114] Daehwan Kim and Daijin Kim. “A fast ICP algorithm for 3-D human body motion tracking”. In: *IEEE* 17.4 (2010), pp. 402–405.

[115] Hangkee Kim et al. “3D reconstruction of leg bones from X-ray images using CNN-based feature analysis”. In: *International Conference on Information and Communication Technology Convergence (ICTC)* (2019), pp. 669–672.

[116] Vladimir G Kim, Yaron Lipman, and Thomas Funkhouser. “Blended intrinsic maps”. In: *ACM Transactions on Graphics (TOG)* 30.4 (2011), p. 79.

[117] Diederik P Kingma and Jimmy Ba. “Adam: A method for stochastic optimization”. In: *arXiv preprint arXiv:1412.6980* (2014).

[118] Wm Kirk Jr and S M Ibrahim. “Fundamental relationship of fabric extensibility to anthropometric requirements and garment performance”. In: *Sage Publications Sage CA: Thousand Oaks, CA* 36.1 (1966), pp. 37–47.

[119] Anke Klepser and Simone Morlock. “4D Scanning-Dynamic View on Body Measurements”. In: *Communications in Development and Assembling of Textile Products* 1.1 (2020), pp. 30–38.

[120] Dejan Knez et al. “Computer-assisted pedicle screw trajectory planning using CT-inferred bone density: A demonstration against surgical outcomes”. In: *Wiley Online Library* 46.8 (2019), pp. 3543–3554.

[121] Patrick M Knupp. “Algebraic mesh quality metrics”. In: *SIAM* 23.1 (2001), pp. 193–218.

[122] Cindy Kohtala. “Addressing sustainability in research on distributed production: an integrated literature review”. In: *Elsevier* 106 (2015), pp. 654–668.

[123] Markus Rafael Konieczny, Hüsseyin Senyurt, and Rüdiger Krauspe. “Epidemiology of adolescent idiopathic scoliosis”. In: *The British Editorial Society of Bone & Joint Surgery London* 7.1 (2013), pp. 3–9.

[124] Y Koren et al. “Open-architecture products”. In: *Elsevier* 62.2 (2013), pp. 719–729.

[125] Yoram Koren. “The global manufacturing revolution: product-process-business integration and reconfigurable systems”. In: *John Wiley & Sons* 80 (2010).

[126] I Krauss et al. “Sex-related differences in foot shape”. In: *Taylor & Francis* 51.11 (2008), pp. 1693–1709.

[127] Inga Krauss et al. “Comparison of female foot morphology and last design in athletic footwear—are men’s lasts appropriate for women?” In: *Taylor & Francis* 18.2 (2010), pp. 140–156.

[128] Felix Sien Shi Kwa. “Design of an accurate and low cost 4D foot scanner for podiatrists”. In: *Delft University of Technology* (2021).

[129] Kwiseok Kwon and Cookhwan Kim. “How to design personalization in a context of customer retention: Who personalizes what and to what extent?” In: *Elsevier* 11.2 (2012), pp. 101–116.

[130] Jeffrey C Lagarias et al. “Convergence properties of the Nelder–Mead simplex method in low dimensions”. In: *SIAM* 9.1 (1998), pp. 112–147.

[131] Hans Lamecker, Thomas H Wenckebach, and H-C Hege. “Atlas-based 3D-shape reconstruction from X-ray images”. In: *18th International Conference on Pattern Recognition (ICPR’06)* 1 (2006), pp. 371–374.

[132] Daniel Lanzoni et al. “Design of Customized Virtual Reality Serious Games for the Cognitive Rehabilitation of Retrograde Amnesia After Brain Stroke”. In: *American Society of Mechanical Engineers Digital Collection* 22.3 (2022).

[133] A Le Bras et al. “Three-dimensional (3D) detailed reconstruction of human vertebrae from low-dose digital stereoradiography”. In: *Springer* 13.2 (2003), pp. 57–62.

[134] Byung-Uk Lee, Chul-Min Kim, and Rae-Hong Park. “An orientation reliability matrix for the iterative closest point algorithm”. In: *IEEE* 22.10 (2000), pp. 1205–1208.

[135] Jeongran Lee and Susan P Ashdoon. “Upper body surface change analysis using 3-D body scanner”. In: *The Korean Society of Clothing and Textiles* 29.12 (2005), pp. 1595–1607.

[136] Sing Chun Lee and Misha Kazhdan. “Computer Graphics Forum”. In: *Wiley Online Library* 38.5 (2019), pp. 27–37.

[137] Sing Chun Lee and Misha Kazhdan. “Executable source code for Dense Point-to-Point Correspondences for Genus-Zero Surfaces”. In: *github* (2019). URL: `https://github.com/mkazhdan/DenseP2PCorrespondences`.

[138] W-S Lee and Nadia Magnenat-Thalmann. “Fast head modeling for animation”. In: *Elsevier* 18.4 (2000), pp. 355–364.

[139] Victor Lempitsky and Yuri Boykov. “Global optimization for shape fitting”. In: *2007 IEEE Conference on Computer Vision and Pattern Recognition* (2007), pp. 1–8.

[140] Frank L Lewis, Draguna Vrabie, and Vassilis L Syrmos. “Optimal control”. In: *John Wiley & Sons* (2012).

[141] Hao Li, Robert W Sumner, and Mark Pauly. “Global correspondence optimization for non-rigid registration of depth scans”. In: *Computer graphics forum* 27.5 (2008), pp. 1421–1430.

[142] Jiaxin Li, Ben M Chen, and Gim Hee Lee. “So-net: Self-organizing network for point cloud analysis”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018), pp. 9397–9406.

[143] Kun Li et al. “Robust non-rigid registration with reweighted position and transformation sparsity”. In: *IEEE* 25.6 (2018), pp. 2255–2269.

[144] Maosen Li et al. “Actional-structural graph convolutional networks for skeleton-based action recognition”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2019), pp. 3595–3603.

[145] Ruixuan Li et al. “A Framework of Real-time Freehand Ultrasound Reconstruction based on Deep Learning for Spine Surgery”. In: *10th Conference on New Technologies for Computer and Robot Assisted Surgery, Date: 2020/09/28-2020/09/30, Location: Barcelona, Spain* (2020).

