

Electronic Thesis and Dissertation Repository

---

2-23-2023 2:00 PM

# Computer Vision-Based Hand Tracking and 3D Reconstruction as a Human-Computer Input Modality with Clinical Application

Tania Banerjee, *The University of Western Ontario*

Supervisor: Roy Eagleson, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Engineering Science degree in Electrical and Computer Engineering

© Tania Banerjee 2023

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Electrical and Computer Engineering Commons](#), and the [Telemedicine Commons](#)

---

## Recommended Citation

Banerjee, Tania, "Computer Vision-Based Hand Tracking and 3D Reconstruction as a Human-Computer Input Modality with Clinical Application" (2023). *Electronic Thesis and Dissertation Repository*. 9173. <https://ir.lib.uwo.ca/etd/9173>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

## Abstract

The recent pandemic has impeded patients with hand injuries from connecting in person with their therapists. To address this challenge and improve hand telerehabilitation, we propose two computer vision-based technologies, photogrammetry and augmented reality as alternative and affordable solutions for visualization and remote monitoring of hand trauma without costly equipment. In this thesis, we extend the application of 3D rendering and virtual reality-based user interface to hand therapy. We compare the performance of four popular photogrammetry software in reconstructing a 3D model of a synthetic human hand from videos captured through a smartphone. The visual quality, reconstruction time and geometric accuracy of output model meshes are compared. Reality Capture produces the best result, with output mesh having the least error of 1mm and a total reconstruction time of 15 minutes. We developed an augmented reality app using MediaPipe algorithms that extract hand key points, finger joint coordinates and angles in real-time from hand images or live stream media. We conducted a study to investigate its input variability and validity as a reliable tool for remote assessment of finger range of motion. The intraclass correlation coefficient between DIGITS and in-person measurement obtained is 0.767-0.81 for finger extension and 0.958–0.857 for finger flexion. Finally, we develop and surveyed the usability of a mobile application that collects patient data medical history, self-reported pain levels and hand 3D models and transfer them to therapists. These technologies can improve hand telerehabilitation, aid clinicians in monitoring hand conditions remotely and make decisions on appropriate therapy, medication, and hand orthoses.

**Keywords:** Computer vision, Photogrammetry, 3D model, Hand telerehabilitation, machine learning, MediaPipe, Range of Motion.

## Summary for Lay Audience

Patients with postoperative hand surgeries and upper extremities issues need regular rehabilitation under a specialist's supervision. COVID-19 has impacted the healthcare delivery system making it difficult for patients to visit a therapist in person leading to the emergence of tele-rehabilitation services. Telerehabilitation is a branch of telemedicine referring to the delivery of rehabilitation services over telecommunication networks. However, remote hand therapy faces challenges in reducing the reliance on in-person evaluation and assessment. The technology and treatment workflow developed in this project provide a way to address those challenges. For proper delivery of hand therapy, therapists require accurate information about the patient's hand condition. A complete 360° scanning of the hand is essential for therapists to understand the hand deformities and the severity of trauma. Most professional 3D Scanners and sensor-based rehabilitation devices are costly, non-portable and require technical knowledge to operate. Our project investigates whether photogrammetry, a non-invasive and cost-effective technique for creating 3D models from 2D images, can accurately reconstruct a 3D model of a human hand for telerehabilitation purposes. We determine if these models are of sufficient quality to support remote hand therapy and improve patient outcomes. To facilitate communication between patients and therapists, we designed an interactive app, HAND SCANS. This app enables patients to self-report pain and sensation level of their injured hand, provide health information, medical history and transfer 3D models of their injured hands digitally to a therapist. We also experiment with a novel machine learning pipeline, MediaPipe to develop a computer vision-based application called DIGITS for tracking hand landmarks in 3D space from 2D images or videos, find joint angles and evaluate finger range of motion. The idea, design and development of mobile/web applications to collect and transfer medical data, patient hand images, 3D hand models and key information during real-time hand movement tracking to the therapists form valuable aspects of hand telerehabilitation method. These cost-effective digital solutions will enable therapists to make clinical decisions on suitable therapy, medication, and orthoses for patients located remotely, improve efficiency and accessibility while bringing healthcare closer to home.

## Co-authorship Statement

1. Comparative Study of Photogrammetry Software in Reconstructing a 3D Model of a Human Hand for Hand Telerehabilitation- sole author

Tania Banerjee was involved in study design, performing experiment ,data collection, data analysis and manuscript writing .

The paper has been submitted to Asian Conference on Computer Vision and Pattern Recognition 2023 and is under review.

2. ‘DIGITS’ app - smartphone augmented reality for hand telerehabilitation- co-authored with Hongdao Dong, Edward Ho, Herbert Shin, Sandrine de Ribaupierrea, Roy Eagleson and Caitlin Symonette, Published online on 16 Nov 2021 by Taylor & Francis Group, DOI:10.1080/21681163.2021.1998927.

Article: Next-Generation Remote Hand Assessments: Cross-Platform DIGITS Web Application, Published on 25 February 2023, Journal of Hand Surgery Global Online DOI: 10.1016/j.jhsg.2023.01.016.

Tania Banerjee,Edward Ho,Herbert Shin was responsible for programming for app development,statistical data analysis and manuscript writing. Hongdao Dong, Dr.Caitlin Symonette helped in clinical study design, clinical data collection from St.Joseph Hospital and Victoria Hospital. Dr. Roy Eagleson and Dr.Sandrine de Ribaupierrea reviewed study design and manuscript review and provided valuable suggestions and recommendations for further improvement.



## Acknowledgments

I would like to sincerely thank my supervisor, Dr. Roy Eagleson for his guidance, meticulous suggestions and encouragement throughout my research period. I would like to express my gratitude to Dr. Louis Ferreira, Department of Mechanical & Materials Engineering at Western University for providing the resources and access to his mechanical engineering lab that made it very convenient for me to perform my experiments. I also want to give special thanks to Dr. Joy McDermid, Western's Bone and Joint Institute, and her group of medical trainees for their valuable feedback on the development of a relevant set of questionnaires for patients with hand injuries. I would like to extend my thanks to the Hand and Upper Limb Clinic (HULC), St. Joseph's Hospital, London Ontario for assisting us with clinical trials. I would also like to thank ORFIT Industries, Belgium for funding this research project. Finally, I would like to thank my parents and my brother for their enormous support throughout this journey and their constant encouragement.

# Table of Contents

<b>Abstract .....</b>	<b>ii</b>
<b>Summary for Lay Audience.....</b>	<b>iii</b>
<b>Co-authorship Statement .....</b>	<b>iv</b>
<b>Acknowledgments .....</b>	<b>v</b>
<b>List of Figures .....</b>	<b>x</b>
<b>List of Tables .....</b>	<b>xiv</b>
<b>List of Appendices.....</b>	<b>xv</b>
<b>Chapter 1 .....</b>	<b>1</b>
<b>1. Introduction.....</b>	<b>1</b>
1.1 Background .....	1
1.2. Problem .....	2
1.3. Literature review .....	3
1.3.1 Existing solutions for upper limb rehabilitation.....	3
1.3.2 3D Scanning techniques .....	5
1.4. Research Objective .....	6
<b>Chapter 2 .....</b>	<b>8</b>
<b>2.Reality Capture Technology .....</b>	<b>8</b>
2.1 Photogrammetry.....	8
2.2 Classification of Photogrammetry (Luhmann 2006) .....	9
2.3. Close-range Photogrammetry .....	10
2.4 Structure from motion algorithm .....	11
2.5. Review of modern photogrammetry software .....	14
2.5.1 Desktop based photogrammetry software: .....	14
2.5.2 Cloud based photogrammetry software.....	16

<b>Chapter 3 .....</b>	<b>19</b>
<b>3.Reconstruction of 3D Hand Models .....</b>	<b>19</b>
3.1 Research materials .....	19
3.1.1 Hardware .....	19
3.1.2 Hand Prosthetic 1 .....	20
3.1.3 Reference data from Artec 3D Scanner.....	21
3.1.4 Choice of open-source photogrammetry software .....	22
3.2 Methodology .....	22
3.2.1 Data Acquisition.....	22
3.2.2 Frames extraction from video.....	23
3.2.3 Camera Alignment .....	28
3.2.4 Photogrammetry workflow.....	29
3.3 Results .....	31
<b>Chapter 4 .....</b>	<b>35</b>
<b>4.Comparison of 3D models using structured light 3D Scanner outputs as a benchmark ...</b>	<b>35</b>
4.1 Materials.....	35
4.1.1 Hand Prosthetic 2 .....	35
4.1.2 Reference Data .....	36
4.1.3 3D models from smartphone images.....	36
4.1.4 Statistical software .....	37
4.2 Evaluation Method .....	38
4.2.1 Scaling .....	38
4.2.2 Alignment and Registration.....	39
4.2.3 Signed distances computation .....	40
4.3 Gaussian distribution.....	42

4.4 Analysis.....	43
4.5 Limitations of the study .....	44
4.5.1 Use of single device for data acquisition.....	44
4.5.2 Constraints for photogrammetry in an indoor environment.....	44
4.5.3 Requisite hardware issues .....	45
<b>Chapter 5 .....</b>	<b>46</b>
<b>5. App for patient data collection .....</b>	<b>46</b>
5.1 App workflow .....	46
5.2 Prototype Development.....	47
5.3 Features .....	48
5.3.1 Patient background information collection .....	49
5.3.2 Health information collection.....	50
5.3.3 In-app Pain and Sensitivity evaluation.....	52
5.3.4 Collecting hand images .....	55
5.4 Advance features.....	56
5.5 KIRI Engine 3D Scanner .....	57
5.6 Hand Scans Usability Survey.....	58
5.6.1 Data collection.....	59
5.6.2 Results .....	59
5.7 Practical implications of the app.....	62
5.8 Limitations of Hand scans App.....	63
<b>Chapter 6 .....</b>	<b>64</b>
<b>6. 3D Hand Tracking In Real-Time.....</b>	<b>64</b>
6.1 MediaPipe .....	64
6.1.1 Palm Detector.....	64

6.1.2 Hand Landmark model .....	65
6.2 MediaPipe Hands Coordinate system .....	66
6.3 Real-time angle measurement .....	71
6.4 DIGITS.....	72
6.4.1 Input variability of DIGITS.....	73
6.4.2 Inference .....	74
6.5 Range of motion (ROM) evaluation of finger joints using DIGITS.....	76
6.5.1 Methodology .....	77
6.5.2 Statistical analysis .....	78
6.5.3 Evaluated Results .....	79
6.6 Discussion .....	80
<b>7. Research Contribution .....</b>	<b>81</b>
<b>8. Future Research Scope.....</b>	<b>83</b>
8.1 Use of digital 3D hand models for virtual fabrication of hand orthosis .....	83
8.2 Research on technology to improve resolution of photogrammetry 3D models .....	83
8.3: DIGITS for tremor detection in hands .....	83
8.4 Combining DIGITS and Hand Scans app .....	83
<b>9. Conclusion .....</b>	<b>84</b>
<b>References.....</b>	<b>85</b>
<b>Appendix.....</b>	<b>94</b>

## List of Figures

Figure 1:Illustrating single and multiple point triangulation and the basis of photogrammetry measurements .....	8
Figure 2:Close-range photogrammetry using a tripod.....	10
Figure 3:An illustration of close-range photogrammetry concept.....	11
Figure 4:Conventional SfM Workflow.....	12
Figure 5:Desktop based software.....	14
Figure 6:Cloud based software .....	14
Figure 7:Node based workflow of Meshroom.....	15
Figure 8:Hand prosthetic 1 with extended fingers.....	20
Figure 9:Artec Space Spider 3D Scanner .....	21
Figure 10:Method to capture images using ASS .....	22
Figure 11:Views captured in a sequential manner a) Palmar b) lateral c) cranial d) dorsal e) lateral f) cranial views of the upper extremity captured from the smartphone video.....	23
Figure 12:Open VLC, select tools then select preferences.....	24
Figure 13>Select Filters settings from Advance Settings in VLC.....	24
Figure 14>Select Scene filter settings to convert video to images with VLC.....	25
Figure 15:Meshroom UI loaded with images .....	26
Figure 16:Process to import videos in 3DF Zephyr.....	27
Figure 17:Process to import videos in Reality Capture .....	27
Figure 18:Process to import videos in Agisoft Metashape.....	28
Figure 19:Aligned camera frames in Meshroom UI.....	28
Figure 20:Aligned camera frames in Agisoft Metashape UI.....	29
Figure 21:Sparse point clouds by 3DF Zephyr.....	29
Figure 22:Dense point clouds by 3DF Zephyr .....	30
Figure 23:Mesh generated by 3DF Zephyr.....	30
Figure 24:Textured Mesh in 3DF Zephyr after cleaning the surrounding.....	31
Figure 25:Comparing results of first three steps of photogrammetry.....	31
Figure 26:Comparing results of last steps of photogrammetry .....	32

Figure 27:Bar showing total processing time taken by four photogrammetry software .....	33
Figure 28:3D model of the hand prosthetic .....	33
Figure 29:Hand prosthetic with flexed fingers .....	35
Figure 30.1:Palmar view of ASS mesh.....	36
Figure 30.2:Lateral view of ASS mesh.....	36
Figure 30.3:Dorsal view of ASS mesh .....	36
Figure 31.1:Palmar view of 3DF Zephyr mesh .....	37
Figure 31.2:Lateral view of 3DF Zephyr mesh .....	37
Figure 31.3:Dorsal view of 3DF Zephyr mesh.....	37
Figure 32:CloudCompare workspace with two imported meshes of different scales .....	38
Figure 33:Matching scales and bounding box centers of the two imported meshes .....	39
Figure 34:Registering two meshes.....	40
Figure 35.1:C2M distance computation between reference mesh and 3DF Zephyr mesh a) dorsal view b) Palmar view .....	41
Figure 35.2:C2M distance computation between reference mesh and Agisoft Metashape mesh a) dorsal view b) Palmar view .....	41
Figure 35.3: C2M distance computation between reference mesh and Reality Capture mesh a) dorsal view b) Palmar view .....	41
Figure 36.1:Histograms of the Gaussian distribution characterizing the distances between the ASS and 3DF Zephyr mesh. ....	42
Figure 36.2:Histograms of the Gaussian distribution characterizing the distances between the ASS and Agisoft Metashape mesh. ....	42
Figure 36.3:Histograms of the Gaussian distribution characterizing the distances between the ASS and Reality Capture mesh.....	42
Figure 37:The images on the left are shadowed images without proper lighting, while the image on the right is an image with homogenous lighting and is considered the correct image for photogrammetry.....	45
Figure 38:HAND SCANS app flowchart .....	47
Figure 39:HAND SCANS homepage a)Webapp interface b)Mobile interfacep interface.....	49
Figure 40:Random value generator widget in Jotform console.....	49
Figure 41:Questions in User Demographic Questionnaire form .....	50

Figure 42:Questions in Upper Extremity Problem form .....	51
Figure 43:Health status form listing comorbidities .....	51
Figure 44:Red dots showing injured areas and pain level of patient.....	52
Figure 45:Pain level (0-10).....	53
Figure 46:Measuring the pain score of all 5 items .....	54
Figure 47:Measuring function items a)specific activities b)usual activities .....	54
Figure 48:Measuring the total PRWE score of all 10 items .....	55
Figure 49:Images and file upload for range of motion estimations.....	55
Figure 50>User interface of Contact form for a) Webapp Application b)Android application .....	56
Figure 51:Progress bar.....	56
Figure 52:Save to continue later .....	57
Figure 53:Green arrows showing the completion and submission of form.....	57
Figure 54:3D model of a) dorsal side of hand prosthetic b)palmer side of hand prosthetic using KIRI Engine.....	58
Figure 55:3D model of the dorsal side of a real hand using KIRI Engine .....	58
Figure 56:Rating distribution of four objective attributes of uMARS namely information quality, aesthetics, functionality, and engagement .....	61
Figure 57:Pi-chart Chart showing rating distribution of all attributes of uMARS .....	62
Figure 58:Visualizing data traffic using Jotform Form Analytics.....	63
Figure 59:Detection of real hands(top) and synthetic hands (bottom) with Mediapipe .....	64
Figure 60:MediaPipe Hands Palm detection model .....	65
Figure 61:Hands detected and pose estimation .....	66
Figure 62:Hand Landmark model showing labels for each Landmark ID .....	66
Figure 63:Image of a right hand used for the experiment .....	67
Figure 64:Image of a right hand used for the experiment .....	68
Figure 65:Output image showing detected landmarks and hand connections.....	68
Figure 66:Showing coordinates for all landmarks (0-20) with timestamps .....	69
Figure 67:Code snippet and map plot of 21 landmarks in default(local) coordinate system .....	70
Figure 68:Diagram of default (local) coordinate system of MediaPipe showing the x , y and z axis.....	70
Figure 69:Diagram of 3D world coordinate system showing x,y and z axis.....	71



Figure 70:Code snippet and map plot of 21 landmarks in world coordinate system .....	71
Figure 71:Showing hand classification and angles between joints of the index finger.....	72
Figure 72:DIGITS user interface .....	72
Figure 73:Stacked line graph showing variation of the wrist joint (landmark 0) along x(top) , y (middle) and z (bottom) coordinates.....	73
Figure 74:Box-Whiskers plot to show variation range of wrist joint (Landmark 0)along a) x , b) y and c) z coordinates .....	74
Figure 75.1:x coordinates mapping for 21 landmarks .....	75
Figure 75.2:y coordinates mapping for 21 landmarks.....	76
Figure 75.3:z coordinates mapping for 21 landmarks.....	76

## List of Tables

Table 1:List of photogrammetry open-source software.....	17
Table 2:Hardware and their description used in the experiment .....	19
Table 3:Mobile Camera specifications .....	20
Table 4:Specifications of Artec Space Spider .....	21
Table 5:Processing time required by four photogrammetry software .....	32
Table 6:Parameters for comparison .....	43
Table 7:uMARS ratings for the HAND SCANS.....	60
Table 8:Showing the mean, standard deviation , minimum and maximum value for 0 landmark. .....	74
Table 9:The average finger joint angles obtained from in-person goniometry measurement of both hands(unit in degrees).....	78
Table 10:The average finger joint angles obtained from DIGITS (unit in degrees) .....	78
Table 11:Intraclass correlation coefficients (ICC) for the reliability of MCP, PIP, and DIP joint angle measurement for both right and left-hand.....	79

## List of Appendices

Appendix 1:Cloud Compare source code in GitHub: <a href="https://github.com/CloudCompare/CloudCompare.git">https://github.com/CloudCompare/CloudCompare.git</a> .....	94
Appendix 2:Statistics for all 21 landmarks (Mean, SD, Min, Median, Max).....	94
Appendix 3:Modified uMARS Sheet used for Hand Scans mobile app usability evaluation .....	95

# Chapter 1

## 1. Introduction

### 1.1 Background

Rehabilitation is the cornerstone of the return of motor function following musculoskeletal trauma or surgery. Traditional rehabilitation treatment requires therapists to provide one-to-one rehabilitation training to patients in person. The COVID-19 pandemic has impacted all aspects of healthcare delivery. To protect healthcare workers and patients across the country from the risk of disease transmission, policies were altered to enable the widespread use of digital technologies instead of in-person clinical visits [1]. The recent pandemic has made it difficult for patients with hand injuries to connect in person with their therapists. The wait time for each clinical visit has increased, and delays in treatment can potentially result in further hand impairments such as prolonged joint stiffness, muscle atrophy, loss of function, and pain syndromes. For these reasons, telemedicine has become more relevant and vital. Hand therapists have implemented telemedicine approaches in their routine practice allowing clinicians to provide care while maintaining social distancing [2]. This has created a fast, convenient, and potentially more efficient flow of information to serve the target patients with upper extremity problems.

Upper extremity rehabilitation aims to stabilize joints, fingers, or wrists and restore joint range of motion, hand strength, and dexterity. Consistent assessment of hand injuries, hand therapy and proper orthosis is linked to improved functional outcomes in terms of recovery of strength and range of motion. Hand therapists usually evaluate the range of motion, sensation, and pain to know what type of medication, therapy, or orthosis is needed based on the results of their evaluation. All these processes are typically done based on a certified clinician's in-person assessment. In the digital transformation of healthcare, there are challenges in reducing this reliance on the in-person evaluation given the complex geometries of the hand, unavailability of high-resolution scanners, orthotics fabrication technicalities and time that may require multiple visits and cost, equipment materials and expertise [3,4]. In this project, we aim to design a workflow to address these challenges and bridge the gap between patients and hand therapists.

The advancement in Human-Computer Interaction (HCI) modalities like computer vision have resulted in the development of more intuitive and interactive systems using digital images. The human-computer interaction paradigm has made a shift from the graphical user interface (GUI) to the virtual reality-based user interface (VRUI). According to Ellis [5], virtual environments (VEs) can “convince users that they are immersed in a synthetic space” and “let users navigate and interact with a computer-generated 3-D environment in real-time”. Digital transformation technology like digital twins [6] set an example of how real-world assets and data can be used to render a 3D virtual representation of physical objects for better visualization. 3D Digital twins have been used for simulation and monitoring real world counterparts in the field of automotives, manufacturing, production and smart city management. We are extending the application of 3D rendering and VRUI to hand therapy and telerehabilitation. Visualization of a physical hand in a virtual environment can provide useful information to clinicians and augment their medical judgments. Photogrammetry is a cost-effective technique to create a 3D model of a physical object. Contemporary smartphone devices are equipped with cameras with high resolution and sensors that are capable of capturing high-resolution images necessary for 3D modeling. Development of real-time hand tracking methods using Machine Learning pipelines can be a possible solution to measure the range of motion of fingers. In this project, we delve into these technologies to reconstruct a 3D model of a human forearm and track hand movements in real-time using machine learning approaches. We also believe that developing an interactive app to collect patients’ medical history and health information about their hand injuries, previous surgeries and medications, and current condition of pain and sensation will relay helpful formation to therapists. Graphical representation of hand trajectories from real-time tracking and a full 360-degree digital view of a 3D hand model will help therapists make proper decisions on suitable therapy, medication, and orthoses.

## 1.2. Problem

Tele-rehab in hand therapy comes with many challenges. It is a mammoth task to reduce the reliance on in-person evaluation given the complex geometries of the hand and the large number of different chronic hand conditions. Hand orthotic fabrication and fitting also requires knowledge about current hand deformity, range of motion, pain and sensation of the injured area and patient’s medical history.

Obtaining a 3D scan of human body parts is possible with the evolution of commercial 3D Scanners. These accurate and high-resolution 3D scanners are helpful for designing suitable customized applications like implants, prosthetics, anatomical models [7] and 3D-printed orthotics. However, most of the 3D scanners are too expensive for most clinical applications [8] as well as for individual purchases. Finally, this sophisticated equipment requires technical knowledge to be able to use and produce 3D scans. Hence, there is a need to develop alternative solutions which should be economical, feasible and convenient for patients as well as useful for hand therapists to obtain patient information. With the goal of improving telerehabilitation for patients with hand injuries, in this project we have collaborated with certified hand therapists, software and mechanical engineers, to investigate possible solutions for remote hand assessment and diagnosis.

### 1.3. Literature review

#### 1.3.1 Existing solutions for upper limb rehabilitation

Upper limb disability is a common health problem in the general population [9]. A study [10] showed eight important categories of complications that lead to upper limb disabilities. Most hand rehabilitation is required for patients who have suffered from stroke (65%). The other categories include Acquired Brain Injury (ABI), Parkinson's Disease, Proximal Humeral Fractures, Spinal Cord Injury, Rheumatoid Arthritis, Cerebral palsy, and Musculoskeletal Disorders (MSD). For most of these conditions, many patients still need rehabilitation services to fully recover from post-operative trauma even after being discharged from the hospital. Telerehabilitation technologies reduce hand rehabilitation costs and in-person clinic visits while increasing the quality of life of patients. Even before the pandemic, there were continuous efforts to investigate the role of telerehabilitation in improving the health condition of patients with upper limb disabilities. We identified various novel research studies and evaluation methods for improving hand telerehabilitation.

Virtual reality (VR) based haptic gloves were designed to rehabilitate hand functions in stroke patients: Rutgers Master-II and Hand Master haptic device [11], CyberGrasp glove (Immersion Inc.) [12], LRP Glove [13] or integrated versions of them [14]. A general description of their architecture can be described as the mechanical units communicating with a host PC in which the results of the exercises are stored to be accessed and processed remotely. YouGrabber, a computer-

enhanced upper limb training system with sensors, was used at home as a motivating training tool to complement daily care and was tested to augment training frequency and intensity in children and adolescents with neuromotor disorders [15]. However, the frequency of technical problems was very high in its trial for feasibility, adoption and usability. MusicGlove [16], a music-based hand rehabilitation device was developed to help people regain hand functions at home. It required the users to practice gripping-like movements and thumb-finger opposition to play highly engaging, music-based video games. The study [17] showed that after six 45-minute sessions, participants improved their ability to grip small objects using the MusicGlove paradigm compared to conventional hand exercises. However, all these devices can be very expensive due to their intrinsic complexity. Moreover, these devices have several force feedback terminals per finger with the forces grounded in the palm or on the back of the hand which makes them heavy, cumbersome, and greatly limits hand movements. Five studies used nonspecific video game systems, the Nintendo Wii™ [18,19], the Microsoft Xbox Kinect™ [20], and the Sony PlayStation EyeToy™ [21] that used the movement of the player to control gameplay. The exercises could be presented as games and such games could include movements from daily activities. Exercise games in upper limb telerehabilitation seem to be effective in promoting post-stroke activity. SCRIPT passive wrist and hand orthosis [22] was developed in 2013 for hand rehabilitation at home, interfacing with gaming, support, and therapeutic software modules. However, robust conclusions could not be drawn for the above method due to the lack of standardized outcome activity measurements, highlighting the need for a gold standard in future research in this area. Video-game-based treatment is still not considered to be more efficient than conventional therapy [23].

Park et al. in 2008 [24] proposed a robotic system that can be operated remotely for evaluating Impaired Elbows in Neurological Disorders. Published in 2016, this paper [25] also presents an advanced robot hand for telerehabilitation and a remote monitoring system. The machinery comprises a hand rehabilitation support system, a robot hand, and a remote monitoring system. Most of the hand and finger motions like adduction/abduction, anteflexion/retroflexion, hand pronation/supination, and palmar flexion/dorsiflexion driven by the support system were found to follow the motion of the robot hand very well. However, the complex technicalities associated with robotic arms are the reason for their low adoption. To eliminate complications associated with mechanical haptic gloves, robotic arms and their associated hardware (compressors, sensors,

controllers, transducers, etc.), an entirely software-based hand telerehabilitation system named “Virtual Glove” was developed [26]. It used only webcams for visual tracking movements and calculating forces. The space occupied by the global assembly was significantly reduced, along with the weight and costs. Moreover, the system could be assembled using WI-FI technology, eliminating any cables between the video cameras and the dedicated PC or between the dedicated PC and the web server. The patient’s hand remained free from wires and cumbersome heavy devices. A similar innovative approach was shown in 2015 [27], where their proposed application, “TeleReh” was based on image processing and web technologies and used to supervise hand exercising remotely. However, these applications are still in their initial stages of testing, and further evidence is required for their effectiveness and robustness.

For most of the methods mentioned above, the main disadvantage is the need for additional specialized devices. It can be observed that there is still a lack of minor applications that provide therapeutic values and at the same time more affordable to people, eliminating the need to purchase additional specialized devices.

### 1.3.2 3D Scanning techniques

Some popular methods in current medicine to obtain detailed and problem-oriented information about human body parts are 3D imaging techniques, such as Computerized Tomography (CT), Cone Beam Computerized Tomography (CBCT), Micro Computerized Tomography (MCT) and Magnetic Resonance Imaging technique (MRI) [28-31]. However, in addition to being expensive, the equipment is heavy, non-portable and delivers high doses of radiation. This makes them the least convenient option for remote hand scanning. 3D laser scanning and Structured-light 3D scanning [32,33] are two modern non-invasive techniques that generate digital 3D models. The models are accurate and able to provide measurements with respect to shape and size [34]. However, most of these professional 3D Scanners are costly for individual purchases and still require technical knowledge to operate and maintain. Technology has changed a lot during the last decade or so more and more people rely on mobile phones for their photography and their cameras are better than ever before. In a study (Kersten T. P. et al., 2016a), various handheld imaging and scanning devices were compared with mobile phone photogrammetry methods to produce a 3D model. Artec Spider, smartSCAN, Structure Sensor, Kinect v1 and Kinect v2 and Google’s Project Tango were some of the few tested systems that were tested against the Smartphone Galaxy Note



1. Several factors like visual quality, the completeness of scanned test objects, and accuracy were also accessed. The test showed the model generated from the images of the mobile phone was incomplete with large holes. Dhonju, et al. (2017) evaluated the feasibility of smartphone photogrammetric modeling for cultural heritage documentation in Nepal. The study assessed the possibilities of using a consumer-grade mobile phone, i.e. Huawei Nexus 6P. The study highlights the pros and cons of such low-cost image-based documentation and concludes the method is not yet ready for documentation purposes but still of great value for visualizing 3D models. However, the performance of the newer generations of mobile phones remain an open question.

Samsung Galaxy S22 Ultra and Samsung Galaxy S20 Ultra are among the new generation smartphones with high-definition cameras. Google Pixel 7 Pro claims to produce more clear inclusive photos using Real Tone technology to enhance skin tones. iPhone 14 Pro/ iPhone 14 Pro Max employing Apple's smart AI optimization is available in the market today, with the company's Deep Fusion technique to tweak your pictures to make them as wonderful as possible. In a recent study (Kanun & Yakar, 2021), CAD data of a boat's hull was generated using the photogrammetric method and the images were captured using Samsung Galaxy S10. The results indicated that mobile phone images could be utilized for quick and precise 3D documentation. The accuracy assessment of the model generated was reported with an error of less than 0.5 cm. And the study calls phone-based close-range photogrammetric survey an effective as well as accurate method. Thus, the future of mobile phone-based photogrammetry seems viable especially given the lower cost, stronger compact lens system, portability, continuous availability, and ease of handling. The studies motivate us to utilize the potential of new-generation smartphone cameras for photogrammetry or 3D modeling.

