MARKERLESS MOTION CAPTURE FOR THE HANDS AND FINGERS

MARKERLESS MOTION CAPTURE FOR THE HANDS AND FINGERS

By NIGEL MAJONI, B. Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

McMaster University MASTER OF SCIENCE (2024) Hamilton, Ontario

(Kinesiology)

TITLE: Markerless motion capture for the hands and fingers

AUTHOR: NIGEL MAJONI, B. Sc. (University of Waterloo)

SUPERVISOR: Dr. Peter J. Keir

NUMBER OF PAGES: 95

THESIS ABSTRACT

Hand postures and movements play a significant role in our daily activities. They are important in the development of upper extremity musculoskeletal disorders, yet they are often underrepresented in biomechanical and ergonomic studies. Improved understanding of hand and finger kinematics is essential for assessing the risk of musculoskeletal disorders. Tracking these movements is challenging due to the limited tools available for tracking the hands and fingers. Traditional marker-based motion capture, while considered the industry standard, faces issues such as occlusion, being time-consuming, and limited to a laboratory setting. Recent advancements in computer vision and machine learning offer potential solutions through markerless motion capture. Current applications have primarily focused on the lower extremities, with limited effort on the hand and fingers. This thesis developed and assessed a markerless motion capture system for tracking hand and finger joint kinematics. A markerless system using four synchronized webcams was developed with camera pairs organized with different angles: Centre/90° (C/90°), Left 45°/Right 45° (L45°/R45°), and Centre/Left 45° (C/L45°). Motion capture was performed with both marker-based and markerless systems. Seventy-seven reflective markers were placed on participants for the marker-based motion capture. Twenty healthy participants performed five dynamic hand tasks, each repeated three times, with and without markers. Three-dimensional joint positions were defined using a musculoskeletal model in OpenSim. The total finger angle was calculated as the sum of MCP, PIP, and DIP joint angles for digits 2-5 and MCP and IP for digit 1. The comparison between markerless and marker-based motion capture systems showed significant interactions between camera orientation and finger movement in several tasks, particularly for the index finger during flexion and typing. No differences were observed between the C/90° and C/L45° markerless camera pairs and the marker-based system, except for the 99th percentile index finger flexion. The L45°/R45° camera pair differed significantly from other markerless pairs in several tasks but agreed with the marker-based system for index finger during flexion. For most of the fingers, no significant differences were found across the different camera pairs, except for some in the index finger flexion task. Correlations and error for concurrent finger flexion revealed high consistency among all camera pairs, with R^2 above 0.90 and RMSD below 10°, though the thumb showed greater variability. The R^2 and RMSD varied depending on the camera comparison and finger for each task. Markerless motion capture for the hands and fingers is possible with little difference to marker-based systems and is dependent on the camera orientation used.

ACKNOWLEDGEMENTS

I would like to start by saying I was very lucky to end up in the situation I was put in at McMaster, and everyone I've met these last three years has been very influential during this journey. First, I would like to thank my supervisor, Dr. Peter Keir. You continued at every step to push me in the right direction, especially when things got tough, and I appreciated the patience and kindness you offered all this time. This process taught me how to be more independent in my learning and take ownership of my work. Thank you for the opportunity to work with you; without this chance, I would not be who I am today, and I would not have learned how to be a better researcher as I will take everything, I learned from you going into my PhD. I would do it again in a heartbeat!

I also want to thank my committee members Dr. Dylan Kobsar, Dr. Alex MacIntosh, and Dr. Aaron Kociolek. You have given me great feedback and perspective on my project, and I truly value the advice you give.

I want to thank the Occupational Biomechanics Lab and the Kobsar Lab members. Dan Mulla, Dan the Man, you are such a special person. When you become a PI, your students are going to be very lucky to work under your supervision, and I was lucky to have you around. You have taught me "the way" because if we follow your lead, we will be successful in anything we do. Without your help during these last three years, I probably would not have been doing this degree, and on behalf of the lab, we appreciate all your assistance, compassion, and patience. Ryan Chhiba, we started this journey three years ago, and I'm so happy to see that you are doing your PhD; the next 4 four years will be awesome. You are such a smart guy, and I can't wait to see what you do! You're also the one person, no matter what the problem is, I can lean on as you will drop everything to be there for me. Paul and Brian, it was a pleasure working with you; you are incredibly smart and talented and a great person overall. Paul, you will do great during your PhD, and I'm excited to see what you do in the future. Josh Keogh and Zaryan Masood, you were my first friends at McMaster, and I truly feel that if it weren't for the two of you, my whole McMaster journey would've been different and not for the better. Thank you for being such good people and supporting me throughout my degree.

I also want to thank Jenna Stone. You were very patient in dealing with me during this degree, especially near the end. You were very compassionate and supportive in helping me prepare to defend my thesis. Thanks for just being there for me overall. Love you!

Finally, I want to thank my family. First, my parents, Memory and Tawanda, have sacrificed everything and continue to sacrifice everything for me to be successful. Without your support, unconditional love, and guidance, I would have NEVER had the opportunities you've been able to give me. Melissa, I always followed your lead and copied everything you did, and I turned out alright and continue to do that. Except I'll be doing my PhD, so I might beat you to being the first Doctor in the family. Terrence, you always bring lightheartedness to any situation, and that is exactly what I needed during this degree. I love you all so much! Thank you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure 4.1: A) Shows the Centre (C) camera view during the concurrent finger flexion task without reflective markers. B) Shows the Centre (C) camera view during the concurrent finger flexion task with reflective markers. C) Shows the Left 45° (L45 °) camera view during the concurrent finger flexion task without reflective markers. D) Shows the Left 45° (L45 °) camera view during the concurrent finger flexion task with reflective markers. These images illustrate when the fingers are in extension during the task.

Figure 4.2: A) Shows the Centre (C) camera view during the concurrent finger flexion task without reflective markers. B) Shows the Centre (C) camera view during the concurrent finger flexion task with reflective markers. C) Shows the Left 45° (L45 °) camera view during the concurrent finger flexion task without reflective markers. D) Shows the Left 45° (L45 °) camera view during the concurrent finger flexion task with reflective markers. These images illustrate when the fingers are in flexion during the task.

Figure A1: Box plot comparison of camera pair R^2 and RMSD across multiple fingers (thumb, index, middle, ring, and little) and tasks (index finger flexion, finger roll flexion, finger flexion, and typing) with the markers off (non-concurrent). The left column presents the R^2 for all tasks and fingers, indicating the correlation between camera pairs, while the right column shows the corresponding RMSD, reflecting the differences in angles measured between camera pairs. The **(blue) box plots** represent the comparison between C/90° and C/L45°, the **(orange) box plots** represent the comparison between C/90° and L45°/R45°, and the **(green) box plots** represent the comparison between C/L45° and L45°/R45°. Each point within the box plots denotes the R^2 or RMSD for each trial for each participant.

Figure A2: Box plot comparison of camera pair R^2 and RMSD across multiple fingers (thumb, index, middle, ring, and little) and tasks (index finger flexion, finger roll flexion, finger flexion, and typing) with the markers on (concurrent). The left column presents the R^2 for all tasks and fingers, indicating the correlation between camera pairs, while the right column shows the corresponding RMSD values, reflecting the differences in angles measured between camera pairs. The **(blue) box plots** represent the comparison between C/90° and C/L45°, the **(orange) box plots** represent the comparison between C/90° and L45°/R45°, and the **(green) box plots** represent the comparison between C/L45° and L45°/R45°. Each point within the box plots denotes the R^2 or RMSD for each trial for each participant.

LIST OF ABBREVIATIONS

APDF – Amplitude probability distribution function

RMSD – Root means square difference

RMSE - Root means square error

MPH – MediaPipe Hands

DLC – DeepLabCut

CNN – Convolutional neural network

2D – Two-dimensional

3D – Three-dimensional

MCP – Metacarpophalangeal

PIP – Proximal interphalangeal

DIP – Distal interphalangeal

IP - Interphalangeal

IMMS - Inertial and magnetic measurement systems

FPS – Frames per second

ROM – Range of motion

CHAPTER ONE: INTRODUCTION

Our hands are our primary tools for interacting with the external environment and are constantly used in our daily lives. The postures and movements of our hands and fingers play significant roles in the development of hand-related musculoskeletal disorders. However, the hands are typically inadequately represented in biomechanical and ergonomic studies (Amell et al., 2001). Improved methods to capture hand and finger movements for use in ergonomic evaluations and studies are needed to better understand the development of hand-related musculoskeletal disorders.

Assessing hand activity is difficult. Tracking upper limb movements, particularly in the hands and fingers, is challenging. Simple hand movements such as grasping depend on several factors ranging from the size and shape of the object to the postures being adopted by the hand and fingers. Furthermore, accurate assessment of hand and finger kinematics is limited by technology. Hand and finger kinematics can be tracked using various motion capture methods, including optical motion capture (active and passive), inertial sensors, and markerless motion capture. In addition to these methods, techniques such as video-based motion capture with specialized software for movement tracking and manual assessment tools, such as electrogoniometers, may also be employed (Cook et al., 2007). While many exist, they are not without limitations. For instance, electrogoniometers and markers are bulky and can interfere with natural movement, and marker-based motion capture is typically limited to a laboratory setting and is not applicable to other settings, including the workplace. Marker-based motion capture systems have been conventionally used for tracking joint and segment kinematics in biomechanical research and are often referred to as the 'gold standard' (Cappozzo et al., 1995; Cheung et al., 2005). Although hand and finger kinematics can be accurately tracked using marker-based motion capture, there are several issues with using marker-based systems to track hand and finger kinematics. One main issue is the number of markers required. Cocchiarella et al. (2015) used 96 reflective markers, consisting of 42 calibration markers and 54 tracking markers. The placement of markers is time-consuming. A significant issue is that these markers can be occluded during multi-planar tasks such as opening and closing the hand, resulting in kinematic errors. Recent technological advancements in computer vision and machine learning have influenced the creation of markerless motion capture systems to track body segments and combat some limitations of marker-based motion capture.

In recent years, there has been an advancement in the evaluation of markerless motion capture. There is a growing interest in investigating the feasibility of markerless motion capture systems as a viable alternative to marker-based systems and exploring their potential for advancing research and practical application in motion analysis. Currently, several applications exist for markerless motion capture: (i) Theia3D (Kanko et al., 2021), (ii) OpenCap (Uhlrich et al., 2022), (iii) OpenCV (OpenCV., 2015), and (iv) DeepLabCut (Mathis et al., 2018; Nath et al., 2019). *Theia*3D (Theia Markerless, Inc., Kingston, Ontario) is currently the leading developer and most used commercial product in markerless motion capture of the human body. It has been validated against conventional marker-based systems, but only for the lower extremity. Validation of markerless systems for the upper extremity is not available. No markerless motion capture system can provide accurate hand and finger kinematics tracking. Consequently, more studies need to determine the accuracy of markerless motion capture systems for hand and finger kinematics compared to marker-based motion capture systems. This thesis aimed to create and assess a markerless motion capture system that can track hand and finger joint kinematics.

CHAPTER TWO: REVIEW OF LITERATURE

2.1. Marker-based motion capture of the hands and fingers

Rash et al. (1999) compared marker-based motion capture to 2-D fluoroscopy. They showed a high correlation between 2-D fluoroscopy and marker-based motion capture measuring index finger kinematics. Correlations for MCP, PIP, and DIP joints were 0.95, 0.98, and 0.94, respectively. Absolute error between the two methods as the range mean (± standard deviation) for the MCP, PIP, and DIP joints was 0.0–7.3° (3.1±1.6°), 0.0–12.9° (4.1±3.2°), and 0.0–6.7° (1.9±1.5°), respectively. Several others have also demonstrated that hand kinematics can be assessed using marker-based motion capture techniques. (Cocchiarella et al., 2016 ; Zhao et al., 2012 ; Metcalf et al., 2020 ; Metcalf et al., 2011 ; Sancho-Bru et al., 2014 ; Fowler et al., 2001).

There are several limitations associated with marker-based motion capture. First, the markers used are often bulky, interfering with natural movement patterns and leading to unnatural movements. Additionally, the occlusion of markers during hand and finger tasks is a challenge with marker-based motion capture. The complexity and size of the hand compared to other body areas also contribute to the challenges of marker-based motion capture, leading to the need for reduced marker sets (Metcalf et al., 2020; Hoyet et al., 2012). While reduced marker sets may be appropriate for some tasks, they may limit the range of movements that can be tracked. Three markers are required to track a segment in three dimensions, and reduced marker sets involve placing a single marker above the joints. Therefore, it is essential to consider selecting a marker set based on the specific tasks of interest when assessing hand and finger kinematics to minimize the limitations and maximize data collection accuracy.

2.2. Inertial motion capture of the hands and fingers

Research has demonstrated the reliability of utilizing inertial and magnetic measurement systems (IMMS) for evaluating segment motion and orientation. IMMS combines inertial sensors (accelerometers and gyroscopes) and magnetic sensors (magnetometers). By integrating the measurements taken from the inertial and magnetic sensors through sensor fusion, we can obtain the orientation and position of each body segment in space. A PowerGlove designed by Kortier et al. (2014) was validated against marker-based motion capture in a study by van den Noort et al. (2016). Three participants were recruited for this study, and the results of this study are limited due to the sample size. The average root mean square difference (RMSD) between the IMMS and the marker-based motion capture system ranged between 3˚ and 8˚. Tasks that required fast movements, such as tapping and circular pointing, showed more significant differences (Table 2.1). When performing slow flexion tasks, finger kinematics were similar to the marker-based motion capture system (van den Noort et al., 2016). The results suggested that the PowerGlove is most comparable during slow tasks.

Table 2.1: Root Mean Square Difference (RMSD) of Joint Angles for MCP, PIP, and DIP Joints across different flexion finger tasks (Van Den Noort et al., 2016).

	Flexion Tasks (54 Trials)			Fast Tasks (20 Trials)			Ab/Adduction Task (5 Trials)	Circular Pointing Task (8 Trials)	
	MCP (deg)	PIP (deg)	DIP (deg)	MCP (deg)	PIP (deg)	DIP (deg)	MCP (deg)	MCP (deg)	
Main Angles of Interest	Flexion/Extension			Flexion/Extension			Ab/Adduction	Flexion/Extension	Ab/Adduction
RMSD	5.0 (3.3)	7.3 (4.3)	5.6 (3.7)	8.4 (3.2)	7.1 (4.7)	7.7 (4.2)	3.3 (2.8)	6.6 (1.9)	7.5 (2.9)

As illustrated in Figure 2.1, the PowerGlove is considerable in size and bulk, with an extensive array of wires and sensors. These components pose a risk of becoming entangled or detached during data collection, potentially compromising the integrity of the collected data. To mount the PowerGlove, tape is needed to wrap around each sensor to prevent it from falling off. These factors might interfere with natural movement patterns and lead to unnatural movements in participants.

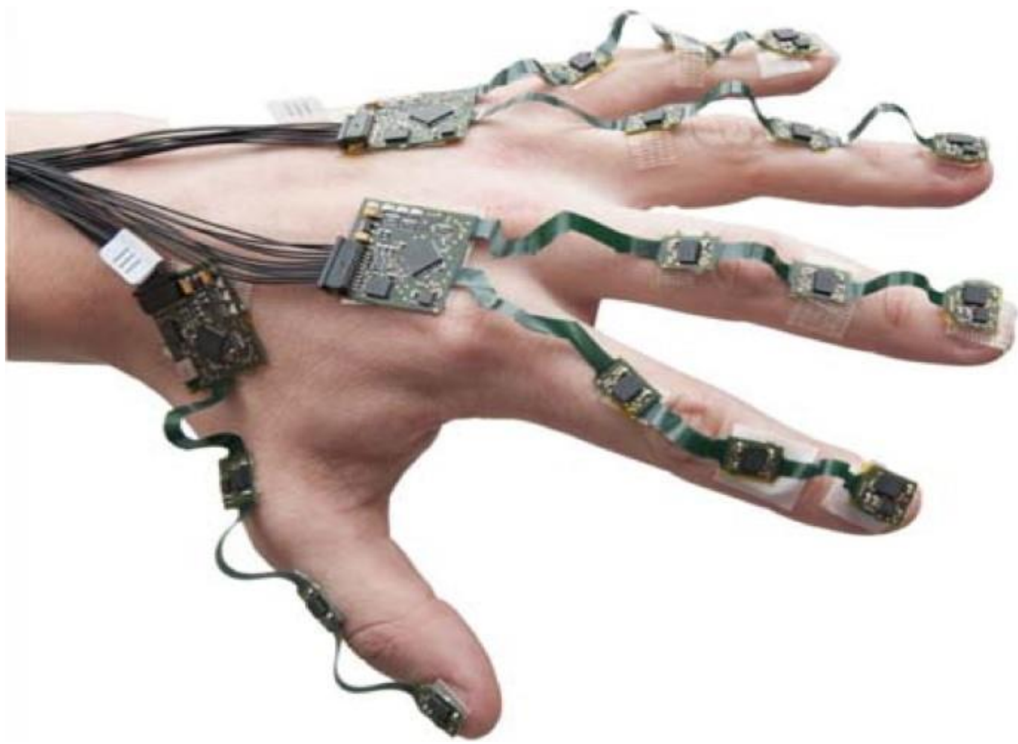

Figure 2.1: Application of the PowerGlove (Kortier et al., 2014).

2.3. Markerless motion capture techniques

2.3.1. Pose estimation

The Microsoft Kinect is a markerless motion capture system that can be used for pose estimation. The Microsoft Kinect V2 (Microsoft Corporation, Redmond, Washington). is a motion-sensing device that leverages a combination of cameras, microphones, and infrared sensors to track human movement (Cai et al. (2019). The Microsoft Kinect V2 has been previously used in many applications, from biomechanical assessments to rehabilitation monitoring (Capecci et al., 2016; Dolatabadi et al., 2014; Lee et al., 2015; Schmitz et al., 2014; Metcalf et al., 2013; Steinbach et al., 2020). Studies have compared the validity of the Microsoft Kinect to that of marker-based motion capture systems (Schmitz et al., 2014; Cai et al., 2019; Guess et al., 2017). In the upper extremity, it was found that the validity of the Kinect V2 is contingent on the task and plane under consideration (Cai et al., 2019). The Kinect V2 demonstrated high correlations in measuring angular waveforms of shoulder and elbow flexion/extension, with a coefficient of multiple correlations (CMC) greater than 0.87. By nature, however, this value relies heavily on the range of motion data, and a smaller range of motion can result in falsely lowered CMC values (Roislien et al., 2012). However, the Microsoft Kinect V2 cannot track hand and finger motion. Multiple Kinect V2 camera setups are not possible, indicating there may be a limitation to synchronizing cameras to combat this limitation, and 3D segment, and joint angles can not be produced (Zhang et al., 2015).