[146] Yang Li et al. “4dcomplete: Non-rigid motion estimation beyond the observable surface”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021, pp. 12706–12716.

[147] Luming Liang et al. “Nonrigid iterative closest points for registration of 3D biomedical surfaces”. In: *Elsevier* 100 (2018), pp. 141–154.

[148] Ruixin Liang et al. “Machine Learning Approaches to Predict Scoliosis”. In: *International Conference on Applied Human Factors and Ergonomics* (2021), pp. 116–121.

[149] Yuxuan Liang et al. “Geoman: Multi-level attention networks for geo-sensory time series prediction.” In: *IJCAI* 2018 (2018), pp. 3428–3434.

[150] Miao Liao et al. “Modeling deformable objects from a single depth camera”. In: *2009 IEEE 12th International Conference on Computer Vision* (2009), pp. 167–174.

[151] Florentin Liebmann et al. “Registration made easy–standalone orthopedic navigation with HoloLens”. In: *arXiv preprint arXiv:2001.06209* (2020).

[152] Anwei Liu and Barry Joe. “Relationship between tetrahedron shape measures”. In: *Springer* 34.2 (1994), pp. 268–287.

[153] Jiaheng Liu and Dong Xu. “GeometryMotion-Net: A strong two-stream baseline for 3D action recognition”. In: *IEEE* 31.12 (2021), pp. 4711–4721.

[154] Xingyu Liu, Charles R Qi, and Leonidas J Guibas. “Flownet3d: Learning scene flow in 3d point clouds”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2019), pp. 529–537.

[155] Xingyu Liu, Mengyuan Yan, and Jeannette Bohg. “Meteornet: Deep learning on dynamic 3d point cloud sequences”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2019), pp. 9246–9255.

[156] Yongcheng Liu et al. “Relation-shape convolutional neural network for point cloud analysis”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2019), pp. 8895–8904.

[157] Christine Loercher, Simone Morlock, and Andreas Schenk. “Design of a motion-oriented size system for optimizing professional clothing and personal protective equipment”. In: *Journal of Fashion Technology* (2018).

[158] Daniel Luft and Volker Schulz. “Pre-Shape Calculus: Foundations and Application to Mesh Quality Optimization”. In: *arXiv preprint arXiv:2012.09124* (2020).

[159] P Lundgren et al. “Invasive in vivo measurement of rear-, mid-and forefoot motion during walking”. In: *Elsevier* 28.1 (2008), pp. 93–100.

[160] Praveen Kumar Reddy Maddikunta et al. “Industry 5.0: A survey on enabling technologies and potential applications”. In: *Elsevier* 26 (Mar. 2022), p. 100257.

[161] Rishi Malhan et al. “Algorithms for Improving Speed and Accuracy of Automated Three-Dimensional Reconstruction With a Depth Camera Mounted on An Industrial Robot”. In: *American Society of Mechanical Engineers* 22.3 (2022), p. 31012.

[162] Samuel Marcos, Jaime Gómez-Garcıa-Bermejo, and Eduardo Zalama. “A realistic, virtual head for human–computer interaction”. In: *Oxford University Press Oxford, UK* 22.3 (2010), pp. 176–192.

[163] José M Mart\'\inez-Mart\'\inez et al. “Use of SOMs for footwear comfort evaluation”. In: *Springer* 28.7 (2017), pp. 1763–1773.

[164] Scott Mason et al. “Heuristic reasoning strategy for automated sensor placement”. In: *[Falls Church, Va.] American Society of Photogrammetry.* 63.9 (1997), pp. 1093–1101.

[165] MATLAB. “Fuzzy membership function in MATLAB”. In: *MATLAB* (2019). URL: https://www.mathworks.com/help/fuzzy/fismf.html.

[166] *MATLAB Convhull Function*. The MathWorks, Natick, MA, USA. (2020).

[167] *MATLAB Deep Learning Toolbox*. The MathWorks, Natick, MA, USA. (2018).

[168] George K Matsopoulos et al. “Thoracic non-rigid registration combining self-organizing maps and radial basis functions”. In: *Elsevier* 9.3 (2005), pp. 237–254.

[169] Xandra van Megen. “Design of a new personalized dental implant and its surgical procedure”. In: *Delft University of Technology* (2019).

[170] R Melnikova, A Ehrmann, and K Finsterbusch. “3D printing of textile-based structures by Fused Deposition Modelling (FDM) with different polymer materials”. In: *IOP conference series: materials science and engineering* 62.1 (2014), p. 012018.

[171] Microsoft. “Synchronize multiple Azure Kinect DK devices”. In: *Microsoft* (2020). URL: https://learn.microsoft.com/en-us/azure/kinect-dk/multi-camera-sync.

[172] Chaobo Min et al. “Adaptive enhanced affine transformation for non-rigid registration of visible and infrared images”. In: *Wiley Online Library* (2021).

[173] A Minnoye et al. “PERSONALIZED PRODUCT DESIGN THROUGH DIGITAL FABRICATION”. In: *ASME* (2022).

[174] Dominic Mischler et al. “Computational optimisation of screw orientations for improved locking plate fixation of proximal humerus fractures”. In: *Elsevier* 25 (2020), pp. 96–104.

[175] Kentaro Mori et al. “Heart failure diagnosis for tagged magnetic resonance images”. In: *2017 International Conference on Machine Learning and Cybernetics (ICMLC)* 1 (2017), pp. 67–70.

[176] Simone Morlock, C Lörcher, and A Schenk. “Entwicklung eines ergonomisch- und bewegungsorientierten Größensystems für Funktionsmaße zur optimierten Gestaltung von Berufs-und Schutzbekleidung”. In: *IGF Project* 18993 (2016).

[177] Fabio Müller et al. “Augmented reality navigation for spinal pedicle screw instrumentation using intraoperative 3D imaging”. In: *Elsevier* 20.4 (2020), pp. 621–628.

[178] A Mundermann, Darren J Stefanyshyn, Benno M Nigg, et al. “Relationship between footwear comfort of shoe inserts and anthropometric and sensory factors”. In: *WILLIAMS & WILKINS* 33.11 (2001), pp. 1939–1945.