#### 1.4. Research Objective

Inspired by the latest research with mobile cameras we attempt to experiment with photogrammetry technology to build 3D models of hands from 2-D images taken at a close range with a smartphone. This thesis aims to reconstruct 3D model of human hand and track hand key points in real-time on smart devices with the main focus on utilizing RGB images and video streams captured in a home setting. We look at existing pipelines developed using Machine Learning that can be trained to capture hand movements, finger joint coordinates and angles in real

time to assess the range of motion. We also design a workflow to bridge the gap between the patient and the therapist by building an app to collect both patient data and 3D hand models and securely export the data to the therapists for assessment and diagnosis. Below is a summary of the specific objectives:

- Explore photogrammetry to develop 3D hand models using smartphone cameras in place of commercial scanners in indoor environments.
- Evaluating photogrammetry software in terms of their processing time and accuracy. Comparing the quality of 3D point clouds and meshes obtained from different photogrammetry software.
- Developing an augmented reality app (DIGITS) for Hand tracking using machine learning pipeline (MediaPipe), evaluate its input variability and establish its validity as a tool for finger range of motion measurement.
- Develop an interactive Android/Web app for patients to collect data about patients' medical history, hand images, 3D hand model files and export them to therapists.

## Chapter 2

### 2.Reality Capture Technology

#### 2.1 Photogrammetry

Photogrammetry is a method of 3D Scanning that uses triangulation or depth perception to create an accurate model of an object from photographs. Photogrammetry scanning can be performed at close-range, via satellite or air (aerial photogrammetry). This technology uses overlapping images considering the same ground points visible in multiple photos and from different vantage points to generate a 3D map through depth perception or triangulation. Figure 1 illustrates single and multiple point triangulation and method of depth perception. The technology works similarly to how our human brain processes information from two eyes to provide depth perception. The high-resolution 3D model that is produced contains not only dimensional information like height and elevation but also texture, shape and color.

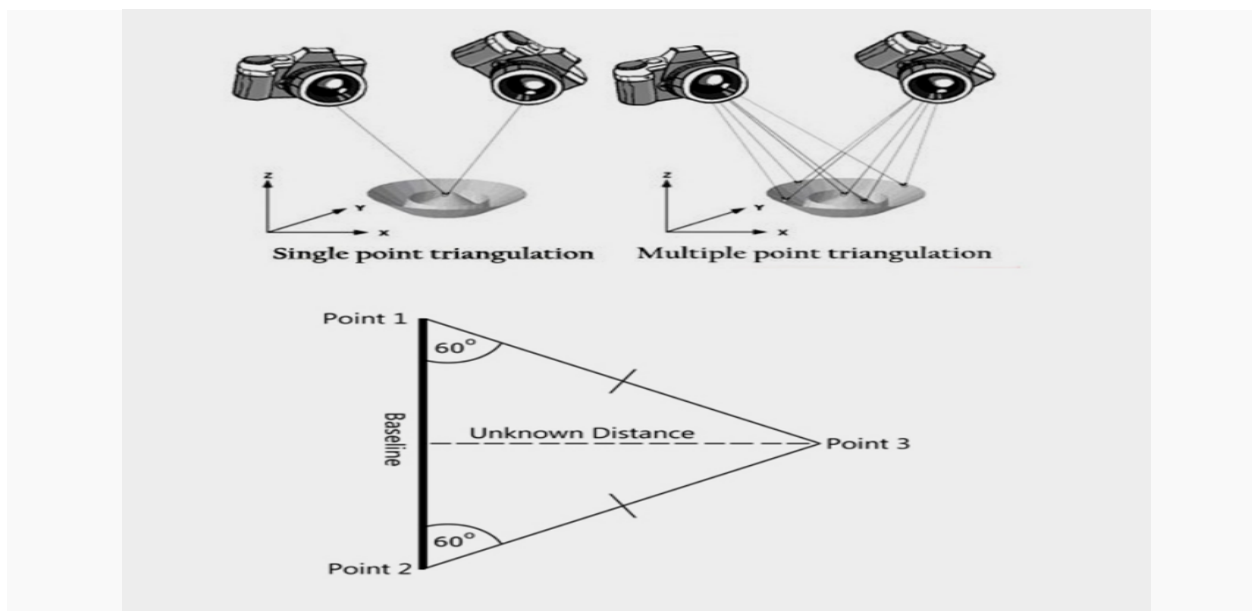


Figure 1: Illustrating single and multiple point triangulation and the basis of photogrammetry measurements

[Picture courtesy: <https://graphicdesignjunction.com/2019/09/how-does-triangulation-in-photogrammetry-work/>]

Photogrammetric projects generally involve two main steps. First is the acquiring of suitable photographs followed by preprocessing of the imagery and second is imagery processing using powerful photogrammetric software (Mikhail et al., 2001). Most of the

modern photogrammetric software implements a structure-from-motion (SfM) algorithm to produce 3D models. SfM involves extracting information from a collection of photographs, including each camera position and other three-dimensional data, such as the scene geometry (Green et al., 2014). The software calculates each camera position by locating matching points across multiple photographs and automatically determining the camera's interior orientation and lens distortion parameters (Green et al., 2014; Granshaw, 2016). Once the camera positions have been calculated, the locations of numerous points can be plotted to create a dense reconstruction of the objects that have been photographed (Green et al., 2014).

This technology can be integrated into design software for correction and editing. Photogrammetry workflows can be so intuitive and user-friendly that any industry professional or novice can navigate quickly and easily without specialized training. The end product is delivered in various formats, which are compatible with 3D printing tools. Data from photogrammetry can be integrated directly with CAD (Computer-Aided Design) and BIM (Building Information Model) design tools such as AutoCAD, Revit, Navisworks, Civil3D, and SolidWorks. CAD software can import the triangular mesh or textured mesh; it also corrects, modifies and optimizes the model by manipulating the parameters. Then the solid geometry of the object can be printed using 3D printers.

## 2.2 Classification of Photogrammetry (Luhmann 2006)

### A. Classification 1

Categorized by camera position and object distance:

- i) Satellite photogrammetry: processing of satellite images where distance between object and camera is 200km.
- ii) Aerial photogrammetry: processing of aerial photographs, where distance between object and camera is 300-200 m.
- iii) Terrestrial photogrammetry: measurements from a fixed terrestrial location .
- iv) Close range photogrammetry: imaging distance where distance between object and camera is less than 300m.
- v) Macro photogrammetry: image scale  $> 1$  (microscope imaging)

## B. Classification 2.

By number of images:

i) Single Image photogrammetry e.g., orthophotos ( $n$ -images,  $n=1$ )

ii) photogrammetry ( $n=2$ )

iii) Multi image photogrammetry ( $n>2$ )

where  $n$  is the number of images.

We will only concentrate on close-range photogrammetry and multi-image photogrammetry by using images obtained using a smartphone camera. We will search for suitable photogrammetry software that is powerful enough to capture thin and complicated objects such as human hands, wrist areas and narrow slender fingers.

### 2.3. Close-range Photogrammetry

Close-range photogrammetry (CRP) is a type of terrestrial photogrammetry used to collect data about physical objects such as shape and position from two or more images taken from different stations using a digital camera mounted on the ground/tripod while facing the object. The main difference from aerial is that in close-range photogrammetry the object-to-camera distance is less than 1000 ft or 300m. The coordinates and other data of the camera are collected at the time at which the photo is taken. Figure 2 shows the method of capturing images of objects from all angles in close-range photogrammetry while figure 3 shows how those images are stitched together by photogrammetry algorithm to form a 3D model of the object.



Figure 2: Close-range photogrammetry using a tripod

[Picture courtesy: [Ahmed Jebur](#), MS in surveying engineering, Thesis for: Master of Technology in Surveying Engineering Technology, Advisor: Asst. Prof. Dr. Fanar M. Abed; Prof. Dr. Mamoun U. Mohammed, Project: [modeling](#), DOI: [10.13140/RG.2.2.11494.06722](#)]

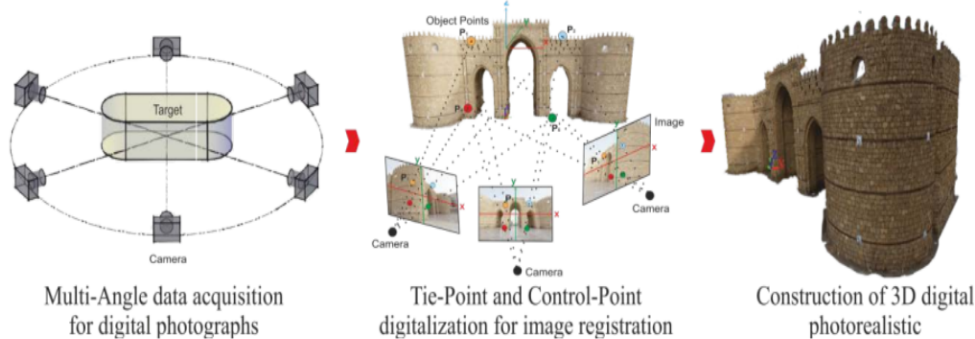


Figure 3: An illustration of close-range photogrammetry concept

[Picture courtesy: [Albourae, Abdullah Taha Ahmed \(2014\)](#): Accuracy assessment of terrestrial laser scanning and digital close range photogrammetry for 3D cultural heritage. Toronto Metropolitan University. Thesis. <https://doi.org/10.32920/ryerson.14644533.v1> ]

The same cadre of products such as digital terrain models, digital orthophotos, vector maps that can be created from traditional aerial photogrammetry can also be generated from Close-range photogrammetry. It is much more efficient to produce animations of small objects and digital-rotatable 3D objects that can be integrated into 3D raster streaming Web technology or uploaded to 3D modeling platform websites like Sketchfab to publish, share, buy and sell 3D VR and AR content.

#### 2.4 Structure from motion algorithm

The Structure from Motion (SfM) technique is the most applied technique in the photogrammetry process. Structure from motion (SfM) is a process of estimating 3D structure from several 2D images of the same structure [35,36]. The sequential algorithm pipeline involves two main phases. The first is known as the matching phase where Feature Extraction, Feature Matching and Geometric Verification are applied. Once relevant and verified correspondences have been found, the second reconstruction phase involving Image Registration, Triangulation and Bundle Adjustment, takes place. In this stage, camera poses, and point depths are iteratively. The implementation of such an algorithm result in forming sparse point clouds whereas dense point clouds can be generated by means of Multiview Stereo matching (MVS). Figure 4 depicts the

complete sequence of phases involved in the conventional SfM method. A description of each phase of SfM pipeline is explained below .

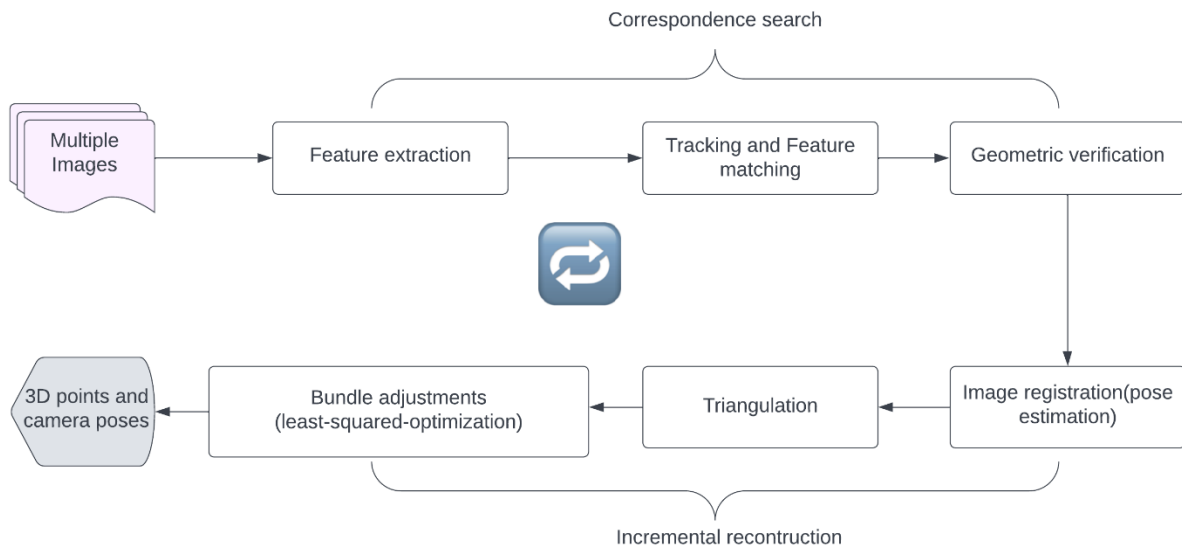


Figure 4: Conventional SfM Workflow

- Feature Extraction & Matching Feature

Extraction is the first operation to be performed in the standard SfM procedure. Each input image is processed to locate points of interest or key points. Typically, Scale Invariant Feature Transform (SIFT) [37] feature descriptor is used in this phase. SIFT descriptors have the advantage of being particularly robust against changes in scale, large variations of viewpoint, and challenging conditions like inconsistent illumination and complete/partial occlusions. For this reason, SIFT is widely popular among the available SfM algorithms. The extracted features are used to determine whether the images have common or partially overlapping parts. Matches are found by comparing key points in different images; if at least a pair of images contains features with the same descriptor, we can say that the images are overlapping. This phase results in a set of overlapping images and correspondence between features.

- Geometric Verification

Geometric verification is the process of geometric transformations of points so that the points are coherent with the geometry of the scene. For this we need the knowledge of fundamental matrix and homography matrix. Fundamental matrix verifies the feature matches map for each image pair [38] while homography matrix is a linear transform between two planes in projective space [39]. However, the feature matches often contain much noise due to image-forming process, feature extraction and representation. In order to remove the influence of noise, the RANdom SAMple Consensus (RANSAC) algorithm [40] is proposed to robustly estimate a map to fit the data and classify the data into inliers (namely, consensus set) and outliers. RANSAC estimates the desired features by eliminating the outliers in an iterative process (MathWorks, 2021). Hence, it is used as an optimization strategy in addition to geometric verification.

- Image Registration

In this phase, camera poses are estimated. A new image is added to the pipeline which is identified as a newly registered image. For this newly registered image the pose of the camera (position and rotation) must again be calculated; this can be achieved using the correspondence with the known 3D points of the reconstruction [41].

- Triangulation

In this step, 3D coordinates of a point is estimated by measuring its projections in two or more images captured from different viewpoints. Newly recorded images keep adding to the pipeline and contribute to the addition of new points and defines the 3D coordinates for those points that had not been reconstructed in the previous step, thus generating a more dense point cloud.

- Bundle Adjustment

Once we have computed all the camera poses and 3D points, we need to refine the poses and 3D points together, initialized by the previous reconstruction by minimizing reprojection errors. Bundle Adjustment (BA) is the optimization process applied in this final phase. It aims to refine the final point cloud by adjusting the error in both camera pose estimation and triangulation step, thus preventing the propagation of the inaccuracies mentioned before. The algorithm used for BA is Levenberg-Marquardt (LM) which is also known as Damped Least-Squares [42].



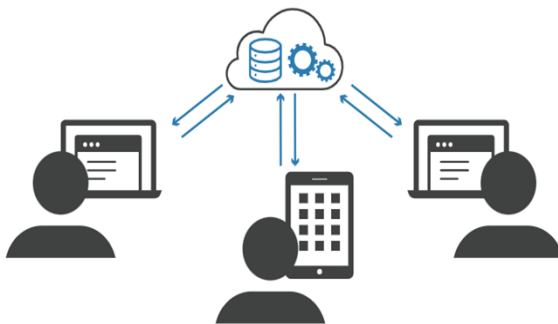
## 2.5. Review of modern photogrammetry software

The SfM method uses multiple algorithms as mentioned in the previous chapter like SIFT, RANSAC and multi-view-stereo (MVS) algorithms. With the advancement in the field of photogrammetry, today commercial photogrammetry software packages are available which follow the SfM workflow and make “reality capture” easy and achievable. There is a large variety of open-source photogrammetry software available. However, we have enlisted some of the new and popular photogrammetry software of 2022 that claims to have the ability to process photographs to make 3D models. We review all the enlisted software and discuss their advantages and disadvantages, feasibility, and hardware requirements. We can broadly divide the latest photogrammetry software into two categories: Desktop-based and Cloud-based software. The difference between these two types of applications can be understood from figure 5 and figure 6.



- Application logic runs on user’s system

Figure 5:Desktop based software



- Application logic runs on remote web servers.
- Servers can be accessed by any type of devices and more than one device at the same time.

Figure 6:Cloud based software

### 2.5.1 Desktop based photogrammetry software:

These software can only be installed on a personal computer and they largely depend on the computer hardware like processors, graphics processing units (GPU) or Graphic Cards and memory storage capacity of the system. We list some examples below.

- Meshroom

Meshroom is free open-source photogrammetry software based on Alice-Vision framework. The Alice-Vision is computer vision framework which provides 3D Reconstruction and Camera Tracking Algorithms. Its documentation shows a simple workflow, mentioning easy uploading of photographs and automatic generation of 3D models and textured meshes utilizing a node-based approach. All the nodes in Meshroom workflow has been shown in the figure 7.



Figure 7:Node based workflow of Meshroom

Meshroom is also available as a free plugin for the 3D modeling and animation software Autodesk Maya. Meshroom required a CUDA-Enabled GPU, with minimum compute capability of 2.0, for generating high-quality mesh. In absence of a CUDA capable NVIDIA video card, the DepthMap and DepthMapFilter steps will not be processed on the PC. Meshroom has its application in forensic pathology education of medical students where photogrammetry has been used as an alternative to CT or MRI scans [\[43\]](#).

- Agisoft Metashape

Agisoft Metashape, formerly known as Agisoft Photoscan, is a standalone application founded by Agisoft LLC, Russia in 2006. It is one of the most widely used photogrammetry programs for 3D modelling. Agisoft offers multiple licensing options (stand-alone, floating, and educational) with prices depending on the license of choice. Photogrammetry triangulation, point cloud data, distance, volume, and area measurements, mesh and texture are just a few of the capabilities available in this software. Agisoft Metashape appears to be a comprehensive program that has been used for numerous terrestrial and close-range photogrammetry, including some amazing

tasks like reconstruction of archeological monuments [44] and human skeletal remains in forest environments [45]. However, the application's basic configuration requirements are 32GB RAM, 4 - 8 core Intel or AMD processor and NVIDIA /AMD GPU with 700+ CUDA cores.

- 3DF Zephyr

3DF Zephyr is a commercial photogrammetry software created by the Italian software house 3DFLOW. There are a few different versions including 3DF Zephyr Lite, Pro and Aerial; all of which vary in price ranging from \$149- \$4200. The Free version limits users to uploading and processing only 50 images at a time. However, Zephyr offers education license which allows users to take advantage of all the advance features. Zephyr 3DF has few applications compared to Agisoft Metashape, however a recent study on 3D face reconstruction using Zephyr [46] motivates us to try this software. The hardware requirements are not much different from Agisoft with dual core 2.0GHz or equivalent processor, 16GB System RAM, hard disk space of 10GB, video card with direct X 9.0c compliant NVIDIA with at least 1GB of RAM; all these only as its minimum requirement.

- Reality Capture

Reality Capture developed by Capturing Reality is a state-of-the-art photogrammetry software solution for producing accurate 3D models from a variety of input media. The software prides itself on a simple and intuitive user interface on top of a rich feature set. It also claims to be the fastest 3D reconstructing software. We can create virtual reality scenes, textured 3D meshes, orthographic projections, geo-referenced maps even by combining both camera images and laser scans. It requires a NVIDIA graphics card for full operation because it uses CUDA for some of the key processing.

### 2.5.2 Cloud based photogrammetry software

Cloud-based photogrammetry solutions do not require users to rely on local computer hardware but use the computing power of Software as a service Cloud model for data processing. The service providers manage the hardware and software, and also ensure the security of the client data. SaaS allows organizations to speed up projects and at minimal upfront cost. Many new photogrammetry applications use cloud storage, servers and GPU which take care of the insufficient computer

power of the local computer. One great advantage of these cloud-based products is that the users will not need to purchase additional hardware services.

- Autodesk ReCap Photo

ReCap Photo is a desktop application that uses cloud services to carry out all the photogrammetry processes in cloud servers. It enables users to upload photos to clouds to generate the final 3D model. Although ReCap automatically registers images, it allows users to pick manually to enhance the overall result. One of the interesting 3D model renderings has been done using Autodesk ReCap Photo for Transtibial Prosthetic Socket Design Development for lower limb [47].

- KIRI Engine

KIRI Engine is a newly developed mobile photogrammetry software created by KIRI Innovations Science and Technology Inc. in 2021. The application features state of the art photogrammetry algorithm coded specifically for smartphones and cloud. KIRI Engine is one of the few available options that is compatible with both Android and iOS platforms and easily be downloaded from Google Play Store and Apple Store for free. In KIRI Engine, the whole process of building point clouds to textured mesh is done in the cloud. The users only need to upload images or move the camera slowly in a circle around the object to capture images from different angles. KIRI claims to be an AI (Artificial Intelligence) powered app that utilizes machine learning to automatically remove noise from the 3D scan.

Table 1: List of photogrammetry open-source software

Name	Type	OS	Output file formats	Price
<u>1.Meshroom</u>	Aerial, Close-Range	Windows, Linux	abc, obj	Free
<u>2.Agisoft Metashape</u>	Aerial, Close-Range	Windows, macOS, Linux	stl, obj, fbx,ply,3ds,wrl,x3d,dae	From \$179(free trial available)
<u>3.3DF Zephyr</u>	Aerial, Close-Range	Windows	ply, obj, fbx, pdf 3D, u3d, dae, pts, ptx, xyz, txt, las, e57	From \$300/month (Free versions and education license available)
<u>4.RealityCapture</u>	Aerial, Close-Range	Windows	obj, ply	\$10 for 3500 PPI credits, or \$3,750 for unlimited access
<u>5.Autodesk ReCap Photo</u>	Aerial, Close-Range	Windows	asc, cl3, clr, e57, fls, fws, isproj, las, pcg, ptg, pts, ptx, rds, txt, xyb, xyz, zfs, zfpri	\$360/year (free for students)
<u>8.KIRI Engine</u>	Aerial, Close-	iOS, Android,	obj,ply,stl	Free (3 downloads

	Range	Browser		per week) or \$49.99/year for unlimited exports
--	-------	---------	--	---

User experience is the most important factor in deciding the selection of the photogrammetry tool. We inspected all the desktop applications and the cloud-based software prior to our experiment. Our focus is on the search, identification, and exploration of an open-source software that is easy to use and offers promising results. The cloud-based software Autodesk ReCap Photo is extremely slow with long wait time to access cloud servers for both uploading and processing of images. It also does not offer choices to users to select the desired quality of output. Hence, we dropped Autodesk ReCap Photo out of the list of further experimentation. KIRI Engine, however, has a relatively shorter wait time to access cloud servers while providing presets for desired model quality but still lacks proper tutorials to start with. The desktop applications showed promising results and hence is selected for further experimentation.

## Chapter 3

### 3.Reconstruction of 3D Hand Models

Our goal is to reconstruct the surface of human hand (finger, palm, wrist) areas that resemble the real anatomical specimen. We are using smartphone Samsung Galaxy S20FE to capture videos of prosthetic hands, from multiple view angles. We test the performance and capability of four most popular photogrammetry open-source software to process image/videos captured through a smartphone device. We compare the outputs ( digital 3D hand models) qualitatively and qualitatively on their visual quality and time required to complete the reconstruction. We analyze the 3D models using CloudCompare software, benchmarking them with scans from 3D Structured Light Scanner(Artec Space Spider).

#### 3.1 Research materials

This section explains the resources used for data acquisition, data processing, and obtaining the results for our experiment.

##### 3.1.1 Hardware

A smartphone camera was used to capture images of the hand prosthetic. A decent hardware configuration is required to process the photographs. Photogrammetric pipelines and the rest of the computation were carried out in two different local computers for experimentation; System 1 is an old model but with a NVIDIA graphics card and system 2 is a new model with an Intel Graphics card. The specifications are given in the table below:

Table 2:Hardware and their description used in the experiment

Material	Name	Description
Mobile phone	Samsung Galaxy S20FE	Version: 2020 Purpose: Image acquisition
System 1	Lenovo Thinkpad	CPU: 11th Gen Intel(R) Core (TM) i7-1165G7, 2.80GHz GPU: Intel® Iris® Xe Graphics 16GB RAM

System 2	Lenovo Ideapad	CPU: Intel Core i5-7200U ,2.71 GHz GPU: Intel HD Graphics 620, NVIDIA GeForce 920MX 16GB RAM
----------	----------------	--

Table 3:Mobile Camera specifications

Specifications	Measures
Main camera	12 MP
Aperture size	F1.8
Focal length	26mm
Sensor size	1/1.76"
Pixel size	1.8 $\mu$ m
Video capture	3840x2160(4K UHD) 60 fps 1920x1080(Full HD) (240fps) 1280x720(HD) (960 fps)

### 3.1.2 Hand Prosthetic 1

We are experimenting with a prosthetic of a left-hand with extended fingers placed vertically (refer to figure 8).



Figure 8:Hand prosthetic 1 with extended fingers

### 3.1.3 Reference data from Artec 3D Scanner

Artec Space Spider is an ultra-high-resolution hand-held 3D scanner intended for precise capturing of small objects and complex details with accuracy up to 0.05 mm (Artec 3D, 2020)(shown in figure 9). It is based on blue-light technology. The Scanner can render complex geometry with intricate details that set the technology apart. It is ideal for CAD users. Latest research on assessment of Space Spider proves that it has superior precision and long-term repeatability in data capture [48].



Figure 9:Artec Space Spider 3D Scanner

[Picture courtesy <https://trimech.com/wp-content/uploads/2021/04/product-hardware-artec-space-spider-1.png>]

Table 4:Specifications of Artec Space Spider

Artec Space Spider	Specifications
3D point accuracy	0.05 mm
3D resolution	0.1 mm
Object size	From 5 mm
Working distance	0.2 – 0.3 m
3D reconstruction rate	8 fps
3D light source	Blue LED
Output formats	STL, OBJ, PLY, BTX
Price	USD \$24,800



The users need to move around the object with the handheld scanner to capture a 360-degree view. It is similar to the method explained in close-range photogrammetry. The models are created in Artec Studio software, which is a powerful desktop tool for engineers and designers.



Figure 10: Method to capture images using ASS

[Picture courtesy: <https://divcomplatformstaging.s3.amazonaws.com/geoweeek.divcomstaging.com/images/2aa3f9a3532e449327bf1ed31c6b6599.png>]

### 3.1.4 Choice of open-source photogrammetry software

We have a lot of open-sourced photogrammetry tools as discussed in Chapter 2. However, choosing the right software to meet the required endpoints remains a challenge. Among the large varieties of solutions, it is significantly important to determine the optimal ones. We are testing the four desktop applications namely, Meshroom, Agisoft Metashape Professional, 3DF Zephyr and Reality Capture to process a video captured using a smartphone. The reason to choose them for our experiment is because their quality and reliability in building complex 3D models has already been verified by a large number of researchers as well as experts from diverse fields.

## 3.2 Methodology

### 3.2.1 Data Acquisition

We are using a prosthetic hand with fingers in an extension position as our reference object. A video sequence of the dummy hand model was captured using mobile device Samsung Galaxy S20FE. The object was fixed vertically, and the video was taken by moving slowly and sequentially in a circle around the object. Starting with a loop at a low angle, then frequently moving at a higher angle to capture the topmost surfaces. The camera path captured the following

sequences of views for a left hand: palmar, lateral, cranial, and dorsal, as we moved around the dummy hand with the whole object taking as much screen space as possible in the camera screen (figure 11).



Figure 11: Views captured in a sequential manner a) Palmar b) lateral c) cranial d) dorsal e) lateral f) cranial views of the upper extremity captured from the smartphone video

We aimed for at least 60-70% overlap between frames, though 80% overlap is considered ideal. The distance between camera and object is 40-60 inches (1-1.5 m). The images from camera are originally uncalibrated data which are auto-calibrated by photogrammetry software. The video specifications for the experiment are of length 1 minute and camera resolution being 30 FHD (1080p). Video size is 125 MB. Images were extracted from the video sequence. The extracted images were RGB images with dimensions of 1920x1080 and maximum size of 128 KB.

### 3.2.2 Frames extraction from video

- Meshroom

Meshroom has no options to upload/import videos. Hence, we are using VLC Media Player to extract images from our video and upload them to Meshroom. VLC can extract video frames and save them as images in any required format in the system. We modified the scene video filter parameters of VLC to get the desired size and number of frames. The stepwise method is shown in figures below.