2.3.2. Pose estimation using computer vision and neural networks

Computer Vision is a field of computer science and artificial intelligence (AI) focused on developing algorithms and systems to understand and interpret the visual world (Szeliski et al., 2022). The most common use of computer vision is pose estimation. Pose estimation determines the position and orientation of an object or person in each image. There are multiple approaches when using computer vision for pose estimation, but the most common approach is based on deep learning, a machine learning subfield based on artificial neural networks (LeCun et al., 2015). Several existing programs and systems are available to be used for markerless motion capture:

I. ***Theia3D Markerless Motion Capture:*** Theia3D (Theia Markerless, Inc., Kingston, Ontario) is a commercial deep learning-based program for markerless motion capture that uses a deep convolutional neural network to recognize humans and body segments in 2D camera footage (Mathis et al., 2020). The neural network is trained on over 500,000 images of humans in various settings, clothing, and activities, with 51 features, such as joint locations and anatomical features, manually labeled by highly trained annotators and quality controlled. This system currently has no application for the hands and fingers.

II. ***DeepLabCut****:* DeepLabCut (DLT) is a toolbox for markerless, deep learning-based video-tracking animal and human movements (Mathis et al., 2018). DLT uses a deep neural network to detect the positions of anatomical landmarks on an animal or human body, which are then used to track the body's movements over time (Mathis et al., 2018). However, this program is time-consuming and takes several images to train the model. But it has potential to be used for the hands and fingers.

III. ***PitchAI™ Markerless Motion Capture****:* PitchAI™ (pitchAI™; 3MotionAI Inc, Oakville, ON, Canada) is a markerless motion capture system that uses a single camera to estimate the 3D joint angles in the sagittal plane. It trained its machine-learning model using marker-based data and tracks the positions of anatomical landmarks. The system calculates joint angles and angular velocities of the upper and lower extremities to produce a standardized biomechanics report based on the pitching motion, breaking down the pitch phases (Dobos et al., 2021). However, PitchAI™ is only used for pitching mechanics and has no hand and finger kinematics application.

IV. ***DARI Motion Capture****:* Dynamic Athletic Research Institute motion capture (DARI Motion, Overland Park, KS) is a markerless motion capture system that creates a full-body skeleton for motion capture using cloud voxels (Martinez et al., 2018). Cloud voxels represent 3D objects and space in computer vision using a cloud of points or voxels (Zhou et al., 2018). Each voxel represents a small volume in 3D space, and through mapping the voxel density, objects can be reconstructed, revealing their shapes, sizes, and locations. DARI motion has two models created with 400,000 human movement data and 55,000 subject data based on age and sex with 55,000 other subject data. Currently, no model exists for the hands and fingers in this system.

V. **Pose2Sim:** Pose2Sim takes OpenPose, a known software that can provide 2D joint coordinate predictions for video data. These points are then used to determine 3D joint position data within OpenSim. It achieves this by allowing the user to create a comprehensive or existing musculoskeletal model, adjust it to fit individual subjects and perform inverse kinematics with adjustable biomechanical restrictions. Pose2Sim also offers additional capabilities, such as computing joint moments or determining individual muscle forces (Pagnon et al., 2022). No model has been used and is available for the hands and fingers.

VI. **OpenCap:** OpenCap (OpenCap, Stanford, CA) is web-based, open-source software that analyzes human movement dynamics by estimating 3D kinematics and kinetics from captured videos. Using computer vision and musculoskeletal simulation advancements, OpenCap can conduct movement analysis and has been demonstrated to have sufficient accuracy in estimating kinematic measures. (Uhlrich et al., 2022). OpenCap is mainly used for whole-body motion and gait analysis; no current model is available for the hands and fingers.

2.4. Markerless vs marker-based motion capture in the lower extremity

Several studies have examined the accuracy of markerless motion capture techniques compared to marker-based systems for the lower extremities (Kanko et al., 2021; Strutzenburger et al., 2021; Pagnon et al., 2022; Sandau et al., 2014; Zhang et al., 2014; Uhlrich et al., 2022). Examining six recent papers published in 2021 and 2022, some inferences can be made regarding the consistency in reporting the relationship between markerless motion capture and marker-based motion capture.

Pagnon et al. (2022) was the only study to report a CMC in their markerless motion capture study. It analyzed three tasks, walking, running, and cycling. However, the study only involved a single participant. Despite this limitation, the study highlights the differences in CMC values among the three tasks. The most striking differences were observed between walking, cycling, and running, particularly in the hip flexion/extension and hip abduction/adduction movements. The correlation value for hip flexion/extension during running was 0.65 and 0.37 for hip abduction/adduction. The remaining correlations across the joints and movements ranged from 0.74 to 1.00, reflecting a high correlation between markerless and marker-based motion capture. It is believed that the lower CMC values in the running task resulted from the movements' faster and less controlled nature. Root mean square errors (RMSE) were also reported in some studies. Firstly, three studies reported internal/external rotation of the hip (Uhlrich et al., 2022; Sandau et al., 2014; Kanko et al., 2021), and two studies reported internal/external rotation of all lower extremity joints. Internal/external rotation had error values consistently larger than all the other joint actions. Kanko et al. (2021) stated that when analyzing internal/external rotation, there was an increase in variability in marker-based motion capture across all participants. Significant errors were reported when comparing the difference in hip flexion/extension in the Kanko et al. (2021) and Strutzenberger et al. (2021) papers, with the RMSE being 11° and 20.6°, respectively. Having a high RMSE for a simple hip flexion/extension movement is not ideal and needs further analysis to understand why this was the result of both studies. However, hip flexion/extension also had the largest RMSE with 6.75° compared to other non-complex joint actions, such as hip abduction/adduction, with an RMSE of 3.17° in the Uhlrich et al. (2022) study, meaning there might be underlying issues when attempting to estimate hip flexion/extension. The studies reviewed show that markerless motion capture can be used for lower extremity kinematics as it closely compares to marker-based motion capture, and a low RMSE is common across most joint actions of the lower extremity. However, markerless motion capture is not only applicable to the lower extremity; its capabilities in the upper extremity are limited and need to be evaluated to determine any gaps in the current state of the art.

2.5. Markerless vs marker-based motion capture for the upper extremity

Upper extremity motion has been assessed using various deep-learning approaches (Lahkar et al., 2022; Geelen et al., 2021; Dobos et al., 2022; Fleisig et al., 2022; Pagnon et al., 2022; Kanko et al., 2021; Mathis et al., 2018; Nath et al., 2019). Limited studies have explored markerless motion capture in the upper extremities and assessing joint kinematics (Fleisig et al., 2022). Dobos et al. (2022) and Fleisig et al. (2022) examined baseball pitching, assessing the shoulder, elbow, trunk, pelvis, and knee kinematics. Lahkar et al. (2022) used Theia3D as their software to assess boxing mechanics by assessing three joint locations of the upper extremity: the shoulder, elbow, and wrist. Geelen et al. (2021) used DeepLabCut to assess the motion of the distal interphalangeal, proximal interphalangeal, and metacarpophalangeal joints of the index finger. Pagnon et al. (2022) assessed shoulder joint kinematics as a secondary measure in their study. The joints captured across all studies were the shoulder and elbow, with Lahkar et al. (2022) being the only study to measure the shoulder, elbow, and wrist. Beginning with the elbow joint, only flexion/extension was reported (excluding Lahkar et al., 2022). Between Pitch AI and Theia3D, R^2 ranged between 0.90 and 0.99 and RMSE from 7.4° to 10.53°. These results are expected when analyzing the correlation between the systems and marker-based motion capture due to the elbow joint having one degree of freedom in flexion and extension. The error, however, is concerning, similar to the hip flexion and extension movements in the lower extremity, where a simple movement has an extensive range of error.

A significant difference in RMSE was reported across joint motions, which included shoulder abduction/adduction, external/internal rotation, and flexion/extension. Pose2Sim (Pagnon et al., 2022) did not correlate well with the other three markerless systems, with correlations reported as low as 0.17 in shoulder external/internal rotation and as high as 0.91 in the shoulder joint. However, this was a case study. Different tasks might dictate the system's effectiveness in tracking joint centers, and this can be attributed to how fast the movement is or if any occlusion occurs during the movement.

The shoulder had a larger RMSE than other joints of the upper extremity (Dobos et al., 2022; Lahkar et al., 2022). Shoulder abduction/adduction between the two studies ranged between 6.6° – 7.3°, shoulder external rotation ranged between 12° – 16.7° and shoulder flexion/extension was 10°, only being reported by Lahkar et al. (2022). This might be attributed to the shoulder's more complex. However, it is essential to note that as they progressed down the distal upper extremity, the correlation between systems was lower than the shoulder and elbow joints. The R^2 at the wrist ranged between 0.31 with an RMSE of 11° and 0.41 with an RMSE of 14° for wrist abduction/adduction and flexion/extension, respectively. This might be attributed to the motion analysis, the speed of the movement, and the difficulty measuring wrist angle with boxing gloves placed on the athletes.

Lastly, Geelen et al. used DeepLabCut (DLC) as a means of markerless motion capture for hand kinematics (Geelen et al., 2021). This study analyzed the index finger and its three joints: the metacarpophalangeal (MCP) joint, the proximal interphalangeal (PIP) joint, and the distal interphalangeal joint (DIP) joint. Index finger motion was also captured using marker-based motion capture to compare joint angles between the two systems. Correlation between the joint angles computed from both systems was not reported. Figure 2.2 is a time series graph created to qualitatively represent the three joint angles to compare their markerless motion capture and Qualisys system. RMSE was reported for all joint angles, resulting in an RMSE of 7.5° for the MCP joint, 3.2 For PIP, and 2.3° for DIP. However, because they only analyzed one finger in one participant, the results of this study need to be reviewed with caution. However, there is potential that markerless motion capture of the hands and fingers can be achieved and is comparable to marker-based systems.

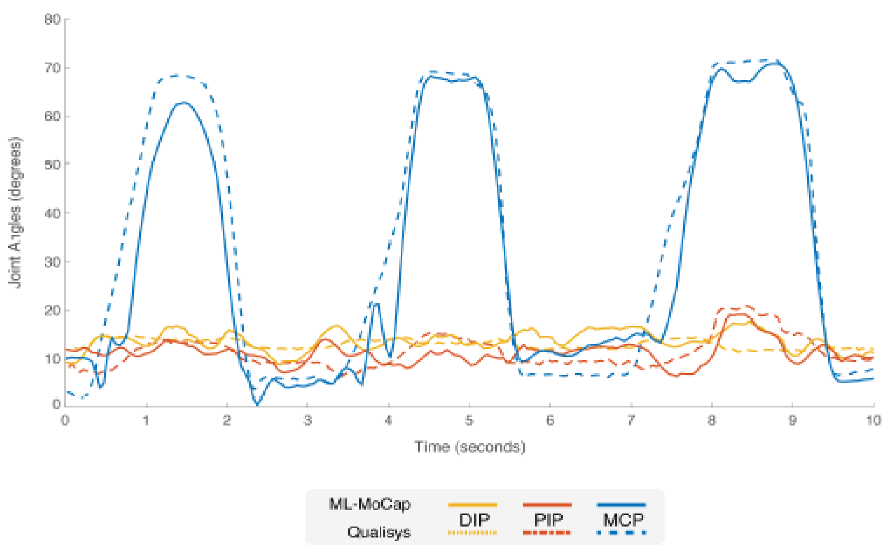

Figure 2.2: Markerless motion capture joint angles of the index finger compared to a marker-based system (Geelen et al., 2021).

2.6. Markerless hand and finger tracking model

MediaPipe Hands (MPH) is a hand and finger tracking solution developed by Google Research (Zhang et al., 2020). MPH can track 21 hand landmarks using a single camera, providing X, Y, and Z coordinates (with the Z-value derived from the image depth map). MPH utilizes two models: a palm detection model and a hand landmark model, working in conjunction. The palm detection model is designed to detect the rigid palm structure, which simplifies the process compared to detecting the entire hand and fingers. This model achieves an average accuracy of 96%. Once the palm is detected, the hand landmark model is introduced, recognizing 21 hand landmarks and outputting joint coordinates for both left and right hands (Figure 2.3). The hand landmark model has been trained on 30,000 images, including rendered synthetic hand models.

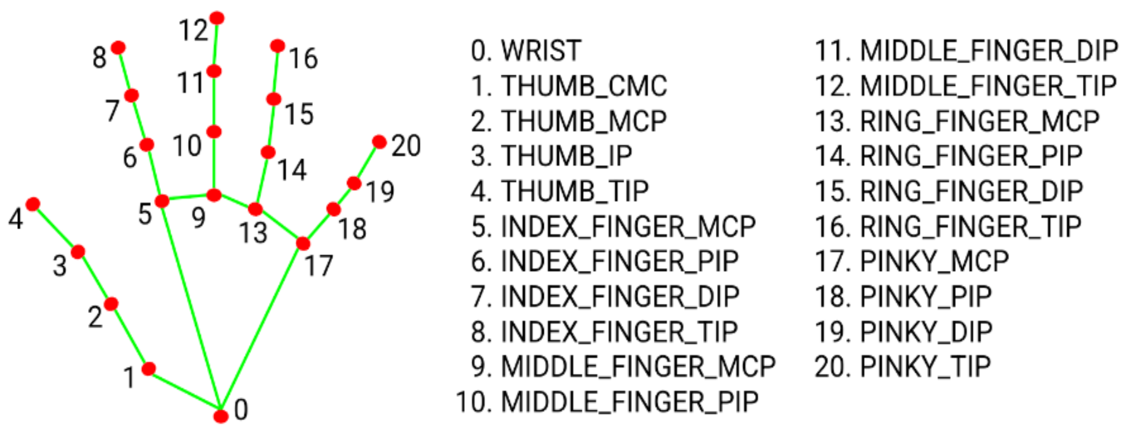

Figure 2.3: The 21 landmarks that MediaPipe Hands can track (MediaPipe Hands, Google Research).

MPH has been evaluated on its capability to track hand movements under various conditions. First, a geographical evaluation was conducted using 700 images, with 50 images from each of the 14 geographical subregions. Next, MPH was evaluated based on skin tone and gender using 420 images, with 35 images from each unique combination of perceived sex and skin tone. Lastly, MPH was also evaluated based on its performance by sex. For a complete overview of these evaluations, refer to MediaPipe Hands' model card for their documentation.

Currently, MPH has yet to be tested in biomechanical assessments. Guney et al. used MPH to measure tremors in Parkinson's patients following intervention to determine whether it would reduce the associated tremors (Guney et al., 2022). Aside from this study, no other work has been conducted using MPH to track hand and finger kinematics. However, it has the potential to be used to create a markerless motion capture program where 3D coordinates can be outputted.

2.7. Limitations to pose estimation

Markerless motion capture using computer vision and pose estimation effectively measures joint kinematics, but that is with limitations. Large volumes of annotated data are required for training markerless motion capture systems, such as Theia3D, which was trained on over 500,000 frames. Theia3D and DARI Motion are commercial programs that have extensively trained their networks and programs. This is the most significant limitation to using computer vision techniques because they inherently work only based on what they have been trained on, unlike marker-based motion capture, where any motion desired can be tracked.

2.8. Pilot Study

My previous pilot work was conducted with DeepLabCut (DLC) for tracking hand and finger kinematics (Majoni et al., 2022; Appendix C). DLC's object recognition paradigm utilizes iterative machine learning, where a training dataset is used to train the CNN model to communicate with the training dataset for a specific number of rounds. The longer this process occurs, the more confident the CNN model becomes in predicting the points of interest labelled within the training dataset (Liu et al., 2017). The model was trained on 1400 digital frames from two of the three participants' trials, with one trial randomly withheld for testing purposes. K-means clustering, a function within DLC, was used to extract 30 frames from each video, the most digitally different frames from the video data. These 30 frames were then manually labelled by trained annotators for a total of 21 points of interest on the hand for each frame, and two annotators were used to control for variability between the labelling of the joints of interest. The training dataset included various postures from each task recorded and various skin tones and hand sizes. The training images were used to train the model, which required 500,000 iterations to reach a plateau in learning (Nath et al., 2019).

DLC provided an avenue to understand the information included and taught in the hand model, thereby ensuring that biases in training do not influence the hand model's performance. However, an issue arose during early trials, where the impact of skin tone was not initially considered. The model was only trained on individuals with fairer skin, and when darker skin complexions were used to test the hand model, it failed to track hand movements or digitize any hand landmarks. Once a diverse group of skin complexions was introduced, the problem was resolved. Nevertheless, the main limitation of this approach is the need for a sufficiently large sample of frames that could be included in the training dataset. The need to manually digitize each frame made the process time-intensive and required significant manpower. In contrast, Theia3D has trained its neural networks on over 500,000 images. Only 1400 images were used in this study, less than 1% of their total images. Therefore, an alternative approach for creating a markerless motion capture system for the hands and fingers had to be pursued.

2.9. Summary

Using markerless motion capture for kinematic analysis has addressed the current issues when using marker-based and other motion capture systems (IMMS, Microsoft Kinect). Several pose estimation algorithms and programs have been developed with the advancements in markerless motion capture. The application of computer vision techniques in joint kinematics estimation has demonstrated efficacy in the lower extremity and not in the upper extremity. The current markerless motion capture systems extend only to the wrist in the upper extremity and not any further. The representation of hand and finger kinematics in current markerless motion capture systems is limited. Implementing a markerless motion capture system capable of tracking hand and finger kinematics can expand its application beyond the laboratory setting, facilitating real-time movements in various environments, including the workplace, and enhancing the capacity for biomechanical and ergonomic assessments.