[179] Andriy Myronenko and Xubo Song. “Executable source code for Coherent Point Drift library for nonrigid registration”. In: *github* (2010). URL: `https://github.com/markeroon/CoherentPointDrift`.

[180] Andriy Myronenko and Xubo Song. “Point set registration: Coherent point drift”. In: *IEEE* 32.12 (2010), pp. 2262–2275.

[181] Troy Robert Nachtigall et al. “Towards ultra personalized 4D printed shoes”. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (2018), CS20.

[182] Ming Shan Ng et al. “Design for digital fabrication: an industry needs analysis of collaboration platforms and integrated management processes”. In: *International Association on Automation and Robotics in Construction (IAARC)* 2020 (2020), pp. 318–325.

[183] William Odom et al. “From research prototype to research product”. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (2016), pp. 2549–2561.

[184] Eshed Ohn-Bar and Mohan Trivedi. “Joint angles similarities and HOG2 for action recognition”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (2013), pp. 465–470.

[185] Ozan Oktay et al. “Stratified decision forests for accurate anatomical landmark localization in cardiac images”. In: *IEEE* 36.1 (2016), pp. 332–342.

[186] Gustavo Olague and Roger Mohr. “Optimal camera placement for accurate reconstruction”. In: *Elsevier* 35.4 (2002), pp. 927–944.

[187] Omar Oreifej and Zicheng Liu. “Hon4d: Histogram of oriented 4d normals for activity recognition from depth sequences”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition* (2013), pp. 716–723.

[188] Ahmed A A Osman et al. “SUPR: A Sparse Unified Part-Based Human Body Model”. In: *European Conference on Computer Vision (ECCV)*. 2022. URL: `https://supr.is.tue.mpg.de`.

[189] Sang Il Park and Jessica K Hodgins. “Capturing and animating skin deformation in human motion”. In: *ACM New York, NY, USA* 25.3 (2006), pp. 881–889.

[190] Saba Pasha et al. “Machine Learning Predicts the 3D Outcomes of Adolescent Idiopathic Scoliosis Surgery Using Patient–Surgeon Specific Parameters”. In: *LWW* 46.9 (2021), pp. 579–587.

[191] A. Pawar et al. “Adaptive FEM-based nonrigid image registration using truncated hierarchical B-splines”. In: *Elsevier* 72.8 (2016), pp. 2028–2040.

[192] A. Pawar et al. “DTHB3D_Reg: Dynamic truncated hierarchical B-spline based 3D nonrigid image registration”. In: *Communications in Computational Physics* 23.3 (2018), pp. 877–898.

[193] A. Pawar et al. “Joint image segmentation and registration based on a dynamic level set approach using truncated hierarchical B-splines”. In: *Elsevier* 78 (2019), pp. 3250–3267.

[194] Li Peng et al. “Prediction of Proximal Junctional Kyphosis After Posterior Scoliosis Surgery With Machine Learning in the Lenke 5 Adolescent Idiopathic Scoliosis Patient”. In: *Frontiers Media SA* 8 (2020), pp. 1–10.

[195] James R Peters et al. “Age-and gender-related changes in pediatric thoracic vertebral morphology”. In: *Elsevier* 15.5 (2015), pp. 1000–1020.

[196] D. Pickup et al. “SHREC’14 track: Shape Retrieval of Non-Rigid 3D Human Models”. In: *Eurographics Association.* EG 3DOR’14 (2014).

[197] B Joseph Pine, Stan Davis, et al. “Mass customization: the new frontier in business competition”. In: *Harvard business school press Boston, MA* (1993).

[198] Buddie Joseph Pine. “Paradigm shift–from mass production to mass customization”. In: *Massachusetts Institute of Technology* (1991).

[199] Ronald Poppe. “A survey on vision-based human action recognition”. In: *Elsevier* 28.6 (2010), pp. 976–990.

[200] Moira K Pryhoda et al. “Alternative upper configurations during agility-based movements: part 1, biomechanical performance”. In: *Taylor & Francis* 13.1 (2021), pp. 91–103.

[201] “Publications Office of the EU”. In: *Regulation (EC) No 1107/2006 of the European Parliament and of the Council of 5 July 2006 concerning the rights of disabled persons and persons with reduced mobility when travelling by air (Text with EEA relevance)* (2006).

[202] Charles R Qi et al. “Pointnet: Deep learning on point sets for 3d classification and segmentation”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017), pp. 652–660.

[203] Charles Ruizhongtai Qi et al. “Pointnet++: Deep hierarchical feature learning on point sets in a metric space”. In: *Advances in neural information processing systems* 30 (2017).

[204] C Do Quang. “Feasibility of Agile Manufacturing for the interior vision of 2025”. In: *Technical University of Delft* (2019).

[205] Pooya Rahimian et al. “Using a virtual environment to study the impact of sending traffic alerts to texting pedestrians”. In: *2016 IEEE Virtual Reality (VR)* (2016), pp. 141–149.

[206] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations”. In: *Journal of Computational Physics* 378 (2019), pp. 686–707.

[207] Cornelius Johannes Frederik Reyneke et al. “Review of 2-D/3-D reconstruction using statistical shape and intensity models and X-ray image synthesis: Toward a unified framework”. In: *IEEE* 12 (2018), pp. 269–286.

[208] Martin G Roberts, Timothy F Cootes, and Judith E Adams. “Automatic location of vertebrae on DXA images using random forest regression”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2012), pp. 361–368.

[209] MG Roberts et al. “Detection of vertebral fractures in DXA VFA images using statistical models of appearance and a semi-automatic segmentation”. In: *Springer* 21.12 (2010), pp. 2037–2046.

[210] Kathleen M Robinette et al. “Civilian american and european surface anthropometry resource (caesar), final report. volume 1. summary”. In: *Sytronics Inc Dayton Oh* (2002).

[211] R Rohan. “Top product customization challenges and ways to overcome them”. In: *https://www.wtpbiz.com/product-customization- challenges/.* (2021).

[212] Roland DG. “7 benefits of product customisation”. In: *Roland DG* (2018). URL: https://www.rolanddg.eu/en/blog/2018/08/13/benefits-of-%20product-customisation.