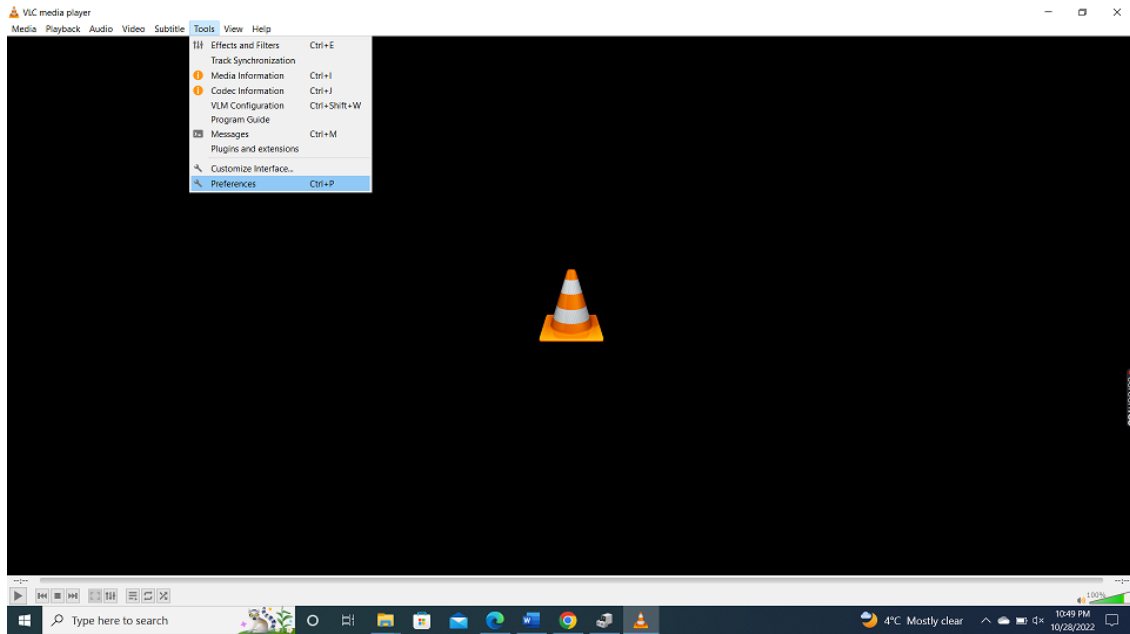


Figure 12: Open VLC, select tools then select preferences

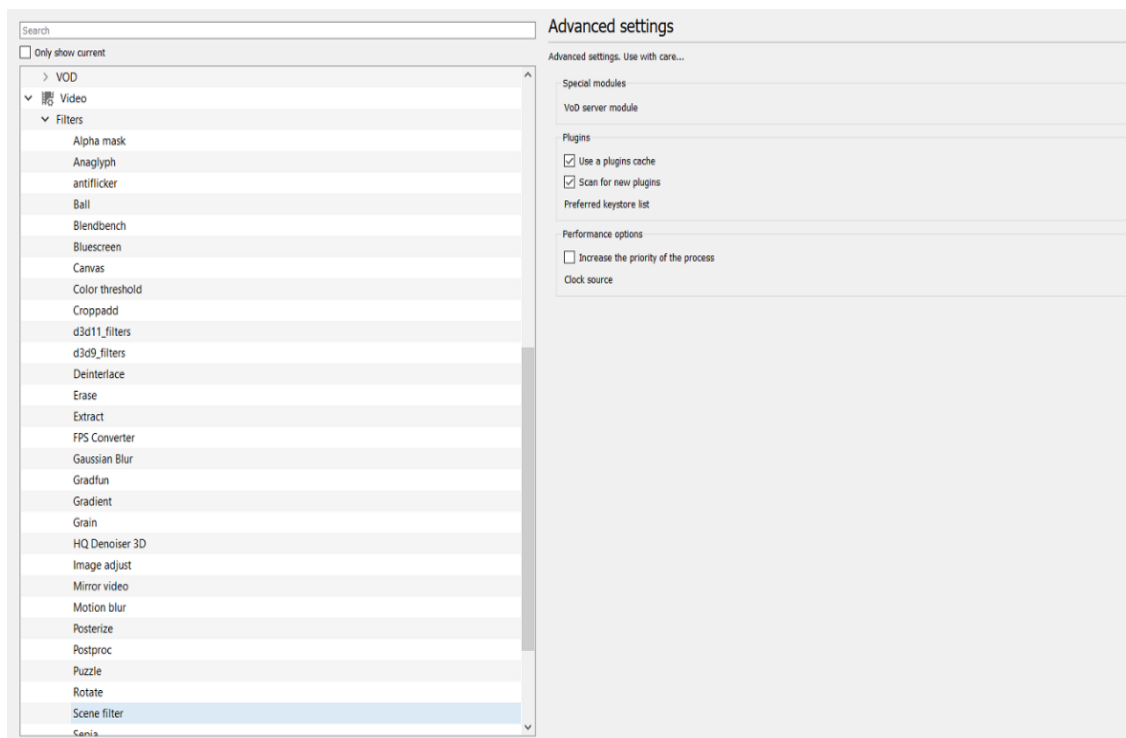


Figure 13: Select Filters settings from Advance Settings in VLC

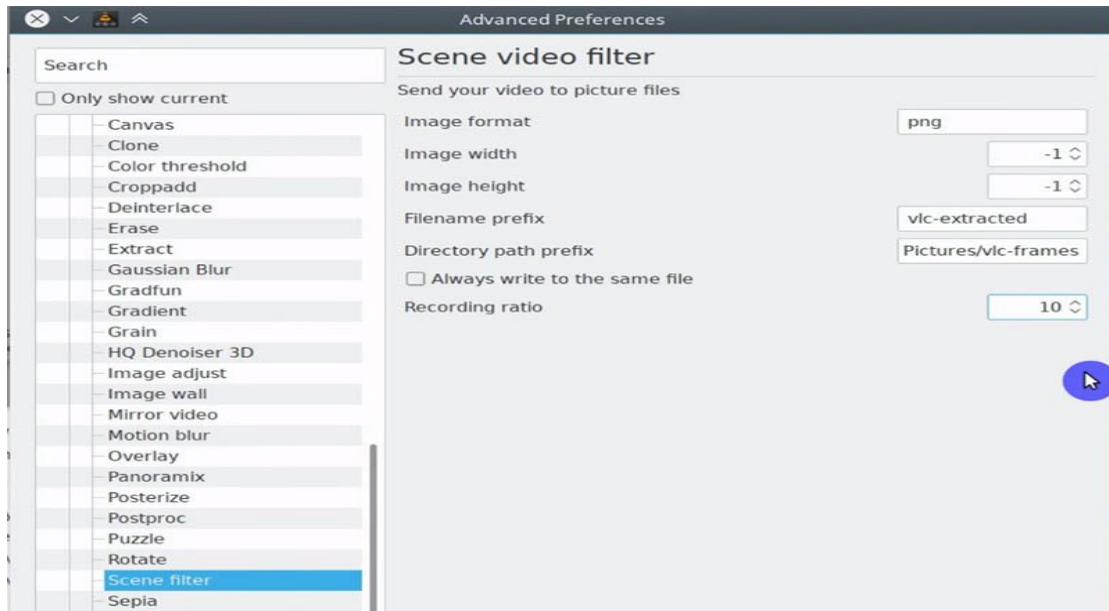


Figure 14: Select Scene filter settings to convert video to images with VLC

We should leave the default image height and image width values as -1. This will enable VLC to create images of the same size and resolution as the video file. We are required to specify the folder where we need to save the generated images, leave unchecked the Always write to the same file box. Otherwise, the VLC will just rewrite the same file. The Recording ratio option defines the frame rate extraction. We can increase or decrease this depending on the number of images we desire. Higher Recording ratio will give us higher number of images.

Once we have obtained a moderate number of images from all angles, images are dragged and dropped to Meshroom workspace. The user interface of the application looks like the figure 15 below.

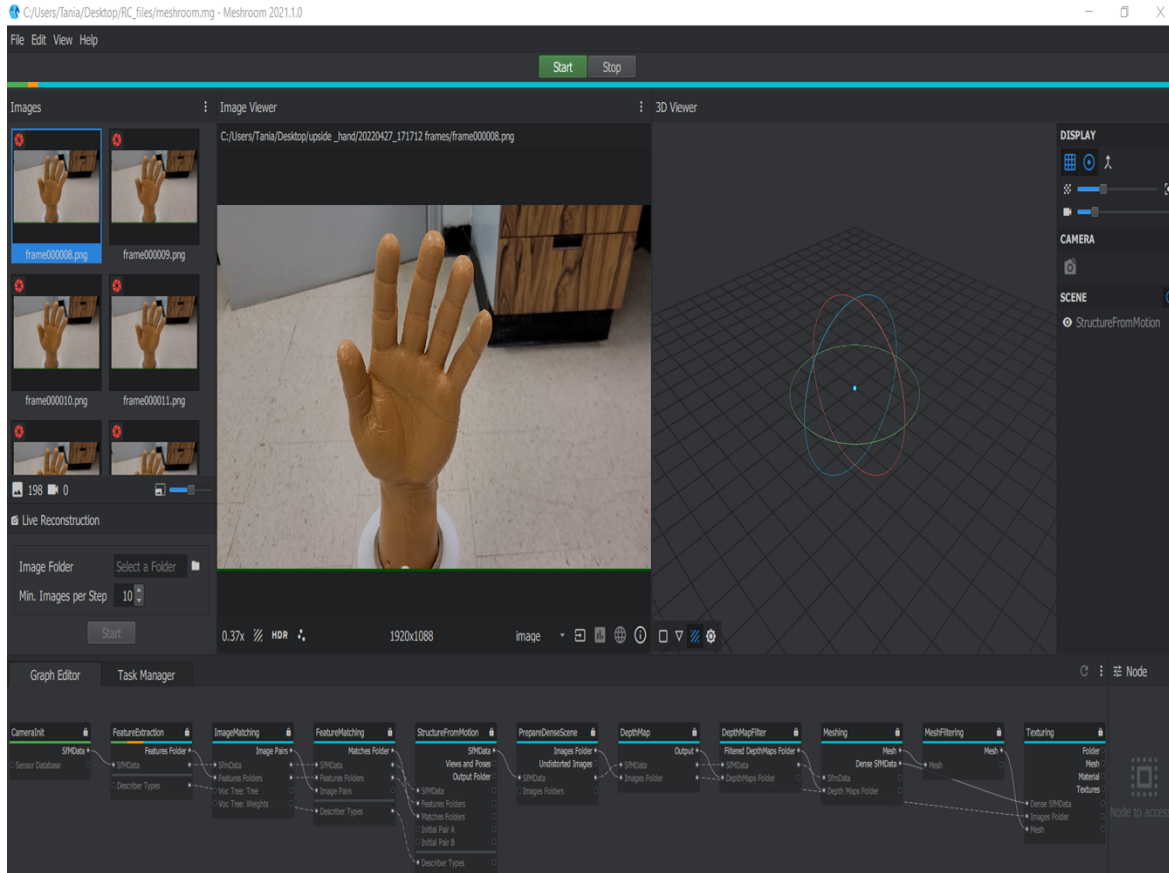


Figure 15: Meshroom UI loaded with images

- 3DF Zephyr, Agisoft Metashape, Reality Capture

The video was uploaded to each of the software. Frames were extracted at a lower rate of 3 fps so that we could obtain a moderate number of images as shown in figure 16,17 and 18. Average number of image frames detected and extracted was 165-180.

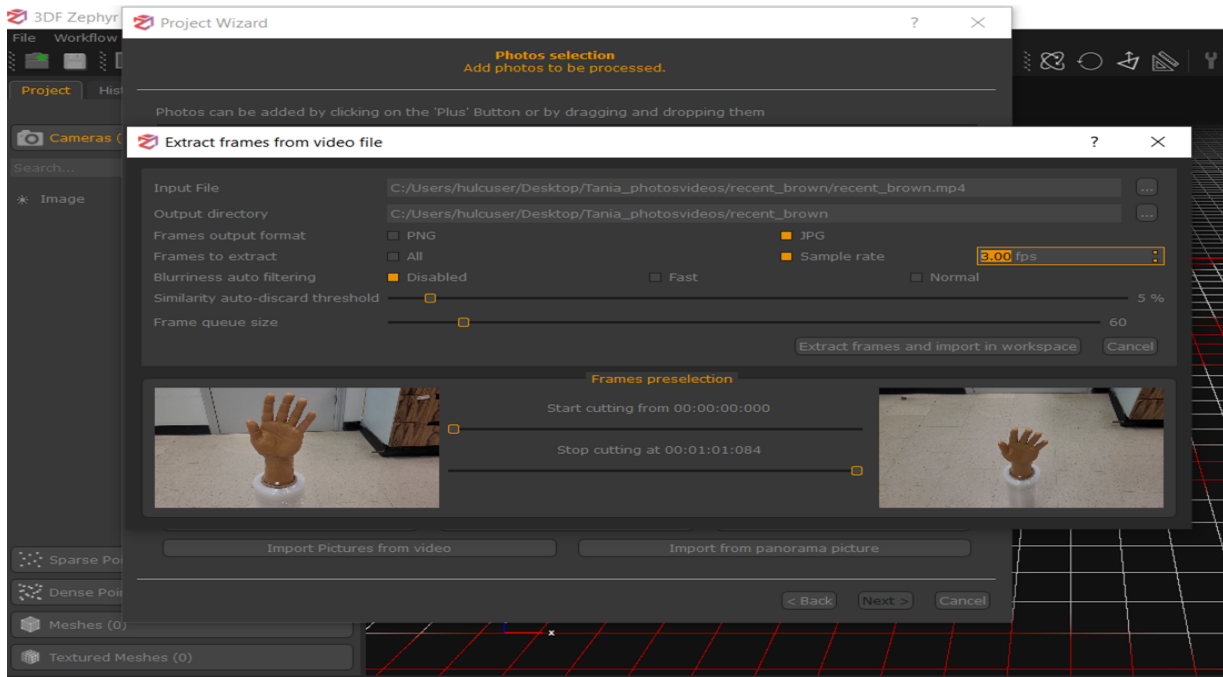


Figure 16: Process to import videos in 3DF Zephyr

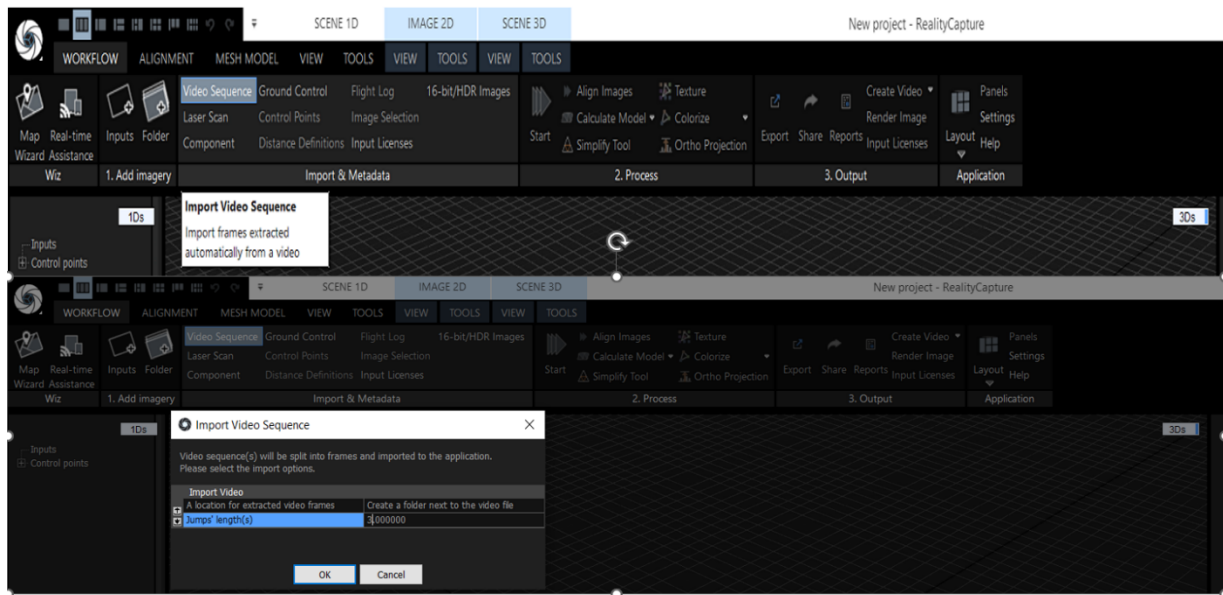


Figure 17: Process to import videos in Reality Capture

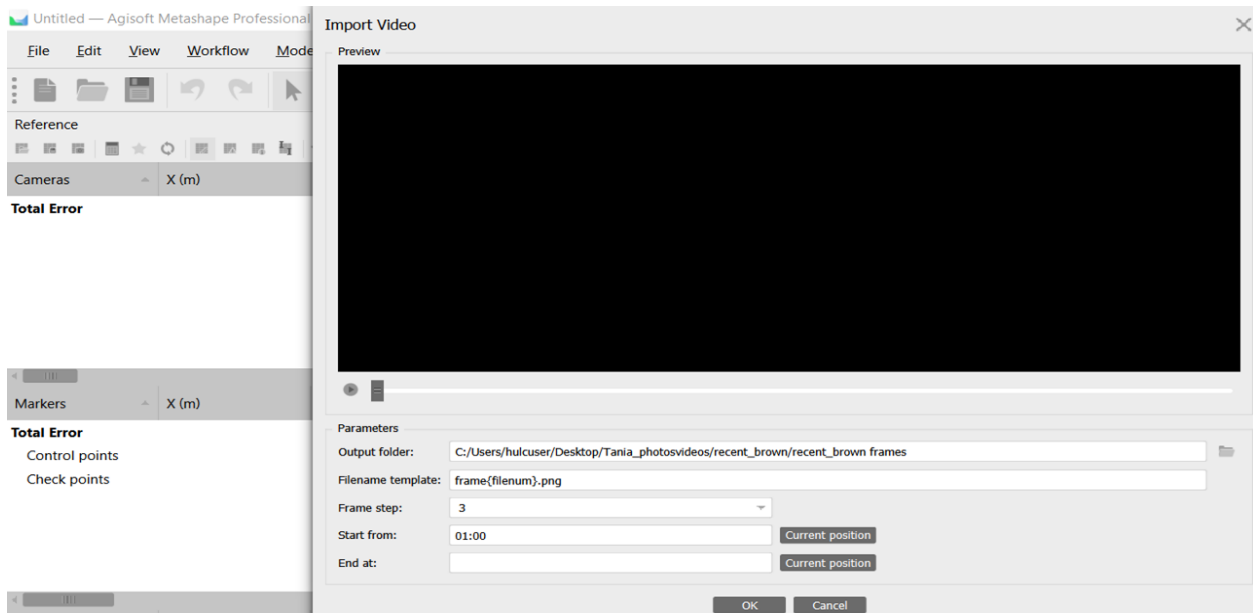


Figure 18: Process to import videos in Agisoft Metashape

### 3.2.3 Camera Alignment

In the first step of photogrammetry, images are aligned. In the figure below, the rectangular structures over the point clouds indicate the detected camera frames from all angles and these give us an estimation of the camera path followed to take images (refer to figure 19 and figure 20).

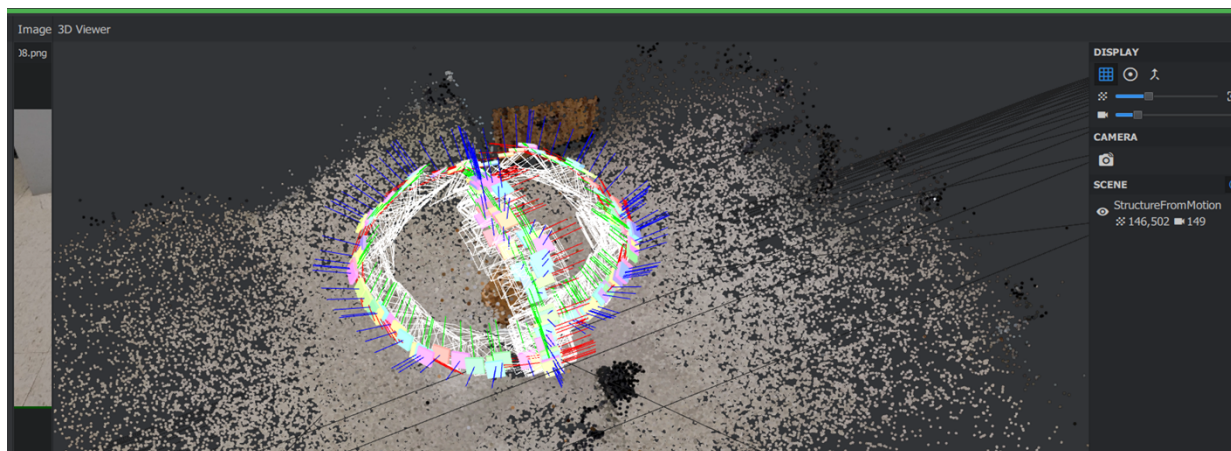


Figure 19: Aligned camera frames in Meshroom UI



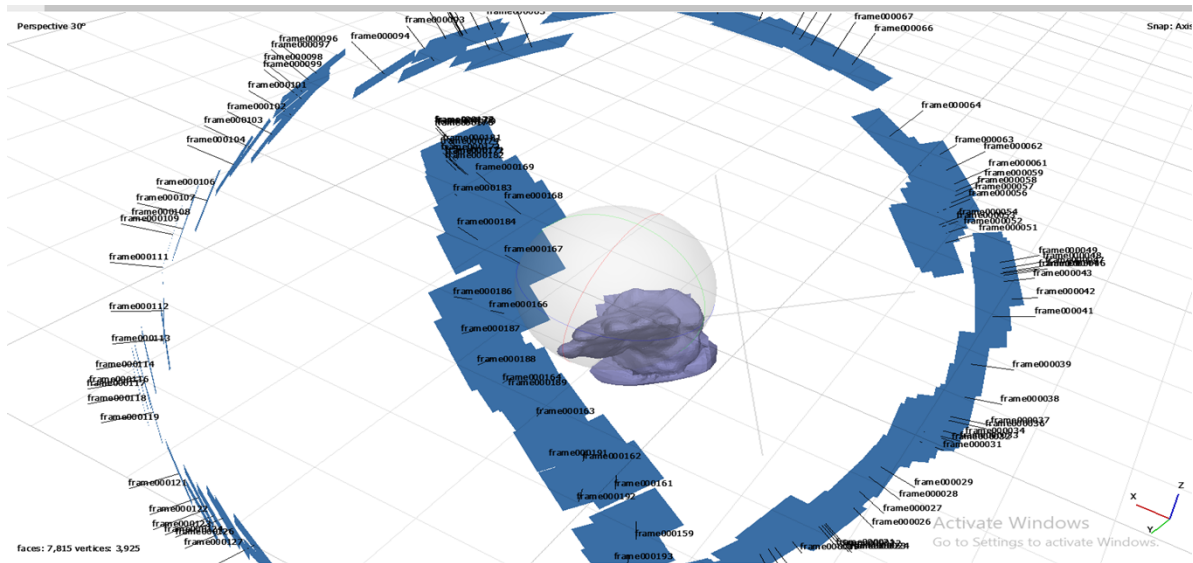


Figure 20: Aligned camera frames in Agisoft Metashape UI

### 3.2.4 Photogrammetry workflow

The entire case study using photogrammetry involves various steps of data processing. The first step is the generation of sparse point cloud followed by generation of dense point cloud, generation of mesh and finally generation of textured 3D model. Figures 21-24 provide a clear depiction of the step-by-step photogrammetry workflow as it was carried out during the experiment.

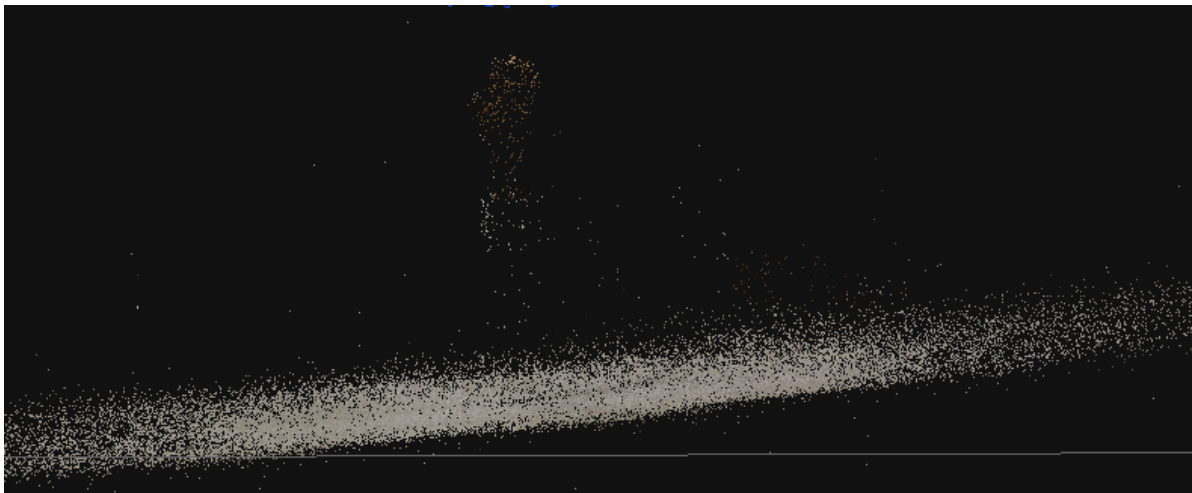


Figure 21: Sparse point clouds by 3DF Zephyr





Figure 22:Dense point clouds by 3DF Zephyr

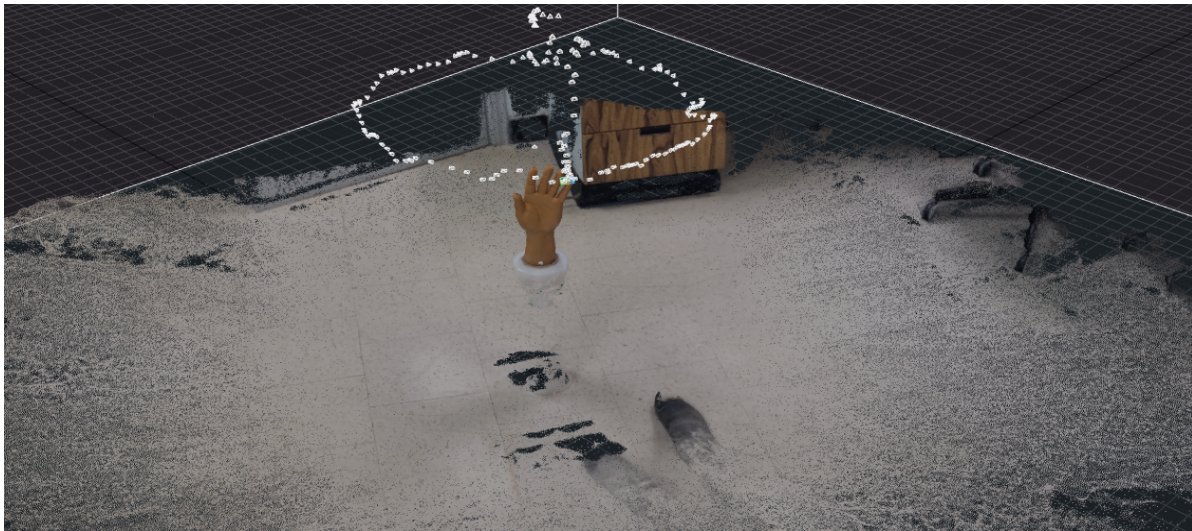


Figure 23:Mesh generated by 3DF Zephyr

All the photogrammetry software except Meshroom has editing tools. Unwanted areas and outliers were manually selected and deleted. The model produced by Meshroom was exported to another open-sourced software Meshlab where the surrounding unwanted structures were removed. Cleaning models is easiest in Reality Capture with minimum effort. We define an area and draw a bounding box around the model. Using the “Reconstruction” tool option, the area outside the bounding box gets removed.

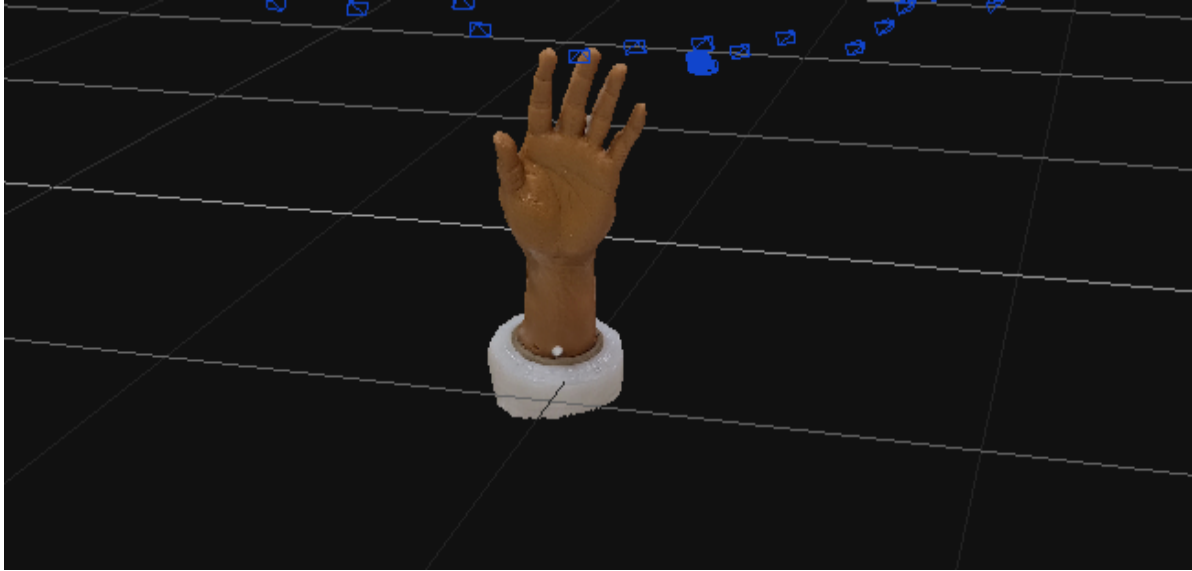


Figure 24: Textured Mesh in 3DF Zephyr after cleaning the surrounding

### 3.3 Results

Figure 25 and 26 shows the results obtained during the experiment, from all the four software, for each step of photogrammetry.