CHAPTER THREE: MANUSCRIPT

MARKERLESS MOTION CAPTURE OF THE HANDS AND FINGERS

Nigel Majoni, Daanish M. Mulla, Peter J. Keir*

Department of Kinesiology, McMaster University, Hamilton, ON, Canada

*Corresponding Author:

Peter J. Keir, PhD

McMaster University

Department of Kinesiology

Ivor Wynne Centre

1280 Main Street West Hamilton, ON, Canada, L8S4K1

Telephone: 905-525-9140 ext. 23543

Email: pjkeir@mcmaster.ca

Prepared for submission to: Journal of Biomechanics

Word Count: 4859

Keywords: Markerless motion capture; Finger kinematics; MediaPipe Hands; Computer vision;

3.1. ABSTRACT

Hand and finger movements are underrepresented in biomechanical studies, primarily due to the challenge of tracking the hands and fingers. Several limitations are associated with marker-based motion capture, including interference with natural movement, and require the tedious, time-consuming application of numerous markers. Advancements in computer vision have led to the development of markerless motion capture systems yet validation of markerless systems for the upper extremities is limited, especially the hand and fingers. The purpose of this study was to develop and assess a markerless motion capture system capable of tracking hand and finger kinematics. A markerless system using four synchronized webcams was developed. Camera pairs were organized in different angles Centre90° (C/90°), Left45°/Right45° (L45°/R45°), and Centre/Left45° (C/L45°). Motion capture was performed with both marker-based and markerless systems. Twenty healthy participants performed five dynamic hand tasks with and without markers. Three-dimensional joint positions were defined using a musculoskeletal model in OpenSim. No significant differences were observed between C/90° and C/L45° markerless camera pairs and the marker-based system. The L45°/R45° camera pair differed significantly from other markerless pairs in several tasks but agreed with the marker-based system for the index finger during flexion. For most of the fingers, no significant differences were found across the different camera pairs. Correlations and error for the concurrent finger flexion task revealed high consistency among all the camera pairs, with R^2 above 0.90 and RMSD below 10°, the thumb showed greater variability. The R^2 and RMSD varied depending on the camera comparison and finger for

each task. Markerless motion capture for the hands and fingers is possible with little difference to marker-based systems and is dependent on the camera orientation used.

3.2. INTRODUCTION

The hands are our primary tools for interacting with the external environment and are at the forefront of our daily activities. Hand and finger movements are often underrepresented in biomechanical and ergonomic studies (Amell et al., 2001). This research aims to provide a methodology to capture the actions of the hands and fingers, with the goal of expanding the scope of assessing the risk of developing hand-related musculoskeletal disorders in various settings.

Tracking upper limb movements, particularly the hands and fingers, is challenging. Hand and finger kinematics have been tracked using various methods, including optical motion capture (both active and passive), inertial sensors, goniometers, and markerless motion capture. However, several limitations are associated with these methods when tracking the hand and fingers. For instance, while marker-based motion capture is considered the industry-standard method, it is prone to marker occlusion, interferes with natural movement, is time-consuming, and requires the tedious application of numerous markers (Cappozzo et al., 1995; Cheung et al., 2005; Cocchiarella et al., 2015). Advancements in computer vision and machine learning have led to the development of markerless motion capture systems.

Currently, several applications exist for markerless motion capture, including (i) Theia3D, (ii) OpenCap, (iii) OpenCV, (iv) OpenPose and (v) DeepLabCut (Kanko et al., 2021; Uhlrich et al., 2022; OpenCV., 2015; Cao et al., 2017; Mathis et al., 2018; Nath et al., 2019). *Theia*3D (Theia Markerless, Inc., Kingston, Ontario) has been validated against conventional marker-based systems (Kanko et al., 2021; Wren et al., 2023;

Steffensen et al., 2023). OpenCap is a two-camera motion capture system that has also been validated against conventional marker-based systems and found little difference in performance between 3 and 5 cameras (Uhlrich et al., 2022; Lima et al., 2023). However, validation of markerless systems for the upper extremity is limited, and no existing markerless motion capture system provides accurate hand and finger kinematics tracking. Consequently, the development and assessment of markerless motion capture systems for hand and finger kinematics is needed. The purpose of this study was to develop a two-camera markerless motion capture system to track 3D hand and finger kinematics, compare its kinematics to a marker-based motion capture system, and evaluate the consistency of joint angles produced by different two-camera setups.

3.3. METHODS

3.3.1. Participants

Twenty healthy participants completed the study (age: 23.6 ± 3.5 years; height: 173.3 ± 10.5 cm; mass: 69.5 ± 11.8 kg). Ten males (age: 25.4 ± 3.7 years; height: 181.2 ± 7.4 cm; mass: 77.4 ± 9.5 kg) and ten females (age: 21.7 ± 2.1 years; height: 165.4 ± 6.2 cm; mass: 61.6 ± 8.0 kg) were recruited from the university and provided written informed consent before participating in the study. McMaster Research and Ethics Board provided clearance for this study (#6200). Exclusion criteria included upper-extremity injury in the last six months before participation.

3.3.2. Development of markerless motion capture system

The markerless motion capture system consisted of four synchronized webcams (C920e Logitech, Newark, CA, USA) with a resolution of 480x640, sampling at 30 Hz (vMix, StudioCoast Pty Ltd.). Our setup was comprised of three pairs of cameras configured as follows: two cameras positioned at 45° to the left and right of the collection space (L45°/R45°); one camera placed in the center of the collection space and one camera at 90° to the collection space (C/90°); and a third pair combining the center camera with the left 45° camera (C/L45°) (Figure 3.1). The markerless motion capture system used an existing trained hand and finger tracking model in conjunction with OpenCV (MediaPipe Hands, Google Research). MediaPipe Hands consists of two existing hand models, a palm detection model, and a hand landmark model, working in tandem (Zhang et al ., 2020). The hand landmark model was trained on 30,000 images and rendered synthetic hand models. MPH has also been evaluated on sex and skin tone.

MPH tracks 21 hand landmarks with a single camera and provides X and Y coordinates for each marker (Figure 3.2).

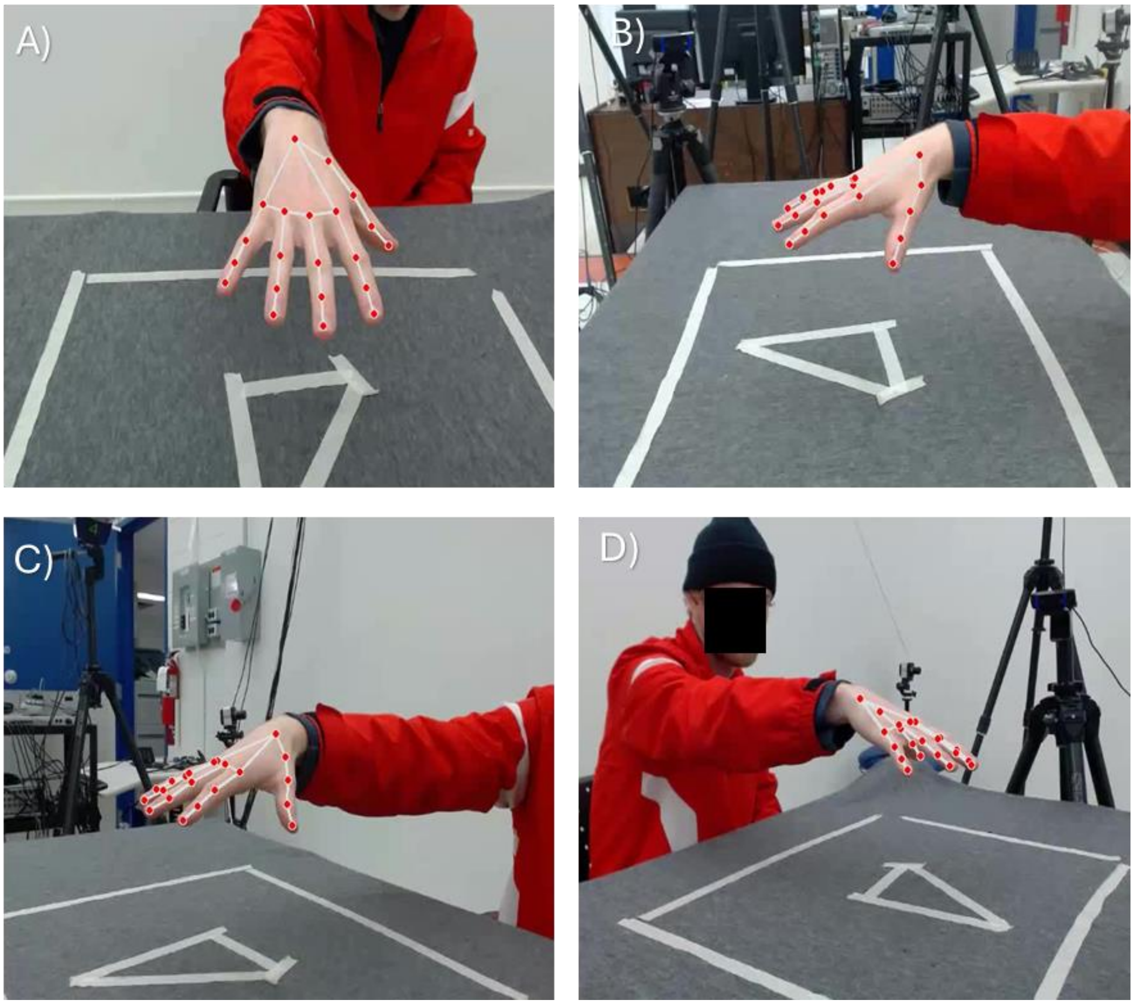

Figure 3.1: The four images on the left illustrate the different perspectives captured by the cameras: A) the Centre camera, B) the 90-degree camera (90°), C) the Left 45-degree (L45°) camera, and D) the Right 45-degree (R45°) camera.

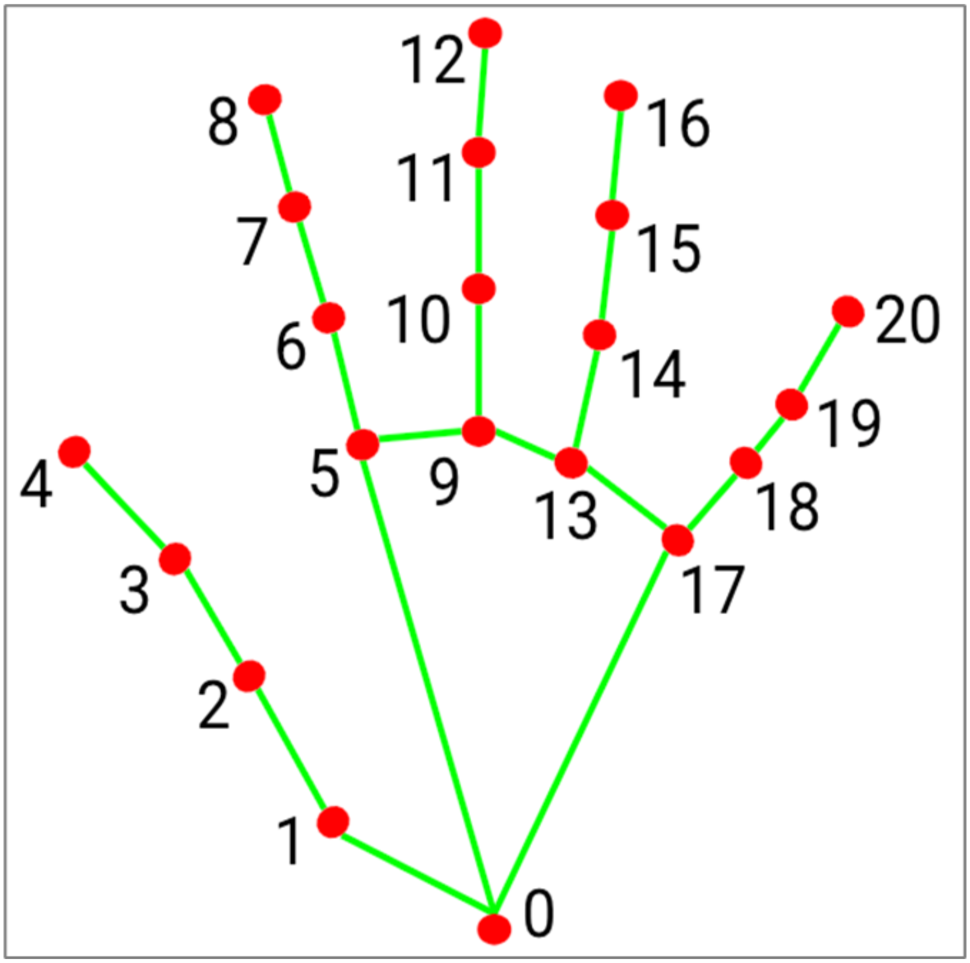

Figure 3.2: The 21 joint landmarks predicted by MediaPipe are highlighted in red. On the right, are the same 21 joint landmarks not on the hand (MediaPipe Hands, Google Research). 1-4 = CMC, MCP, IP, TIP of the thumb. 5-8 = MCP, PIP, DIP, TIP of the index finger. 9-12 = MCP, PIP, DIP, TIP of the middle finger. 13-16 = MCP, PIP, DIP, TIP of the ring finger. 17-20 = MCP, PIP, DIP, TIP of the little finger.

To obtain three-dimensional coordinates, a series of calibration steps were performed. For each camera, we determined the internal characteristics (intrinsic parameters by capturing twenty still images of a 4-row by 7-column checkerboard. The checkerboard was shown to each camera individually and moved to different positions within the frame to ensure it was clearly visible. Capturing twenty images was intended to provide sufficient frames with a clearly visible checkerboard, allowing us to exclude any poor-quality frames due to any movement. To determine if we had a good calibration, an RMSE was outputted for each camera. The RMSE represents the pixel projection error, quantifying how accurately the calibration aligns the projected points with the actual points on the checkerboard. Our target RMSE was less than 0.50; any calibration resulting in an RMSE above this threshold was repeated until the desired accuracy was achieved. Next, for each camera pair (L45°/R45°, C/90°, and C/L45°), we determined the relative position and orientation (extrinsic parameters) by holding the checkerboard in view of both cameras and capturing 20 simultaneous paired images. These paired images were used to calculate the rotation matrix and translation vector, which describe how one camera is positioned and oriented relative to another to derive a global coordinate system. An RMSE threshold of less than 0.50 was used to determine good accuracy. Three-dimensional coordinates are then determined using a direct linear transform (DLT).

3.3.3. Experimental setup and protocol

Motion capture was completed using two systems: (i) standard motion capture and (ii) markerless motion capture. Twelve Raptor-4 cameras (Motion Analysis Corporation, Santa Rosa, CA, USA) provided the hardware for the standard motion capture system. Marker trajectories of the marker-based system were collected at 30 Hz to match the markerless motion capture system capture rate at 30 frames per second (FPS). Before placing the reflective markers on our participants, we recorded their anthropometric measurements, including height, weight, age, and sex. Additionally, we measured their hand dimensions, including hand length, width, breadth, and the lengths of each finger. A total of seventy-seven reflective markers were used: 19 calibration markers and 58 tracking markers, placed on the finger joints and segments (Figure 3.3).

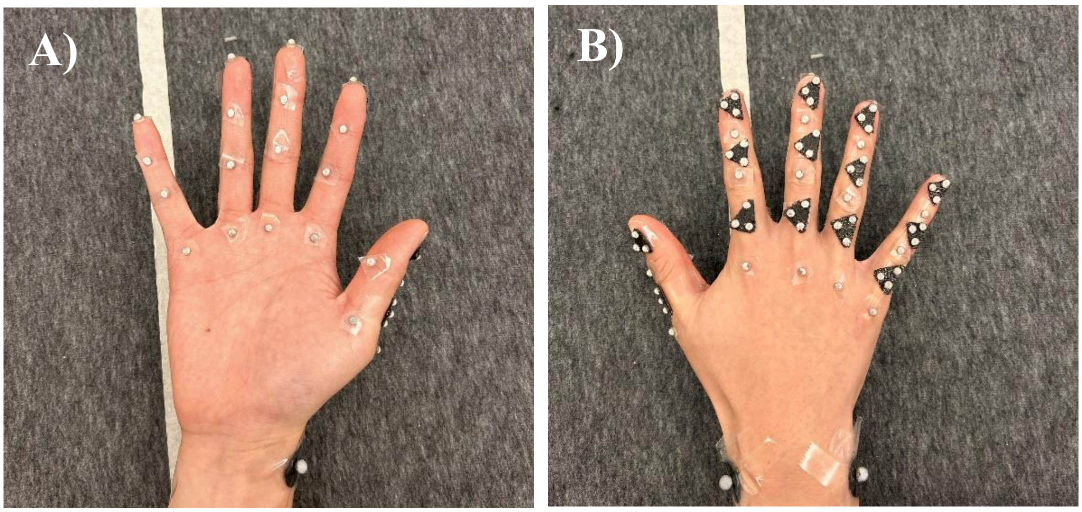

Figure 3.3: Marker set used to obtain motion using Motion Analysis. 77 hemispherical markers (4 mm diameter) consisting of A) 19 Calibration markers (Palmar surface) and B) 58 Tracking markers (Dorsum) were outfitted on the right hand.

Data was collected in one 2-hour session for each participant. All participants were seated with their right arm fully extended into the capture volume, ensuring their hand was positioned above the origin space. They were instructed to maintain this posture while completing each task to ensure consistency across all measurements. Participants then performed a series of five dynamic hand tasks, each 30 seconds in duration and repeated three times. The 5 tasks included (i) index finger flexion, (ii) concurrent finger flexion of digits 1-5, (iii) sequential finger flexion (i.e. flexion of digits 1-5), (iv) typing of a set text passage, and (v) object manipulation task (pinch grip in different postures - Jenga^{TM} task). During tasks i, and ii, participants followed the beat of a metronome at 60 BPM and 100 BPM during task iii. For tasks iv (typing) and v (object manipulation), participants were given 30 seconds to complete the task. Participants were given a 1-

minute (minimum) break after each task to mitigate any fatiguing effects. Participants performed the five tasks with (concurrent) and without (non-concurrent) reflective markers.