[213] Lars Rossing et al. “Bonding between silicones and thermoplastics using 3D printed mechanical interlocking”. In: *Elsevier* 186 (2020), p. 108254.

[214] Mohammad Rouhani, Edmond Boyer, and Angel D Sappa. “Non-rigid registration meets surface reconstruction”. In: *2014 2nd International Conference on 3D Vision* 1 (2014), pp. 617–624.

[215] RUSSIAN3DSCANNER. “Wrap 3.4”. In: *R3DS* (2019).

[216] Sourav Saha et al. “Hierarchical Deep Learning Neural Network (HiDeNN): An artificial intelligence (AI) framework for computational science and engineering”. In: *Elsevier* 373 (2021), p. 113452.

[217] Yusuf Sahillioğlu. “Recent advances in shape correspondence”. In: *Springer* 36.8 (2020), pp. 1705–1721.

[218] John N Sanders-Reed. “Error propagation in two-sensor three-dimensional position estimation”. In: *International Society for Optics and Photonics* 40.4 (2001), pp. 627–636.

[219] Nazli Sarkalkan, Harrie Weinans, and Amir A Zadpoor. “Statistical shape and appearance models of bones”. In: *Elsevier* 60 (2014), pp. 129–140.

[220] U Schmid and J Mecheels. “Kräfte an Textilien und Nähten der Kleidung in Abhängigkeit von Körperbewegungen und Kleidungsschnitt’”. In: *Bekleidung und Wäsche* 33 (1981), pp. 77–82.

[221] G W Schmid-Schönbein et al. “Passive mechanical properties of human leukocytes”. In: *Elsevier* 36.1 (1981), pp. 243–256.

[222] Gregory C Sharp, Sang W Lee, and David K Wehe. “ICP registration using invariant features”. In: *IEEE* 24.1 (2002), pp. 90–102.

[223] Lei Shi et al. “Skeleton-based action recognition with directed graph neural networks”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2019), pp. 7912–7921.

[224] Shahram Shirani. “Content-based multiple description image coding”. In: *IEEE* 8.2 (2006), pp. 411–419.

[225] Chenyang Si et al. “An attention enhanced graph convolutional lstm network for skeleton-based action recognition”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2019), pp. 1227–1236.

[226] Adil Masood Siddiqui, Asif Masood, and Muhammad Saleem. “A locally constrained radial basis function for registration and warping of images”. In: *Elsevier* 30.4 (2009), pp. 377–390.

[227] Luciano Silva, Olga Regina Pereira Bellon, and Kim L Boyer. “Precision range image registration using a robust surface interpenetration measure and enhanced genetic algorithms”. In: *IEEE* 27.5 (2005), pp. 762–776.

[228] sketchfab. “Thoracic-vertebrae 3D models”. In: *Sketchfab* (2020). URL: `https://sketchfab.com/search?q=tag$%5C%$3Athoracic-vertebrae$%5C&$sort$%5C_$by=-likeCount$%5C&$type=models`.

[229] Jean-Jacques E Slotine, Weiping Li, et al. “Applied nonlinear control”. In: *Prentice hall Englewood Cliffs, NJ* 199.1 (1991).

[230] Y Song and Peter Vink. “On the objective assessment of comfort”. In: *3rd International Comfort Congress 2021* (2021).

[231] Y Song et al. “Parameter-driven freeform deformations.” In: *diglib.eg.org* (2003).

[232] Michael Sony. “Industry 4.0 and lean management: a proposed integration model and research propositions”. In: *Taylor and Francis Ltd.* 6.1 (Jan. 2018), pp. 416–432.

[233] Olga Sorkine and Marc Alexa. “As-rigid-as-possible surface modeling”. In: *Symposium on Geometry processing* 4 (2007), pp. 109–116.

[234] IAF Stokes. “Mechanical effects on skeletal growth”. In: *Hylonome* 2.3 (2002), pp. 277–280.

[235] CCM Stolwijk and Matthijs Punter. “Going Digital: Field labs to accelerate the digitization of the Dutch Industry”. In: *Den Haag: TNO* (2018).

[236] Mervyn Stone. “Cross-validatory choice and assessment of statistical predictions”. In: *Wiley Online Library* 36.2 (1974), pp. 111–133.

[237] Hang Su et al. “Splatnet: Sparse lattice networks for point cloud processing”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018), pp. 2530–2539.

[238] Robert W Sumner and Jovan Popović. “Deformation transfer for triangle meshes”. In: *ACM* 23.3 (2004), pp. 399–405.

[239] Zehua Sun et al. “Human Action Recognition From Various Data Modalities: A Review”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022), pp. 1–20.

[240] Jochen Süßmuth, Marco Winter, and Günther Greiner. “Reconstructing animated meshes from time-varying point clouds”. In: *Computer Graphics Forum* 27.5 (2008), pp. 1469–1476.

[241] Farzam Tajdari. “Adaptive time-delay estimation and control of optimized Stewart robot”. In: *SAGE Publications Sage UK: London, England* (2022), pp. 1–21.

[242] Farzam Tajdari and Naeim Ebrahimi Toulkani. “Implementation and intelligent gain tuning feedback–based optimal torque control of a rotary parallel robot”. In: *SAGE Publications Sage UK: London, England* (2021), p. 10775463211019177.

[243] Farzam Tajdari, Toon Huysmans, and Yu Song. “Non-Rigid Registration Via Intelligent Adaptive Feedback Control”. In: *IEEE Transactions on Visualization and Computer Graphics* (2023), pp. 1–17.

[244] Farzam Tajdari and Claudio Roncoli. “Adaptive traffic control at motorway bottlenecks with time-varying fundamental diagram”. In: *Elsevier* 54.2 (2021), pp. 271–277.

[245] Farzam Tajdari, Claudio Roncoli, and Markos Papageorgiou. “Feedback-based ramp metering and lane-changing control with connected and automated vehicles”. In: *IEEE* (2020).

[246] Farzam Tajdari, Mahsa Tajdari, and Amin Rezaei. “Discrete time delay feedback control of stewart platform with intelligent optimizer weight tuner”. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)* (2021), pp. 12701–12707.