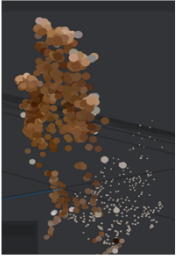


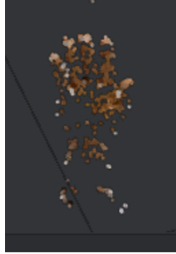


	MESHROOM	AGISOFT METASHAPE	3DF ZEPHYR	REALITY CAPTURE
Sparse Point Clouds				Sparse points clouds for Reality Capture was not displayed
Dense Point Clouds				Dense point clouds for Reality Capture was not displayed

Figure 25: Comparing results of first three steps of photogrammetry



Figure 26:Comparing results of last steps of photogrammetry

Table 5:Processing time required by four photogrammetry software

	Average Number of image inputs	Average Tie points	Reprojection error	Average reconstruction time	
				With only CPU	With NVIDIA GPU
MESHROOM	175	9800	0.96239 px	340 mins (without depth maps and mesh generation)	240mins
AGISOFT METASHAPE	175	2402	0.46899 px	45 mins	23 mins 17 secs
3DF ZEPHYR	175	3045	0.53581 px	42 mins	21 mins 50 secs
REALITY CAPTURE	175	2256	0.45230 px	-	15 min 19 sec

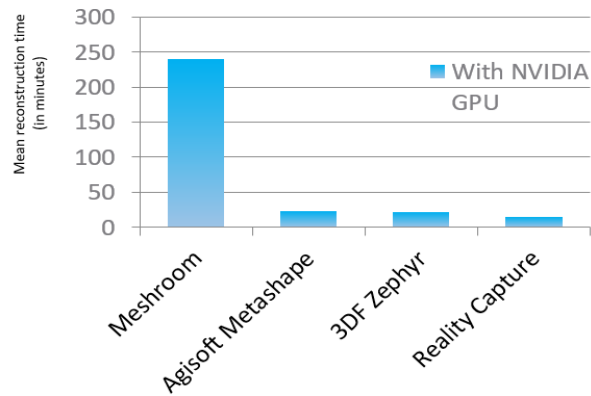


Figure 27: Bar showing total processing time taken by four photogrammetry software

Output :3D hand model by Reality Capture



Figure 28:3D model of the hand prosthetic

We have tested four different open-source photogrammetry software for the close-range photogrammetry using multi-images extracted from videos. Smartphone camera resolution has a noticeable impact on the final output. High-definition images capture minute details of the object, however, increase the data processing time. Formally, mean re-projection errors of less than one pixel is acceptable [49], whereas above one pix indicate a noisier reconstruction. The mean reprojection error for all the tested software is within one pix. Reconstruction of a 3D solid model of a human prosthetic hand could be implemented best using Reality Capture. Reality Capture lives up to its claim of being the fastest processing photogrammetry software. It takes a minimum time of around 15 mins for the whole reconstruction process as compared to other software. It also has the least reprojection error. The editing of the model using Reality Capture requires less effort compared to other tools. We only need to define a boundary box around the object of interest, and it automatically removes the points outside the boundary box.

3DF Zephyr has difficulty in reconstructing sharp edges and contours of the object. In the model generated by Agisoft, there are unwanted areas in between the fingers of the hand. Since fingers could not be separated it may not be the best model. The quality of the model generated by Meshroom is the worst when compared to others. It fails to generate a proper mesh and provide texture to it. Agisoft and 3DF Zephyr both take an almost equal amount of time to process the images, approximately 20-22 mins. Meshroom takes relatively longer hours just to produce point clouds. The availability of a NVIDIA GPU accelerates the reconstruction process even if the laptop/PC does not have recent configurations.

The differences between Reality Capture ,Agisoft and 3DF Zephyr as inferred from the experiment are summarized as follows:

1. User interface: The user interface of Reality Capture is more intuitive and user-friendly than the other two. The other two software have a steeper learning curve and can be difficult for beginners.
2. Performance: Reality Capture has relatively faster processing speed .
3. Output result: Reality capture produce sharper 3D textured mesh,it has least reprojection error hence more accurate output than other two software.
4. Output formats: Agisoft and Zepyr supports a wide range of outputs than Reality Capture.The later only supports OBJ and PLY.
5. Featues: Reality Capture has advanced mesh editing tools compared to Agisoft and Zephyr.

## Chapter 4

### 4. Comparison of 3D models using structured light 3D Scanner outputs as a benchmark

This chapter reports a comparative study of 3D models generated from smartphone images through photogrammetry and a commercial 3D Scanner. The 3D Scanner that we are using here is Artec Space Spider (ASS). We are using a different hand prosthetic with a more complex posture, having flexed fingers. We take ASS models as a benchmark for accessing the quality of the models produced by Agisoft Metashape, 3DF Zephyr and Reality Capture during the experimentation. The assessment is based on the geometric accuracy of 3D dense point clouds and triangular meshes. We employ signed distance and error metrics to evaluate them.

#### 4.1 Materials

##### 4.1.1 Hand Prosthetic 2

We are experimenting with a secondhand prosthetic with flexed fingers as shown in figure 29.



Figure 29: Hand prosthetic with flexed fingers

#### 4.1.2 Reference Data

An existing mesh from Artec Space Spider serves as a ground reality. It is used as a reference for accessing the quality of the models produced during the experimentation with this complex object. The reference model created in Artec Studio 17, looks like figures 30.1,30.2 and 30.3 .



Figure 30.1:Palmar view of ASS mesh



Figure 30.2: Lateral view of ASS mesh



Figure 30.3: Dorsal view of ASS

Average reconstruction time taken to produce the above scans in Artec Studio is approximately 30 mins.

#### 4.1.3 3D models from smartphone images

The curled fingers of the prosthetic cast shadows. Shadowed images are automatically discarded by photogrammetry software. Hence the input data faces some disparities due to missing images which ultimately leads to difficulties in image interpretation, feature matching and change detection, eventually delivering a poor model. Output models obtained from 3DF Zephyr looks like figures 31.1, 31.2 and 31.3.

~~the model is not as good as the reference model, the model is not as good as the reference model, the model is not as good as the reference model, for~~





Figure 31.1: Palmar view of 3DF Zephyr mesh



Figure 31.2: Lateral side of 3DF Zephyr mesh



Figure 31.3: Dorsal view of 3DF Zephyr mesh

- Average Reconstruction time of 3DF Zephyr for the above scan is about 20-25 mins.
- Average Reconstruction time of Agisoft Metashape for the above scan is about 25-30 mins.
- Average Reconstruction time of Reality Capture for the above scan is about 15-20 mins.

#### 4.1.4 Statistical software

CloudCompare (CC) is an open-source project designed to perform comparisons between two 3D points clouds or two triangular meshes. It can deal with massive point clouds (between 10-20 million points). Here we use CC to register, scale and find signed distances between scans from ASS and other photogrammetry software; 3DF Zephyr, Agisoft Metashape, Reality Capture. CC Source code repository is given in Appendix 1 .

We import two meshes, one generated by Artec Studio and the other one by photogrammetry software to the CC workspace. We realize that the two scans may have different scales; hence we need to scale them first.



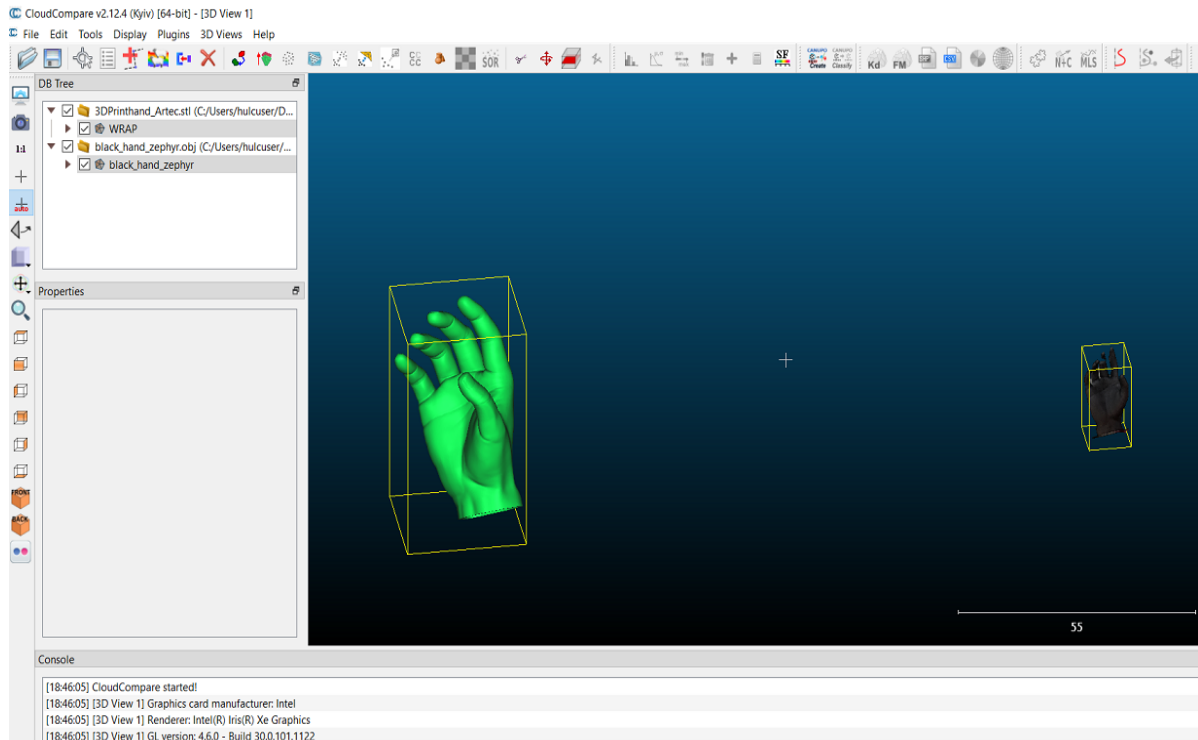


Figure 32: CloudCompare workspace with two imported meshes of different scales

In figure 32, the green-colored mesh is generated by Artec Studio (image captured by Artec Space Spider) while the black colored mesh in the above diagram is generated from 3DF Zephyr (Image captured by smartphone camera).

## 4.2 Evaluation Method

### 4.2.1 Scaling

As we can see in figure 32 above, the two scans have different sizes. CC offers the option to match scale and match bounding box centers. We keep the ASS scan as the ‘reference’ mesh (i.e., all other entities will be scaled to match this one) and scale it to the absolute scale of the ground truth. The algorithm used for matching scale is the principal dimension (deduced by Principal Component Analysis). The smaller mesh was rescaled by a factor of 19 to match the reference scale. The results of rescaling are shown in figure 33 below. In this figure, the bounding-box center comes at the same place as the center of the referenced entity.

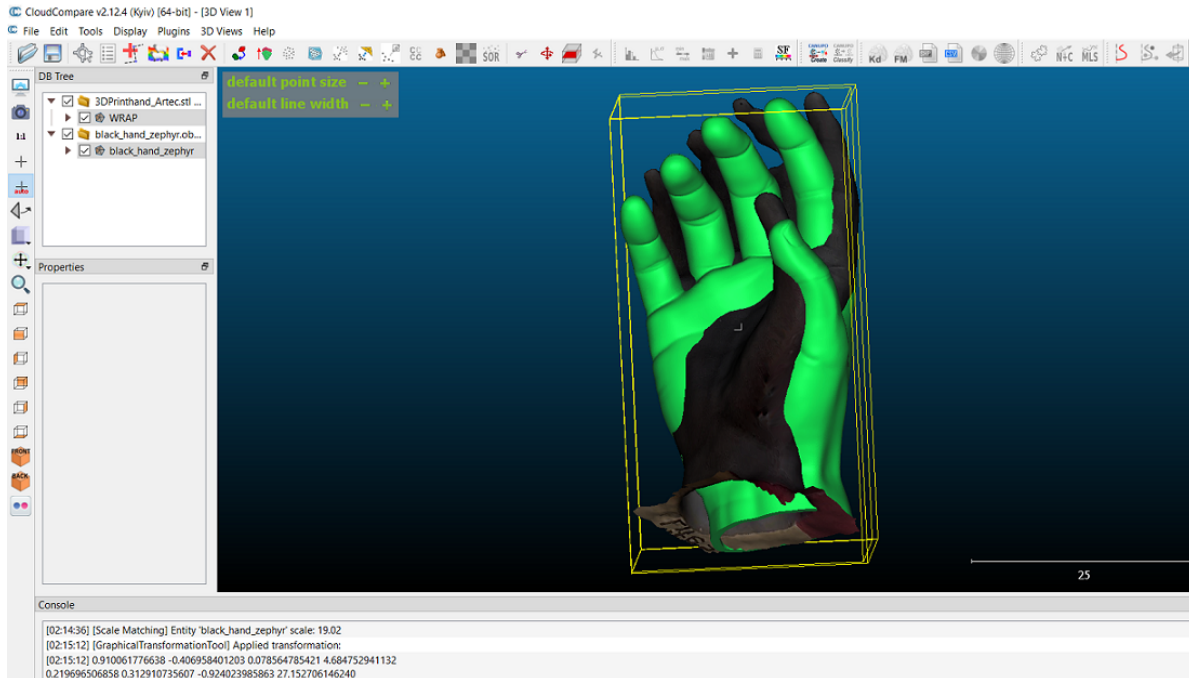


Figure 33: Matching scales and bounding box centers of the two imported meshes

#### 4.2.2 Alignment and Registration

- ICP Registration:

The Iterative Closest Point (ICP) algorithm represents the gold standard registration method for 3D shapes (Low, 2004). ICP is an iterative process explained by Besl and McKay [50]. It works by minimizing the difference between the points in two data sets. The algorithm keeps one of the point clouds fixed as a reference while moving or transforming the other to minimize an error metric and to best fit the models. In our experiment, we consider the ASS mesh as a reference and attempt to transform the other mesh.

- Fine Alignment in CloudCompare:

There are two fundamental assumptions for fine alignments in CC. Both clouds/mesh should be roughly aligned and both clouds/mesh should represent the same object or at least have the same shape (at least in their overlapping parts). CC uses original ICP algorithm denominations for fine alignment process. During this process, the registration error slowly decreases until the meshes converge and the final overlap shows a 100% convergence. Figure 34 shows the results after completion of registration process. We see in the 3D view that the color for two entities, ASS mesh and photogrammetry meshes turn yellow and black respectively. They correspond to the

'Reference Data' (yellow) and 'Aligned Data' (black) colors respectively. The meshes have been registered to the absolute scale of the ground truth and the 4x4 transformation matrix for the registered mesh is shown in the console.

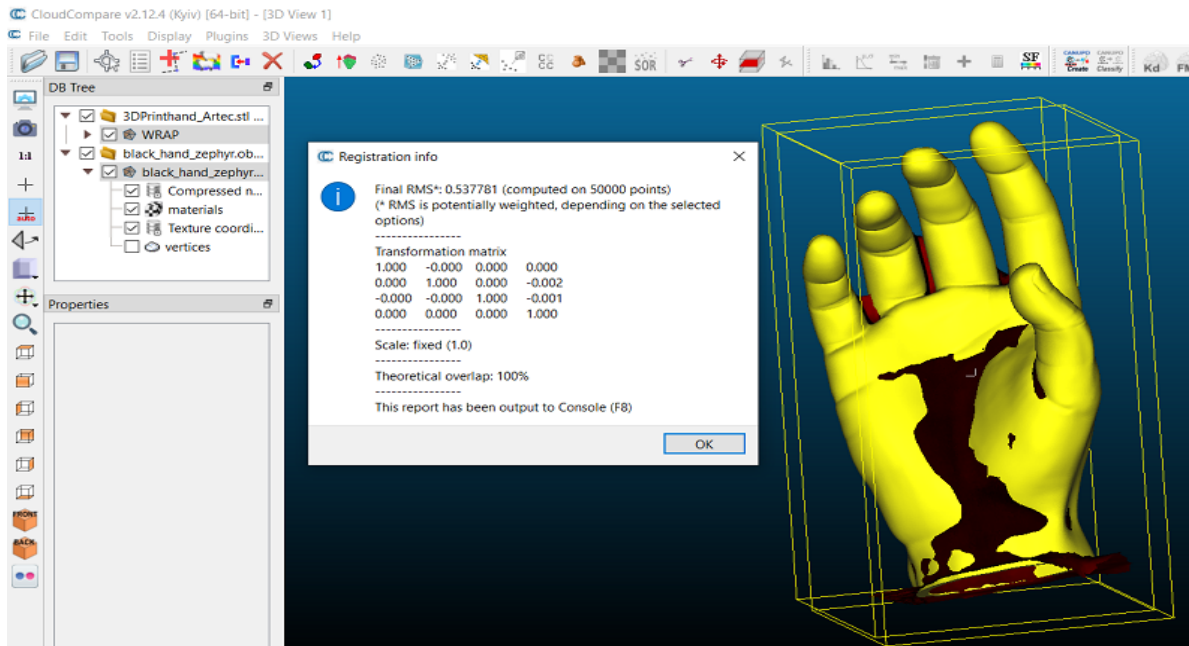


Figure 34: Registering two meshes

#### 4.2.3 Signed distances computation

Cloud-to-Mesh (C2M) distance comparisons have been previously used to detect a change between point cloud and a mesh [51,52]. However, C2M can also be used to detect changes between two meshes. In C2M, the vertices of the triangular meshes are considered and the distance for each vertex is computed relative to that of the reference mesh. Once the models are registered, using the normal of the meshes the distances between the triangle vertices are calculated. These signed distances are visualized as pseudo color heat map. We have filtered the maps in a range of  $\pm 2.6$  cm and displayed them in colors for easy visualization. Saturation is also adjusted within the same range where the two extremes are displayed in red and blue color.

From the distances, the mean and standard deviation are calculated. Pseudo color distance maps between the ASS mesh and the photogrammetry meshes are shown in the figures 35.1,35.2 and

35.3 below. Red colors indicate distances above the ground truth, blue colors indicate distances below the ground truth and green colors indicate where the surfaces coincide.

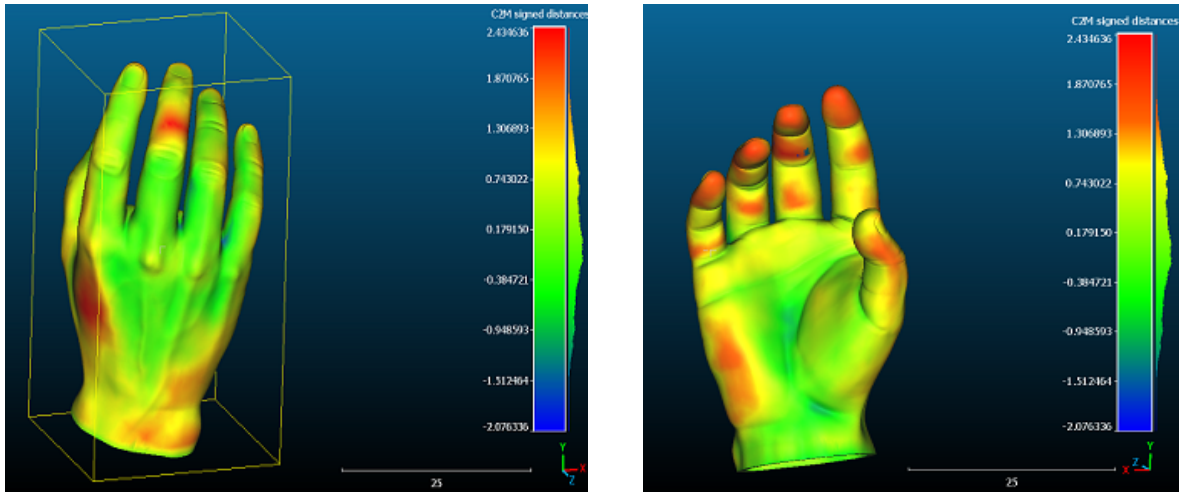


Figure 35.1: C2M distance computation between reference mesh and 3DF Zephyr mesh a) dorsal view b) Palmar view

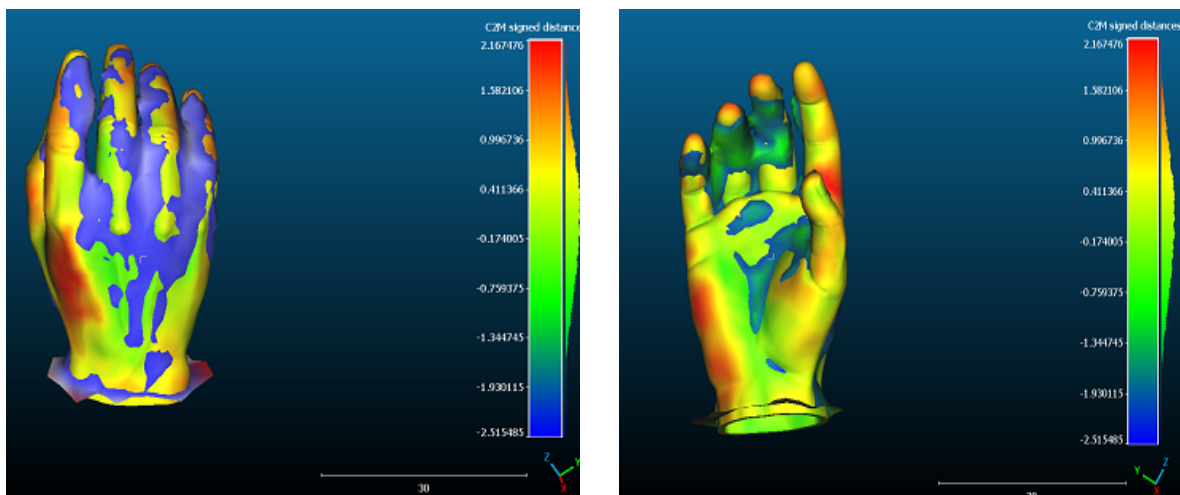


Figure 35.2: C2M distance computation between reference mesh and Agisoft Metashape mesh a) dorsal view b) Palmar view

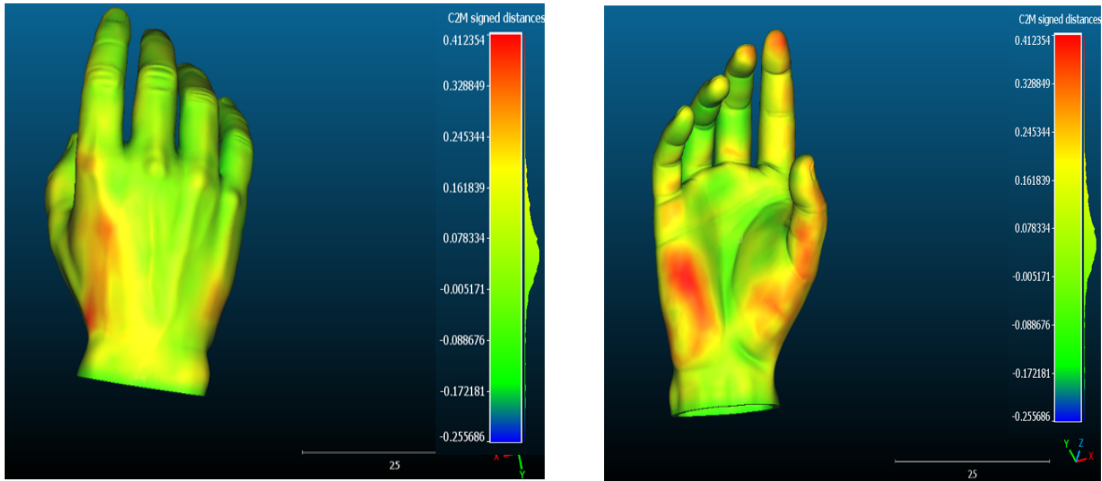


Figure 35.3: C2M distance computation between reference mesh and Reality Capture mesh a) Dorsal view b) Palmar

### 4.3 Gaussian distribution

A Gaussian normal distribution is assumed for the modeling of the C2M distance distribution between the reference model and the compared reconstruction. The distributions are shown in histograms below, in figures 36.1, 36.2 and 36.3. The mean, standard deviation and root-mean-square error from the calculations are summarized in the table 6.

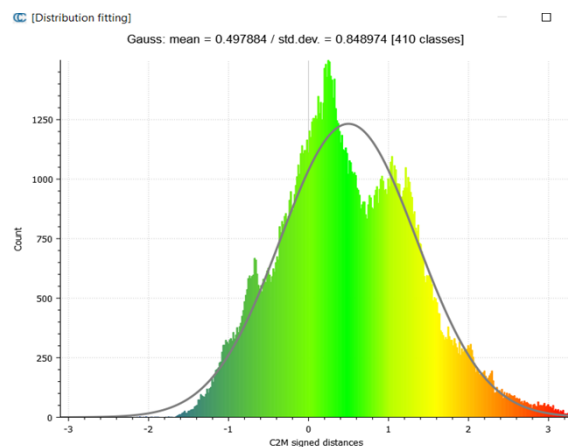


Figure 36.1: Histograms of the Gaussian distribution characterizing the distances between the ASS and 3DF Zephyr mesh.

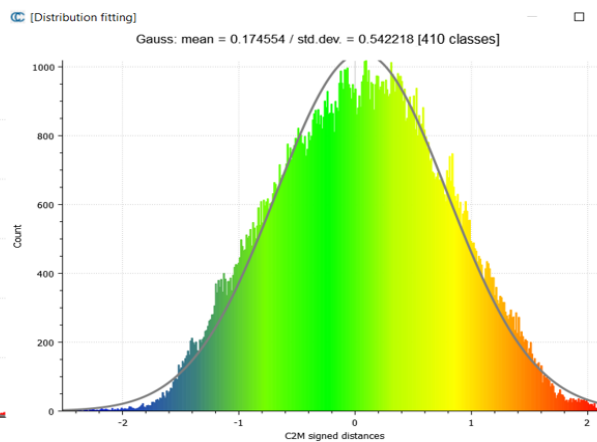


Figure 36.2: Histograms of the Gaussian distribution characterizing the distances between the ASS and Agisoft Metashape mesh.

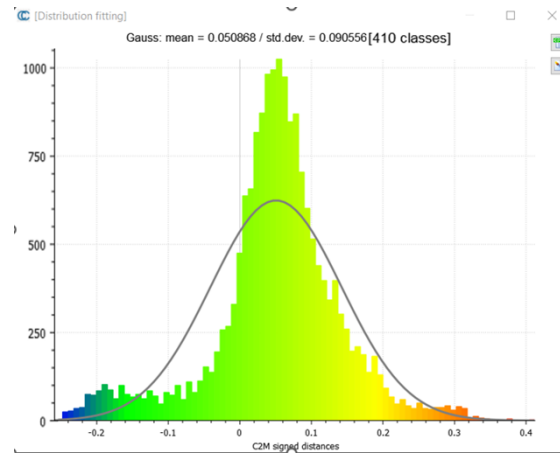


Figure 36.3: Histograms of the Gaussian distribution characterizing the distances between the ASS and Reality Capture mesh.

Table 6: Parameters for comparison

	Mesh face counts	Gaussian mean(cm)	Standard deviation(cm)	RMSE (cm)
Agisoft Metashape	22,732	0.17455	0.544	0.658
3DF Zephyr	29,270	0.49788	0.8489	0.884
Reality Capture	99,821	0.05086	0.0905	0.103

#### 4.4 Analysis

The models from 3DF Zephyr, Agisoft Metashape and Reality Capture have been analyzed, keeping outputs from ASS as a standard measure. The geometric inaccuracies of photogrammetry models can be seen in figures 35.1, 35.2 and 35.3. 3DF Zephyr could not reconstruct the flexed finger properly hence the signed distance measures are far above the ground truth near the fingertips (as indicated by red coloured areas). The model by Agisoft Metashape has unwanted adjoining areas between four fingers; the poorly constructed mesh areas have been marked by a contrasting blue color to indicate the differences from the standard mesh. The Reality Capture model has more green areas which means having more points with signed values closer to the ground truth.

This is also proved by the histogram obtained for signed distances against vertex count, shown in the figures 36.1, 36.2 and 36.3. The Gaussian distribution of Reality Capture mesh shows there are a greater number of points whose values lie closer to ground truth and there is a lesser number of extreme values on either side of the gaussian mean. It has a narrow distribution which indicates a smaller standard deviation. A smaller standard deviation for Reality Capture has been verified

from the values in table 6. The distribution of Agisoft Metashape and 3DF Zephyr has a wider normal distribution. High values of the standard deviation for the normal distribution of the cloud-mesh distances indicate that data points are widely spread around the mean and mesh has a poor quality. Reality Capture mesh model has the lowest standard deviation, thus the highest quality mesh. The error metric is also lowest for Reality Capture model with RMSE being 1.03 mm. Agisoft and 3DF Zephyr perform moderately when reconstructing a flexed prosthetic hand, with RMSE of value 6.5 mm and 8.8 mm respectively, while the difference in the RMSE error between their models being less than 3mm. Hence, Reality Capture can be an alternative solution for ASS as it produces a reliable model with a moderate mesh face count and geometric accuracies, lower errors and deviations and an average reconstruction time of less than 20 mins when compared to other software.

#### 4.5 Limitations of the study

The scope of this study is not an ideal solution and has its limitations and constraints. There are several rooms for further research and improvement in the future.

##### 4.5.1 Use of single device for data acquisition

We used the same mobile phone camera to capture videos/images for the experiment. But one should also consider the possibility of the input data being contributed from several users using various devices (Agarwal et al., 2011). This study does not showcase the effect of using images from multiple devices or a variety of sources. Latest Apple smartphones like iPhone 12 Pro, iPhone 13 Pro, iPhone 14 Pro Max are equipped with LIDAR (Light Detection and Ranging) sensors [53,54] that produce high quality pictures. Using an input data by combining normal RGB and LIDAR images from different or same mobile devices may produce a different result.

##### 4.5.2 Constraints for photogrammetry in an indoor environment

When performing photogrammetry in an indoor environment we had to consider the following constraints. Camera resolution played an important role in determining the quality of models. Mid-resolution to high-resolution photos yielded good 3D models. However, very high-resolution photos led to longer processing time. A moderate number of images (more than 50) with a high degree of overlap (60%) were compulsory for a decent model reconstruction. A very large number

of images, however, reduced data processing speed and took long hours to complete. Hence, we recommend a moderate number of images (100-200) for this type of close-range photogrammetry. Users should be careful about the lighting condition of the room where photographs are captured. The low light intensity makes the object look featureless, which results in erroneous image alignment. Hence, the lighting in the room should be uniform as shown in figure 37. Images with blurriness and shadows were automatically rejected by photogrammetry software.