3.3.4. Data analysis

Three-dimensional joint positions were defined in the markerless motion capture system defined the joint positions (MCP, PIP, DIP for digits 2 to 5, and MCP and IP for digit 1) A second-order, 6 Hz dual-pass Butterworth filter was applied to the motion capture data. Joint angles for both the markerless and marker-based data were then computed using OpenSim's inverse kinematics function, using a musculoskeletal model of the hand and wrist (McFarland et al., 2022). Two models were used to calculate our kinematics: one for the markerless motion capture data and one for the marker-based motion capture data. The only difference between them was the number of markers on the model due to the different marker sets—58 reflective markers for the marker-based system and 21 virtual markers for the markerless system. All the joint centres were in the same location except for the TIP marker for each of the digits exclusive to the markerless system. Each model was scaled to the participant's hand measurements, which were obtained by manually measuring hand length, width, and digit lengths. Constraints were added to the models to ensure the computation of realistic joint angles. For digits 1-5, the MCP joint was constrained to a range of -5° extension to 90° flexion. For digits 2-5, the PIP joints were constrained to a flexion range of 0° to 100°, and the DIP joints were constrained to a flexion range of 0° to 80°. For digit 1, the IP (thumb) joint was constrained to a flexion range of 0° to 100°.

A total finger angle (θ_{finger}) was calculated as the sum of the MCP, PIP, and DIP joint angles for digits 2-5 and MCP and IP for digit 1 [Equation 1]. We utilized the total finger angle to condense the data, providing a representative measure of the finger's overall movement rather than individual joint angles. This approach focuses on a single, comprehensive angle per finger.

$$\theta_{\text{finger}} = \theta_{\text{MCP}} + \theta_{\text{PIP}} + \theta_{\text{DIP}} \qquad [1]$$

Due to the interference from the reflective markers on our markerless joint predictions during our pilot work we found that the predictions were not correctly annotated on the participants' hands, and the hand model appeared smaller with shorter segment lengths. We made a strategic decision to primarily analyze descriptive statistics for the session without reflective markers. We used our concurrent marker-based kinematics and compared those to our non-concurrent markerless kinematics. Task (v), the object manipulation task (pinch grip in different postures – Jenga™ task) was not analyzed due to actual and virtual marker loss caused by occlusion in both systems. Amplitude probability distribution functions (APDFs) of the total finger angles were generated for each task, with the 1st (minimum), 50th (median), and 99th (maximum) percentiles. To ensure an equal comparison between marker-based and markerless systems, five participants were excluded from the analysis due to hardware issues and noisy data from the marker-based motion capture, resulting in 15 participants being analyzed. For the analysis comparing the marker-based and markerless systems, we also excluded these participants' markerless data to maintain equal numbers. The APDFs were calculated from all three trials per participant. R^2 and RMSD between each of the three camera pairs, and in this analysis, data from all twenty participants was included.

A factorial ANOVA was conducted to assess finger angles for the three markerless camera pairs and marker data. For each task and percentile (4 tasks x 3 percentiles), a two-way ANOVA was employed with camera pair orientation and finger as factors. A factorial ANOVA was performed to examine differences in R^2 and RMSD across the markerless motion capture camera pairs, reported as mean [95% confidence interval]. For each task and measurement (4 tasks x 2 measurements), a two-way ANOVA with camera pair and finger as factors was utilized. Post-hoc pairwise comparisons were conducted using Tukey's method for all tests with significant interaction effects. An alpha level of 0.05 was set for all analyses. Statistical analyses were performed using RStudio (v4.3.3).

3.4. RESULTS

3.4.1. Markerless motion capture vs marker-based motion capture

Sample trials for the concurrent finger flexion, sequential finger flexion, index finger flexion and typing tasks of what the index finger angle motion is for each task can be observed in Figure 3.4. The representation of the APDFs for one participant for each finger is depicted, showing how the 1^{st}, 50^{th}, and 99^{th} percentiles were calculated (Figure 3.5). This analysis was extended to include all fingers, with similar results on the middle, ring, little and thumb provided in the supplementary documentation (Table S1). The results for the index finger will be reported for each task, comparing the concurrent marker-based kinematics with the non-concurrent markerless kinematics results. Significant interactions ($p < 0.001$) between camera orientation and finger were identified in several conditions: the 50^{th} percentile sequential finger flexion task, the 50^{th} percentile concurrent finger flexion task, the 50^{th} and 99th percentile index finger flexion task, and the 50^{th} percentile typing task (Table S1). Significant main effects were observed for camera orientation ($p < 0.001$) and finger ($p < 0.001$) for all percentiles (1^{st}, 50^{th}, and 99^{th}) and all tasks, except for sequential finger flexion at the 1^{st} percentile, where only a significant main effect of finger was found ($p < 0.001$).

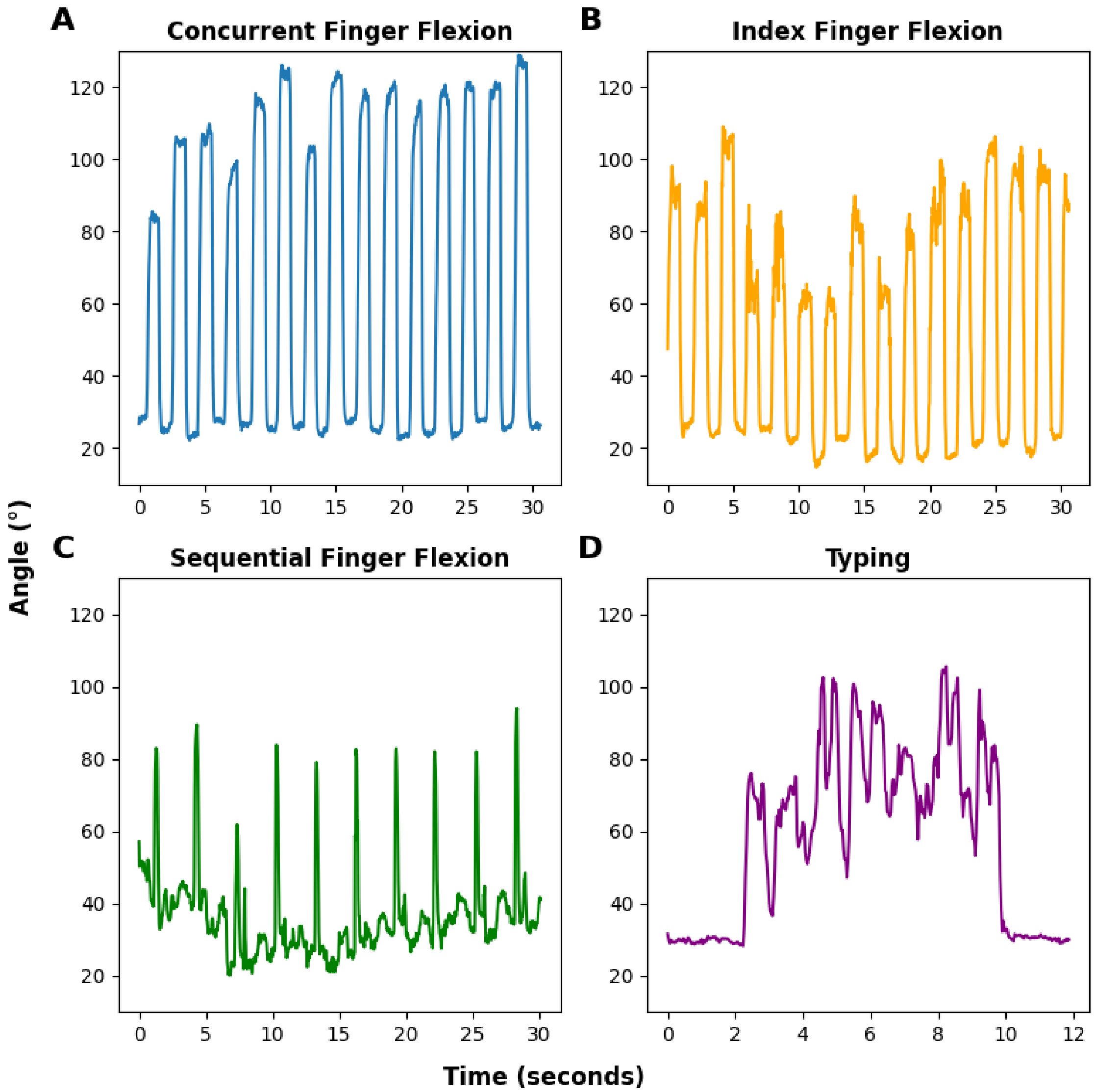

Figure 3.4: Subplots of the index total finger angle for the 4 tasks analyzed for our markerless camera pairs: (A) Concurrent Finger Flexion, (B) Index Finger Flexion, (C) Sequential Finger Flexion, and (D) Typing.

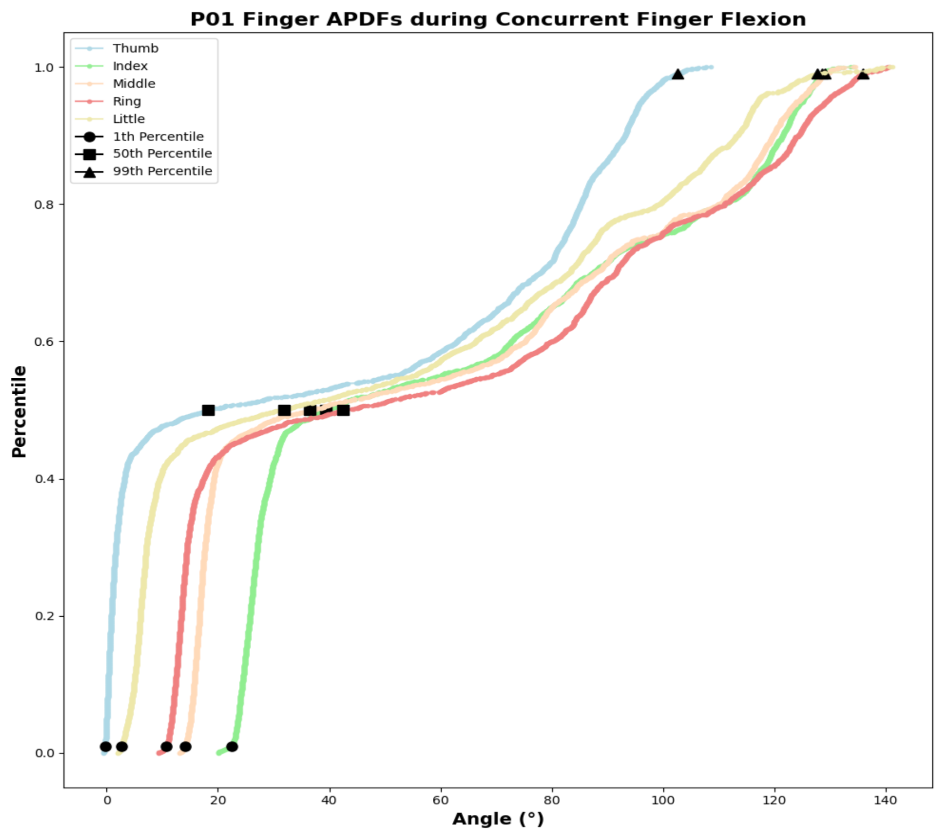

Figure 3.5: Amplitude probability distribution functions (APDFs) for the concurrent finger flexion task for Participant P01.

The means [95% confidence interval] of our concurrent marker-based kinematic with the non-concurrent markerless kinematics result is reported, highlighting the 1st, 50th, and 99th percentiles difference between the camera pairs and the marker-based system for the index finger for each task (Table 3.1). There was no statistically significant difference between the C/90° and C/L45° camera pairs when compared to the marker-based system. This is true for the concurrent finger flexion, sequential finger flexion, and typing tasks, except for index finger flexion at the 99th percentile where the index finger exhibited a mean finger angle of 97.9° [85.6, 110.3°] for the concurrent marker-based measurements, differing from the C/90° and C/L45° camera orientations, which showed angles of 122.3°

[102.5, 128.4°] and 124.1° [112.2, 136.1°], respectively ($p < 0.001$). The L45°/R45° camera pair exhibited statistically significant differences across all tasks, except for sequential finger flexion, when compared to the C/L45° and C/90° markerless camera pairs. Specifically, for the 99th percentile in index finger flexion, the L45°/R45° camera pair differed from the other two markerless pairs, with the index finger angle recorded as 88.3° [76.4°, 100.2°]. This angle did not show a statistically significant difference from the concurrent marker-based measurements, indicating agreement between the L45°/R45° camera pair and the marker-based measurements in measuring the index finger angle during index finger flexion. Statistically significant differences were observed at the 50th percentile for the index finger when comparing the L45°/R45° camera pair to the C/90° and C/L45° camera pairs during typing and concurrent finger flexion tasks. During the concurrent finger flexion task, the 50th percentile index finger angle for the L45°/R45° camera pair was 60.4° [51.6°, 69.1°], compared to 44.1° [35.4°, 52.8°] for the C/90° camera pair and 46.6° [37.9°, 55.4°] for the C/L45° camera pair. Similar trends were observed during the typing task, with the L45°/R45° camera pair showing greater joint angles than the C/L45° and C/90° camera pairs. For the middle, ring, and little fingers, no significant differences were found across all tasks, except for some differences in index finger flexion and the thumb exhibited statistically significant differences, similar to the index finger with the L45°/R45° camera pair. The differences between the marker-based and markerless systems arise because the marker-based system measured finger motions are not identical to those measured by the markerless system, whereas the markerless systems evaluated identical motions for each task.

Table 3.1: Estimated marginal means [95% confidence intervals] for the 1st, 50th, and 99th percentiles (°) for the index finger and all tasks, across different camera orientations. The orientations C/90°, C/L45°, and L45°/R45° refer to the markerless motion capture camera pairs, and Motion Analysis refers to the marker-based

Index					
		Concurrent Finger Flexon	**Sequential Finger Flexion**	**Typing**	**Index Finger Flexion**
C/90°	**1st**	19.24 [14.90, 23.57]	14.90 [9.38, 20.42]	29.76 [24.75, 34.78]	20.83 [16.22, 25.44]
	50th	44.10 [35.37, 52.84]	32.63 [22.22, 43.03]	40.62 [32.04, 49.20]	47.63 [40.94, 54.33]
	99th	116.08 [106.42, 125.74]	115.45 [102.46, 128.43]	115.78 [97.76, 133.80]	122.25 [110.32, 134.17]
C/L45°	**1st**	16.57 [12.23, 20.90]	11.22 [5.70, 16.74]	25.32 [20.31, 30.33]	19.34 [14.73, 23.94]
	50th	46.63 [37.90, 55.37]	30.80 [20.39, 41.20]	37.98 [29.40, 46.56]	55.90 [49.21, 62.60]
	99th	115.39 [105.73, 125.05]	109.30 [96.31, 122.28]	97.56 [79.54, 115.58]	124.14 [112.21, 136.06]
L45°/R45°	**1st**	23.39 [19.05, 27.72]	16.90 [11.38, 22.42]	30.78 [25.77, 35.79]	25.42 [20.81, 30.02]
	50th	**60.39 ***** **[51.65, 69.12]**	37.79 [27.39, 48.20	**59.44 ɣ** **[50.86, 68.02]**	44.96 [38.27, 51.66]
	99th	125.14 [115.48, 134.80]	88.59 [75.60, 101.57]	119.82 [101.80, 137.84]	**88.29 #** **[76.36, 100.21]**
Motion Analysis	**1st**	19.31 [14.83, 23.80]	20.58 [14.86, 26.29]	30.65 [25.26, 36.03]	21.43 [16.67, 26.20]
	50th	43.32 [34.29, 52.36]	42.46 [31.69, 53.23]	43.74 [34.52, 52.96]	52.49 [45.56, 59.42]
	99th	81.34 [71.34, 91.34]	97.88 [84.44, 111.32]	86.66 [67.30, 106.02]	**97.97 Ⱶ** **[85.63, 110.32]**

3.4.2 Markerless motion capture camera pairs comparison

The C/L45°, C/90°, and L45°/R45° camera pairs were compared to each other to measure how they agree with one another using the R^2 and RMSD. This analysis was conducted for each trial for the concurrent finger flexion, index finger flexion, sequential finger flexion, and typing tasks. A sample trial for one participant for the index finger during the concurrent finger flexion task showcasing this analysis comparing the C/L45°, C/90°, and L45°/R45° camera pairs can be observed in Figure 3.6.

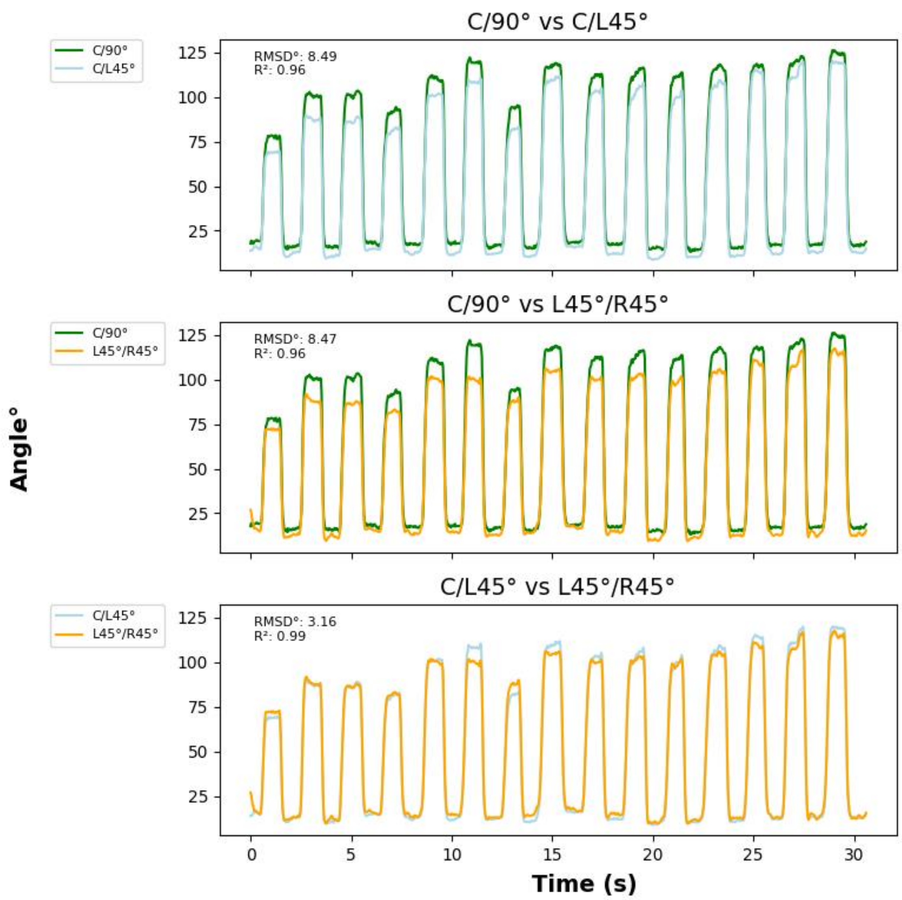

Figure 3.6: Comparison of the index finger angles during the concurrent finger flexion across different camera pairs for P01 during Trial 1. Top: Comparison between C90° and C/L45° camera pairs. Middle: Comparison between C/90° and L45°/R45° camera pairs. Bottom Plot: Comparison between C/L45° and L45°/R45° camera pairs.