[247] Farzam Tajdari, Naeim Ebrahimi Toulkani, and Maral Nourimand. “Intelligent architecture for car-following behaviour observing lane-changer: Modeling and control”. In: *2020 10th International Conference on Computer and Knowledge Engineering (ICCKE)* (2020), pp. 579–584.

[248] Farzam Tajdari, Naeim Ebrahimi Toulkani, and Nima Zhilakzadeh. “Intelligent optimal feed-back torque control of a 6dof surgical rotary robot”. In: *2020 11th Power Electronics, Drive Systems, and Technologies Conference (PEDSTC)* (2020), pp. 1–6.

[249] Farzam Tajdari, Naeim Ebrahimi Toulkani, and Nima Zhilakzadeh. “Semi-real evaluation, and adaptive control of a 6dof surgical robot”. In: *2020 11th Power Electronics, Drive Systems, and Technologies Conference (PEDSTC)* (2020), pp. 1–6.

[250] Farzam Tajdari et al. “Dynamic 3D mesh reconstruction based on nonrigid iterative closest-farthest points registration”. In: *International design engineering technical conferences and computers and information in engineering conference* (2022), pp. 1–9.

[251] Farzam Tajdari et al. “Feature preserving non-rigid iterative weighted closest point and semi-curvature registration”. In: *IEEE* 31 (2022), pp. 1841–1856.

[252] Farzam Tajdari et al. “Fuzzy control of anticipation and evaluation behaviour in real traffic flow”. In: *2019 7th International Conference on Robotics and Mechatronics (ICRoM)* (2019), pp. 248–253.

[253] Farzam Tajdari et al. “Optimal position of cameras design in a 4d foot scanner”. In: *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* 2022 (2022), pp. 1–9.

[254] Farzam Tajdari et al. “Robust control of a 3-DOF parallel cable robot using an adaptive neuro-fuzzy inference system”. In: *2017 Artificial Intelligence and Robotics (IRANOPEN)* (2017), pp. 97–101.

[255] Farzam Tajdari et al. “Simultaneous intelligent anticipation and control of follower vehicle observing exiting lane changer”. In: *IEEE* 70.9 (2021), pp. 8567–8577.

[256] Mahsa Tajdari et al. “2D to 3D Volumetric Reconstruction of Human Spine for Diagnosis and Prognosis of Spinal Deformities”. In: *Conference: 16th US National Congress on Computational Mechanics* (2021).

[257] Mahsa Tajdari et al. “Image-based modelling for adolescent idiopathic scoliosis: mechanistic machine learning analysis and prediction”. In: *Elsevier* 374 (2021), p. 113590.

[258] Mahsa Tajdari et al. “Next-generation prognosis framework for pediatric spinal deformities using bio-informed deep learning networks”. In: *Springer* 38.5 (2022), pp. 4061–4084.

[259] F Tajdati et al. “4D Feet: Registering Walking Foot Shapes Using Attention Enhanced Dynamic-Synchronized Graph Convolutional LSTM Network”. In: *arXiv preprint arXiv:2307.12377* (2023).

[260] Chak Hin Tam et al. “Drop-on-demand printing of personalised orodispersible films fabricated by precision micro-dispensing”. In: *Elsevier B.V.* 610 (Dec. 2021).

[261] Changbai Tan et al. “Incorporating customer personalization preferences in open product architecture design”. In: *Elsevier B.V.* 56 (July 2020), pp. 72–83.

[262] Sarah Tang and Afzal Godil. “An evaluation of local shape descriptors for 3D shape retrieval”. In: *Three-Dimensional Image Processing (3DIP) and Applications II* 8290 (2012), 82900N.

[263] Ulla Hellstrand Tang et al. “Foot anthropometrics in individuals with diabetes compared with the general Swedish population: Implications for shoe design”. In: *Foot Ankle Online J* 10.3 (2017), p. 1.

[264] M Lamine Tazir et al. “CICP: Cluster Iterative Closest Point for sparse–dense point cloud registration”. In: *Elsevier* 108 (2018), pp. 66–86.

[265] Mohamed Lamine Tazir et al. “Cluster ICP: Towards Sparse to Dense Registration”. In: *International Conference on Intelligent Autonomous Systems* (2018), pp. 730–747.

[266] *Temporal-3dMD Systems (4D)*. URL: https://3dmd.com/.

[267] Daniele Tomassoni, Enea Traini, and Francesco Amenta. “Gender and age related differences in foot morphology”. In: *Elsevier* 79.4 (2014), pp. 421–427.

[268] I Toschi et al. “Quality-based registration refinement of airborne LiDAR and photogrammetric point clouds”. In: *Elsevier* 172 (2021), pp. 160–170.

[269] TU Delft. “Flying V project”. In: *Technical University of Delft* (2021). URL: https://www.tudelft.nl/lr/flying-v.

[270] TUDelft. “DINED / Anthropometry in design”. In: *DINED* (2020). URL: https://dined.io.tudelft.nl/en.

[271] Sebastien Valette, Jean Marc Chassery, and Remy Prost. “Generic remeshing of 3D triangular meshes with metric-dependent discrete Voronoi diagrams”. In: *IEEE* 14.2 (2008), pp. 369–381.

[272] R Van Wijngaarden. “Design of customizable sunglasses based on additive manufacturing techniques”. In: *Technical University of Delft* (2014).

[273] Ewoud Veltmeijer. “Development of a 3D printed patient specific Ankle Foot Orthosis”. In: *Delft University of Technology* (2019).

[274] Matthias Vestner et al. “Efficient deformable shape correspondence via kernel matching”. In: *2017 International Conference on 3D Vision (3DV)* (2017), pp. 517–526.

[275] Matthias Vestner et al. “Executable source code for implementation of Efficient Deformable Shape Correspondence via Kernel Matching”. In: *github* (2017). URL: https://github.com/zorah/KernelMatching.

[276] Neil J Vickers. “Animal communication: when i’m calling you, will you answer too?” In: *Elsevier* 27.14 (2017), R713–R715.

[277] Antonio Visioli. “Practical PID control”. In: *Springer Science & Business Media* (2006).