Figure 37: The images on the left are shadowed images without proper lighting, while the image on the right is an image with homogenous lighting and is considered the correct image for photogrammetry.

#### 4.5.3 Requisite hardware issues

The 3D reconstruction processes are computationally demanding and require specific hardware configurations. An important prerequisite to run Meshroom and Reality Capture is Nvidia graphics card, Compute Unified Device Architecture (CUDA) and an enabled graphical processing unit (GPU) with a minimum compute capability of 2.0. Without these, Meshroom proceeds to run but stops at DepthMap node leaving behind an error message and a model that is not quite recognizable. Reality Capture is fast and easy but it never proceeds to run the photogrammetry algorithm without the requisite hardware. Agisoft Metashape and Zephyr run without CUDA but take a long processing time. Some cloud based solutions may offer possible solutions to run photogrammetry pipelines for desktop applications on public clouds. The company behind Reality Capture, CapturingReality has recently developed a cloud based app named RealityScan that follows the technology behind Reality Capture. However, it is still in its beta stage and only available to iOS users. The cloud computing instances mentioned above have not been tested and remain outside the scope of this study. However, they may affect the results by significantly altering the processing time and quality of the 3D model.



## Chapter 5

### 5. App for patient data collection

Web apps and mobile apps are advantageous for collecting and sharing patient data where patients can fill out questionnaires from the comfort of their homes. These applications significantly increase access to point-of-care tools, which have been shown to support better clinical decision-making and improved patient outcomes [55,56]. They also offer an affordable platform that reaches a large audience with possible positive implications for health promotion and prevention strategies [57]. There are few apps available online for hand telerehabilitation; however, almost none of them contains a comprehensive set of questionnaires to self-report and self-assess hand injury.

Our primary objective is to build a comprehensive data-collecting tool, an eHealth app where patients can self-report their medical history, pain and sensation score, and share images and 3D models of their injured hands with a consulting hand therapist for diagnosis. We developed a set of relevant questions that can help therapists to understand the hand conditions of patients. Questionnaires for evaluation are based on the Ten Test Scale for Sensation assessment [58], Patient-Rated Wrist Evaluation (PRWE) [59] and the Patient-Rated Wrist and Hand Evaluation (PRWHE) [60]. We combined the developed questionnaire with a cloud-based photogrammetry 3D scanning app, KIRI Engine that can produce a 3D model of the dorsal and palmer region of the hand. Our second objective is to use the user version of the Mobile Application Rating Scale (uMARS) [61] to investigate the usability of the Hand Scans app and demonstrate the utility of the uMARS to assess the quality and functionality of our app. The Mobile Application Rating Scale (MARS) was developed by Stoyanov et al [62] and till date, it is considered the reference scale for healthcare professionals in scientific literature.

#### 5.1 App workflow

The flowchart in figure 38 illustrates the workflow of the developed app. The users can open the app on their personal devices, log in with their email address and complete a series of questionnaires digitally using the application. They need to perform various hand and finger postures and upload the images as instructed for evaluation of range of motion. Finally, they need to use the photogrammetry-based 3D Scanning app KIRI Engine, permit device camera access and

move the camera in a circle around the hand at a slow pace to capture hand images from different angles. Photos should be taken sequentially, aiming for at least 60-70% overlap between images. The 3D scanning app will process the images and produce a 3D model of the hand. Data exports from the application come in the form of a csv file and a stl file that are sent to the therapist.

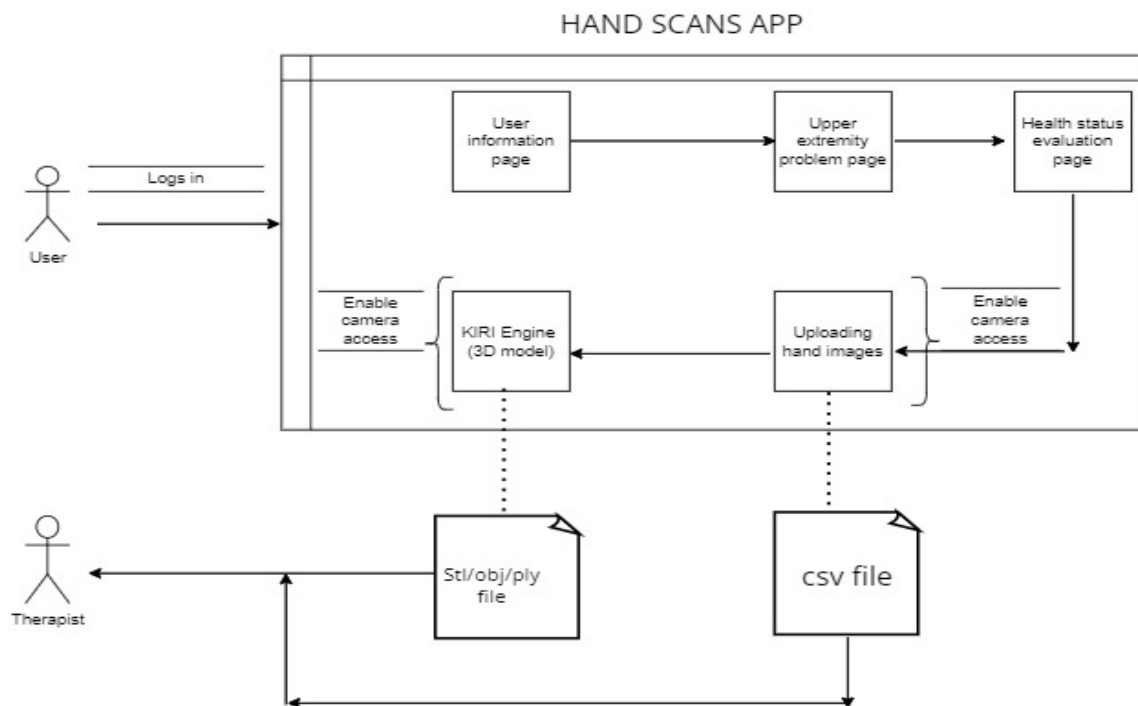


Figure 38:HAND SCANS app flowchart

## 5.2 Prototype Development

When designing the app, one of our main considerations is to ensure the confidentiality of the data provided by the patients. Second, the system needs to be user-friendly with ease for quick data entry (not exceeding >2 minutes) by patients. Finally, the medical data must be presented in a way that can be easily read and interpreted. The patient collection app has been developed using two online app builders: Jotforms for building the web application and AppyPie for the Android application. Both app builders offer hybrid platforms where we can write our own codes as well as use some impressive ready-made widgets. We used HTML, JavaScript and React Native framework for the front-end. We designed four online forms to collect data and synced the

responses to spreadsheets, using Google Sheets API, keeping all information in one secure collaboration workspace. We have connected the Sheets with the Google account of the hand therapist. We collect the data fields in a Google spreadsheet connected to the therapist's account.

Source code GitHub repository: <https://github.com/tbanerj2/Hand-Scans.git>

Webapp url:

[https://app.jotform.com/extremity-problem\\_photoapp/HandScan\\_io](https://app.jotform.com/extremity-problem_photoapp/HandScan_io)

Apk for downloading in Android mobile devices:

<https://d2wuvg8krwnvon.cloudfront.net/appfile/697c309030a0.apk>

### 5.3 Features

The app home page is carefully designed to cater to users with upper limb injuries (see figure 39). It features a logo and four forms dedicated to questions related to the users' condition and patients' medical history. Additionally, the page includes a link to a photogrammetry based 3D scanner, KIRI Engine, which scans and exports 3D hand models (refer to Figure 39). To ensure accessibility for colorblind users, the font colors and icons are chosen with great care. Prior to accessing the home page, users are required to log in with their email addresses. For any queries or concerns, a contact form is available to connect users with the developers or technicians.

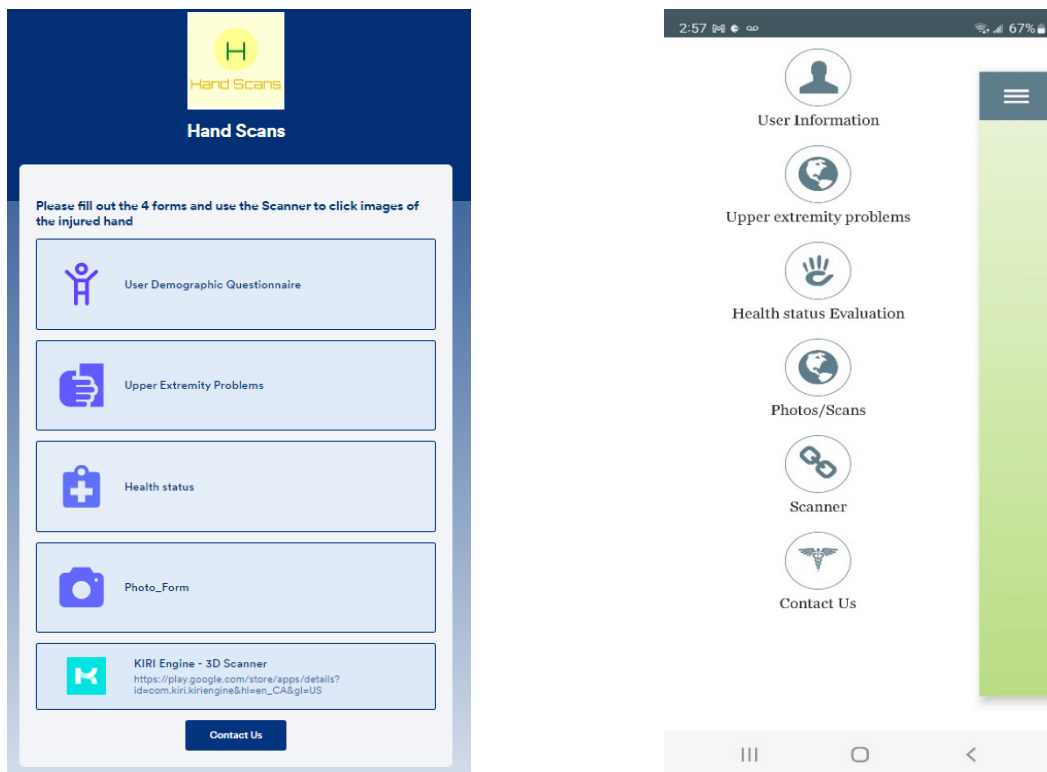


Figure 39:HAND SCANS homepage a)Webapp interface b)Mobile interface interface

### 5.3.1 Patient background information collection

The first form deals with demographic questions and requires users to fill in their personal details and basic questions related to their hand injury (see figure 41). We generate a unique ID with a combination of random numbers and alphanumeric characters to identify each patient. We use a free widget from Jotform console to build the input field (see figure 40).

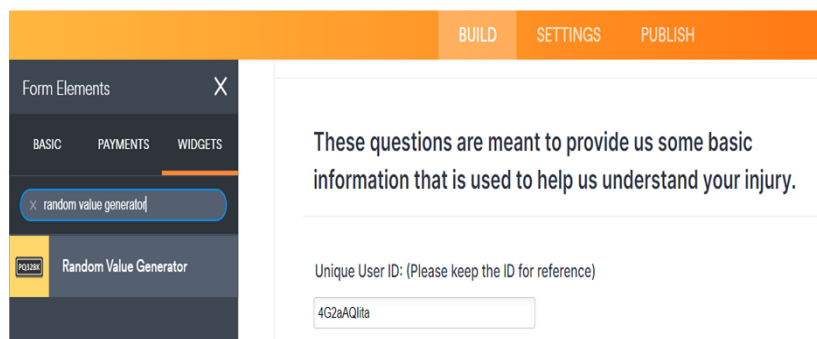



Figure 40:Random value generator widget in Jotform console

Figure 41: Questions in User Demographic Questionnaire form

### 5.3.2 Health information collection

The second form is designed to gather detailed information related to hand injuries, including the precise cause and time of the injury. To streamline the process, the form employs a "branching" logic feature where selecting the "Yes" option for the cause of the hand/arm problem will prompt additional questions related to the injury. The figure 42 shows the user interface and questions from the second form. Comorbidity information is essential for therapists to accurately diagnose hand injuries. It helps them understand how multiple medical conditions or diseases may be interacting and affecting the patient's condition. For instance, a patient with diabetes may have a slower healing process and may require specialized treatment for their hand injury. To facilitate the collection of comorbidity information, the third form features a table that lists common comorbidities and allows users to select multiple fields. Figure 43 illustrates all the comorbidities included in the prototype for reference.

## Upper Extremity Problems



1. Did an injury cause your hand/arm problem?

Yes  
 No

1.a. Select the cause of injury from the list

1.b. What was the date you were injured?

Date

2. Did your problem develop over a period of time?

Yes  
 No

2. a. You selected YES. Check all that applies.

Nerve compression (CTS)  
 Arthritis  
 Tendon/muscle strains  
 Osteoporosis

2. b. How long did you have the problem?

Less than three months  
 More than 3 months

3. Other than the problem being treated now, do you have other problems or injuries that limit the function of your hand or upper limb?

Yes  
 No

4. Do you take pain medication for your upper extremity problem?

never  
 daily  
 occasionally  
 several times a day

5. How would you describe your usual activity level (e.g. pre-injury)?

Always active  
 Occasionally active  
 Rarely active  
 Never active

6. If working what is your current occupation?

Figure 42: Questions in Upper Extremity Problem form

## HEALTH PROBLEMS

I. The following is a list of common health problems. Please select if you currently have that problem listed below. If you do not have that problem, skip to the next problem.

1. Please select all those apply for the problems

	Yes I have this problem?	Yes I receive treatment for it?	Yes it limits my activities?
Diabetes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cancer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Depression	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Osteoarthritis, degenerative arthritis	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rheumatoid arthritis	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other medical problems	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. Please list any other medical problems

Figure 43: Health status form listing comorbidities



Area 1

Select the sensitivity level

0 1 2 3 4 5 6 7 8

Low

9 10

High

OK

Figure 45:Pain level (0-10)

The patient-rated wrist evaluation (PRWE) was developed by Dr. MacDermid to measure pain and disability in patients with distal radius fracture (DRF) [63]. Subsequent research has reported the PRWE to be a reliable, valid, and responsive tool for assessing self-reported pain and disability in patients with DRF and certain other upper limb injuries [64,65]. The reliability of the PRWE was assessed using the intraclass correlation coefficient (ICC) in previous studies and varied within a range of 0.78 to 0.94, suggesting good reliability [66-68]. The question set in PWRE and PWRHE has been adapted and used in this app. The pain subscale consists of five items inquiring about the characteristics of wrist pain. The function subscale has ten items and is further divided into “specific activities” and “usual activities”. Each item on both subscales is rated on a visual analog scale of 0 (no pain/no difficulty) to 10 (worst ever pain/unable to do the activity). The total score range from 0 to 100 and is calculated as  $(\text{Usual activity score} + \text{Specific activity score})/2 + \text{Pain score}$ . 0 is evaluated as the best score and 100 as the worst score. Greater value indicates worse hand condition.

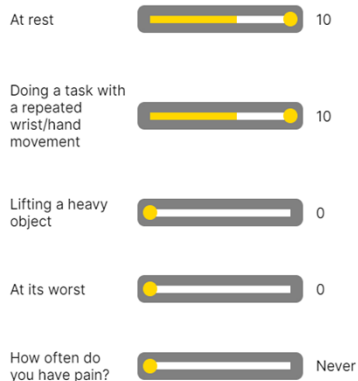
The figures 46, 47 and 48 provide an example of how the PRWE score is measured using the app. In figure 46, the illustration displays the five items used for pain score estimation and a user's selection of the pain scale led to a pain score calculation of 20. Figure 47 showcases the ten item questions used for estimating the function of an injured hand, where the user's special and usual activity scores both resulted in a score of 20. Finally, in figure 48, the total calculated PRWE score is shown to be 30, obtained by applying the formula  $[(\text{Usual activity score} + \text{Specific activity score})/2 + \text{Pain score explained}]$  as explained earlier.



## RATING PAIN LEVEL

(Rate how difficult it was doing the things listed below, this week. A zero (0) means it was not difficult at all and a ten (10) means it was so difficult you were unable to do it).

Rate your pain during the following :



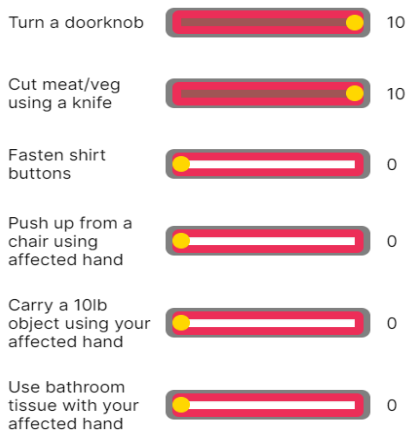
The total pain score is **20**

Figure 46: Measuring the pain score of all 5 items

## FUNCTIONS

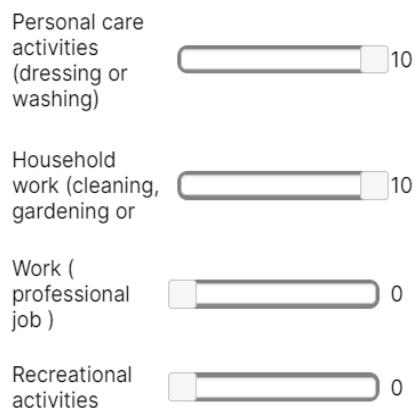
(Rate how difficult it was doing the things listed below, this week. A zero (0) means it was not difficult at all and a ten (10) means it was so difficult you were unable to do it).

### SPECIFIC ACTIVITIES



The specific activity score is **20**

### USUAL ACTIVITIES



The total usual activity score is **20**

Figure 47: Measuring function items a) specific activities b) usual activities

Your TOTAL score is:

30

Figure 48: Measuring the total PRWE score of all 10 items

### 5.3.4 Collecting hand images

To evaluate the Range of Motion we ask the user to follow the picture guides in the fourth form and make similar hand postures and upload the image (refer to figure 49). The photo widget can gain access to the system's camera, both front and rear. It also gives an option for a retake. In case of retake the previous image gets overwritten by the new one. There is a field to upload any other document the user may feel is important, for example, previous medical prescription.

**RANGE OF MOTION**  
Instructions: To evaluate we need to have pictures of your injured hand in different positions. Please follow the steps: 1. Position your hand in all requested pictorial guides, 2. Click a photo of your injured hand using your phone camera.

**HAND POSTURE 1**  
Make this hand posture and click a photo

Take Photo  
Take Photo

**HAND POSTURE 2**  
Make this hand posture and click a photo

Take Photo  
Take Photo

**HAND POSTURE 3**  
Make these hand postures and click photos

Take Photo(first image) Take Photo(2nd image)  
Take Photo Take Photo

**HAND POSTURE 5**  
Make these hand postures and click photos

Take Photo (first image) Take Photo(2nd image)  
Take Photo Take Photo

**WRIST POSTURE**  
Make these hand postures and click photos

Take Photo(first image) Take Photo(2nd image)  
Take Photo Take Photo

Take Photo(3rd image)  
Take Photo

Upload any other files you want to add

Upload File  
Choose Images

Save Submit

Figure 49: Images and file upload for range of motion estimations

## 5.4 Advance features

### 1. Contact form

Users can fill in a query form in case they need assistance. The query form looks like the figure 50.

The figure displays two versions of a 'Contact Us' form. On the left (a) is a desktop web application interface. It features a title 'Contact Us' and a sub-header 'We will try to respond as soon as possible'. Below this are four input fields: 'Name \*' with a placeholder 'Enter your name', 'Email \*' with a placeholder 'Enter your email', 'Phone Number' with a placeholder 'Enter your phone number', and 'Inquiry' with a placeholder 'Please share additional contact details...'. A dark green 'Submit' button is positioned at the bottom center. On the right (b) is an Android application interface. It has a dark blue header with a hamburger menu icon, the title 'Contact Us', and a three-dot menu icon. Below the header is the same sub-header. The form fields are stacked vertically: 'Patient Name \*' (with a person icon), 'Phone' (with a phone icon), 'Email \*' (with an envelope icon), and 'Your Inquiry \*' (with a document icon). A dark green 'Submit' button is at the bottom of the form. A blue 'Chat' button with a speech bubble icon is located at the bottom right of the screen. The Android interface also shows a status bar at the top with the time 5:26 and 21% battery, and a navigation bar at the bottom.

Figure 50:User interface of Contact form for a) Webapp Application b)Android application

### 2. Progress bar

Each page has a progress bar fixed to the top of the page. A progress bar is particularly helpful for longer surveys where a respondent is more likely to get frustrated and anxious about how much of their time they will require to invest to complete the survey.

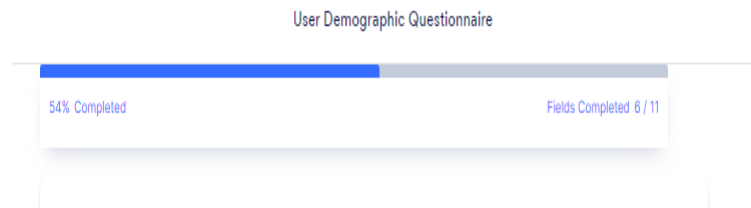


Figure 51:Progress bar

### 3. Save workspace and continue later

A user can save his progress with the help of the SAVE button. He will receive a link in his email which he can use to complete the forms at a later time.

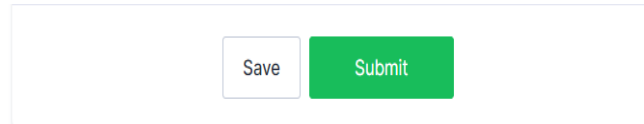


Figure 52: Save to continue later

Once a form gets submitted successfully, the completed forms are indicated by small green arrows beside the respective forms. Users can view which sections are yet to be completed.

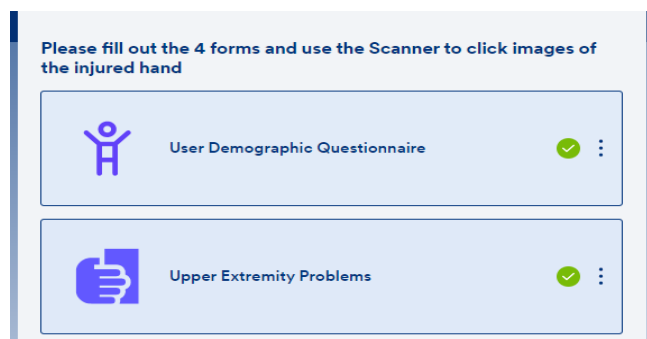


Figure 53: Green arrows showing the completion and submission of form

## 5.5 KIRI Engine 3D Scanner

KIRI Engine is a 3D scanning app compatible with both Android and iOS platforms. It uses photogrammetry algorithm on the cloud to produce high-quality 3D scans which can be downloaded or shared over the internet. It does away with the system hardware and storage requirements typically required by photogrammetry desktop software like Metashape and Zephyr 3DF. The Free version allows only for 70 photos per project and 3 free exports per week. Beyond that users are charged a small amount for exports. Premium allows for 200 photos per project, unlimited scan exports, and the possibility to upload photos from your PC or mobile gallery. The users however need to select the quality of scans they desire (low/medium/high) and the export file format (PLY/OBJ / STL) . The figures 54 and 55 depict an excellent example of the hand 3D models generated by the KIRI Engine.



Figure 54:3D model of a) dorsal side of hand prosthetic b)palmer side of hand prosthetic using KIRI Engine



Figure 55:3D model of the dorsal side of a real hand using KIRI Engine

## 5.6 Hand Scans Usability Survey

We used a modified user version of the Mobile App Rating Scale (uMARS) to critically appraise the mobile version of our Hand Scans app. uMARS provides preliminary evidence for its internal consistency and test-retest reliability [60]. uMARS comprises 20 questions, containing 4 objective quality subscales—1) User engagement, 2)App Functionality, 3)Aesthetics, and 4)Information quality—and a subjective quality rating. We chose not to include the subjective qualities or the perceived impact items of the uMARS, however, the users may decide to fill them if they desired.

According to the study performed by Stoyanov et.al [69], the uMARS has good internal consistency ( $\alpha=.90$ ) and high inter-rater reliability, thus presenting uMARS as a reliable tool for quality ratings of apps by health care personnel. For all the uMARS sections, items are rated on a 5-point scale (1-inadequate, 2-poor, 3-acceptable, 4-good, and 5-excellent).

#### 5.6.1 Data collection

A certified hand therapist and a group of 27 medical trainees from The Bone and Joint Institute at Western University evaluated the app using the uMARS tool. We asked them whether they think our app is useful and convenient from a user's point of view and whether it contains important questions that a therapist would ask patients if they visit their clinic in person. The prototype was examined for 15 mins and then independently rated by the trainees. App review using the modified uMARS was completed by 23 of 28 respondents, where 18 users were running the app on iPad and 5 on smartphones devices. The uMARS questionnaire used for this survey is given in Appendix 3.

#### 5.6.2 Results

The ratings obtained for each attribute have been listed in Table 7 while figures 56 and 57 offer an insightful graphical representation of the rating distribution of all uMARS attributes.

Engagement criteria were evaluated based on users' interest, customization, and interactivity. Customization refers to allowing the users to customize settings and preferences such as content and notifications. Interactivity means whether the app allows user input, provides feedback, and contains prompts such as alerts, sharing options, and notifications. Our app received a relatively low rating of 3.2. The main reasons for the low score on engagement were due to availability of only basic interactive features, and the lack of customizable settings according to the user preferences.

Functionality was evaluated based on the app's performance, ease of use, navigation, and gestural design. We received a higher mean value of 3.99 which indicates the prototype components (button, menu, functions) work accurately and the user interface is simple, intuitive, and easy to use. The app has all the necessary links to navigate between screens along with consistent gestural designs.

Evaluation of aesthetics attribute is based on the layout, graphics, and visual appeal of the prototype. Our app scored 3.51 which indicates it has acceptable aesthetic appeal. We can improve the stylization in the future by including more graphics and animations.

We received the highest rating of 3.94 for information quality, which signifies that the content is highly relevant, appropriate and credible.

The Subjective quality score ranged from 2.5 to 4 with a mean score of 3.42. The questions regarding Recommending to others received a mean high score of 4 while paying for the app received a low mean score of 2.

The overall mean App quality score is 3.61 with a standard deviation of 0.26, with most features of the app scoring >3 out of 5 on uMARS domains. From the results, we can infer that Hand Scans is a good quality mHealth app with relevant content for remote surveillance of patients with hand injuries. The future versions of the prototype should have improved User Interface and interactive and engaging features.