Statistically significant interactions between camera pairs and fingers were evident for all tasks and measurements (R^2 and RMSD) (($p < 0.001$). Significant main effects were observed for both camera pairs ($p < 0.001$) and fingers ($p < 0.001$) across all tasks and measurements (R^2 and RMSD). For each task, significant differences were found when comparing different camera pairs and fingers.

In this study, we examined the correlations for concurrent finger flexion of the index, middle, ring, and little fingers across three camera pairs. Overall, correlations were high, during this concurrent finger flexion task ranging from 0.91 to 0.99. For the thumb, the correlation was high when comparing C/90° to C/L45°, with an R^2 of 0.97. However, lower correlations were observed for the thumb when comparing C/90° to L45°/R45° and C/L45° to L45°/R45°, with R^2 of 0.71 and 0.74, respectively. The consistency of the different camera pairs in measuring finger angles varied depending on the task performed. When observing all the fingers during the sequential finger flexion, typing, and index finger flexion tasks the correlation between the camera pairs varied. In the sequential finger flexion task, the index finger had an R^2 of 0.77 when comparing the C/90° to C/L45° camera pairs but drops to 0.34 and 0.38 in the comparison between C/90° vs C/L45 and C/L45° vs. L45°/R45°, respectively. A similar trend is present when also looking at the thumb during the sequential finger flexion task where the thumb had correlations of 0.88 when comparing the C/90° to C/L45° camera pairs but drops to 0.17 and 0.16 in the comparison between C/90° vs C/L45 and C/90° vs. L45°/R45°, respectively. For the middle, ring, and little fingers during the sequential finger flexion task, a similar trend is present, but instead of the C/90° to C/L45° camera pairs performing the best, it is the C/L45° to L45°/R45° camera pairs. The correlation for the ring finger when comparing the C/L45° to L45°/R45° camera pairs was 0.76 but drops to 0.50 and 0.44 in the comparison between C/90° vs C/L45 and C/90° vs. L45°/R45°, respectively. This trend is similar for the typing and index finger flexion tasks, where the index finger and thumb perform the best during the C/90° vs C/L45° comparison, and the middle, ring, and little fingers perform the best during the C/90° vs. L45°/R45° comparison (Figure 3.7).

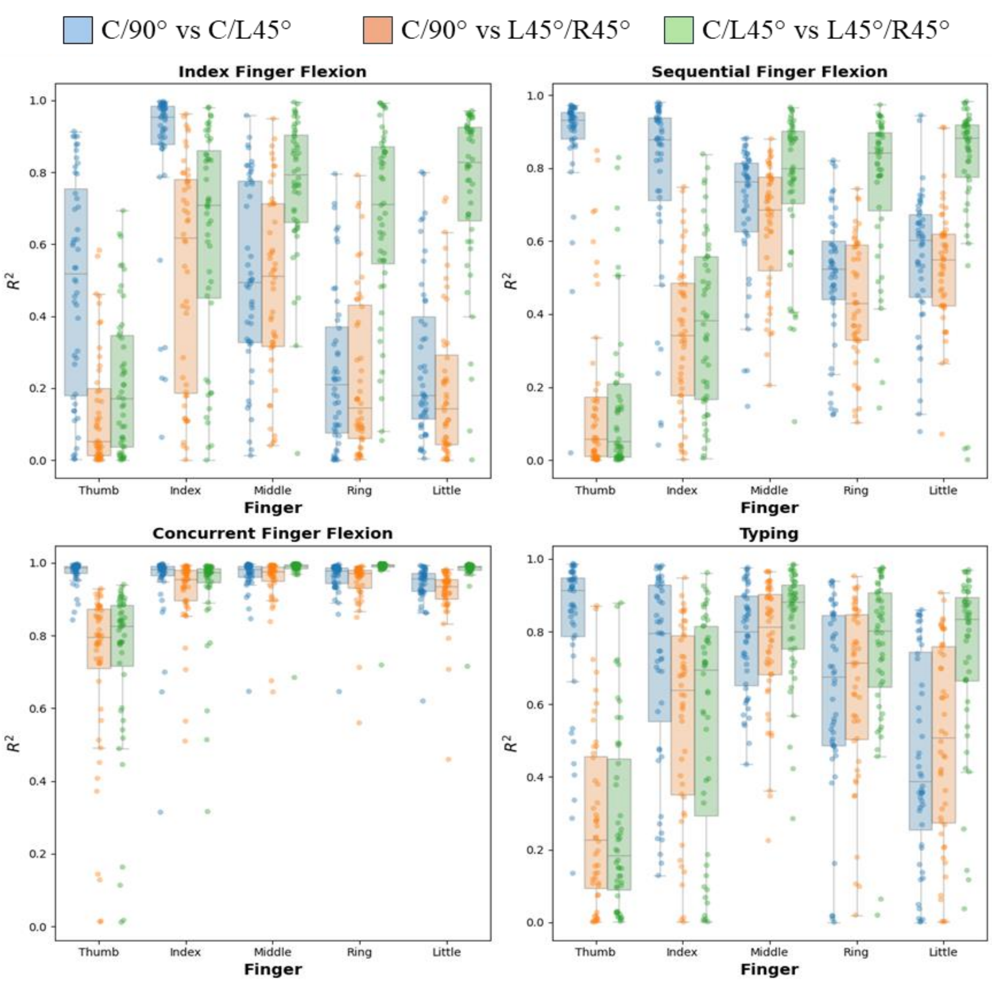

Figure 3.7: Box plot comparison of camera pair R^2 across the thumb, index, middle, ring, and little finger and tasks (index finger flexion, finger roll flexion, finger flexion, and typing) with the markers off (non-concurrent).

For the concurrent finger flexion tasks, the finger angle RMSD for all the camera pair comparisons for the index, middle, ring, and little fingers ranged from 5.7° to 16.8° but the RMSD was closer to 5.7° being ~ 9-10° and for the thumb it showed an RMSD of 6.1° in the comparison for C/90° vs C/L45° but a larger RMSD of 22.4° and 22.1° for the comparison between C/90° vs L45°/R45°and C/L45° vs. L45°/R45°, respectively. When observing the fingers during the sequential finger flexion, typing, and index finger flexion tasks, the RMSD between the camera pairs varied. In the sequential finger flexion task, for the index finger, when comparing the C/90° vs C/L45° camera pairs, the RMSD was 13.1° but increased to 20° and 22° in the comparison between C/90° vs L45°/R45° and C/L45° vs. L45°/R45° camera pairs, respectively. Similar results are present when looking at the thumb during the sequential finger flexion task where the thumb had an RMSD of 10.3° when comparing the C/90° VS C/L45° camera pairs but increases significantly to 39.6° and 38.5° in the comparison between C/90° vs L45°/R45° and C/L45° vs. L45°/R45° camera pairs, respectively. This is similar when observing the RMSD for the middle, ring, and little fingers in the concurrent finger flexion task but similar to the correlation analysis. Instead of the C/90° to C/L45° camera pairs performing the best, it is the C/L45° to L45°/R45° camera pairs. This trend is common, and the RMSD is similar for the index finger flexion and typing tasks where the index finger and thumb perform the best during the C/90° vs C/L45° comparison, and the middle, ring and little fingers perform the best during the C/90° vs. L45°/R45° comparison (Figure 3.8).

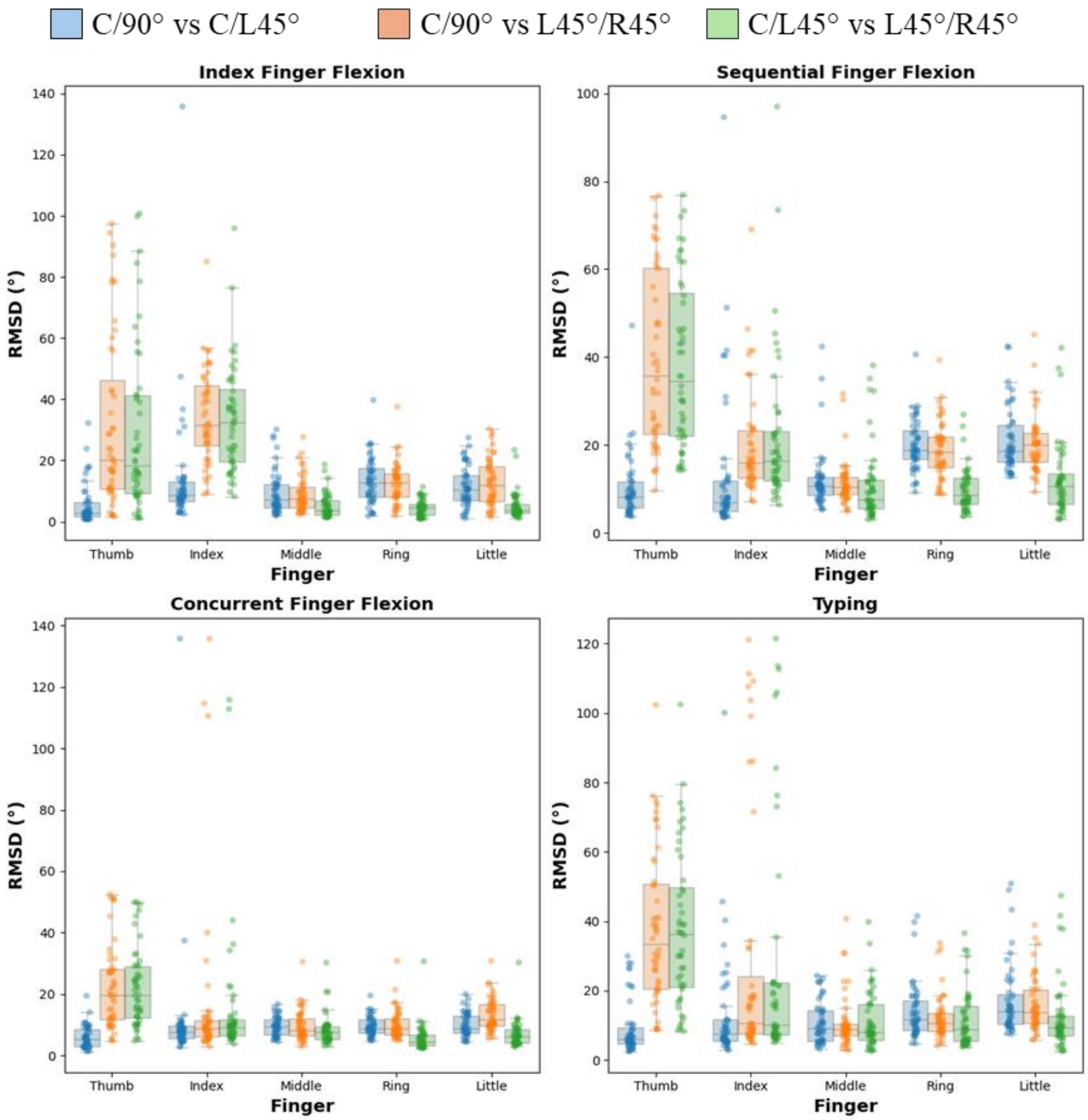

Figure 3.8: Box plot comparison of camera pair RMSD across the thumb, index, middle, ring, and little and tasks (index finger flexion, finger roll flexion, finger flexion, and typing) with the markers off (non-concurrent).

3.5. DISCUSSION

This study developed a markerless motion capture system capable of tracking hand and finger kinematics, filling a gap in biomechanical analyses. The joint angles obtained from the markerless system were comparable to those from marker-based motion capture during the concurrent and sequential finger flexion, index finger flexion, and typing tasks. Although a direct comparison was not conducted, the non-concurrent markerless finger angles closely matched the concurrent marker-based angles. Our findings also highlight the consistency of joint angles across different camera pairs, with variations depending on the specific camera pair and finger being analyzed.

3.5.1. Markerless vs marker-based motion capture

Due to issues with reflective markers, we were unable to directly compare the three markerless camera pairs with the marker-based data. The reflective markers interfered with our markerless joint predictions at the hand. A similar problem was reported in a study examining whole-body motion in which where reflective markers negatively impacted pose estimation using the Kinect system, resulting in inaccurate joint location identification and interference with markerless predictions (Naeemabadi et al., 2018). Consequently, we focused on comparing descriptive statistics.

Analysis of the C/90°, C/L45°, and L45°/R45° camera pairs revealed no significant differences in the non-concurrent markerless index finger angles across all tasks when compared to the concurrent marker-based measurements. While these data were not the same trials, they were performed within moments of each other. This suggests that the markerless system can provide measurements comparable to those obtained with marker-based techniques for the index finger across various tasks. There was a wide range of differences, from as little as 1° to large differences such as 19° in the total finger angle for our index, middle, ring and little fingers but most of the finger angle differences between the markerless camera pairs and the concurrent marker-based measurements were small for all the tasks, typically ranging from 1°-10° at the 50th percentile. By focusing on the 50th percentile, we gain insight into the central tendency of our joint angle measurements. By looking at the 50th percentile, we can see if the joint angles are similar between our markerless camera pairs and the concurrent marker-based finger angles. Our findings for our index, middle, little, and ring fingers are comparable to the joint angle differences of the index finger when using a markerless motion capture system on a trained model from DeepLabCut compared to a marker-based motion capture system where they found the MCP joint to have an RMSE of 7.5°, the PIP joint 2.3° and the DIP joint 3.2°, if summed, would be 13°, aligning with our results (Geelen et al., 2021). These findings are also comparable to a previous study that validated the Kinect sensor against a marker-based system, where the average differences for the MCP, PIP, and DIP joints were 2.4° ± 10°, 4.8° ± 12°, and 4.8° ± 11°, respectively (Metcalf et al., 2013). The individual joint angles in that study sum to align with our results. It is noteworthy that an error of 7°- 9° is considered clinically acceptable and the standard for determining individual joint angles in the hands and fingers using goniometry (Ellis et al., 2002; Ballan et al., 2008). Our total finger angles are similar to this threshold.

For our maximum finger angles, we used the 99th percentiles, where we saw the largest differences for the index, middle, ring, and little fingers between the markerless camera pairs and the marker-based finger angles. Even though we found no statistical differences for the C/90° and C/L45° camera pairs the absolute difference of these joint angles was large. For instance, the index finger angle difference ranged from 34° to 35° during concurrent finger flexion, 9 to 17° during sequential finger flexion, and 10 to 34° during the typing task. We did, however, find statistically significant differences during the index finger flexion task with differences ranging from 24 to 26°. It is important to note that the comparison of the markerless camera pairs on the finger angles compared to the marker-based system was not completed at the same time. The 99th percentiles are the largest values of finger angles, and we should expect these larger variations. The markerless camera pair finger angles not being compared at the same time to the marker-based system is the biggest factor that influences the large variation at the 99th percentiles. Another factor is that the reflective marker used in the marker-based system needed to be placed firmly to prevent movement and ensure its accuracy during data collection but because of this, participants reported that during data collection the reflective marker impeded their full finger movement, affecting their range of motion during all the tasks (Debril et al., 2009; Lucchetti et al., 1998). This observation is consistent with previous literature indicating that reflective markers can influence natural movement (Das et al., 2024). Although our findings comparing our markerless camera pairs and the marker-based system are not a direct comparison, the results of this analysis still indicate that a two-camera markerless motion capture system can be a tool used to assess hand and finger kinematics comparable to a marker-based motion capture system.

3.5.2. Markerless motion capture camera pairs

In our study, we compared joint angles derived from three camera pairs (C/90°, C/L45°, and L45°/R45°) by analyzing R^2 and RMSD. Despite differences in the tasks performed and the different anatomy our markerless hand and finger motion capture system produced R^2 comparable to those of the OpenCap system, which achieved R^2 of 0.65-0.80 for dynamic measures of the lower extremity (Uhlrich et al., 2023). Unlike prior studies focusing on single-camera setups for hand and finger kinematics (Gionfrida et al., 2022; Metcalf et al., 2013; Sridhar et al., 2014), our work evaluates different two-camera setups for markerless motion capture. Our results indicated that certain camera pair comparisons exhibited higher correlations than others. Specifically, the C/90° to C/L45° camera pair showed higher correlations for the index finger and thumb compared to the C/90° to L45°/R45° and C/L45° to L45°/R45° comparisons. Conversely, the C/L45° to L45°/R45° camera pair demonstrated higher correlations for the middle, ring, and little fingers compared to the C/90° to L45°/R45° and C/90° to C/L45° camera comparisons. Several factors could at least potentially explain these differences, as all camera pairs tracked the same motion simultaneously. The higher correlations for the index finger and thumb in the C/90° to C/L45° comparison may be attributed to the clear visibility of these fingers by both the 90° and L45° cameras. In contrast, the middle, ring, and little fingers did not perform as well in this comparison because the 90° and L45° cameras could not consistently see the 3 fingers despite the centre camera's visibility. For accurate joint position tracking, it is essential that both cameras in a pair have a clear view of the corresponding joints. This likely explains why the C/L45° to L45°/R45°

camera pair, had better visibility of the middle, ring, and little fingers and showed higher correlations for these fingers. Also, the difference between the camera pairs also might likely be due to the overlapping fields of view in the comparisons. For example, the C/90° to C/L45° both share the centre camera, and the C/L45° to L45°/R45° share the L45° camera. The camera pairs C/90°, and L45°/R45° do not share a camera and have different views from one another. This comparison performed had the lowest correlations compared to the other two camera pair comparisons.