[278] Limin Wang et al. “Temporal segment networks: Towards good practices for deep action recognition”. In: *European conference on computer vision* (2016), pp. 20–36.

[279] Pichao Wang et al. “Depth pooling based large-scale 3-d action recognition with convolutional neural networks”. In: *IEEE* 20.5 (2018), pp. 1051–1061.

[280] Pichao Wang et al. “RGB-D-based human motion recognition with deep learning: A survey”. In: *Elsevier* 171 (2018), pp. 118–139.

[281] Shubo Wang and Jing Na. “Parameter estimation and adaptive control for servo mechanisms with friction compensation”. In: *IEEE* 16.11 (2020), pp. 6816–6825.

[282] Yancheng Wang et al. “3dv: 3d dynamic voxel for action recognition in depth video”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2020), pp. 511–520.

[283] Yi Wang et al. “Industry 4.0: a way from mass customization to mass personalization production”. In: *Springer* 5.4 (2017), pp. 311–320.

[284] Dwi Basuki Wibowo et al. “Effect of in-shoe foot orthosis contours on heel pain due to calcaneal spurs”. In: *MDPI AG* 9.3 (Jan. 2019).

[285] Gregory C Wiggins et al. “Pediatric spinal deformities”. In: *American Association of Neurological Surgeons* 14.1 (2003), pp. 1–14.

[286] Hans-Joachim Wilke, Volker Michael Betz, and Annette Kienle. “Morphometry of the kangaroo spine and its comparison with human spinal data”. In: *Wiley Online Library* 238.3 (2021), pp. 626–642.

[287] Billy M Williams and Lester A Hoel. “Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results”. In: *American Society of Civil Engineers* 129.6 (2003), pp. 664–672.

[288] Robert L. Williams and Douglas A. Lawrence. “Linear State-Space Control Systems”. In: *John Wiley & Sons, Inc.* (2007).

[289] Jeannette M. Wing. “Computational thinking and thinking about computing”. In: *Royal Society* 366.1881 (Oct. 2008), pp. 3717–3725.

[290] Peter Wolf et al. “Functional units of the human foot”. In: *Elsevier* 28.3 (2008), pp. 434–441.

[291] Chenglei Wu et al. “Shading-based dynamic shape refinement from multi-view video under general illumination”. In: *2011 International Conference on Computer Vision* (2011), pp. 1108–1115.

[292] J Wu et al. “Advances in generative design”. In: *Elsevier Ltd* 116 (Nov. 2019).

[293] Jiajun Wu et al. “Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling”. In: *proceedings.neurips.cc* (2016).

[294] John Jun Wu, Rajeev Sharma, and Thomas S Huang. “Analysis of uncertainty bounds due to quantization for three-dimensional position estimation using multiple cameras”. In: *International Society for Optics and Photonics* 37.1 (1998), pp. 280–292.

[295] Jun Wu, Ole Sigmund, and Jeroen P. Groen. “Topology optimization of multi-scale structures: a review”. In: *Springer Science and Business Media Deutschland GmbH* 63.3 (Mar. 2021), pp. 1455–1480.

[296] Jun Wu et al. “Infill Optimization for Additive Manufacturing-Approaching Bone-Like Porous Structures”. In: *IEEE Computer Society* 24.2 (Feb. 2018), pp. 1127–1140.

[297] Liyang Wu et al. “Robust non-rigid registration algorithm based on local affine registration”. In: *Ninth International Conference on Graphic and Image Processing (ICGIP 2017)* 10615 (2018), 106151Q.

[298] Roshna E Wunderlich and Peter R Cavanagh. “Gender differences in adult foot shape: implications for shoe design.” In: *Medicine and science in sports and exercise* 33.4 (2001), pp. 605–611.

[299] P Xi, WS Lee, and C Shu. “In graphics, interface”. In: *Analysis of segmented human body scans* (2007), pp. 19–26.

[300] Yang Xiao et al. “Action recognition for depth video using multi-view dynamic images”. In: *Elsevier* 480 (2019), pp. 287–304.

[301] Lei Xiong et al. “Robust Non-rigid Registration Based on Affine ICP Algorithm and Part-Based Method”. In: *Springer* 48.3 (2018), pp. 1305–1321.

[302] Shuping Xiong et al. “Foot deformations under different load-bearing conditions and their relationships to stature and body weight”. In: *The Anthropological Society of Nippon* 117.2 (2009), pp. 77–88.

[303] Jun Xu et al. “Computational Design for Digitally Fabricated 3D Inductive Power Transfer Coils”. In: *ASME International* 22.3 (June 2022).

[304] Zhiqiang Xu and Guoliang Xu. “Discrete schemes for Gaussian curvature and their convergence”. In: *Elsevier* 57.7 (2009), pp. 1187–1195.

[305] Sijie Yan, Yuanjun Xiong, and Dahua Lin. “Spatial temporal graph convolutional networks for skeleton-based action recognition”. In: *Thirty-second AAAI conference on artificial intelligence* (2018).

[306] Jingyu Yang et al. “Sparse Non-rigid Registration of 3D Shapes”. In: *Computer Graphics Forum* 34.5 (2015), pp. 89–99.

[307] Xiaodong Yang and YingLi Tian. “Super normal vector for activity recognition using depth sequences”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition* (2014), pp. 804–811.

[308] Xiaodong Yang and YingLi Tian. “Super normal vector for human activity recognition with depth cameras”. In: *IEEE* 39.5 (2016), pp. 1028–1039.

[309] Yusheng Yang et al. “Identify dominant dimensions of 3D hand shapes using statistical shape model and deep neural network”. In: *Elsevier* 96 (2021), p. 103462.

[310] Yusheng Yang et al. “Posture-Invariant 3D Human Hand Statistical Shape Model”. In: *American Society of Mechanical Engineers* 21.3 (2021), p. 031006.

[311] Yusheng Yang et al. “Posture-invariant three dimensional human hand statistical shape model”. In: *American Society of Mechanical Engineers (ASME)* 21.3 (June 2021).

[312] Yusheng Yang et al. “The development of a low-cost photogrammetry-based 3D hand scanner”. In: *Elsevier Ltd* 10 (Oct. 2021).