Table 7:uMARS ratings for the HAND SCANS

Reviewer ID	A. Engagement (Average score)	B.Functionality (Average score)	C.Aesthetics (average score)	D. Information quality (average score)	E. Subjective quality (Average score)	Overall App quality score (A+B+C+D)/4
1	3.6	4	3	4.2	2.5	3.7
2	4.2	5	4.3	4	2.8	4.4
3	4.2	4	4	3.7	3.8	3.97
4	2.6	4	3.3	3.5	4	3.35
5	3	3	4	4	3.3	3.5
6	3	3.5	3.3	4.5	4	3.6
7	2.8	3.5	3.7	3.8	4.8	3.45
8	2.6	4.3	3	3.5	4	3.35
9	3.4	4	4	3	3	3.6
10	4.2	3.5	3	4.2	2.8	3.72
11	3.6	3.5	4	4	2.8	3.77
12	3	4.5	3.7	3.5	3.3	3.67
13	3	4	3	4.5	4	3.62
14	2	4.5	4.3	4	3.8	3.7
15	3.2	4	4	3.3	4	3.62
16	3	4	3.3	3.8	3.3	3.52
17	2.6	3.5	2.7	4.8	4	3.4
18	2.6	4	3	4	2.8	3.4
19	2.8	4.8	4	4.5	3	4.02

20	4.2	4	3.3	3.8	3	3.82
21	3.4	4.3	4	4.8	2.5	4.12
22	2.8	4	3	4	4	3.45
23	3.8	4	3	3.3	3.3	3.53
<b>Mean (S.D)</b>	<b>3.2 (0.61)</b>	<b>3.99 (0.45)</b>	<b>3.51 (0.51)</b>	<b>3.94 (0.475)</b>	<b>3.42 (0.62)</b>	<b>3.61(0.26)</b>

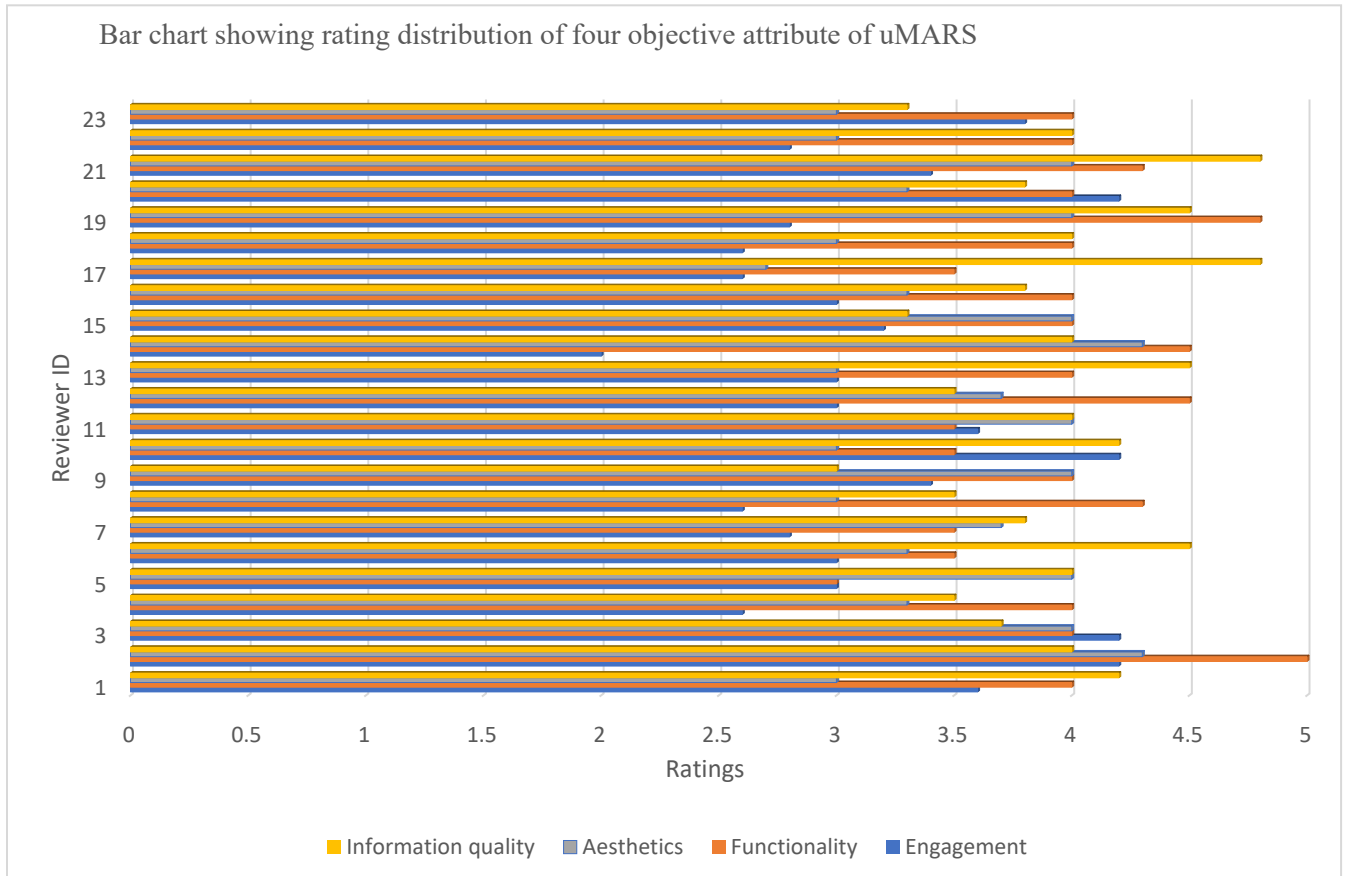


Figure 56:Rating distribution of four objective attributes of uMARS namely information quality, aesthetics, functionality, and engagement



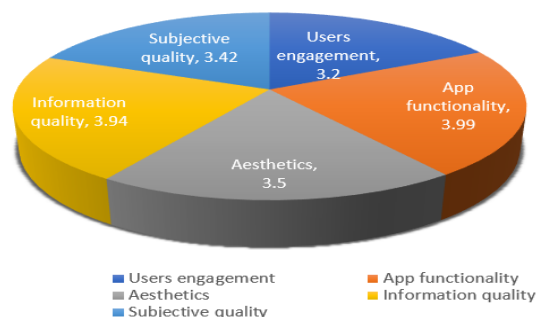


Figure 57:Pi-chart Chart showing rating distribution of all attributes of uMARS

### 5.7 Practical implications of the app

HAND SCANS app should be incorporated in clinical settings to understand pain and sensitivity level of the trauma. It will help provide timely care to patients from remote communities. The 3D model files collected through the app can be reviewed to detect any cardinal signs of infections like partial flexion of the digit at rest. Therapists can suggest medications or wrist/finger braces for urgent cases. The health data collected through the app will help monitor patients' responses to a given treatment and improve analysis of trends, patterns, and risk factors. Patients may be asked by the clinicians to use the app for the second time as a follow-up to see any improvements.. It will improve patient compliance as it will eliminate the costs and time involved in conventional physical consultation with therapists. The relevant set of questionnaires and information related to hand injury will act as a foundation for future app development for remote surveillance and assessment of hand injury. In the future, Hand Scans will be commercialized by including an additional feature to process payments. We will implement a monthly or annual subscription plan with a 2-week limited-time free trial for users. Jotforms has built-in Form Analytics, which enables developers to visualize the data traffic, how many people viewed the forms and how many responded. Figure 58 shows sorted data traffic from 22<sup>nd</sup> September to 22<sup>nd</sup> November. The graph gives us an estimation of how many times our app has been viewed over the two months. The traffic panel in the figure below shows where the visitors originate from and which device they use to interact with the app.

We can get more insights into the customer type, choices and response time which will boost further improvements of our prototype. Alternatively, we can publish our app on the Google Play store to reach a wider audience.

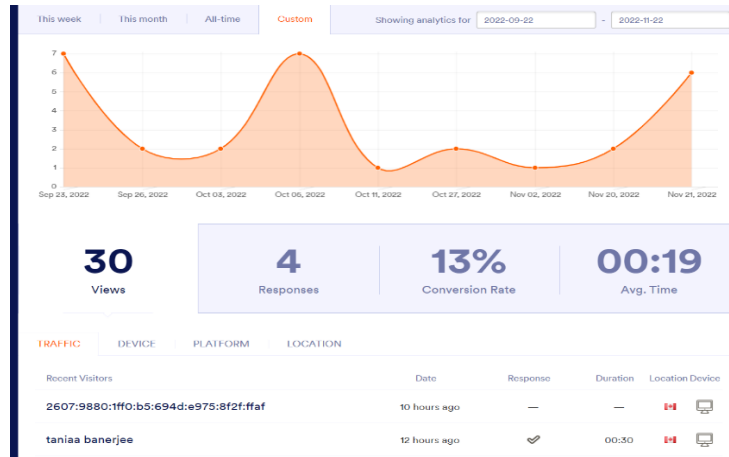


Figure 58: Visualizing data traffic using Jotform Form Analytics

## 5.8 Limitations of Hand scans App

For range of motion measurement, HAND SCANS has only a feature to upload and share hand images with therapists. However, relying solely on two-dimensional images does not provide sufficient information to fully assess the range of motion of human hand, which possess 27 degrees of freedom [70].

The app still needs extensive clinical trials with patients for establishing its content validity and acceptability among users. The app's content validity refers to the extent to which it accurately measures the health outcome compared to in-person assessments. It can only be determined through rigorous testing and validation using a diversified data sample. Similarly, acceptability also needs to be evaluated to ensure that patients prefer the application for injury assessment and that it integrates well into the existing healthcare workflows.

## Chapter 6

### 6. 3D Hand Tracking In Real-Time

#### 6.1 MediaPipe

Google MediaPipe technology, launched in 2019, is an open-source framework that provides solutions for computer vision technologies such as object detection, hand detection, hand tracking, pose estimation and face detection. The solutions work on Web, Desktop PCs, Android, and iOS platforms. MediaPipe Hands pipeline can predict 21 3D key points from a single RGB camera frame (Zhang et al. 2021) (refer to figure 59). We will explore the potential of MediaPipe algorithms to obtain hand coordinates from a live video stream.

MediaPipe Hands employs machine learning algorithms to track multiple hands in real-time. The pipeline consists of two neural network pipelines working together, a palm detection model and a hand landmark model [71] to predict 2D and 3D hand landmarks on an image or video sequence. The images below show how hand and finger detection work through the MediaPipe algorithm.

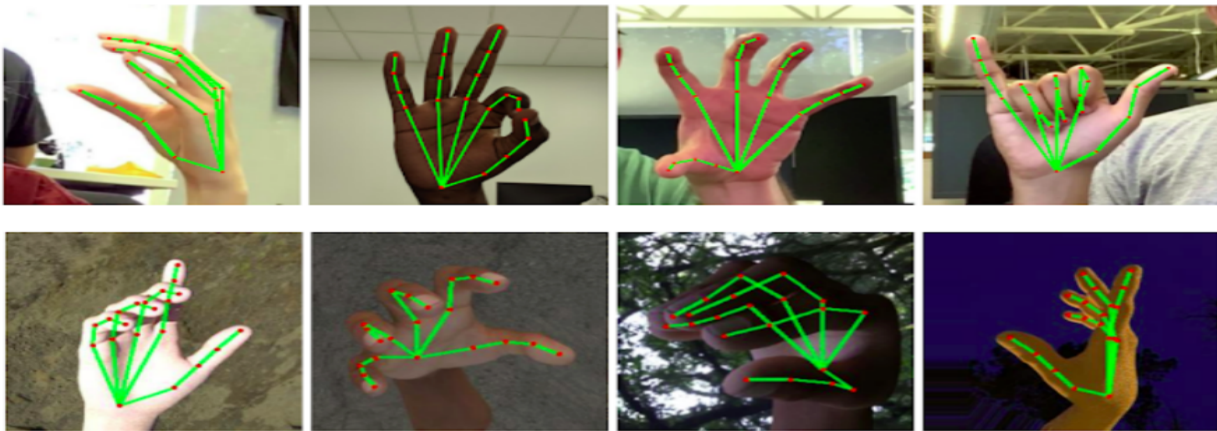


Figure 59: Detection of real hands(top) and synthetic hands (bottom) with Mediapipe  
 [Picture courtesy: Google. Home – <https://google.github.io/mediapipe/solutions/hands>]

##### 6.1.1 Palm Detector

Detection of hands is a complex task compared to others given the variety of hand sizes and absence of apparent distinguishing features. The contrast patterns and shapes around the eyes, noses and mouth region can be used as distinct features for face detection. However, such features

are missing in the case of hand-detection tasks. As a result, a different strategy is adopted; to detect and recognize a palm in real-time, they use a single-shot detector (SSD) model [72], a feed-forward convolutional neural network for detection and training methodology for image recognition. A palm detector is trained for estimating bounding boxes around rigid objects like the palm and fist (as shown in figure 60). Then, encoder-decoder feature extractor is used for larger scene context awareness and finally focal loss is minimized to support large-scale variance [73]. According to MediaPipe documentation, with the above techniques, they could achieve an average precision of 95.7% in palm detection. Using a regular cross entropy loss and no decoder provides a baseline of just 86.22% [74].

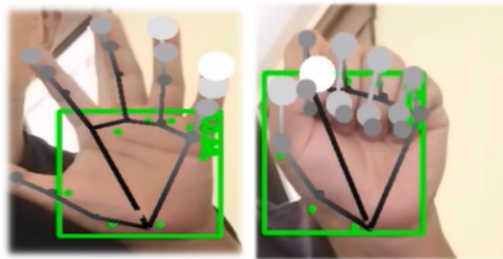


Figure 60:MediaPipe Hands Palm detection model

### 6.1.2 Hand Landmark model

After completion of the palm detection process, our subsequent hand landmark model performs precise landmark localization of 21 2.5D coordinates inside the detected hand areas (as shown in figure 61) via regression modeling technique. MediaPipe data has been trained on partially visible hands as well as self-occlusions, hence the model is a consistent in tracking hand postures. The algorithm weighs the probability of hand presence and the handedness to know whether left or right hand are present in the image frame. All 21 landmarks of a detected hand with their unique ID s have been shown in figure 62.

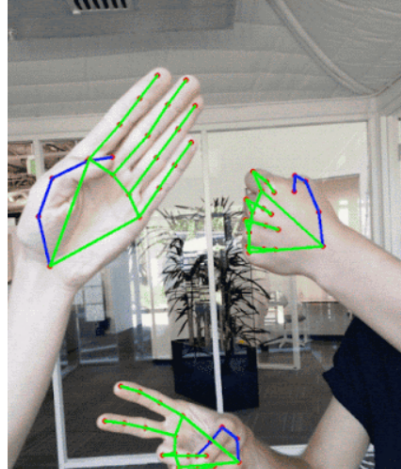


Figure 61: Hands detected and pose estimation  
 [Picture courtesy: <https://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html>]

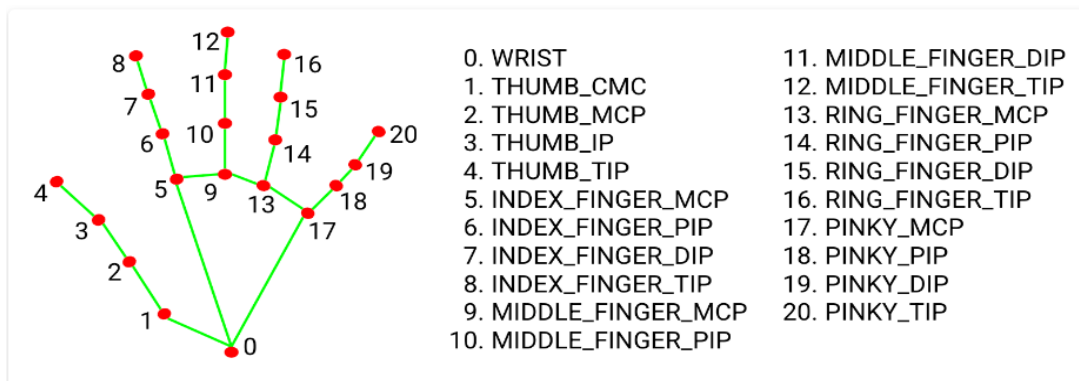


Figure 62: Hand Landmark model showing labels for each Landmark ID

## 6.2 MediaPipe Hands Coordinate system

In this section, we build our hand detection model using a single right hand input image and MediaPipe as the base library, draw the landmarks on the output image, print the coordinates of all the 21 landmarks with timestamps detected in the input image frame and plot all the coordinates in 3D coordinate system. This will help us to understand the orientation of the axes and the relationships between points in a 3D space.

Input image:



Figure 63: Image of a right hand used for the experiment

Pseudo code for hand landmark detection:

Arguments used in the code are as follows.

- `STATIC_IMAGE_MODE`: If the input is single image, we set this to true, otherwise we set false.
- `MAX_NUM_HANDS`: Maximum number of hands in frame, the default value is 2.
- `MIN_DETECTION_CONFIDENCE`: Detections confidence which we will set to 0.5, which means that if the confidence level drops below 50% then the hands will not be detected at all.
- `MIN_TRACKING_CONFIDENCE`: If tracking frames, then tracking confidence in each frame.
- `MULTI_HAND_LANDMARKS`: Detection or tracked landmarks as a list.
- `MULTI_HAND_WORLD_LANDMARKS`: Maps into real-world 3D coordinates.
- `MULTI_HANDEDNESS`: This detects the hand as a left or right hand with a score.

```

with mp_hands.Hands(
    static_image_mode=True,
    max_num_hands=1,min_detection_confidence=0.5,min_tracking_confidence=0.5) as hands:
for name, image in images.items():

    results = hands.process(cv2.cvtColor(image, cv2.COLOR_BGR2RGB))

    # Print handedness (left v.s. right hand).
    print(f'Handedness of {name}:')
    print(results.multi_handedness)

    if not results.multi_hand_landmarks:
        continue
    # Draw hand landmarks of hand.
    print(f'Hand coordinates of {name} are below:')
    image_height, image_width, _ = image.shape
    annotated_image = image.copy()
    for hand_landmarks in results.multi_hand_landmarks:
        for id,lm in enumerate(hand_landmarks.landmark):
            print("\nTime stamp", time.strftime('%H:%M:%S'))
            print("Landmark", id)
            print(f"x: {lm.x}",f"y: {lm.y}",f"z: {lm.z}")
            mp_drawing.draw_landmarks(
                annotated_image,
                hand_landmarks,
                mp_hands.HAND_CONNECTIONS,
                mp_drawing.DrawingSpec(color=(121, 22, 76), thickness=2, circle_radius=4),
                mp_drawing.DrawingSpec(color=(250, 44, 250), thickness=2, circle_radius=2), )

resize_and_show(annotated_image)

```

Figure 64: Image of a right hand used for the experiment

Output:

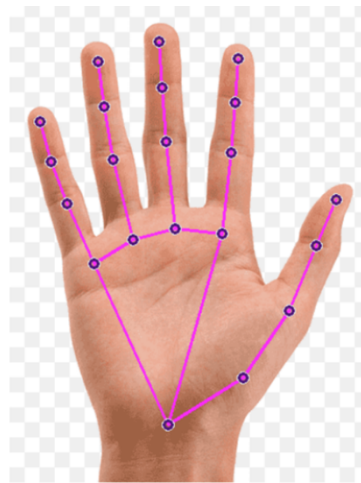


Figure 65: Output image showing detected landmarks and hand connections

Each detected landmark is composed of x, y and z coordinates. MediaPipe Hands uses a local coordinate system where x and y are normalized to [0,1] by the image width and height

respectively.  $z$  represents the landmark depth (distance of detected object from camera), with the depth at the wrist being the origin, and the value decreases as one move closer to the camera [74]. Figure 66, below shows the landmark coordinates obtained for 21 landmarks. A careful observation of the results shows that the value of both  $x$  and  $y$  coordinates lies between 0 and 1.

Time stamp 22:45:06 Landmark 0 x: 0.44150209426879883 y: 0.8796089887619019 z: 1.1035868965336704e-06	Time stamp 22:45:06 Landmark 8 x: 0.6358009576797485 y: 0.11747828125953674 z: -0.1759696900844574
Time stamp 22:45:06 Landmark 1 x: 0.6489231586456299 y: 0.7819097638130188 z: -0.0775880292057991	Time stamp 22:45:06 Landmark 9 x: 0.4588448405265808 y: 0.47180190682411194 z: -0.04668908566236496
Time stamp 22:45:06 Landmark 2 x: 0.7759615182876587 y: 0.6417551636695862 z: -0.09917569905519485	Time stamp 22:45:06 Landmark 10 x: 0.43424683809280396 y: 0.2910120487213135 z: -0.09749007225036621
Time stamp 22:45:06 Landmark 3 x: 0.8506774306297302 y: 0.5067577958106995 z: -0.11796978861093521	Time stamp 22:45:06 Landmark 11 x: 0.4237833023071289 y: 0.17729699611663818 z: -0.14941294491291046
Time stamp 22:45:06 Landmark 4 x: 0.9078612923622131 y: 0.41083183884620667 z: -0.13776840269565582	Time stamp 22:45:06 Landmark 12 x: 0.41488784551620483 y: 0.08196043968200684 z: -0.18864087760448456
Time stamp 22:45:06 Landmark 5 x: 0.5895074009895325 y: 0.48153290152549744 z: -0.04355768859386444	Time stamp 22:45:06 Landmark 13 x: 0.3426052927970886 y: 0.4950079917907715 z: -0.061665162444114685
Time stamp 22:45:06 Landmark 6 x: 0.6142444610595703 y: 0.3134039640426636 z: -0.09409549832344055	Time stamp 22:45:06 Landmark 14 x: 0.28779712319374084 y: 0.32711824774742126 z: -0.11970046162605286
Time stamp 22:45:06 Landmark 7 x: 0.6264801621437073 y: 0.20893767476081848 z: -0.13972057402133942	Time stamp 22:45:06 Landmark 15 x: 0.26196080446243286 y: 0.2179761826992035 z: -0.16790048778057098
Time stamp 22:45:06 Landmark 16 x: 0.24472251534461975 y: 0.12419414520263672 z: -0.20092488825321198	
Time stamp 22:45:06 Landmark 17 x: 0.23542839288711548 y: 0.5456691384315491 z: -0.08423905819654465	
Time stamp 22:45:06 Landmark 18 x: 0.15922433137893677 y: 0.41956204175949097 z: -0.14372938871383667	
Time stamp 22:45:06 Landmark 19 x: 0.11628851294517517 y: 0.3324128985404968 z: -0.17685620486736298	
Time stamp 22:45:06 Landmark 20 x: 0.08552321791648865 y: 0.24887539446353912 z: -0.19781550765037537	

Figure 66: Showing coordinates for all landmarks (0-20) with timestamps

The local coordinate system used by Mediapipe is different from the conventional 3D cartesian coordinates system. Here, the origin (0, 0) is located at the top-left corner of the image frame, with positive  $x$ -axis extending to the right and positive  $y$ -axis increasing downwards. This type of



coordinate system is often commonly used in computer graphics and image processing applications.

The results from figure 66 confirms this conjecture, landmark 20 (small finger tip) is located more towards the left side of the image frame (see figure 65). The x coordinate of landmark 20 is close to 0, while landmark 4 (Thumb tip) which lies towards the right side of the image frame has x coordinate value close to 1 (see results from figure 66). Likewise, with respect to the y axis, landmark 20 is situated towards the upper part of the frame, whereas landmark 0 (wrist) is situated towards the bottom. Y value increases as one moves from top to bottom of the frame, landmark 20 has a y-value close to 0 while landmark 0 has a y-value close to 1 (see results from figure 66).

Figure 67 shows the position of all the coordinates when mapped in 3D. The negative sign along the y axis is only an indication of a reversed or flipped y axis. Figure 68 is the diagram for the conventions used in this default coordinate system.

```
#plotting local coordinates
mp_drawing.plot_landmarks(hand_landmarks, mp_hands.HAND_CONNECTIONS, azimuth=2)
```

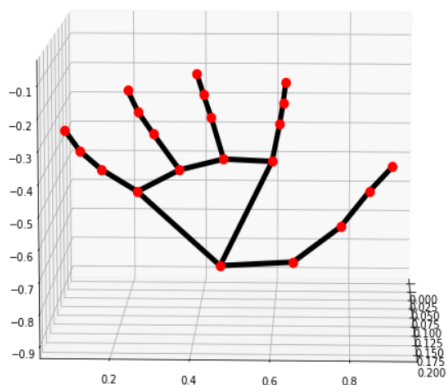


Figure 67: Code snippet and map plot of 21 landmarks in default(local) coordinate system

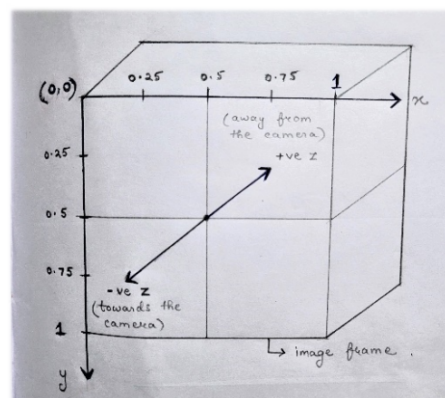


Figure 68: Diagram of default (local) coordinate system of MediaPipe showing the x, y and z axis.

The x, y and z coordinates determined in the previous step can be transformed into real-world 3D coordinates [75]. In world coordinate system, the x, y and z values are in meters. Figure 69 shows the results after mapping the resultant landmark coordinates in 3D world coordinate system. From

the graph we can determine that the origin (0,0,0) is at the hand's approximate geometric center (close to landmark 9). Figure 70 provides clarifications on the conventions used to represent points in world coordinate system.

```
#plotting world coordinates
mp_drawing.plot_landmarks( hand_world_landmarks,mp_hands.HAND_CONNECTIONS,azimuth=2)
```

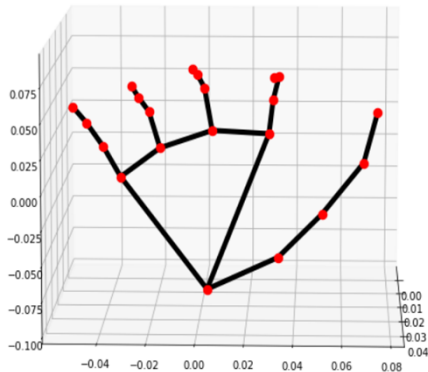


Figure 70:Code snippet and map plot of 21 landmarks in world coordinate system

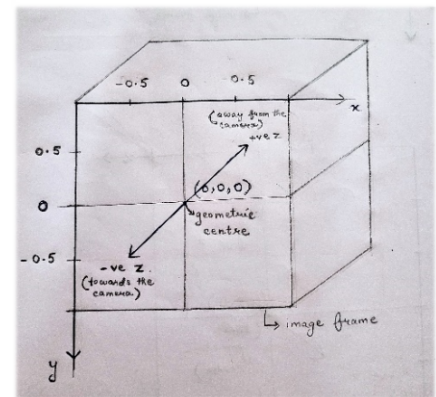


Figure 69:Diagram of 3D world coordinate system showing x,y and z axis

### 6.3 Real-time angle measurement

We modified the MediaPipe Hands algorithm to calculate angles between a list of detected landmarks from a live web stream.

Pseudo Code: Function for calculating angles between index finger joints

```
def draw_finger_angles(image, results, joint_list):
    # Loop through hands
    for hand in results.multi_hand_landmarks:
        # Loop through joint sets
        for joint in joint_list:
            a = np.array([hand.landmark[joint[0]].x, hand.landmark[joint[0]].y]) # First coord
            b = np.array([hand.landmark[joint[1]].x, hand.landmark[joint[1]].y]) # Second coord
            c = np.array([hand.landmark[joint[2]].x, hand.landmark[joint[2]].y]) # Third coord

            radians = np.arctan2(c[1] - b[1], c[0] - b[0]) - np.arctan2(a[1] - b[1], a[0] - b[0])
            angle = np.abs(radians * 180.0 / np.pi)

            if angle > 180.0:
                angle = 360 - angle

            cv2.putText(image, str(round(angle, 2)),tuple(np.multiply(b, [640, 480]).astype(int)),
                        cv2.FONT_HERSHEY_SIMPLEX, 0.5, (255, 255, 255), 2, cv2.LINE_AA)

    return image
```

Full code repository in GitHub: <https://github.com/tbanerj2/Mediapipe-hand-tracking.git>

Output

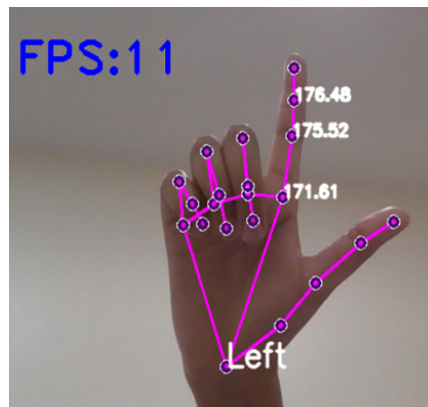


Figure 71: Showing hand classification and angles between joints of the index finger

## 6.4 DIGITS

The MediaPipe Hands machine learning pipeline is modified and developed into a stand-alone web application named DIGITS. It is coded in JavaScript and React libraries. The landmark coordinates and timestamps logged in real-time are collected into a comma-separated values (CSV) text file. The interface shows the real-time coordinates of the wrist joint, offers a feature to reset the camera resolution and provides options to download coordinates of all the landmarks in a .csv files and visualize joint angles in the web console. The user interface of DIGITS is displayed in figure 72.

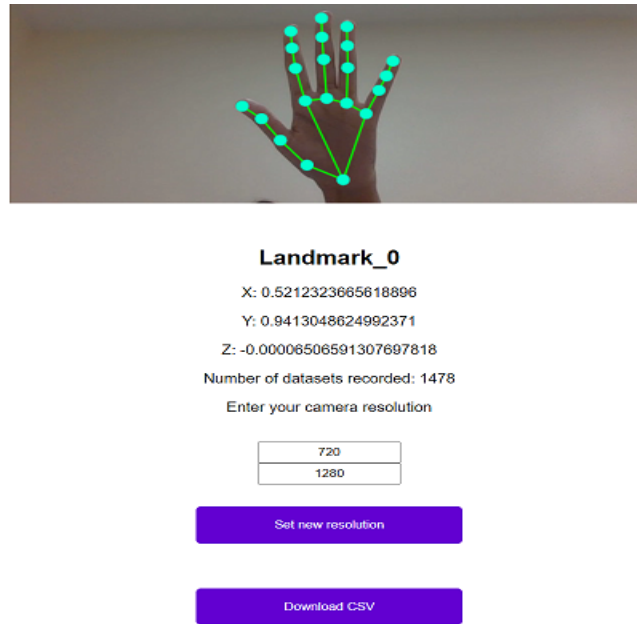


Figure 72: DIGITS user interface

DIGITS webapp url: <https://handtracker-af91b.web.app/hand-tracker>

Copy of DIGITS app code repository: <https://github.com/HerbertShin/Hand-Tracking-React-App.git>

#### 6.4.1 Input variability of DIGITS

DIGITS extracts the x,y, and z coordinates of 21 landmarks of hands. We attempted to find the mean and dispersion (1 standard deviation) of the spectrum of the resting hand (dominant hand) recorded by DIGITS. To test the variability of distances along the x, y and z axis, the application is evaluated with 3 users (between 20 and 30 years of age), who don't have a medical history of hand tremor disease. The distance between the camera and the hand is between 10-20 inches and the illumination is bright. The right hand is fixed vertically in an extension position while the wrist is supported on a table. The average data capture rate is 15 frames per second. Data was sub-sampled with a sampling time of 20 secs for quantification and visualization. The total number of records from a single user can be approximately 300 samples. We average the x y and z coordinates for each landmark obtained from 3 users to understand the variability of the data produced by DIGITS. Figure 73 demonstrates the variation of x, y, and z coordinates of the wrist joint with time, and the box-whiskers plot in figure 74 illustrates the variation range (mean, median and standard deviation). Table 8 displays the descriptive statistics of a 20 secs input registration sample

for 0 landmark (averaged value from 3 users), detected by DIGITS. Appendix 2 shows all the statistics for other 20 landmarks.