The camera pair error results are consistent with the R^2 results; the index and thumb showed better performance with lower RMSD in the C/90° to C/L45° camera comparisons, while the middle, ring and little fingers performed better in the C/L45° to L45°/R45° comparison. Specifically for the concurrent finger flexion task, C/L45° to L45°/R45° had higher RMSD for the index finger with 15.6° and the thumb with 22.1°. In contrast, the C/90° to C/L45° comparison had slightly higher RMSD for the middle finger with 9.8°, the ring finger with 9.5°, and the little finger with 10.1°, compared to the C/L45° to L45°/R45° comparison for the middle finger with 8.4°, 5.7° for the ring finger and 6.9° for the little finger. This trend was consistent across all tasks. Additionally, an inverse relationship was observed between R^2 and RMSD: higher correlations resulted in lower RMSD and vice versa. These findings indicate that both R^2 and RMSD are influenced by the camera pairs being compared for each task and finger.

3.5.3. Limitations

There are several limitations to our study. First, a direct comparison between markers and markerless motion capture was not feasible in our setup. Other studies have successfully conducted concurrent analyses of markerless and marker-based systems for hand kinematics, as markers were not reported to influence the markerless system (Gionfrida et al., 2022; Metcalf et al., 2013; Geelen et al., 2021). Secondly, due to hardware issues and noisy data from the marker-based motion capture, we had to exclude five participants from our markerless vs. marker-based assessment. This underscores the potential advantages of markerless systems for hand and finger kinematics. Lastly, we did not include tasks involving both hands. Most hand tasks require the use of both hands, but our study did not assess this. Future work should incorporate a broader range of tasks and include assessments involving both hands.

3.5.4. Conclusions

In conclusion, we developed a two-camera markerless motion capture system using two cameras to track and analyze 3D hand and finger kinematics with open-source tools. Our findings indicate no significant differences between the markerless motion capture system and the marker-based system. Additionally, our results show that different markerless camera pairs generally produced similar predictions of joint angles with little or no significant differences in most cases. Certain camera pairs demonstrated better agreement than others, and this agreement was based on the camera pairs showing better consistency across different camera comparisons than others when for each task and finger. Future work should aim to evaluate markerless hand tracking with a wider range of tasks.

Conflict of Interest Statement

The authors declare no conflicts of interest related to the manuscript.

Acknowledgements

This work was supported by research funding from the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery Grant RGPIN-2023-05473 to PJK). We would like to thank the members of the McMaster Occupational Biomechanics Laboratory (Lara Matloub, Walid Elcheikh Hassan, Ryan Chhiba) for their assistance in data collection.

3.6. REFERENCES

Amell, T., & Kumar, S. (2001). Work-related musculoskeletal disorders: design as a prevention strategy. A review. *Journal of occupational rehabilitation*, *11*, 255-265

Ballan, L., & Cortelazzo, G. M. (2008, June). Marker-less motion capture of skinned models in a four camera set-up using optical flow and silhouettes. 3DPVT.

Cappozzo, A., Catani, F., Della Croce, U., & Leardini, A. (1995). Position and orientation in space of bones during movement: anatomical frame definition and determination. *Clinical biomechanics*, *10*(4), 171-178.

Cheung, G. K. M., S. Baker, and T. Kanade. Shape-from-silhouette across Time Part II: Applications to human modeling and markerless motion tracking. *Int. J. Comp. Vis*. 63(3):225–245, 2005.

Cocchiarella, D. M., Kociolek, A. M., Tse, C. T., & Keir, P. J. (2016). Toward a realistic optoelectronic-based kinematic model of the hand: representing the transverse metacarpal arch reduces accessory rotations of the metacarpophalangeal joints. *Computer methods in biomechanics and biomedical engineering*, *19*(6), 639-647.

Das, K., de Paula Oliveira, T., & Newell, J. (2023). Comparison of markerless and marker-based motion capture systems using 95% functional limits of agreement in a linear mixed-effects modelling framework. *Scientific Reports*, *13*(1), 22880.

Debril, J. F., Pudlo, P., Gorce, P., & Lepoutre, F. X. (2009). Experimental platform to assess the weight relief lifting and sitting pivot transfer movements. *IFAC Proceedings Volumes*, *42*(12), 419-424.

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

Kanko, R. M., Laende, E. K., Davis, E. M., Selbie, W. S., & Deluzio, K. J. (2021). Concurrent assessment of gait kinematics using marker-based and markerless motion capture. *Journal of biomechanics*, *127*, 110665.

Kanko, R. M., Laende, E., Selbie, W. S., & Deluzio, K. J. (2021). Inter-session repeatability of markerless motion capture gait kinematics. *Journal of Biomechanics*, *121*, 110422.

Lima, Y., Collings, T., Hall, M., Bourne, M., & Diamond, L. (2023). Assessing lower-limb kinematics via OpenCap during dynamic tasks relevant to anterior cruciate ligament injury: a validity study. *Journal of Science and Medicine in Sport*, *26*, S105.

Lucchetti, L., Cappozzo, A., Cappello, A., & Della Croce, U. (1998). Skin movement artefact assessment and compensation in the estimation of knee-joint kinematics. *Journal of biomechanics*, *31*(11), 977-984.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature neuroscience*, *21*(9), 1281-1289.

McFarland, D. C., Binder-Markey, B. I., Nichols, J. A., Wohlman, S. J., De Bruin, M., & Murray, W. M. (2022). A musculoskeletal model of the hand and wrist capable of simulating functional tasks. *IEEE Transactions on Biomedical Engineering*, *70*(5), 1424-1435.

Metcalf, C. D., Robinson, R., Malpass, A. J., Bogle, T. P., Dell, T. A., Harris, C., & Demain, S. H. (2013). Markerless motion capture and measurement of hand kinematics: validation and application to home-based upper limb rehabilitation. *IEEE Transactions on Biomedical Engineering*, *60*(8), 2184-2192.

Naeemabadi, M., Dinesen, B., Andersen, O. K., & Hansen, J. (2018). Influence of a marker-based motion capture system on the performance of Microsoft Kinect v2 skeleton algorithm. *IEEE Sensors Journal*, *19*(1), 171-179

Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature protocols*, *14*(7), 2152-2176.

Sridhar, S., Oulasvirta, A., & Theobalt, C. (2014). Fast tracking of hand and finger articulations using a single depth camera.

Steffensen, M. S., Magalhaes PhD, F., Knarr PhD, B. A., & Kingston PhD, D. C. (2023). Comparison of markerless and marker-based motion capture in cerebral palsy and chronic stroke gait kinematics.

Uhlrich, S. D., Falisse, A., Kidziński, Ł., Muccini, J., Ko, M., Chaudhari, A. S., . & Delp, S. L. (2022). OpenCap: 3D human movement dynamics from smartphone videos. *bioRxiv*, 2022-07.

Wren, T. A., Isakov, P., & Rethlefsen, S. A. (2023). Comparison of kinematics between theia markerless and conventional marker-based gait analysis in clinical patients. *Gait & Posture, 104*, 9-14.

Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C. L., & Grundmann, M. (2020). Mediapipe hands: On-device real-time hand tracking. *arXiv preprint arXiv:2006.10214*.

SUPPLEMENTARY DOCUMENTATI

Table S1: Estimated marginal means [95% confidence intervals] for the 1st, 50th, and 99th percentiles (°) for each finger and task, across different camera orientations. The orientations C/90°, C/L45°, and L45°/R45° refer to the markerless motion capture cameras, while Motion Analysis refers to the marker-based system.

		C/90°			C/L45°			L45°/R45°			Motion Analysis		
		1st	**50th**	**99th**	**1st**	**50th**	**99th**	**1st**	**50th**	**99th**	**1st**	**50th**	**99th**
Index	**Concurrent Finger Flexon**	19.24 [14.90, 23.57]	44.10 [35.37, 52.84]	116.08 [106.42, 125.74]	16.57 [12.23, 20.90]	46.63 [37.90, 55.37]	115.39 [105.73, 125.05]	23.39 [19.05, 27.72]	**60.39 *** [51.65, 69.12]**	125.14 [115.48, 134.80]	19.31 [14.83, 23.80]	43.32 [34.29, 52.36]	81.34 [71.34, 91.34]
	Sequential Finger Flexion	14.90 [9.38, 20.42]	32.63 [22.22, 43.03]	115.45 [102.46, 128.43]	11.22 [5.70, 16.74]	30.80 [20.39, 41.20]	109.30 [96.31, 122.28]	16.90 [11.38, 22.42]	37.79 [27.39, 48.20	88.59 [75.60, 101.57]	20.58 [14.86, 26.29]	42.46 [31.69, 53.23]	97.88 [84.44, 111.32]
	Typing	29.76 [24.75, 34.78]	40.62 [32.04, 49.20]	115.78 [97.76, 133.80]	25.32 [20.31, 30.33]	37.98 [29.40, 46.56]	97.56 [79.54, 115.58]	30.78 [25.77, 35.79]	**59.44 ɣ [50.86, 68.02]**	119.82 [101.80, 137.84]	30.65 [25.26, 36.03]	43.74 [34.52, 52.96]	86.66 [67.30, 106.02]
	Index Finger Flexion	20.83 [16.22, 25.44]	47.63 [40.94, 54.33]	122.25 [110.32, 134.17]	19.34 [14.73, 23.94]	55.90 [49.21, 62.60]	124.14 [112.21, 136.06]	25.42 [20.81, 30.02]	44.96 [38.27, 51.66]	**88.29 # [76.36, 100.21]**	21.43 [16.67, 26.20]	52.49 [45.56, 59.42]	**97.97 Ͱ [85.63, 110.32]**
Middle	**Concurrent Finger Flexon**	12.96 [8.62, 17.29]	50.11 [41.37, 58.84]	124.63 [114.97, 134.29]	8.79 [4.45, 13.12]	44.78 [36.05, 53.51]	118.50 [108.84, 128.16]	12.23 [7.90, 16.57]	50.98 [42.24, 59.71]	124.97 [115.31, 134.63]	12.16 [7.67, 16.64]	41.04 [32.00, 50.08]	86.15 [76.15, 96.15]
	Sequential Finger Flexion	16.51 [10.99, 22.03]	34.24 [23.83, 44.64]	95.54 [82.56, 108.52]	13.28 [7.76, 18.80]	31.93 [21.52, 42.33]	91.13 [78.15, 104.12]	16.51 [10.99, 22.03]	36.68 [26.28, 47.09]	89.18 [76.20, 102.17]	19.84 [14.13, 25.55]	47.90 [37.13, 58.67]	101.36 [87.92, 114.80]
	Typing	25.49 [20.48, 30.51]	39.31 [30.72, 47.89]	103.99 [85.96, 122.01]	21.69 [16.68, 26.71]	34.41 [25.83, 42.99]	102.23 [84.21, 120.25]	25.48 [20.47, 30.49]	41.54 [32.96, 50.12]	99.89 [81.87, 117.91]	23.72 [18.34, 29.11]	33.58 [24.36, 42.80]	78.70 [59.34, 98.06]
	Index Finger Flexion	19.34 [14.74, 23.95]	29.56 [22.87, 36.26]	59.22 [47.30, 71.15]	15.85 [11.24, 20.45]	25.21 [18.51, 31.90]	40.11 [28.19, 52.04]	18.35 [13.74, 22.95]	28.90 [22.21, 35.59]	44.06 [32.13, 55.98]	15.71 [10.95, 20.48]	25.35 [18.42, 32.27]	41.47 [29.13, 53.81]
Ring	**Concurrent Finger Flexon**	13.88 [9.54, 18.21]	53.28 [44.54, 62.01]	124.56 [114.90, 134.22]	10.76 [6.43, 15.10]	48.07 [39.33, 56.80]	119.59 [109.93, 129.25]	12.63 [8.29, 16.96]	51.55 [42.82, 60.28]	122.41 [112.75, 132.07]	7.43 [2.94, 11.92]	**35.17 ¥ [26.13, 44.21]**	102.81 [92.81, 112.81]
	Sequential Finger Flexion	20.08 [14.56, 25.60]	46.41 [36.00, 56.81]	112.50 [99.52, 125.48]	14.76 [9.24, 20.28]	33.75 [23.35, 44.16]	97.93 [84.94, 110.91]	17.29 [11.77, 22.80]	38.07 [27.66, 48.47]	91.90 [78.91, 104.88]	16.56 [10.84, 22.27]	51.03 [40.26, 61.80]	116.70 [103.26, 130.14]
	Typing	24.56 [19.54, 29.57]	41.08 [32.50, 49.66]	102.93 [84.91, 120.96]	21.16 [16.14, 26.17]	35.00 [26.42, 43.58]	102.34 [84.32, 120.36]	23.66 [18.65, 28.67]	40.97 [32.39, 49.56]	95.31 [77.29, 113.34]	14.47 [9.09, 19.86]	25.10 [15.88, 34.32]	68.02 [48.66, 87.38]
	Index Finger Flexion	20.18 [15.58, 24.79]	32.42 [25.73, 39.12]	62.14 [50.22, 74.07]	17.03 [12.42, 21.63]	24.32 [17.63, 31.02]	40.42 [28.50, 52.35]	18.29 [13.68, 22.90]	26.74 [20.04, 33.43]	44.38 [32.45, 56.30]	10.83 [6.06, 15.60]	**16.40 + [9.47, 23.33]**	**26.98 ¿ [14.64, 39.33]**
Little	**Concurrent Finger Flexon**	9.35 [5.01, 13.68]	43.49 [34.75, 52.22]	115.01 [105.35, 124.67]	10.80 [6.46, 15.13]	44.41 [35.68, 53.15]	115.68 [106.02, 125.34]	10.97 [6.64, 15.31]	46.87 [38.14, 55.61]	116.57 [106.91, 126.23]	11.76 [7.27, 16.25]	32.31 [23.27, 41.35]	70.00 [60.00, 80.00]
	Sequential Finger Flexion	11.27 [5.75, 16.79]	37.35 [26.95, 47.75]	122.89 [109.91, 135.87]	9.68 [4.16, 15.20]	29.77 [19.37, 40.18]	121.61 [108.62, 134.59]	10.65 [5.13, 16.17]	31.98 [21.58, 42.39]	115.14 [102.16, 128.13]	12.15 [6.43, 17.86]	49.27 [38.50, 60.04]	117.26 [103.82, 130.70]
	Typing	13.84 [8.83, 18.86]	31.37 [22.78, 39.95]	99.98 [81.95, 118.00]	12.62 [7.60, 17.63]	28.63 [20.05, 37.21]	111.86 [93.84, 129.89]	14.01 [9.00, 19.02]	32.80 [24.21, 41.38]	94.81 [76.79, 112.83]	11.22 [5.83, 16.60]	19.72 [10.50, 28.94]	50.90 [31.54, 70.26]
	Index Finger Flexion	13.99 [9.38, 18.59]	26.24 [19.55, 32.94]	55.00 [43.08, 66.93]	13.06 [8.45, 17.66]	21.16 [14.47, 27.85]	48.44 [36.51, 60.36]	12.94 [8.34, 17.55]	21.71 [15.02, 28.41]	57.36 [45.44, 69.29]	11.74 [6.97, 16.51]	15.29 [8.36, 22.22]	**20.94 ¢ [8.60, 33.29]**
Thumb	**Concurrent Finger Flexon**	8.00 [3.67, 12.34]	42.55 [33.82, 51.29]	94.19 [84.53, 103.85]	7.23 [2.89, 11.56]	42.39 [33.66, 51.12]	95.38 [85.72, 105.04]	9.27 [4.93, 13.60]	47.96 [39.23, 56.70]	88.84 [79.19, 98.50]	-1.30 [-5.79, 3.19]	**9.57 ‡ [0.53, 18.61]**	54.21 [44.21, 64.21]
	Sequential Finger Flexion	5.51 [-0.01, 11.03]	20.31 [9.91, 30.71]	100.32 [87.33, 113.30]	5.54 [0.02, 11.06]	23.55 [13.15, 33.96]	105.07 [92.09, 118.06]	7.48 [1.96, 13.00]	**55.43** [45.03, 65.83]**	98.37 [85.39, 111.36]	-0.75 [-6.46, 4.96]	16.74 [5.97, 27.51]	82.62 [69.18, 96.06]
	Typing	5.97 [0.96, 10.98]	15.86 [7.28, 24.44]	90.86 [72.84, 108.88]	5.80 [0.78, 10.81]	16.07 [7.49, 24.65]	90.19 [72.16, 108.21]	12.06 [7.05, 17.08]	**49.54 ¡ [40.96, 58.13]**	103.17 [85.15, 121.19]	-1.30 [-6.68, 4.09]	3.54 [-5.67, 12.76]	52.81 [33.45, 72.17]
	Index Finger Flexion	8.11 [3.51, 12.72]	14.85 [8.16, 21.55]	33.53 [21.61, 45.46]	9.52 [4.92, 14.13]	17.84 [11.14, 24.53]	41.79 [29.87, 53.72]	9.19 [4.59, 13.80]	**41.85 * [35.15, 48.54]**	**72.33 • [60.40, 84.25]**	0.60 [-4.16, 5.37]	5.93 [-1.00, 12.86]	**13.44§ [1.10, 25.79]**

CHAPTER FOUR: THESIS DISCUSSION

The overarching goal of this thesis was to develop and assess a markerless motion capture for the hands and fingers. This thesis filled a gap in the literature by providing a tool needed to assess hand and finger kinematics. This research journey involved overcoming various challenges, implementing innovative methodologies, and refining techniques to ensure accurate and reliable data from our markerless motion capture program.

We initially explored the development of a markerless motion capture system utilizing DeepLabCut (DLC), an open-source tool (Mathis et al., 2018; Nath et al., 2019). The goal was to train to produce hand and finger kinematics. DLC appeared highly promising due to the flexibility it offered in the model input. However, this flexibility presented challenges, particularly the extensive time required for manual image labelling and model training. Despite these challenges, I managed to annotate over 1400 images to evaluate model accuracy in predicting joint positions with virtual markers. Unfortunately, this number of images is insufficient for achieving accurate and consistent results in markerless motion capture. Nonetheless, the extensive time to annotate frames and the limited image dataset prompted a shift in our approach. This transition led us to MediaPipe Hands (MPH), which offered a pre-trained hand-tracking model. This solution significantly streamlined the process, allowing us to focus more on tracking 3D hand and finger kinematics without training our model.