[313] Zongwei Yao et al. “Point cloud registration algorithm based on curvature feature similarity”. In: *Elsevier* 177 (2021), p. 109274.

[314] N Zakaria. “Infants and children: Understanding sizing, body shapes and apparel requirements for infants and children”. In: *Elsevier* (2014), pp. 95–131.

[315] Karl E Zelik and Eric C Honert. “Ankle and foot power in gait analysis: Implications for science, technology and clinical assessment”. In: *Elsevier* 75 (2018), pp. 1–12.

[316] Chunjia Zhang et al. “Robust 3D point set registration using iterative closest point algorithm with bounded rotation angle”. In: *Elsevier* 120 (2016), pp. 777–788.

[317] Ju Zhang, David Ackland, and Justin Fernandez. “Point-cloud registration using adaptive radial basis functions”. In: *Taylor & Francis* 21.7 (2018), pp. 498–502.

[318] Pengfei Zhang et al. “View adaptive neural networks for high performance skeleton-based human action recognition”. In: *IEEE* 41.8 (2019), pp. 1963–1978.

[319] Y. J. Zhang. “Geometric Modeling and Mesh Generation from Scanned Images”. In: *Chapman and Hall/CRC* (2016).

[320] Eric Zivot and Jiahui Wang. “Vector autoregressive models for multivariate time series”. In: *Springer* (2006), pp. 385–429.

ACKNOWLEDGEMENTS

Gratitude fills my heart as I extend my sincere appreciation to the numerous individuals whose invaluable contributions have played a pivotal role in the development and realization of this doctoral thesis.

Foremost, I express my profound gratitude to my esteemed Ph.D. promoters, **Dr. Yu (Wolf) Song** and **Dr. Toon Huysmans**. Without their unwavering support and academic expertise, this work would not have come to fruition. Throughout my research journey, their mentorship has been a beacon of guidance and inspiration. Dear **Wolf** and Dear **Toon**, our collaboration was marked by celebrations and challenges, and I am deeply moved by the sentiment you expressed: "Let's see what is the best for your future Farzam, as we are a family." I cherished every moment of being part of this academic family, and I extend my heartfelt thanks.

I extend my deepest gratitude to the **Dutch Research Council (NWO)** for their invaluable financial and resource-based support, which has been instrumental in allowing me to dedicate my full attention to this significant scholarly pursuit.

The journey I embarked upon would not have been possible without the support of **Delft University of Technology**, a vibrant and nurturing environment that not only provided me with valuable knowledge but also offered the necessary facilities for conducting comprehensive research.

Collaboration has been a cornerstone of this work, and I am indebted to several remarkable individuals who have significantly enriched and elevated this project at every stage. My sincere appreciation goes to **Prof. dr. Wing Kam Liu** at Northwestern University, **Dr. Yongjie Jessica Zhang** at Carnegie Mellon University, **Dr. Mirwais Wardak** at Stanford University, **Dr. Aishwarya Pawar** at Iowa State University, **Prof. Daan van Eijk** at Delft University of Technology, **Mr. Christiaan Versteegh** at BATA Company, **Mrs. Lili-Marjan Boelens** at PHILIPS Company, **Dr. Lyè Goto** at Delft University of Technology, **Mrs. Laura Ahsmann** at TNO Company, and special thanks to **Dr. Mahsa Tajdari** at Northwestern University for her creative contributions in extending the applicability of this work. Moreover, I express gratitude to **Mrs. Erin Boogaard** and **Mrs. Alitta Song** for their assistance in composing the summary (Samenvatting) of this thesis in the Dutch language.

My sincerest gratitude is extended to the examiners of this thesis, who have invested their precious time and expertise in reviewing my work. Their thoughtful analysis and precise feedback have undeniably shaped my work for the better.

Completing this thesis would have been an insurmountable task without the unwavering support of my dear sister, **Dr. Farnaz Tajdari**, and brother-in-law, **Dr. Farshad Farzannasab**, who encouraged and motivated me throughout my Ph.D. journey. I extend heartfelt thanks to my mother, **Dr. Batool Eshraghi**, and my father, **Dr. Majid Tajdari**, for instilling in me the determination to pursue my goals and find joy in the challenges. This thesis is dedicated to my soon-to-arrive niece, affectionately referred to as "**Baby**" (yet to be named), whose imminent arrival fills me with eager anticipation.

Additionally, I express my gratitude to my wonderful friends and colleagues who have played a significant role in supporting me on this academic expedition: **Mr. Hengyang Li**, **Dr. Ayesha Maqsood**, **Mr. Emmett Cleary**, **Mr. Sourav Saha**, **Mr. Pouyan Shirzadian**, **Prof. John Francis Sarwark**, **Mr. Chanwook Park**, **Dr. Yusheng Yang**, **Dr. Xinhe Yao**, **Mr. Jun Xu**, **Dr. Tianyun Yuan**, **Dr. Willemijn S Elkhuizen**, **Mr. Felix Kwa**, **Ir. A(Sander) LM Minnoye**, **Dr. E(Zjenja) L Doubrovski**, **Dr. Jun Wu**, **Mr. Christiaan Eijck**, **Dr. Amir Ebrahimi Fard**, **Dr. Majid Mohammadi**, **Dr. Javanshir Fouladvand**, **Mr. Timon Staal**, **Mrs. Svetlana Mironcika**, **Mr. Robbert-Jan Torn**, **Mr. Elvis Borges**, **Mrs. Qing Sheng**, and **Mr. Naeim Ebrahimi Toulkani**. Your collective support has been instrumental in shaping the course of this academic endeavor.

Farzam Tajdari
Delft, November 2023

LIST OF PUBLICATIONS

JOURNAL PAPERS

1. [*]**F. Tajdari**, T. Huysmans, Y. Yang, and Y. Song *Feature Preserving Non-Rigid Iterative Weighted Closest Point and Semi-Curvature Registration*, IEEE Transactions on Image Processing **31**, 1841-1856 (2022).

2. [*]**F. Tajdari**, T. Huysmans, and Y. Song *Non-rigid registration via intelligent adaptive feedback control*, IEEE Transactions on Visualization and Computer Graphics, 1-17 (2023).