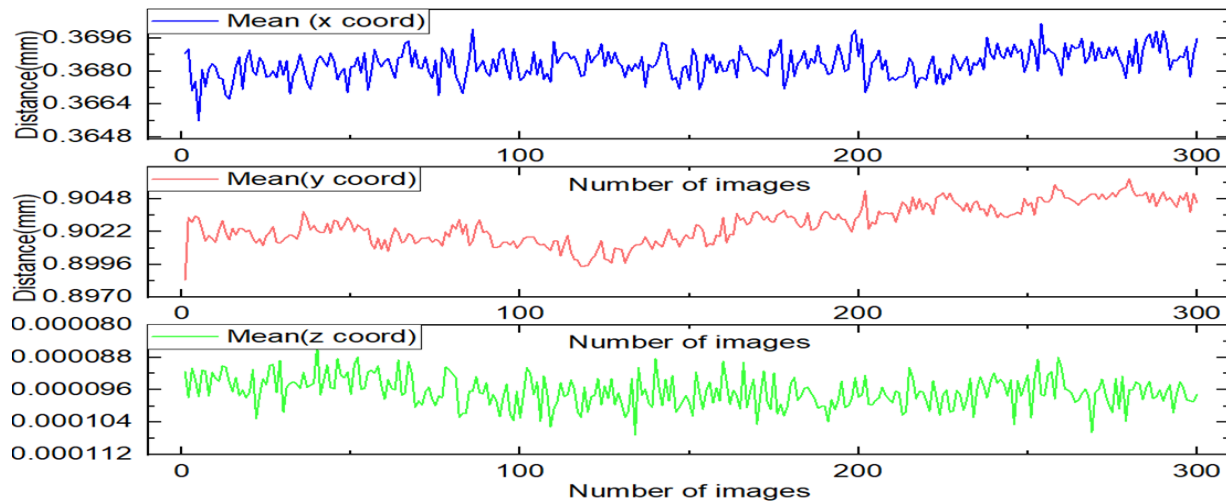


Figure 73: Stacked line graph showing variation of the wrist joint (landmark 0) along x (top) , y (middle) and z (bottom) coordinates.

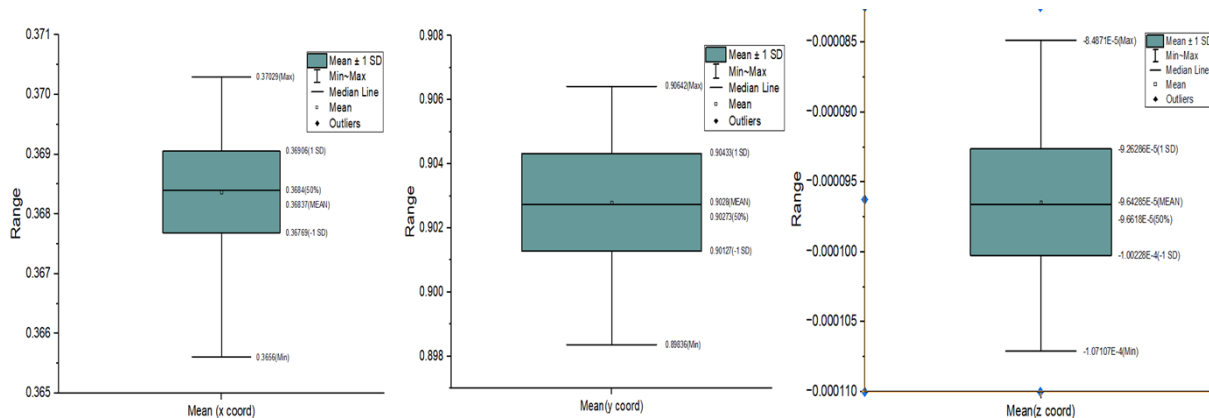


Figure 74: Box-Whiskers plot to show variation range of wrist joint (Landmark 0) along a) x , b) y and c) z coordinates

Table 8: Showing the mean, standard deviation , minimum and maximum value for 0 landmark.

	Landmark 0		
	$\bar{X}$	$\bar{Y}$	$\bar{Z}$
<u>Mean (mm)</u>	0.368	0.9028	-0.000096
<u>S.D (mm)</u>	6.85E-4	0.00153	0.0000038
<u>Min (mm)</u>	0.3656	0.8936	-0.000107
<u>Median (mm)</u>	0.3684	0.90273	-0.000097
<u>Max (mm)</u>	0.3703	0.9062	-0.000085

### 6.4.2 Inference

The box plots in figure 75.1, 75.2, 75.3 show the statistical measures of x, y and z coordinates of all the 21 landmarks while the hand is at rest for 20 secs.

The variation along z axis is greatest, with the larger standard variation obtained by inspecting 8<sup>th</sup> landmark (S.D= 0.00668) followed by 12<sup>th</sup>(S.D= 0.00638) , 7<sup>th</sup>(S.D= 0.00637 ) , 16<sup>th</sup>(S.D= 0.00631) , 4<sup>th</sup>(S.D= 0.0059) and 11<sup>th</sup>(S.D= 0.0057),6<sup>th</sup>( S.D= 0.00557) and 15<sup>th</sup> ( S.D= 0.00547) landmark in descending order.

Along y axis the order is as follows: 4<sup>th</sup> landmark (S.D= 0.0052) >8<sup>th</sup> landmark(S.D= 0.00391)>16<sup>th</sup> landmark(S.D= 0.00378)>15<sup>th</sup> landmark(S.D= 0.00376)>7<sup>th</sup> landmark(S.D= 0.00366).>20<sup>th</sup> landmark(S.D= 0.00365).

Along x axis the order is as follows:8 landmark(S.D=0.00459)>12<sup>th</sup> landmark(S.D= 0.00441)>16<sup>th</sup> landmark(S.D=0.00395)>7<sup>th</sup> landmark(S.D=0.00394)>11<sup>th</sup> landmark (S.D=0.00378)>20<sup>th</sup> landmark(S.D= 0.00344)>6<sup>th</sup> landmark(S.D= 0.00337)>15<sup>th</sup> landmark(S.D= 0.00325).

Our results tally with the box plots below. The size of the box plots for 8<sup>th</sup> ,12<sup>th</sup>,7<sup>th</sup> and 16<sup>th</sup> landmark are relatively larger, showing larger variation of the data. The least variation is seen in landmark 0 and landmark 1 along all the three axes, with S.D (0.00068, 0.00153, 0.0000037) and S.D(0.001044, 0.00212,0.00126) respectively and can be verified by the box plots below.

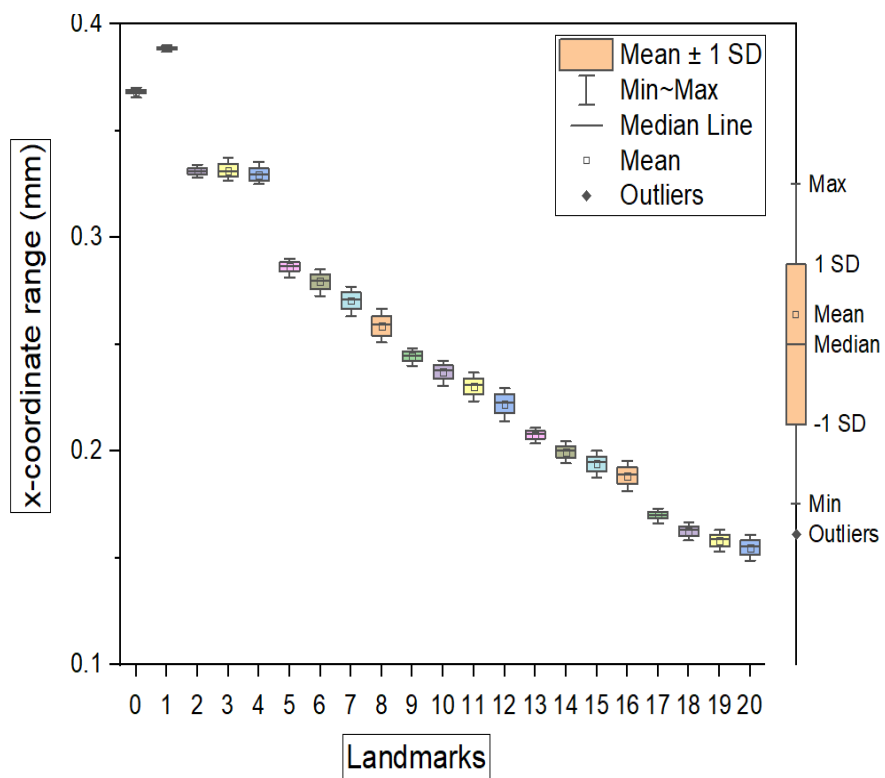


Figure 75.1: x coordinates mapping for 21 landmarks

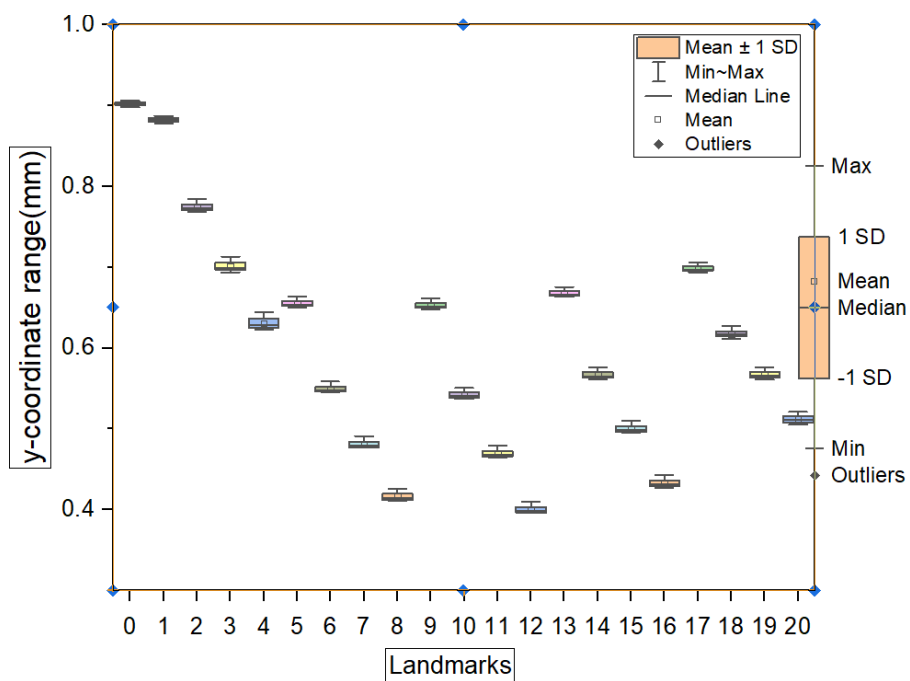


Figure 75.2: y coordinates mapping for 21 landmarks

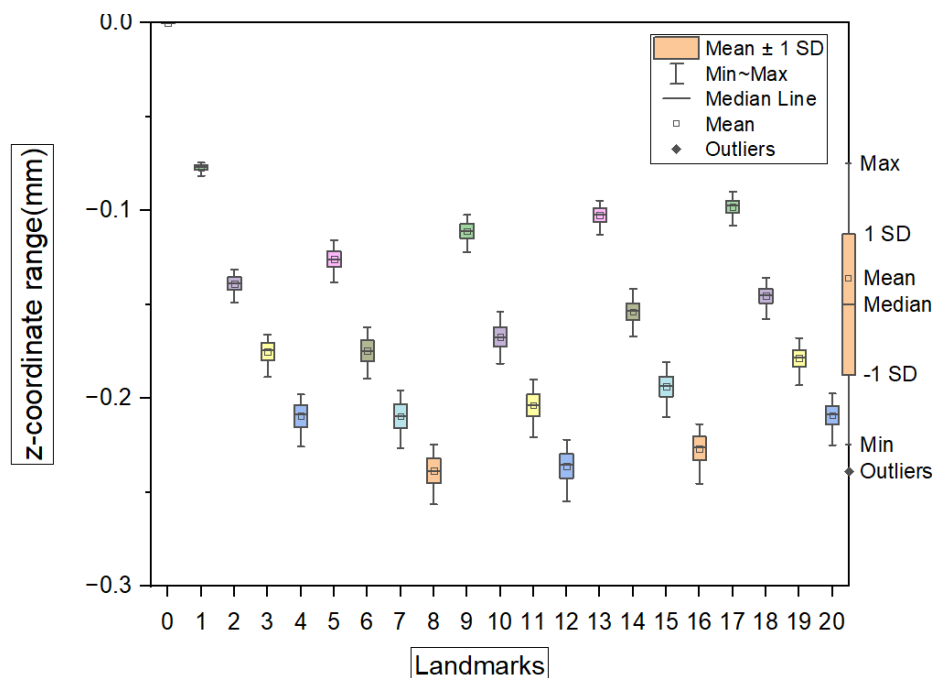


Figure 75.3: z coordinates mapping for 21 landmarks

### 6.5 Range of motion (ROM) evaluation of finger joints using DIGITS

A clinical study is carried out to validate DIGITS as a stable, reliable, efficient, and accurate method of remote assessment tool for finger joint range of motion. The vectors between each hand landmark are used to calculate the angle between each finger and hand segment, yielding the corresponding joint angle. Given two adjacent segments (landmarks)  $x_1$  and  $x_2$ , the angle  $\theta$  was calculated as:

$$\theta = \arccos\left(\frac{\vec{x}_1 \cdot \vec{x}_2}{|\vec{x}_1| |\vec{x}_2|}\right)$$

.....Equation 1

In this study, we perform real-time on-device measurements of our range of motion in the small joints of the hand namely metacarpophalangeal Joints (MCP), proximal interphalangeal joints (PIP) and distal interphalangeal joints (DIP), for flexion and extension arcs for the index, middle, ring and small finger. We compare the results obtained digitally from DIGITS and in-person measurements obtained from a certified hand therapist. We determine the reliability and validity of the DIGITS application by computing intraclass correlation coefficient (ICC) analysis [76,77]



to assess whether this application can accurately measure finger range of motion just as effectively as in-person goniometric measurements.

### 6.5.1 Methodology

Manual goniometry measurements for the range of motion (ROM) for flexion and extension are assessed by a certified hand therapist during three separate sessions with a goniometer using a scale range of 0–180 degrees. The goniometer is placed along the dorsal surface of each affected joint, and measurements from MCP, PIP and DIP joints in the right hand of a dominant right-hand male are recorded. All assessments have taken place at the paediatric or adult plastics clinic at LHSC Victoria Hospital and the Hand and Upper Limb Clinic (HULC) at St. Joseph’s Health Care.

The landmarks of each finger joint of the hand are identified and digitally tracked in real-time by DIGITS, and the angles between adjacent segments are calculated to estimate the range of motion endpoints across the MCP, PIP, and DIP finger joints of the hand. The distance between the camera and the object is between 10-20 inches and the light setting is bright. The datasets are sampled at a frequency of 15 per second, and the mean sampling time for each data set is 30 seconds. We collected 54 individual datasets for both flexion and extension for the right hand and 9 sets each for flexion and extension for the left hand. Values from all 126 datasets are averaged. The collected data and patient consent forms are stored in Lawson REDCap. Measurements by certified hand therapists and DIGITS of the flexion and extension angles are recorded in tables 9 and 10 below.

Table 9: The average finger joint angles obtained from in-person goniometry measurement of both hands (unit in degrees)

	Index MCP	Index PIP	Index DIP	Middle MCP	Middle PIP	Middle DIP	Ring MCP	Ring PIP	Ring DIP	Small MCP	Small PIP	Small DIP
Right Flexion	87.0	101.0	81.7	89.3	102.3	82.0	87.7	103.7	80.0	94.7	93.0	80.3
Right Extension	23.0	26.3	0.0	0.0	33.0	0.0	0.0	29.7	0.0	18.7	15.7	2.7
Left flexion	86.0	100.0	84.0	80.0	101.0	87.0	85.0	105.0	71.0	90.0	90.0	81.0
Left Extension	8.0	6.0	5.0	14.0	15.0	5.0	13.0	14.0	0.0	18.0	1.0	20.0

Table 10: The average finger joint angles obtained from DIGITS (unit in degrees)

	Index MCP	Index PIP	Index DIP	Middle MCP	Middle PIP	Middle DIP	Ring MCP	Ring PIP	Ring DIP	Small MCP	Small PIP	Small PIP
Right Flexion	85.37	100.06	68.58	90.38	98.52	79.82	87.49	102.39	77.09	87.10	91.76	77.69
Right Extension	13.65	11.04	6.25	7.69	23.88	9.79	2.62	16.72	9.27	21.17	12.84	7.63
Left flexion	86.80	101.82	83.17	92.21	103.0	81.2	88	101.30	76.57	88.07	96.13	77.31
Left Extension	13.95	8.68	4.69	20.6	11.22	6.74	23.44	13.39	5.56	22.71	6.68	14.15

### 6.5.2 Statistical analysis

ICCs are obtained between the 2 assessments (in-person and using DIGITS) and are calculated for both flexion and extension data sets separately. The ICC value below 0.5 indicates poor reliability, between 0.5 and 0.75 indicates moderate reliability, between 0.75 and 0.9 indicates good reliability, and any value above 0.9 indicates excellent reliability [77]. Statistical analysis was conducted using the SPSS Statistics software, version 29 (IBM Corp., Armonk, New York). The ICC model we use here is the two-way random-effects, absolute agreement, single rater/measurement [ICC<sub>(2,1)</sub>] [77] with 95% confidence intervals (CI) and alpha being 0.05.

We compare the two different measurements for each finger joint. ICC can be calculated using the formula:

$$ICC = \frac{k\sigma_m^2 - \sigma^2}{(k-1)\sigma^2} \dots\dots\dots \text{Equation 2}$$

(Picture courtesy: <https://allthingsstatistics.com/inferential-statistics/intraclass-correlation/>)

where  $\sigma^2$  denotes the variance of all data values and  $\sigma_m^2$  denotes the variance of the means of the group which has a size k. We calculated ICC in SPSS software by following these commands: Analyze>Scale>Reliability Analysis>Intraclass Correlation Coefficient. The model selected is two-way random effects model with 95% confidence intervals and 0.5 alpha value.

Values closer to +1 indicate higher positive intraclass correlation whereas a value close to 0 implies that the data points within that group are independent of each other.

### 6.5.3 Evaluated Results

The average ICC measure shows excellent reliability of our results for flexion (0.958 for Right hand and 0.925 for left hand) and good reliability for extension (0.767 for Right hand, 0.857 for Left hand), as can be verified from the results from Table 11.

Table 11: Intraclass correlation coefficients (ICC) for the reliability of MCP, PIP, and DIP joint angle measurement for both right and left-hand

#### Right hand flexion

	Intraclass Correlation Coefficient						
	Intraclass Correlation <sup>b</sup>	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.919 <sup>a</sup>	.743	.976	23.577	11	11	<.001
Average Measures	.958	.853	.988	23.577	11	11	<.001

Two-way random effects model where both people effects and measures effects are random.

#### Right hand extension

	Intraclass Correlation Coefficient						
	Intraclass Correlation <sup>b</sup>	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.622 <sup>a</sup>	.105	.874	4.293	11	11	.012
Average Measures	.767	.191	.933	4.293	11	11	.012

Two-way random effects model where both people effects and measures effects are random.

#### Left-hand flexion

	Intraclass Correlation Coefficient						
	Intraclass Correlation <sup>b</sup>	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.860 <sup>a</sup>	.586	.958	13.307	11	11	<.001
Average Measures	.925	.739	.978	13.307	11	11	<.001

Two-way random effects model where both people effects and measures effects are random.

#### Left-hand extension

	Intraclass Correlation Coefficient						
	Intraclass Correlation <sup>b</sup>	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.749 <sup>a</sup>	.335	.921	6.976	11	11	.002
Average Measures	.857	.502	.959	6.976	11	11	.002

Two-way random effects model where both people effects and measures effects are random.

## 6.6 Discussion

DIGITS provides precision and reliability for monitoring hand movements. It is advantageous to clinicians for capturing key points from hand regions and finger joint angles for assessing finger range of motion. It can be used remotely with a standard 2D camera rather than specialist depth attached to a smartphone, tablet or laptop. To further develop and improve this application, it will be necessary to evaluate our system in a larger population, across all ages and genders. Our future endeavor is to take similar measurements from patients with hand pathologies such as trauma and arthritis using DIGITS. The implementation of this intervention will be studied in terms of its effect on the speed of post-trauma recovery of the full range of motion. Future research can also focus on many other uses in clinical settings like detecting tremors, for example Parkinson's, dystonic and Essential tremors. Before this system becomes a clinically useful tool, it will also be important to investigate whether variables such as ambient lighting, camera resolution and skin color affect the results.

## 7. Research Contribution

In the challenging field of hand telerehabilitation, photogrammetry 3D models and augmented reality have not yet been used for remote clinical diagnosis. This thesis presents a foundational work to highlight the capability of photogrammetry to produce 3D models of hands from videos and images captured through a smartphone camera that can be shared with a therapist over the internet and visualized remotely, without requiring any costly equipment. Our main contribution is testing photogrammetry as a possible method to reconstruct a 3D hand using only smartphone images. Our study provides insights into the performance of four open-source photogrammetry applications, namely Meshroom, Agisoft Metashape, 3DF Zephyr and Reality Capture, for close-range photogrammetry in an indoor environment. We find that Reality Capture produces the best textured and most accurate 3D models in terms of C2C signed distance computation with lowest

errors in minimum time. However, photogrammetry models from smartphone images have lower resolution compared to commercial scanners. Also, the 3D reconstruction processes are computationally demanding and require specific hardware configurations. Therefore, our work has opened up opportunities for future research to address these limitations and develop techniques for effectively utilizing photogrammetry for remote 3D hand reconstruction.

Our second objective is to develop digital platforms to collect patient data to facilitate remote treatment of the upper limb injuries. To achieve this goal, we developed a stand-alone app called HAND SCANS using a low-code app development platform, Jotforms. The app contains a meaningful question set and 3D rendering feature using photogrammetry technique. It also includes self-reporting and self-assessing features and allows patients to share the assessment score, images and 3D models of their injured hands with a consulting hand therapist for diagnosis. A survey conducted using uMARS evaluated the usability of the mobile version of the app and showed its effectiveness from the healthcare professionals' perspective. The survey results validate HAND SCANS as a quality eHealth tool containing useful content for remote surveillance of patients with hand injuries. It can be further improved by improving the user interface and adding more interactive and engaging features. However, its limitation lies in its usage of 2D images for evaluation hand range of motion which encourage further research into possible techniques for hand tracking in 3D using 2D images.

We conducted an investigation into the capabilities of MediaPipe, a newly released and novel Machine Learning pipeline, to extract 3D landmarks and key features from hand images. Our study explores the potential of applications of MediaPipe in clinical studies for hand therapy. To gain a deep understanding of MediaPipe's coordinate orientations, we carry out an extensive study. Using MediaPipe Hands algorithm, we create a custom application named DIGITS for hand telerehabilitation. DIGITS is a web application developed using JavaScript and React libraries. It can track hand keypoints in 3D space, generate log files containing the x,y,z coordinates for each of the 21 landmarks, as well as measure the angles between finger joints which could be visualized in the console. Range of variability of input data detected by DIGITS is very small with only 8th, 12th, 7th and 16th landmark show slightly larger variations in ideal condition (hand at rest in extension position) compared to other landmarks. The least variation is seen in landmark 0 and landmark 1 along all the three axes.

We conducted a clinical study to validate DIGITS as a reliable tool to measure finger range of motion. For this study, angles between MCP, PIP and DIP joints, for flexion and extension arcs for the index, middle, ring and small finger of patient's hands are measured. We take in-person measurements with goniometer, compare them with DIGITS measurements. We compute ICC values between these two sets of data, to assess the reliability and consistency of measurement across observers. The study proves that DIGITS has excellent reliability for flexion (0.958 for Right hand, 0.925 for left hand) and good reliability for extension (0.767 for Right hand, 0.857 for Left hand). DIGITS application proves to be a robust, reliable augmented reality app that can be used for remote assessment of finger range of motion. The development of HAND SCANS and DIGITS along with the findings of this research, represent a significant contribution towards improving telerehabilitation of upper limb. These tools are useful for healthcare professionals and patients alike, paving the way for more accessible and cost-effective hand telerehabilitation solutions.

## 8.Future Research Scope

### 8.1 Use of digital 3D hand models for virtual fabrication of hand orthosis

3D hand models can be further utilized for virtual fitting of hand splints and 3D printing of orthosis in the future. In recent years, designing customized hand orthosis has been possible using Computer-Aided Design (CAD) software [78]. This research provides the motivation to look at photogrammetry 3D models along with CAD integration for developing remote orthotic fabrication platforms for the upper limb that will be clinically usable and approved. The idea of design and development of a customized fabrication of hand splint using 3D hand models can form valuable aspects of hand telerehabilitation method. This method will improve efficiency and reduce costs of upper limb orthosis treatments.

## 8.2 Research on technology to improve resolution of photogrammetry 3D models

New research techniques can evolve around converting the low resolution of 3D photogrammetry models to high-resolution 3D models with impressive accuracy. The techniques may involve optimization of different external parameters, adding more vertices to the mesh structure to create a more complex structure ,surface details and intricate features.

## 8.3: DIGITS for tremor detection in hands

DIGITS has the ability to provide objective and quantitative measurements of hand movements in real time. Our experiment revolves around finding its input variability in healthy patients in ideal conditions. Further experiments can be performed to incorporate a threshold that detects any type of tremors in hand. Healthcare providers can use the app to objectively measure and monitor tremors in patients, which can aid in diagnosis, treatment and ongoing care.

## 8.4 Combining DIGITS and Hand Scans app

Hand Scans and DIGITS apps can be combined and developed into a comprehensive stand-alone mobile app for data collection that is compatible with both Android and iOS platforms. It can be collectively used to record patient medical history, hand coordinates, joint angles and 3D models of the hand. However, secured encryption of the data needs to be incorporated to maintain confidentiality.

# 9. Conclusion

The rapid advancements in smartphone technology and computer vision have captivated the world. Additionally, the evolution of open-source photogrammetric tools has resulted in easy 3D reconstruction solutions. Together, these advancements have opened up a vast array of possibilities for photogrammetric 3D reconstruction from mobile phone images. In our research, we extend the application of 3D rendering and augmented reality to hand therapy and telerehabilitation .This thesis concludes that photogrammetry, using smartphones, has the potential to produce 3D models of hands that can be shared with a therapist over the internet for remote diagnosis of hand injuries. Reality Capture was found to be the best application for close-range photogrammetry, although it still produced lower resolution models compared to commercial scanners.

During the COVID-19 pandemic, professionals across various fields had to rely heavily on digital solutions for virtual interactions. Consequently, the field of teleconsultation for hand rehabilitation also required the development of applications to connect patients and hand therapists virtually. To meet this need, we created HAND SCANS, a cross-platform app that includes a comprehensive questionnaire meaningful to clinicians for analyzing the severity of hand trauma. A review of this app from the clinicians' perspective was important to understand whether the app aligns with best clinical practices, is clinically valid, safe and appropriate for specific patient populations. The app's limitation lied in its usage of 2D images for evaluating hand range of motion, which led to the investigation of the capabilities of MediaPipe for extracting 3D landmarks and key features from hand images and videos using only the webcam of devices. The custom application, DIGITS, was developed using MediaPipe Hands package and proved to be a reliable tool for measuring finger range of motion in clinical studies. These computer-vision based modalities can form a low-cost alternative to expensive methods of hand scanning and movement tracking. In future research, DIGITS application has the potential to aid in studying the speed of post-trauma recovery for the full hand range of motion. It can also be useful in detecting hand tremors, such as those associated with Parkinson's, dystonic and Essential tremors.

## References

1. Gajarawala, S. N., & Pelkowski, J. N. (2021). Telehealth Benefits and Barriers. *The journal for nurse practitioners: JNP*, 17(2), 218–221. <https://doi.org/10.1016/j.nurpra.2020.09.013>.
2. Szekeres, M., & Valdes, K. (2022). Virtual health care & telehealth: Current therapy practice patterns. *Journal of Hand Therapy*, 35(1), 124-130. <https://doi.org/10.1016/j.jht.2020.11.004>.
3. Patterson, R. M., Salatin, B., Janson, R., Salinas, S. P., & Mullins, M. J. S. (2020). A current snapshot of the state of 3D printing in hand rehabilitation. *Journal of hand therapy : official journal of the American Society of Hand Therapists*, 33(2), 156–163. <https://doi.org/10.1016/j.jht.2019.12.018>.



4. Schwartz, J.K., et al., Methodology and feasibility of a 3D printed assistive technology intervention. *Disability and Rehabilitation: Assistive Technology*, 2020. 15(2): p. 141-147.
5. S. R. Ellis, What Are Virtual Environments? ,*IEEE Computer Graphics and Applications* 14 (1994)17-22.
6. Hao Wang, Xiaowei Chen, Fu Jia, Xiaojuan Cheng.(2023).Digital twin-supported smart city: Status,challenges and future research directions.*Expert Systems with Applications*,Volume 217,119531,<https://doi.org/10.1016/j.eswa.2023.119531>.
7. Javaid, Mohd & Haleem, Abid. (2018). Current status and challenges of Additive manufacturing in orthopaedics: An overview. *Journal of Clinical Orthopaedics and Trauma*. 10. 10.1016/j.jcot.2018.05.008.
8. Paola Volonghi, Gabriele Baronio & Alberto Signoroni (2018) 3D scanning and geometry processing techniques for customised hand orthotics: an experimental assessment, *Virtual and Physical Prototyping*, 13:2, 105-116, doi: 10.1080/17452759.2018.1426328.
9. Almomani F, Alghwiri AA, Alghadir AH, Al-Momani A, Iqbal A. Prevalence of upper limb pain and disability and its correlates with demographic and personal factors. *J Pain Res*. 2019;12:2691–700.
10. Moulaei, K., Sheikhtaheri, A., Nezhad, M.S. et al. (2022).Telerehabilitation for upper limb disabilities: a scoping review on functions, outcomes, and evaluation methods. *Archives of Public Health* 80, 196 . <https://doi.org/10.1186/s13690-022-00952-w>.
11. Bouzit, M., Burdea, G., Popescu, G., & Boian, R. (2002). The Rutgers Master II-New Design force-feedback glove. *IEEE/ASME Transactions on Mechatronics*, 7, 256-263. <https://doi.org/10.1109/TMECH.2002.1011262>.
12. Aiple, M., & Schiele, A. (2013). Pushing the limits of the CyberGrasp™ for haptic rendering. 2013 *IEEE International Conference on Robotics and Automation*, 3541-3546. doi: 10.1109/ICRA.2013.6631073.