MPH proved effective because of its pre-trained model tracking 21 hand landmarks. This model utilizes two distinct models: a hand landmark model and a palm detection model. An interesting feature of MediaPipe is its ability to handle both hands simultaneously. However, to simplify our process and reduce the number of required markers, we chose to track only one hand instead of two. Initially, I planned to use four cameras in the markerless system, necessitating the triangulation of these cameras. This process, however, was challenging and complex. Dute The real-time calibration process involved streaming video data from each camera and capturing still images, which posed several difficulties. Although capturing still images for individual cameras and obtaining their internal parameters was straightforward, streaming video from all four cameras simultaneously to obtain their rotation and translation to one another, overwhelmed our hardware. The computer lagged and eventually crashed, unable to handle the multiple video feeds. Even reducing the video resolution to 480p from 720p did not alleviate the lag. Switching to a more powerful computer with GPU integration also failed to resolve the issue. Even when reducing the number of cameras from four to three, the same problems persisted. At the time, it seemed that Python and OpenCV could not effectively manage multiple video feeds, leading to the computer's crash and lag. Due to time constraints and needing to complete our data collection we had to move on to a different approach. We used three camera pairs to evaluate how views might influence the angle predictions. This solution allowed me to continue my research without the limitations imposed by the initial hardware and software constraints. By doing so, we were able to get effective tracking of the hands and fingers when we did this. To ensure the joint predictions were tracking the hands we measured the distance between the joint predictions and matched them to real-world measurements of the hands and fingers. The segment lengths analyzed from MediaPipe matched within 0.5 cm, indicating these were accurate predictions.

A significant factor that influenced the markerless joint predictions was the reflective markers used in the marker-based system. For MediaPipe to work effectively, the model needs to detect a palm in the frame first and then detect the 21 joint landmarks on the hand. The reflective markers impacted this detection because they were placed directly on the joints near the landmarks MPH uses to track the hand fingers. Additionally, the segment marker triads were intrusive, covering most of the participant's fingers. This was anticipated since any obstruction would prevent MediaPipe from detecting the hand within the frame, highlighting how the model was trained to detect the skin for the hand landmark predictions. This highlights a key moment during this thesis because a concurrent frame-by-frame assessment of the marker-based motion capture system was not feasible due to prediction errors, which would not accurately represent the feasibility of the markerless motion capture system. Therefore, when participants came into the lab, we collected markerless data concurrently with the marker-based data to understand and address these errors. Then, participants performed the series of tasks again without any reflective markers on. During these errors, the predictions were not correctly annotated on the participants' hands, and the hand model appeared smaller with shorter segment lengths. This issue was noticeable in static/extended postures. The results of the analysis with the markers during collection can be seen in Figure A2 to see how the markers affected the results. Also, as participants performed their tasks, the segment lengths changed significantly, resulting in a complete failure to predict joint locations accurately. This is illustrated in Figures 4.1 and 4.2, where in the top right, you can see the predictions are of the participant's hands, and the finger lengths are smaller.

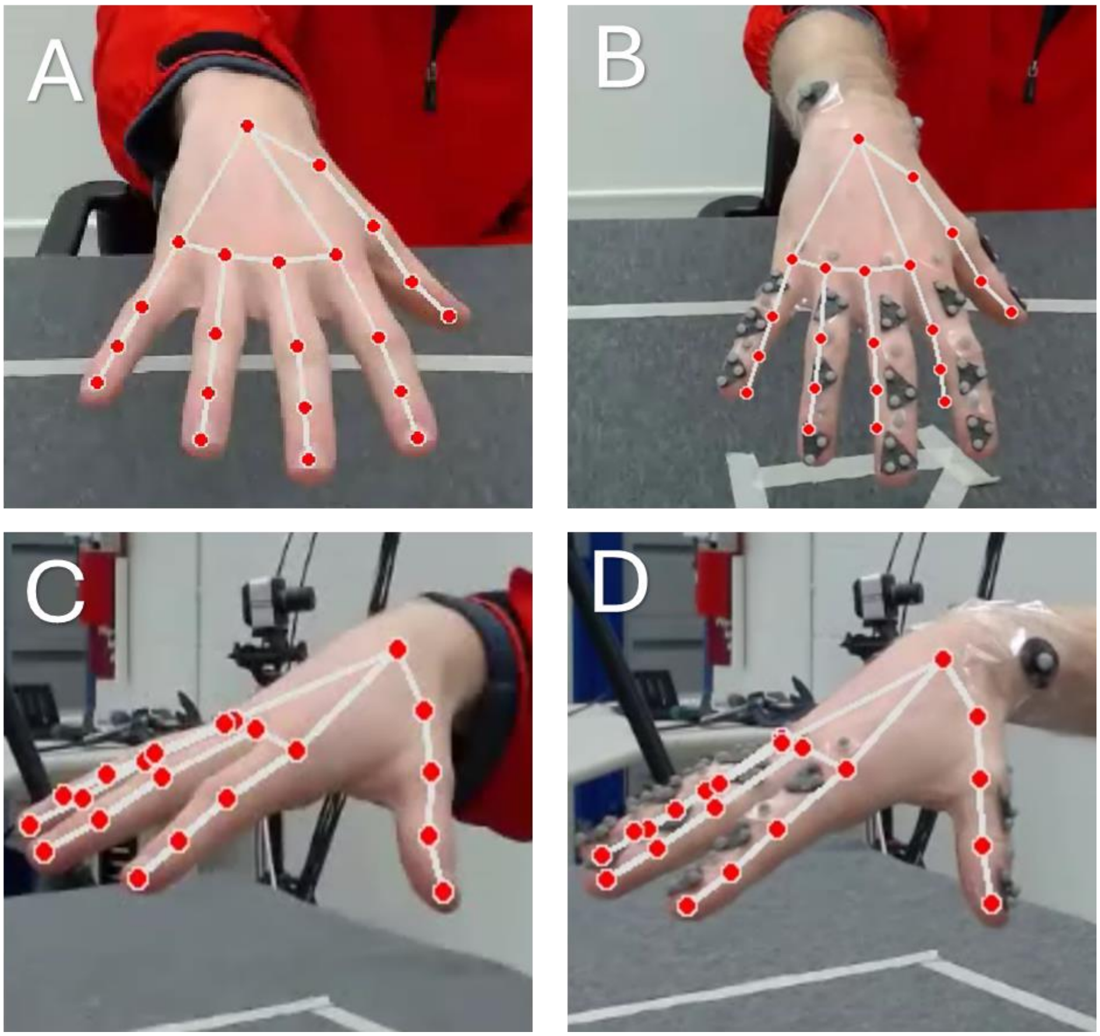

Figure 4.1: A) Shows the Centre (C) camera view during the concurrent finger flexion task without reflective markers. B) Shows the Centre (C) camera view during the concurrent finger flexion task with reflective markers. C) Shows the Left 45° (L45 °) camera view during the concurrent finger flexion task without reflective markers. D) Shows the Left 45° (L45 °) camera view during the concurrent finger flexion task with reflective markers. These images illustrate when the fingers are in extension during the task.

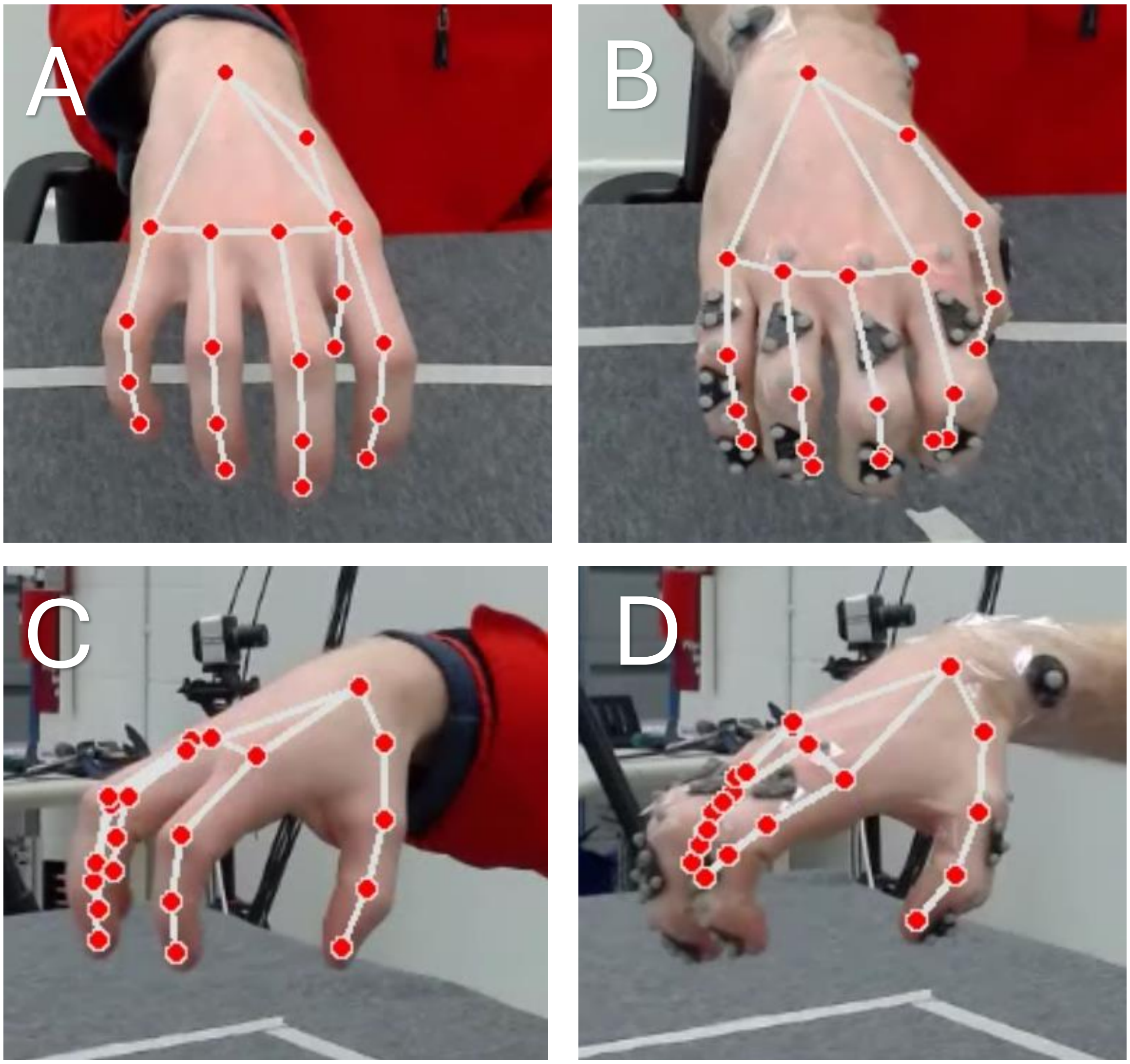

Figure 4.2: A) Shows the Centre (C) camera view during the concurrent finger flexion task without reflective markers. B) Shows the Centre (C) camera view during the concurrent finger flexion task with reflective markers. C) Shows the Left 45° (L45 °) camera view during the concurrent finger flexion task without reflective markers. D) Shows the Left 45° (L45 °) camera view during the concurrent finger flexion task with reflective markers. These images illustrate when the fingers are in flexion during the task.

Using OpenSim to calculate joint kinematics for markerless data also presented challenges. To scale our models accurately and ensure the precision of joint kinematics, we manually inputted participants' hand anthropometrics and uniformly scaled the hand model for each participant. Initial inverse kinematics results showed significant extension angles (up to 15 degrees) in the MCP and PIP joints, which were not observed during the trials. We made sure to add these constraints to our marker-based data to ensure we can confidently compare the results. Another issue we saw was problems with the DIP joints in the index, middle, ring, and little fingers. During flexion, the lengths of distal phalange segments changed, affecting the inverse kinematics. To resolve this, we aligned the tip marker with the other finger joints, which were centred in the participant's joints when scaling our models. This adjustment was crucial, resulting in a fully constrained model for each participant without these issues. Modelling with OpenSim was challenging due to its complexity, but it allowed us to explore abduction and adduction angles that influence flexion and extension at the joints. This was important because when I performed my own joint angle calculations for the MCP, PIP, and DIP joints for all the fingers, it focused solely on flexion and extension by measuring the segment lengths between the MCP, PIP, and DIP joints and using the segment vectors to calculate the joint angles. However, my joint angle calculations could not be compared to the marker-based kinematics since the calculation of obtaining the joint angles was different from the inverse kinematics calculation in OpenSim. Overall, although it was challenging OpenSim helped overcome the issue of the changing segment lengths in the markerless data.

4.1. Future directions

There are several potential directions for this research. First, we need to analyze more than one hand to evaluate the performance of the markerless motion capture system with multiple hands in the scene to obtain 3D motion and joint angles. Most tasks require the use of both hands, and this would be beneficial for others using this system for data collection. Next, the model needs to be improved by adding more data to the training set. Currently, there are errors when joints are obstructed. For instance, Theia 3D can predict joints through clothing because it has been trained on over 500,000 images, compared to MediaPipe's 30,000 images. Increasing the training data would significantly enhance the system's ability to capture hand and finger kinematics and potentially resolve issues with reflective markers interfering with predictions. This would also enable a more comprehensive concurrent assessment with the marker-based motion capture system. Lastly, expanding the model to include the elbow could allow us to obtain wrist kinematics. Currently, the wrist is represented as a single point, but including the elbow could provide more detailed wrist kinematics.

4.2. Contributions

Currently, no markerless motion capture system can track 3D hand and finger kinematics. This thesis fills that gap, being the first of its kind to create and assess a markerless motion capture system using open-source tools. We will share our code with the scientific community, allowing others to expand or improve upon our work. This system will also benefit researchers who wish to use it in their own studies. By developing more tools, we enhance the field's ability to collect and analyze hand and finger motion capture data. This can be used in different applications ranging from in-lab and out-of-lab collections and potentially clinical settings.

REFERENCES

Amell, T., & Kumar, S. (2001). Work-related musculoskeletal disorders: design as a prevention strategy. A review. *Journal of occupational rehabilitation*, *11*, 255-265

Armstrong, T. J. (1986). Ergonomics and cumulative trauma disorders. *Hand clinics*, *2*(3), 553-565.

Barr, A. E., Barbe, M. F., & Clark, B. D. (2004). Work-related musculoskeletal disorders of the hand and wrist: epidemiology, pathophysiology, and sensorimotor changes. *Journal of orthopaedic & sports physical therapy*, *34*(10), 610-627.

Cai, L., Ma, Y., Xiong, S., & Zhang, Y. (2019). Validity and reliability of upper limb functional assessment using the Microsoft Kinect V2 sensor. *Applied bionics and biomechanics*, *2019*.

Cappozzo, A., Catani, F., Leardini, A., Benedetti, M. G., & Della Croce, U. (1996). Position and orientation in space of bones during movement: experimental artifacts. *Clinical biomechanics*, *11*(2), 90-100.

Cappozzo, A., Catani, F., Della Croce, U., & Leardini, A. (1995). Position and orientation in space of bones during movement: anatomical frame definition and determination. *Clinical biomechanics*, *10*(4), 171-178.

Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7291-7299).

Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, *40*(4), 834-848.

Cheung, G. K. M., S. Baker, and T. Kanade. Shape-from-silhouette across Time Part II: Applications to human modeling and markerless motion tracking. *Int. J. Comp. Vis.* 63(3):225–245, 2005.

Cao, W., Zhong, J., Cao, G., & He, Z. (2019). Physiological function assessment based on kinect v2. IEEE Access, 7, 105638-105651.

Cao, Z., Simon, T., Wei, S. E., & Sheikh, Y. (2017). Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7291-7299).

Cocchiarella, D. M., Kociolek, A. M., Tse, C. T., & Keir, P. J. (2016). Toward a realistic optoelectronic-based kinematic model of the hand: representing the transverse metacarpal arch reduces accessory rotations of the metacarpophalangeal joints. *Computer methods in biomechanics and biomedical engineering*, *19*(6), 639-647.

Cook, J. R., Baker, N. A., Cham, R., Hale, E., & Redfern, M. S. (2007). Measurements of wrist and finger postures: a comparison of goniometric and motion capture techniques. *Journal of applied biomechanics*, *23*(1), 70-78.

Das, K., de Paula Oliveira, T., & Newell, J. (2023). Comparison of markerless and marker-based motion capture systems using 95% functional limits of agreement in a linear mixed-effects modelling framework. Scientific Reports, 13(1), 22880.

Debril, J. F., Pudlo, P., Gorce, P., & Lepoutre, F. X. (2009). Experimental platform to assess the weight relief lifting and sitting pivot transfer movements. IFAC Proceedings Volumes, 42(12), 419-424.

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

Dobos, T. J., Bench, R. W., McKinnon, C. D., Brady, A., Boddy, K. J., Holmes, M. W., & Sonne, M. W. (2022). Validation of pitchAITM markerless motion capture using marker-based 3D motion capture. *Sports Biomechanics*, 1-21.

Dolatabadi, E., Taati, B., & Mihailidis, A. (2016). Concurrent validity of the Microsoft Kinect for Windows v2 for measuring spatiotemporal gait parameters. *Medical engineering & physics*, *38*(9), 952-958.

Fleisig, G. S., Slowik, J. S., Wassom, D., Yanagita, Y., Bishop, J., & Diffendaffer, A. (2022). Comparison of marker-less and marker-based motion capture for baseball pitching kinematics. *Sports Biomechanics*, 1-10.

Fowler, N. K., & Nicol, A. C. (2001). Functional and biomechanical assessment of the normal and rheumatoid hand. *Clinical biomechanics*, *16*(8), 660-666.

Fuller, J., Liu, L. J., Murphy, M. C., & Mann, R. W. (1997). A comparison of lower-extremity skeletal kinematics measured using skin-and pin-mounted markers. *Human movement science*, *16*(2-3), 219-242.

Geelen, J. E., Branco, M. P., Ramsey, N. F., Van Der Helm, F. C., Mugge, W., & Schouten, A. C. (2021, November). MarkerLess Motion Capture: ML-MoCap, a low-cost modular multi-camera setup. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 4859-4862). IEEE.