3. [*]**F. Tajdari**, T. Huysmans, X. Yao, J. Xu, and Y. Song *4D Feet: Registering Walking Foot Shapes Using Attention Enhanced Dynamic-Synchronized Graph Convolutional LSTM Network*, Submitted in IEEE Open Journal of the Computer Society, and is available in arXiv preprint arXiv:2307.12377.

4. [*]M. Tajdari, **F. Tajdari**, P. Shirzadian, A. Pawar, M. Wardak, S. Saha, C. Park, T. Huysmans, Y. Song, J. Zhang, J. Sarwark, and W. Liu *Next-generation prognosis framework for pediatric spinal deformities using bio-informed deep learning networks*, Engineering with Computers **38**, 4061-4084 (2023).

5. **F. Tajdari**, and N. Ebrahimi *Implementation and intelligent gain tuning feedback–based optimal torque control of a rotary parallel robot*, Journal of Vibration and Control **28**, 2678-2695 (2022).

6. **F. Tajdari** *Adaptive time-delay estimation and control of optimized Stewart robot*, Journal of Vibration and Control **0**, 0-0 (2022).

7. M. Tajdari, A. Pawar, H. Li, **F. Tajdari**, A. Maqsood, E. Cleary, S. Saha, J. Zhang, J. Sarwark, and W. Liu *Image-based modelling for adolescent idiopathic scoliosis: mechanistic machine learning analysis and prediction*, Computer Methods in Applied Mechanics and Engineering **374**, 1-30 (2020).

8. Y. Yang, T. Yuan, T. Huysmans, WS. Elkhuizen, **F. Tajdari**, Y. Song *Posture-invariant three dimensional human hand statistical shape model*, Journal of Computing and Information Science in Engineering **21**, 1-12 (2021).

9. **F. Tajdari**, C. Roncoli and M. Papageorgiou *Feedback-based ramp metering and lane-changing control with connected and automated vehicles*, IEEE Transactions on Intelligent Transportation Systems **23**, 939-951 (2020).

---

[*] The dataset and the code are available in: `10.4121/3a5eb5a8-bbae-4dd9-9a8d-d621bc1e36d2`

10. **F. Tajdari**, A. Golgouneh, A. Ghaffari, A. Khodayari, A. Kamali and N. Hosseinkhani *Simultaneous intelligent anticipation and control of follower vehicle observing exiting lane changer*, IEEE Transactions on Vehicular Technology **70**, 8567-8577 (2021).

11. **F. Tajdari** and C. Roncoli *Online set-point estimation for feedback-based traffic control applications*, IEEE Transactions on Intelligent Transportation Systems , 1-13 (2023).

CONFERENCE PAPERS

12. ***F. Tajdari**, C. Eijck, F. Kwa, C. Versteegh, T. Huysmans, and Y. Song *Optimal position of cameras design in a 4D foot scanner*, International design engineering technical conferences and computers and information in engineering conference - American Society of Mechanical Engineers **86212**, V002T02A044 (2022).

13. ***F. Tajdari**, F. Kwa, C. Versteegh, T. Huysmans, and Y. Song *Dynamic 3d mesh reconstruction based on nonrigid iterative closest-farthest points registration*, International design engineering technical conferences and computers and information in engineering conference - American Society of Mechanical Engineers **86212**, V002T02A051 (2022).

14. ASL. Minnoye ,**F. Tajdari**, EZL. Doubrovski, J. Wu, F. Kwa, WS. Elkhuizen, T. Huysmans, and Y. Song *Personalized product design through digital fabrication*, International design engineering technical conferences and computers and information in engineering conference - American Society of Mechanical Engineers **86212**, V002T02A054 (2022).

15. **F. Tajdari**, M. Tajdari, and A. Rezaei *Discrete time delay feedback control of stewart platform with intelligent optimizer weight tuner*, IEEE International Conference on Robotics and Automation (ICRA) , 12701-12707 (2021).

16. **F. Tajdari**, NE. Toulkani, and N. Zhilakzadeh *Intelligent optimal feed-back torque control of a 6dof surgical rotary robot*, IEEE International Conference on Power Electronics, Drive Systems, and Technologies Conference (PEDSTC) , 1-6 (2020).

17. **F. Tajdari**, NE. Toulkani, and N. Zhilakzadeh *Semi-real evaluation, and adaptive control of a 6dof surgical robot*, IEEE International Conference on Power Electronics, Drive Systems, and Technologies Conference (PEDSTC) , 1-6 (2020).

18. M. Tajdari, **F. Tajdari**, A. Pawar, J. Zhang and WK. Liu *2D to 3D volumetric reconstruction of human spine for diagnosis and prognosis of spinal deformities*, Conference: 16th US national congress on computational mechanics , 1-2 (2021).

19. **F. Tajdari**, H. Ramezanian, S. Paydarfar, A. Lashgari and S. Maghrebi *Flow metering and lane-changing optimal control with ramp-metering saturation*, IEEE International Symposium on Real-Time and Embedded Systems and Technologies (RTEST) , 1-6 (2022).

20. **F. Tajdari**, NE. Toulkaniand and M. Nourimand *Intelligent architecture for car-following behaviour observing lane-changer: Modeling and control*, IEEE International Conference on Computer and Knowledge Engineering (ICCKE) , 579-584 (2020).
21. **F. Tajdari**, A. Ghaffari, A. Khodayari, A. Kamali, N. Zhilakzadeh and NE. Toulkaniand *Fuzzy control of anticipation and evaluation behaviour in real traffic flow*, IEEE International Conference on Robotics and Mechatronics (ICRoM) , 248-253 (2019).
22. **F. Tajdari** and C. Roncoli *Adaptive traffic control at motorway bottlenecks with time-varying fundamental diagram*, IFAC-PapersOnLine **54**, 271-277 (2021).
23. **F. Tajdari**, C. Roncoli, N. Bekiaris-Liberis and M. Papageorgiou *Integrated ramp metering and lane-changing feedback control at motorway bottlenecks*, IEEE European Control Conference (ECC) **54**, 3179-3184 (2019).