13. Bouzit, M. (1996). Design, implementation and testing of a data glove with force feedback for virtual and real objects telemanipulation. Ph.D. Thesis, Paris, France.
14. Luo, X., Kline, T., Fischer, H., Stubblefield, K., Kenyon, R., & Kamper, D. (2005). Integration of augmented reality and assistive devices for post-stroke hand opening rehabilitation. Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 6855-6858. <https://doi.org/10.1109/IEMBS.2005.1616080>.
15. Gerber, C. N., Kunz, B., & van Hedel, H. J. (2016). Preparing a neuropediatric upper limb exergame rehabilitation system for home-use: a feasibility study. *Journal of NeuroEngineering and Rehabilitation*, 13, 33. <https://doi.org/10.1186/s12984-016-0141-x>.
16. Friedman, N., Chan, V., Zondervan, D., Bachman, M., & Reinkensmeyer, D. J. (2011). MusicGlove: motivating and quantifying hand movement rehabilitation by using functional grips to play music. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference, 2011, 2359–2363. <https://doi.org/10.1109/IEMBS.2011.6090659>.
17. Friedman, N., Chan, V., Reinkensmeyer, A.N. et al. Retraining and assessing hand movement after stroke using the MusicGlove: comparison with conventional hand therapy and isometric grip training. *J NeuroEngineering Rehabil* 11, 76 (2014). <https://doi.org/10.1186/1743-0003-11-76>.
18. Adie, K., Schofield, C., Berrow, M., Wingham, J., Humfryes, J., Pritchard, C., James, M., & Allison, R. (2017). Does the use of Nintendo Wii Sports™ improve arm function? Trial of Wii™ in Stroke: a randomized controlled trial and economics analysis. *Clinical rehabilitation*, 31(2), 173–185. <https://doi.org/10.1177/0269215516637893>.
19. McNulty, P. A., Thompson-Butel, A. G., Faux, S. G., Lin, G., Katrak, P. H., Harris, L. R., & Shiner, C. T. (2015). The efficacy of Wii-based Movement Therapy for upper limb rehabilitation in the chronic poststroke period: a randomized controlled trial. *International journal of stroke : official journal of the International Stroke Society*, 10(8), 1253–1260. <https://doi.org/10.1111/ijss.12594>.

20. Da Gama, A., Fallavollita, P., Teichrieb, V., & Navab, N. (2015). Motor Rehabilitation Using Kinect: A Systematic Review. *Games for health journal*, 4(2), 123–135. <https://doi.org/10.1089/g4h.2014.0047>.
21. Neil, A., Ens, S., Pelletier, R., Jarus, T., & Rand, D. (2012). Sony PlayStation EyeToy elicits higher levels of movement than the Nintendo Wii: Implications for stroke rehabilitation. *European journal of physical and rehabilitation medicine*, 49.
22. Ates, S., Lobo-Prat, J., Lammertse, P., Kooij, H., & Stienen, A. (2013). SCRIPT Passive Orthosis: Design and technical evaluation of the wrist and hand orthosis for rehabilitation training at home. *IEEE International Conference on Rehabilitation Robotics Proceedings*, 1-6. <https://doi.org/10.1109/ICORR.2013.6650401>.
23. 30. Laver, K. E., Lange, B., George, S., Deutsch, J. E., Saposnik, G., & Crotty, M. (2017). Virtual reality for stroke rehabilitation. *The Cochrane database of systematic reviews*, 11(11), CD008349. <https://doi.org/10.1002/14651858.CD008349.pub4>.
24. Park, H. S., Peng, Q., & Zhang, L. Q. (2008). A portable telerehabilitation system for remote evaluations of impaired elbows in neurological disorders. *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, 16(3), 245–254. <https://doi.org/10.1109/TNSRE.2008.920067>.
25. Tetsuya Mouri, Haruhisa Kawasaki, Takaaki Aoki, Yutaka Nishimoto, Satoshi Ito, Satoshi Ueki.(2009). Telerehabilitation for Fingers and Wrist Using a Hand Rehabilitation Support System and Robot Hand, *IFAC Proceedings Volumes*, 42(16), 603-608,<https://doi.org/10.3182/20090909-4-JP-2010.00102>.
26. Placidi G. (2007). A smart virtual glove for the hand telerehabilitation. *Computers in biology and medicine*, 37(8), 1100–1107. <https://doi.org/10.1016/j.combiomed.2006.09.011>.
27. Staszuk, A., Wiatrak, B., Tadeusiewicz, R., Karuga-Kuźniewska, E., & Rybak, Z. (2016). Telerehabilitation approach for patients with hand impairment. *Acta of Bioengineering and Biomechanics*, 18(4), 55–62.

28. Herman GT. Fundamentals of computerized tomography: Image reconstruction from projection. 2nd ed. New York: Springer; 2009. p. 1-17.
29. Scarfe WC, Farman AG, Sukovic P. Clinical applications of cone beam computed tomography in dental practice. *J Can Dent Assoc.* 2006;72:75-80.
30. Ritman EL. Micro computed tomography current status and developments. *Annu Rev Biomed Eng.* 2004;6:185-208.
31. Edelman RR, Hesselink J, Zlatkin M. Clinical magnetic resonance imaging: 3-volume set. Saunders; 2005.
32. Paddock SW, Eliceiri KW. Laser scanning confocal microscopy: History, applications, and related optical sectioning techniques. *Methods Mol Biol* 2014;1075:9-47.
33. Jason Geng, "Structured-light 3D surface imaging: a tutorial," *Adv. Opt. Photon.* 3, 128-160 (2011).
34. Liberadzki P, Adamczyk M, Witkowski M, Sitnik R. Structured-Light-Based System for Shape Measurement of the Human Body in Motion. *Sensors.* 2018; 18(9):2827. <https://doi.org/10.3390/s18092827>.
35. Barron, JL, and Eagleson, Roy (1996) "Recursive Estimation of Time-Varying Motion and Structure Parameters" *Journal of Pattern Recognition*, VMay, 29(5):797-818.
36. Barron, JL, and Eagleson, Roy (1997) "Computation of Time-Varying Motion and Structure Parameters from Real Image Sequences", pp.181-190.
37. Lowe, David. "Object Recognition from Local Scale Invariant Features". In: *Proceedings of the IEEE International Conference on Computer Vision 2* (Jan. 2001).
38. Qingshan Xu, Jie Li, Wenbing Tao, Delie Ming, Efficient large-scale geometric verification for structure from motion, *Pattern Recognition Letters*, Volume 125, 2019, Pages 166-173, <https://doi.org/10.1016/j.patrec.2018.09.028>.
39. Akhloufi, Moulay & Tong, W. & Polotski, Vladimir & Cohen (1998). Estimating the Fundamental Matrix for a Stereoscopic System from Planar Surfaces.

40. Fischler, R. and Bolles, M. "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". In: *Commun ACM* 24 (Jan. 1981), pp. 619–638.
41. Bianco, Simone & Ciocca, Gianluigi & Marelli, Davide. (2018). Evaluating the Performance of Structure from Motion Pipelines. *Journal of Imaging*. 4. 98. 10.3390/jimaging4080098.
42. Lourakis, Manolis & Argyros, Antonis. (2005). Is Levenberg-Marquardt the Most Efficient Optimization Algorithm for Implementing Bundle Adjustment? *Proceedings of the Tenth IEEE International Conference on Computer Vision*. 2. 1526-1531. 10.1109/ICCV.2005.128.
43. Tóth, D., Petrus, K., Heckmann, V., Simon, G. and Poór, V.S. (2021), Application of photogrammetry in forensic pathology education of medical students in response to COVID-19. *J Forensic Sci*, 66: 1533-1537. <https://doi.org/10.1111/1556-4029.14709>.
44. Abbas Khalaf, Tariq Ataiwe, Israa Mohammed, and Ali Kareem, 3D Digital modeling for archeology using close range photogrammetry, *MATEC Web of Conferences* 162, 03027 (2018), <https://doi.org/10.1051/mateconf/201816203027>
45. Documenting Skeletal Scatters in Obstructed Wooded Environments Using Close-Range Photogrammetry, Ferrell, Morgan J. ;*Forensic Anthropology*; Gainesville Vol. 4, Iss. 2, (2021): 88-103. DOI: 10.5744/fa.2020.0044.
46. Andrea F. Abate, Luigi De Maio, Riccardo Distasi, Fabio Narducci, Remote 3D face reconstruction by means of autonomous unmanned aerial vehicles, *Pattern Recognition Letters*, Volume 147, 2021, Pages 48-54, ISSN 0167-8655, <https://doi.org/10.1016/j.patrec.2021.04.006>.
47. R. B. Taqriban, R. Ismail, M. Ariyanto and A. F. Yaya Syah Putra, "3D Model of Photogrammetry Technique for Transtibial Prosthetic Socket Design Development," 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 2019, pp. 456-461, doi: 10.1109/ISRITI48646.2019.9034670.

48. Abellán, A.; Jaboyedoff, M.; Oppikofer, T.; Vilaplana, J. Detection of millimetric deformation using a terrestrial laser scanner: Experiment and application to a rockfall event. *Nat. Hazards Earth Syst. Sci.* 2009, 9, 365–372.
49. Pix4D. Reprojection error. Pix4D Support. Retrieved March 22, 2023, from <https://support.pix4d.com/hc/en-us/articles/202559369-Reprojection-error>.
50. Besl, P. J., & McKay, N. D. (1992). Method for registration of 3-D shapes. In *Robotics-DL tentative* (pp. 586-606). International Society for Optics and Photonics.
51. Abellán, A., Calvet, J., Vilaplana, J., & Blanchard, J. (2010). Detection and spatial prediction of rockfalls by means of terrestrial laser scanner monitoring. *Geomorphology*, 119(1-2), 162-171.
52. Olsen, M. J., Kuester, F., Chang, B. J., & Hutchinson, T. C. (2010). Terrestrial laser scanning-based structural damage assessment. *Journal of Computing in Civil Engineering*, 24(3), 264-272.
53. Geospatial World. What is LiDAR Technology and How Does it Work? Retrieved from <https://www.geospatialworld.net/prime/technology-and-innovation/what-is-lidar-technology-and-how-does-it-work/>.
54. Sacyr. LiDAR, el nuevo ojo láser de los teléfonos móviles [LiDAR, the new laser eye of mobile phones]. Retrieved from <https://www.sacyr.com/en/-/lidar-el-nuevo-ojo-laser-de-los-telefonos-moviles>.
55. Aungst T. D. (2013). Medical applications for pharmacists using mobile devices. *The Annals of pharmacotherapy*, 47(7-8), 1088–1095. <https://doi.org/10.1345/aph.1S035>.
56. Divall, P., Camosso-Stefinovic, J., & Baker, R. (2013). The use of personal digital assistants in clinical decision making by health care professionals: a systematic review. *Health informatics journal*, 19(1), 16–28. <https://doi.org/10.1177/1460458212446761>.
57. Zhao, J., Freeman, B., & Li, M. (2016). Can Mobile Phone Apps Influence People's Health Behavior Change? An Evidence Review. *Journal of medical Internet research*, 18(11), e287. <https://doi.org/10.2196/jmir.5692>.

58. Durrant, N. (2014). The Reliability and Validity of the Ten Test and Exploring a New, Visual Version.
59. MacDermid J. C. (1996). Development of a scale for patient rating of wrist pain and disability. *Journal of hand therapy : official journal of the American Society of Hand Therapists*, 9(2), 178–183. [https://doi.org/10.1016/s0894-1130\(96\)80076-7](https://doi.org/10.1016/s0894-1130(96)80076-7).
60. MacDermid J. C. (2019). The PRWE/PRWHE update. *Journal of hand therapy: official journal of the American Society of Hand Therapists*, 32(2), 292–294. <https://doi.org/10.1016/j.jht.2019.01.001>.
61. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H Development and Validation of the User Version of the Mobile Application Rating Scale (uMARS) *JMIR Mhealth Uhealth* 2016;4(2):e72 doi: <https://mhealth.jmir.org/2016/2/e72/.5849> PMID: 27287964 ,PMCID: 4920963.
62. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps *JMIR Mhealth Uhealth* 2015;3(1):e27, doi: <https://doi.org/10.2196/mhealth.3422>, PMID: 25760773, PMCID: 4376132.
63. Goldhahn, J., Angst, F., & Simmen, B. R. (2008). What counts: Outcome assessment after distal radius fractures in aged patients. *Journal of Orthopaedic Trauma*, 22(8 Suppl), S126-S130.
64. Changulani, M., Okonkwo, U., Keswani, T., Kalairajah, Y. (2008). Outcome evaluation measures for wrist and hand: Which one to choose? *International Orthopaedics*, 32(1), 1-6.
65. John, M., Angst, F., Awiszus, F., Pap, G., MacDermid, J. C., & Simmen, B. R. (2008). The patient-rated wrist evaluation (PRWE): Cross-cultural adaptation into German and evaluation of its psychometric properties. *Clinical and Experimental Rheumatology*, 26, 1047-1058.
66. Hemelaers L, Angst F, Drerup S, Simmen BR, WoodDauphinee S. Reliability and validity of the German version of “the patient-rated wrist evaluation (PRWE)” as an outcome measure of wrist pain and disability in patients with acute distal radius fractures. *J Hand Ther.* 2008;21(4):366–76.

67. Bodian, C. A., Freedman, G., Hossain, S., Eisenkraft, J. B., & Beilin, Y. (2001). The visual analog scale for pain: Clinical significance in postoperative patients. *Anesthesiology*, 95(6), 1356-1361. doi: 10.1097/00000542-200112000-00013.
68. Batista, M. A., & Gaglani, S. M. (2013). The future of smartphones in health care. *The virtual mentor: VM*, 15(11), 947–950. <https://doi.org/10.1001/virtualmentor.2013.15.11.stas1-1311>.
69. Batista MA, Gaglani SM. The future of smartphones in health care. *Virtual Mentor* 2013 Nov 01;15(11):947-950. [doi: 10.1001/virtualmentor.2013.15.11.stas1-1311] [Medline: 24257085].
70. ElKoura, G., & Singh, K. (2003). Handrix: Animating the human hand. In D. Breen & M. Lin (Eds.), *Eurographics/SIGGRAPH Symposium on Computer Animation* (pp. 97-105).
71. Google LLC. Hands. <https://google.github.io/mediapipe/solutions/hands.html>.
72. Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, ChengYang Fu, and Alexander Berg. Ssd: Single shot multibox detector. In *SSD: Single Shot MultiBox Detector*, volume 9905, pages 21–37, 10 2016.
73. Fan Zhang, Valentin Bazarevsky, Andrey Vakunov, Andrei Tkachenka, George Sung, Chuo-Ling Chang, and Matthias Grundmann. Mediapipe hands: On-device real-time hand tracking, 06 2020.
74. <https://google.github.io/mediapipe/solutions/hands.html>.
75. Chua, E. H. 3D Graphics with OpenGL: Basic Theory. Retrieved from [https://www3.ntu.edu.sg/home/ehchua/programming/opengl/CG\\_BasicsTheory.html](https://www3.ntu.edu.sg/home/ehchua/programming/opengl/CG_BasicsTheory.html).
76. Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420.
77. Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>.



78. Patias, P., & Peipe, J. (2000). Photogrammetry and CAD/CAM in culture and industry: An ever changing paradigm.

## Appendix

Appendix 1: Cloud Compare source code in GitHub:

<https://github.com/CloudCompare/CloudCompare.git>

Appendix 2: Statistics for all 21 landmarks (Mean, SD, Min, Median, Max)

		Mean (mm)	S.D(mm)	Min(mm)	Median(mm)	Max(mm)
	x 0	0.36837	6.84941E-4	0.3656	0.3684	0.37029
landmark 0	y 0	0.9028	0.00153	0.89836	0.90273	0.90642
	z 0	-9.64E-05	3.80E-06	-1.07E-04	-9.66E-05	-8.49E-05
	x 1	0.38953	1.04E-03	0.3871	0.38884	0.56458
landmark 1	y 1	0.88258	0.00212	0.87812	0.8821	0.88779
	z 1	-0.07678	0.00126	-0.08152	-0.0767	-0.0685
	x 2	0.33119	0.00139	0.32829	0.33113	0.33424
landmark 2	y 2	0.77426	0.0035	0.7682	0.77344	0.78426
	z 2	-0.13888	0.00338	-0.14918	-0.13882	-0.13122
	x 3	0.33141	0.00287	0.32675	0.33109	0.33735
landmark 3	y 3	0.70128	0.00448	0.69393	0.6997	0.71346
	z 3	-0.17504	0.00469	-0.18864	-0.17457	-0.16606
	x 4	0.32956	0.00285	0.32515	0.32988	0.33542
landmark 4	y 4	0.63085	0.0052	0.62315	0.62877	0.64437
	z 4	-0.20929	0.0059	-0.2256	-0.20849	-0.19782
	x 5	0.28632	0.00212	0.28148	0.28668	0.29019
landmark 5	y 5	0.65507	0.00307	0.65005	0.65391	0.66415
	z 5	-0.1258	0.00434	-0.13826	-0.12586	-0.11582
	x 6	0.27944	0.00337	0.27275	0.27983	0.28503
landmark 6	y 6	0.54917	0.00325	0.54491	0.54784	0.55846
	z 6	-0.17475	0.00557	-0.18947	-0.1748	-0.16211
	x 7	0.27053	0.00394	0.26328	0.27106	0.27718
landmark 7	y 7	0.4808	0.00366	0.4765	0.47933	0.49091
	z 7	-0.20954	0.00637	-0.22654	-0.20947	-0.1959
	x 8	0.25868	0.00459	0.25084	0.25961	0.26652
landmark 8	y 8	0.4158	0.00391	0.41091	0.41428	0.42609
	z 8	-0.23869	0.00668	-0.25652	-0.23863	-0.22453
	x 9	0.24445	0.00198	0.23999	0.24489	0.24814
landmark 9	y 9	0.65261	0.00288	0.64822	0.65142	0.66116
	z 9	-0.11087	0.00386	-0.12194	-0.11088	-0.10225
	x 10	0.23704	0.00316	0.23067	0.2377	0.24245
landmark 10	y 10	0.54163	0.0033	0.53723	0.54026	0.55072
	z 10	-0.16718	0.00499	-0.1814	-0.16735	-0.15384
	x 11	0.23037	0.00378	0.22347	0.2311	0.23678
landmark 11	y 11	0.46935	0.0035	0.4649	0.46794	0.47914
	z 11	-0.20368	0.0057	-0.22071	-0.20337	-0.18991
	x 12	0.22211	0.00441	0.21411	0.22286	0.22946
landmark 12	y 12	0.39974	0.00351	0.39566	0.39837	0.40986
	z 12	-0.23609	0.00638	-0.25485	-0.23552	-0.22198
	x 13	0.20759	0.00176	0.20349	0.20793	0.21085
landmark 13	y 13	0.66734	0.00289	0.66362	0.66623	0.67553
	z 13	-0.10239	0.00345	-0.11279	-0.102	-0.09453
	x 14	0.19965	0.00261	0.19432	0.20024	0.20455
landmark 14	y 14	0.56658	0.00352	0.5614	0.5652	0.57567
	z 14	-0.15381	0.00451	-0.16699	-0.1535	-0.14154
	x 15	0.19397	0.00325	0.18772	0.1947	0.2001
landmark 15	y 15	0.49993	0.00376	0.49465	0.49851	0.50996
	z 15	-0.19377	0.00547	-0.21021	-0.19326	-0.18074
	x 16	0.18845	0.00395	0.18138	0.18916	0.19552
landmark 16	y 16	0.43258	0.00378	0.42694	0.43119	0.44291
	z 16	-0.2268	0.00631	-0.24553	-0.22596	-0.21379

	x 17	0.17	0.00159	0.16624	0.17013	0.17318
landmark 17	y 17	0.69833	0.00299	0.69372	0.69732	0.70606
	z 17	-0.09786	0.00318	-0.10802	-0.09739	-0.08992
	x 18	0.16242	0.00228	0.15807	0.16297	0.16671
landmark 18	y 18	0.61772	0.00345	0.61204	0.61663	0.62675
	z 18	-0.14539	0.00383	-0.15774	-0.14519	-0.13569
	x 19	0.15808	0.00279	0.15286	0.15877	0.16298
landmark 19	y 19	0.56687	0.00355	0.56093	0.56578	0.57594
	z 19	-0.17853	0.00442	-0.19306	-0.17825	-0.16788
	x 20	0.15478	0.00344	0.14868	0.15545	0.16086
landmark 20	y 20	0.51187	0.00365	0.50528	0.5108	0.5211
	z 20	-0.20906	0.00493	-0.22487	-0.2086	-0.19743

Appendix 3: Modified uMARS Sheet used for Hand Scans mobile app usability evaluation

## Modified Mobile Application Rating Scale: user version (uMARS)

Raters should:

1. Use the app and trial it thoroughly for at least 10 minutes;
2. Determine how easy it is to use, how well it functions and does it do what it purports to do;
3. Review app settings, developer information, external links, security features, etc.

### Scoring

A: Engagement Mean Score = \_\_\_\_\_

B: Functionality Mean Score = \_\_\_\_\_

C: Aesthetics Mean Score = \_\_\_\_\_

D: Information Mean Score\* = \_\_\_\_\_

E: Subjective quality Mean Score = \_\_\_\_\_

\* Exclude questions rated as "N/A" from the mean score calculation.

App overall quality mean score \_\_\_\_\_ =  $A + B + C + D / 4$

The *App subjective quality* scale can be reported as individual items or as a mean score, depending on the aims of the research.

App Name: **HAND SCANS**

Circle the number that most accurately represents the quality of the app you are rating. All items are rated on a 5-point scale from "1.Inadequate" to "5.Excellent". Select N/A if the app component is irrelevant.

## App Quality Ratings

### SECTION A

---

**Engagement – fun, interesting, customisable, interactive, has prompts (e.g. sends alerts, messages, reminders, feedback, enables sharing)**

1. **Entertainment: Is the app fun/entertaining to use? Does it have components that make it more fun than other similar apps?**
  - 1 Dull, not fun or entertaining at all
  - 2 Mostly boring
  - 3 OK, fun enough to entertain user for a brief time (< 5 minutes)
  - 4 Moderately fun and entertaining, would entertain user for some time (5-10 minutes total)
  - 5 Highly entertaining and fun, would stimulate repeat use
  
2. **Interest: Is the app interesting to use? Does it present its information in an interesting way compared to other similar apps?**
  - 1 Not interesting at all
  - 2 Mostly uninteresting
  - 3 OK, neither interesting nor uninteresting; would engage user for a brief time (< 5 minutes)
  - 4 Moderately interesting; would engage user for some time (5-10 minutes total)
  - 5 Very interesting, would engage user in repeat use
  
3. **Customisation: Does it allow you to customise the settings and preferences that you would like to (e.g. sound, content and notifications)?**
  - 1 Does not allow any customisation or requires setting to be input every time
  - 2 Allows little customisation and that limits app's functions
  - 3 Basic customisation to function adequately
  - 4 Allows numerous options for customisation
  - 5 Allows complete tailoring the user's characteristics/preferences, remembers all settings
  
4. **Interactivity: Does it allow user input, provide feedback, contain prompts (reminders, sharing options, notifications, etc.)?**
  - 1 No interactive features and/or no response to user input
  - 2 Some, but not enough interactive features which limits app's functions
  - 3 Basic interactive features to function adequately
  - 4 Offers a variety of interactive features, feedback and user input options
  - 5 Very high level of responsiveness through interactive features, feedback and user input options

5. **Target group: Is the app content (visuals, language, design) appropriate for the target audience?**
- 1 Completely inappropriate, unclear or confusing
  - 2 Mostly inappropriate, unclear or confusing
  - 3 Acceptable but not specifically designed for the target audience. May be inappropriate/ unclear/confusing at times
  - 4 Designed for the target audience, with minor issues
  - 5 Designed specifically for the target audience, no issues found

A. Engagement Mean Score = \_\_\_\_\_

## **SECTION B**

---

**Functionality – app functioning, easy to learn, navigation, flow logic, and gestural design of app**

6. **Performance: How accurately/fast do the app features (functions) and components (buttons/menus) work?**
- 1 App is broken; no/insufficient/inaccurate response (e.g. crashes/bugs/broken features, etc.)
  - 2 Some functions work, but lagging or contains major technical problems
  - 3 App works overall. Some technical problems need fixing, or is slow at times
  - 4 Mostly functional with minor/negligible problems
  - 5 Perfect/timely response; no technical bugs found, or contains a 'loading time left' indicator (if relevant)
7. **Ease of use: How easy is it to learn how to use the app; how clear are the menu labels, icons and instructions?**
- 1 No/limited instructions; menu labels, icons are confusing; complicated
  - 2 Takes a lot of time or effort
  - 3 Takes some time or effort
  - 4 Easy to learn (or has clear instructions)
  - 5 Able to use app immediately; intuitive; simple (no instructions needed)
8. **Navigation: Does moving between screens make sense; Does app have all necessary links between screens?**
- 1 No logical connection between screens at all /navigation is difficult
  - 2 Understandable after a lot of time/effort
  - 3 Understandable after some time/effort
  - 4 Easy to understand/navigate
  - 5 Perfectly logical, easy, clear and intuitive screen flow throughout, and/or has shortcuts
9. **Gestural design: Do taps/swipes/pinches/scrolls make sense? Are they consistent across all components/screens?**
- 1 Completely inconsistent/confusing
  - 2 Often inconsistent/confusing
  - 3 OK with some inconsistencies/confusing elements
  - 4 Mostly consistent/intuitive with negligible problems
  - 5 Perfectly consistent and intuitive

B. Functionality Mean Score = \_\_\_\_\_

## SECTION C

---

### Aesthetics – graphic design, overall visual appeal, colour scheme, and stylistic consistency

10. **Layout: Is arrangement and size of buttons, icons, menus and content on the screen appropriate?**
- 1 Very bad design, cluttered, some options impossible to select, locate, see or read
  - 2 Bad design, random, unclear, some options difficult to select/locate/see/read
  - 3 Satisfactory, few problems with selecting/locating/seeing/reading items
  - 4 Mostly clear, able to select/locate/see/read items
  - 5 Professional, simple, clear, orderly, logically organised
11. **Graphics: How high is the quality/resolution of graphics used for buttons, icons, menus and content?**
- 1 Graphics appear amateur, very poor visual design - disproportionate, stylistically inconsistent
  - 2 Low quality/low resolution graphics; low quality visual design – disproportionate
  - 3 Moderate quality graphics and visual design (generally consistent in style)
  - 4 High quality/resolution graphics and visual design – mostly proportionate, consistent in style
  - 5 Very high quality/resolution graphics and visual design - proportionate, consistent in style throughout
12. **Visual appeal: How good does the app look?**
- 1 Ugly, unpleasant to look at, poorly designed, clashing, mismatched colours
  - 2 Bad – poorly designed, bad use of colour, visually boring
  - 3 OK – average, neither pleasant, nor unpleasant
  - 4 Pleasant – seamless graphics – consistent and professionally designed
  - 5 Beautiful – very attractive, memorable, stands out; use of colour enhances app features/menus

C: Aesthetics Mean Score = \_\_\_\_\_

## SECTION D

---

### Information – Contains high quality information (e.g. text, feedback, measures, references) from a credible source

13. **Quality of information: Is app content correct, well written, and relevant to the goal/topic of the app?**
- N/A There is no information within the app
- 1 Irrelevant/inappropriate/incoherent/incorrect
  - 2 Poor. Barely relevant/appropriate/coherent/may be incorrect
  - 3 Moderately relevant/appropriate/coherent/and appears correct
  - 4 Relevant/appropriate/coherent/correct
  - 5 Highly relevant, appropriate, coherent, and correct
14. **Quantity of information: Is the information within the app comprehensive but concise?**
- N/A There is no information within the app
- 1 Minimal or overwhelming
  - 2 Insufficient or possibly overwhelming
  - 3 OK but not comprehensive or concise
  - 4 Offers a broad range of information, has some gaps or unnecessary detail; or has no links to more information and resources
  - 5 Comprehensive and concise; contains links to more information and resources

**15. Visual information: Is visual explanation of concepts – through charts/graphs/images/videos, etc. – clear, logical, correct?**

N/A There is no visual information within the app (e.g. it only contains audio, or text)

- 1 Completely unclear/confusing/wrong or necessary but missing
- 2 Mostly unclear/confusing/wrong
- 3 OK but often unclear/confusing/wrong
- 4 Mostly clear/logical/correct with negligible issues
- 5 Perfectly clear/logical/correct

**16. Credibility of source: does the information within the app seem to come from a credible source?**

N/A There is no information within the app

- 1 Suspicious source
- 2 Lacks credibility
- 3 Not suspicious but legitimacy of source is unclear
- 4 Possibly comes from a legitimate source
- 5 Definitely comes from a legitimate/specialised source

D: Information Mean Score = \_\_\_\_\_

## App subjective quality

### SECTION E

---

**17. Would you recommend this app to people who might benefit from it?**

- |              |   |
|--------------|---|
| 1 Not at all | I would not recommend this app to anyone                |
| 2            | There are very few people I would recommend this app to |
| 3 Maybe      | There are several people I would recommend this app to  |
| 4            | There are many people I would recommend this app to     |
| 5 Definitely | I would recommend this app to everyone                  |

**18. How many times do you think you would use this app in the next 12 months if it was relevant to you?**

- 1 None
- 2 1-2
- 3 3-10
- 4 10-50
- 5 >50

**19. Would you pay for this app?**

- 1 Definitely not
- 2 No
- 3 Maybe
- 4 Yes
- 5 Definitely yes

**20. What is your overall (star) rating of the app?**

- |         |                                 |
|---------|---------------------------------|
| 1 ★     | One of the worst apps I've used |
| 2 ★★    |                                 |
| 3 ★★★   | Average                         |
| 4 ★★★★  |                                 |
| 5 ★★★★★ | One of the best apps I've used  |