Guess, T. M., Razu, S., Jahandar, A., Skubic, M., & Huo, Z. (2017). Comparison of 3D joint angles measured with the kinect 2.0 skeletal tracker versus a marker-based motion capture system. *Journal of applied biomechanics*, *33*(2), 176-181.

Güney, G., Jansen, T. S., Dill, S., Schulz, J. B., Dafotakis, M., Hoog Antink, C., & Braczynski, A. K. (2022). Video-based hand movement analysis of parkinson patients before and after medication using high-frame-rate videos and MediaPipe. *Sensors*, *22*(20), 7992.

Hoyet, L., Ryall, K., McDonnell, R., & O'Sullivan, C. (2012, March). Sleight of hand: perception of finger motion from reduced marker sets. In *Proceedings of the ACM SIGGRAPH symposium on interactive 3D graphics and games* (pp. 79-86).

Hoyt, W. R. (1984). Carpal tunnel syndrome: Analysis and prevention. *Professional Safety*, *29*(11), 16-21.

Hu, B., Ding, T., Peng, Y., Liu, L., & Wen, X. (2020). Flexible and attachable inertial measurement unit (IMU)-based motion capture instrumentation for the characterization of hand kinematics: A pilot study. *Instrumentation Science & Technology*, *49*(2), 125-145.

Kanko, R. M., Laende, E. K., Davis, E. M., Selbie, W. S., & Deluzio, K. J. (2021). Concurrent assessment of gait kinematics using marker-based and markerless motion capture. *Journal of biomechanics*, *127*, 110665.

Kanko, R. M., Laende, E., Selbie, W. S., & Deluzio, K. J. (2021). Inter-session repeatability of markerless motion capture gait kinematics. *Journal of Biomechanics*, *121*, 110422.

Kortier, H. G., Sluiter, V. I., Roetenberg, D., & Veltink, P. H. (2014). Assessment of hand kinematics using inertial and magnetic sensors. *Journal of neuroengineering and rehabilitation*, *11*(1), 1-15.

Lahkar, B. K., Muller, A., Dumas, R., Reveret, L., & Robert, T. (2022). Accuracy of a markerless motion capture system in estimating upper extremity kinematics during boxing. *Frontiers in sports and active living*, *4*.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444.

Majoni, N., Mulla, DM, & Keir, PJ. (2022, August). *Markerless Motion Capture of the Hands and Fingers*. NACOB 2022: Ottawa, Ontario, Canada

Martinez, H. R., Garcia-Sarreon, A., Camara-Lemarroy, C., Salazar, F., & Guerrero-González, M. L. (2018). Accuracy of Markerless 3D Motion Capture Evaluation to Differentiate between On/Off Status in Parkinson's Disease after Deep Brain Stimulation. *Parkinson's Disease*, *2018*.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature neuroscience*, *21*(9), 1281-1289.

Maw, J., Wong, K. Y., & Gillespie, P. (2016). Hand anatomy. *British Journal of Hospital Medicine*, *77*(3), C34-C40.

McFarland, D. C., Binder-Markey, B. I., Nichols, J. A., Wohlman, S. J., De Bruin, M., & Murray, W. M. (2022). A musculoskeletal model of the hand and wrist capable of simulating functional tasks. *IEEE Transactions on Biomedical Engineering*, *70*(5), 1424-1435.

Metcalf, C. D., Robinson, R., Malpass, A. J., Bogle, T. P., Dell, T. A., Harris, C., & Demain, S. H. (2013). Markerless motion capture and measurement of hand kinematics: validation and application to home-based upper limb rehabilitation. *IEEE Transactions on Biomedical Engineering*, *60*(8), 2184-2192.

Metcalf, C. D., C. Phillips, A. Forrester, J. Glodowski, K. Simpson, C. Everitt, A. Darekar, L. King, D. Warwick, A. S. Dickinson. (2020) Quantifying Soft Tissue Artefacts and Imaging Variability in Motion Capture of the Fingers. *Annals of Biomedical Engineering* 48:5, pages 1551-1561

Metcalf, C. D., and S. V. Notley. Modified kinematic technique for measuring pathological hyperextension and hypermobility of the interphalangeal joints. *IEEE Trans. Biomed. Eng.* 58(5):1224–1231, 2011

Metcalf, C. D., S. V. Notley, P. H. Chappell, J. H. Burridge, and V. T. Yule. Validation and application of a computational model for wrist and hand movements using surface markers. *IEEE Trans. Biomed. Eng.* 55(3):1199–1210, 2008

Moore, K. L., & Dalley, A. F. (2018). *Clinically oriented anatomy*. Wolters kluwer india Pvt Ltd.

Muggleton, J. M., Allen, R., & Chappell, P. H. (1999). Hand and arm injuries associated with repetitive manual work in industry: a review of disorders, risk factors and preventive measures. *Ergonomics*, *42*(5), 714-739.

Naeemabadi, M., Dinesen, B., Andersen, O. K., & Hansen, J. (2018). Influence of a marker-based motion capture system on the performance of Microsoft Kinect v2 skeleton algorithm. *IEEE Sensors Journal*, *19*(1), 171-179

Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature protocols*, *14*(7), 2152-2176.

Nordquist, R. E., Meijer, E., van der Staay, F. J., & Arndt, S. S. (2017). Pigs as model species to investigate effects of early life events on later behavioral and neurological functions. In *Animal models for the study of human disease* (pp. 1003-1030). Academic Press.

Pagnon, D., Domalain, M., & Reveret, L. (2022). Pose2Sim: an end-to-end workflow for 3D markerless sports kinematics—part 2: accuracy. *Sensors*, *22*(7), 2712.

Peters, A., Galna, B., Sangeux, M., Morris, M., & Baker, R. (2010). Quantification of soft tissue artifact in lower limb human motion analysis: a systematic review. *Gait & posture*, *31*(1), 1-8.

Prasad, R., Bhamidi, A., Rajeswaran, A., Muthukumar, S., Kothandaraman, P., & Sivaraj, M. (2014). Epidemiology and sequelae of workplace hand injuries at a tertiary trauma care centre. *Surgical Science*, *2014*.

Rash, G. S., P. P. Belliappa, M. P. Wachowiak, N. N. Somia, and A. Gupta. A demonstration of the validity of a 3-D video motion analysis method for measuring finger flexion and extension. *J. Biomech.* 32(12):1337–1341, 1999.

Roetenberg, D., Luinge, H. J., Baten, C. T., & Veltink, P. H. (2005). Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation. *IEEE Transactions on neural systems and rehabilitation engineering*, *13*(3), 395-405.

Salhi, S., & Queen, N. M. (2004). A hybrid algorithm for identifying global and local minima when optimizing functions with many minima. *European Journal of Operational Research*, *155*(1), 51-67.

Sancho-Bru, J. L., N. J. Jarque-Bou, M. Vergara, and A. Pérez-González. Validity of a simple videogrammetric method to measure the movement of all hand segments for clinical purposes. *Proc. Inst. Mech. Eng. H* 228(2):182–189, 2014.

Sandau, M., Koblauch, H., Moeslund, T. B., Aanæs, H., Alkjær, T., & Simonsen, E. B. (2014). Markerless motion capture can provide reliable 3D gait kinematics in the sagittal and frontal plane. *Medical Engineering & Physics*, *36*(9), 1168-1175.

Schmitz, A., Ye, M., Shapiro, R., Yang, R., & Noehren, B. (2014). Accuracy and repeatability of joint angles measured using a single camera markerless motion capture system. *Journal of biomechanics*, *47*(2), 587-591.

Schröder, M., Waltemate, T., Maycock, J., Röhlig, T., Ritter, H., & Botsch, M. (2018). Design and evaluation of reduced marker layouts for hand motion capture. *Computer Animation and Virtual Worlds*, *29*(6), e1751.

Steinebach, T., Grosse, E. H., Glock, C. H., Wakula, J., & Lunin, A. (2020). Accuracy evaluation of two markerless motion capture systems for measurement of upper extremities: Kinect V2 and Captiv. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *30*(4), 291-302.

Steffensen, M. S., Magalhaes PhD, F., Knarr PhD, B. A., & Kingston PhD, D. C. (2023). Comparison of markerless and marker-based motion capture in cerebral palsy and chronic stroke gait kinematics.

Strutzenberger, G., Kanko, R., Selbie, S., Schwameder, H., & Deluzio, K. (2021). Assessment of kinematic CMJ data using a deep learning algorithm-based markerless motion capture system. *ISBS Proceedings Archive*, *39*(1), 236.

Szeliski, R. (2022). *Computer vision: algorithms and applications*. Springer Nature.

Tharsanee, R. M., Soundariya, R. S., Kumar, A. S., Karthiga, M., & Sountharrajan, S. (2021). Deep convolutional neural network–based image classification for COVID-19 diagnosis. In *Data Science for COVID-19* (pp. 117-145). Academic Press.

Uhlrich, S. D., Falisse, A., Kidziński, Ł., Muccini, J., Ko, M., Chaudhari, A. S., ... & Delp, S. L. (2022). OpenCap: 3D human movement dynamics from smartphone videos. *bioRxiv*, 2022-07.–

van den Noort, J. C., Kortier, H. G., Beek, N. V., Veeger, D. H., & Veltink, P. H. (2016). Measuring 3D hand and finger kinematics—a comparison between inertial sensing and an opto-electronic marker system. *PLoS One*, *11*(11), e0164889.

Valevicius, A. M., Jun, P. Y., Hebert, J. S., & Vette, A. H. (2018). Use of optical motion capture for the analysis of normative upper body kinematics during functional upper limb tasks: A systematic review. *Journal of electromyography and kinesiology*, *40*, 1-15.

Wei, S. E., Ramakrishna, V., Kanade, T., & Sheikh, Y. (2016). Convolutional pose machines. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 4724-4732).

Wren, T. A., Isakov, P., & Rethlefsen, S. A. (2023). Comparison of kinematics between theia markerless and conventional marker-based gait analysis in clinical patients. *Gait & Posture*, *104*, 9-14.

Yang, L., Zhang, L., Dong, H., Alelaiwi, A., & El Saddik, A. (2015). Evaluating and improving the depth accuracy of Kinect for Windows v2. *IEEE Sensors Journal*, *15*(8), 4275-4285

Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C. L., & Grundmann, M. (2020). Mediapipe hands: On-device real-time hand tracking. *arXiv preprint arXiv:2006.10214*.

Zhang, J. T., Novak, A. C., Brouwer, B., & Li, Q. (2013). Concurrent validation of Xsens MVN measurement of lower limb joint angular kinematics. *Physiological measurement*, *34*(8), N63.

Zhao, W., Chai, J., & Xu, Y. Q. (2012, July). Combining marker-based mocap and rgb-d camera for acquiring high-fidelity hand motion data. In *Proceedings of the ACM SIGGRAPH/eurographics symposium on computer animation* (pp. 33-42).

Zhou, Y., & Tuzel, O. (2018). Voxelnet: End-to-end learning for point cloud based 3d object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4490-4499).

APPENDIX A – SUPPLEMENTARY TABLES

Table A1: Estimated marginal means [95% confidence intervals] for the $\mathbf{R^2}$ for each finger and task across different camera pairs. Pair 1 refers to C/90° to C/L45°, Pair 2 refers to C/90° to L45°/R45°, and Pair 3 refers to C/L45° to L45°/R45°.

	Index			Middle			Ring			Little			Thumb		
	Pair 1	Pair 2	Pair 3	Pair 1	Pair 2	Pair 3	Pair 1	Pair 2	Pair 3	Pair 1	Pair 2	Pair 3	Pair 1	Pair 2	Pair 3
Concurrent finger Flexion	0.95 [0.94, 0.95]	0.92 [0.92, 0.93]	0.93 [0.92, 0.94]	0.96 [0.96, 0.97]	0.95 [0.94, 0.96]	0.98 [0.98, 0.99]	0.96 [0.95, 0.97]	0.94 [0.94, 0.95]	0.99 [0.98, 0.99]	0.94 [0.93, 0.95]	0.91 [0.91, 0.92]	0.98 [0.97, 0.99]	0.97 [0.97, 0.98]	0.71 [0.71, 0.72]	0.74 [0.73, 0.74]
Sequential Finger Flexion	0.77 [0.76, 0.78]	0.34 [0.33, 0.36]	0.38 [0.36, 0.39]	0.69 [0.68, 0.71]	0.64 [0.63, 0.66]	0.76 [0.74, 0.77]	0.50 [0.49, 0.51]	0.44 [0.42, 0.45]	0.76 [0.75, 0.77]	0.55 [0.54, 0.57]	0.54 [0.52, 0.55]	0.80 [0.79, 0.81]	0.88 [0.87, 0.90]	0.17 [0.15, 0.18]	0.16 [0.15, 0.18]
Typing	0.71 [0.69, 0.73]	0.56 [0.55, 0.58]	0.55 [0.53, 0.57]	0.78 [0.76, 0.79]	0.77 [0.75, 0.78]	0.83 [0.81, 0.84]	0.62 [0.61, 0.64]	0.65 [0.63, 0.66]	0.75 [0.73, 0.77]	0.45 [0.43, 0.47]	0.50 [0.48, 0.51]	0.74 [0.72, 0.76]	0.82 [0.80, 0.83]	0.27 [0.26, 0.29]	0.27 [0.25, 0.29]
Index Finger Flexion	0.86 [0.84, 0.88]	0.52 [0.51, 0.54]	0.62 [0.60, 0.64]	-	-	-	-	-	-	-	-	-	-	-	-

Table A2: Estimated marginal means [95% confidence intervals] for the RMSD (°) for each finger and task, across different camera pairs. Pair 1 refers to C/90° to C/L45°, Pair 2 refers to C/90° to L45°/R45°, and Pair 3 refers to C/L45° to L45°/R45°.

	Index			Middle			Ring			Little			Thumb		
	Pair 1	Pair 2	Pair 3	Pair 1	Pair 2	Pair 3	Pair 1	Pair 2	Pair 3	Pair 1	Pair 2	Pair 3	Pair 1	Pair 2	Pair 3
Finger Flexion	10.80 [9.95, 11.64]	16.79 [15.94, 17.63]	15.63 [14.79, 16.48]	9.47 [8.62, 10.31]	9.76 [8.92, 10.60]	8.41 [7.56, 9.25]	9.51 [8.67, 10.36]	10.27 [9.42, 11.11]	5.72 [4.87, 6.56]	10.09 [9.25, 10.93]	13.35 [12.50, 14.19]	6.99 [6.15, 7.83]	6.06 [5.21, 6.90]	22.44 [21.59, 23.28]	22.03 [21.19, 22.87]
Finger Roll Flexion	13.18 [12.38, 13.98]	19.97 [19.17, 20.77]	22.04 [21.24, 22.83]	12.10 [11.30, 12.90]	11.44 [10.64, 12.24]	10.89 [10.09, 11.69]	20.08 [19.28, 20.88]	19.06 [18.26, 19.86]	9.89 [9.09, 10.68]	21.09 [20.29, 21.89]	20.30 [19.51, 21.10]	12.17 [11.37, 12.97]	10.33 [9.53, 11.13]	39.62 [38.82, 40.41]	38.49 [37.69, 39.29]
Typing	13.09 [11.90, 14.28]	28.74 [27.55, 29.93]	27.52 [26.33, 28.71]	10.76 [9.57, 11.95]	10.80 [9.61, 11.99]	11.48 [10.29, 12.67]	14.11 [12.92, 15.30]	12.27 [11.08, 13.47]	11.62 [10.43, 12.81]	17.01 [15.82, 18.20]	15.94 [14.75, 17.13]	12.44 [11.25, 13.63]	9.25 [8.06, 10.44]	37.94 [36.75, 39.13]	37.82 [36.63, 39.01]
Index Flexion	14.13 [13.16, 15.10]	34.34 [33.37, 35.31]	33.40 [32.44, 34.37]	-	-	-	-	-	-	-	-	-	-	-	-

Table A3: Pairwise differences in means [95% confidence intervals] for the R^2 for each finger and task, across different camera pairs. The values should be read as (row-column). For instance, the R^2 of 0.023 at the finger flexion row, column (Pair 1, Pair 2) is interpreted as a 0.023 higher average for all participants when comparing the Pair 1 R^2 to the Pair 2 R^2. Pair 1 refers to C/90° to C/L45°, Pair 2 refers to C/90° to L45°/R45°, and Pair 3 refers to C/L45° to L45°/R45°.

		Index		Middle		Ring		Little		Thumb	
		Pair 2	**Pair 3**	**Pair 2**	**Pair 3**	**Pair 2**	**Pair 3**	**Pair 2**	**Pair 3**	**Pair 2**	**Pair 3**
Concurrent Finger Flexion	**Pair 1**	0.023 [0.010, 0.037]	0.019 [0.006, 0.032]	0.014 [0.001, 0.027]	-0.018 [-0.032, -0.005]	0.016 [0.003, 0.029]	-0.027 [-0.040, -0.014]	0.026 [0.012, 0.039]	-0.040 [-0.053, -0.026]	0.259 [0.246, 0.273]	0.237 [0.224, 0.250]
	Pair 2	-	-0.004 [-0.017, 0.009]	-	-0.032 [-0.045, -0.019]	-	-0.043 [-0.057, -0.030]	-	-0.065 [-0.079, -0.052]	-	-0.023 [-0.036, -0.009]
Index Finger Flexion	**Pair 1**	0.337 [0.308, 0.365]	0.240 [0.212, 0.269]	0.021 [-0.008, 0.050]	-0.254 [-0.283, -0.225]	0.014 [-0.015, 0.043]	-0.423 [-0.452, -0.394]	0.068 [0.039, 0.096]	-0.485 [-0.514, -0.456]	0.350 [0.321, 0.378]	0.270 [0.241, 0.299]
	Pair 2	-	-0.096 [-0.125, -0.067]	-	-0.275 [-0.304, -0.246]	-	-0.437 [-0.466, -0.408]	-	-0.552 [-0.581, -0.523]	-	-0.080 [-0.108, -0.051]
Typing	**Pair 1**	0.147 [0.119, 0.176]	0.158 [0.130, 0.187]	0.010 [-0.019, 0.038]	-0.049 [-0.078, -0.021]	-0.023 [-0.052, 0.005]	-0.125 [-0.154, -0.097]	-0.048 [-0.077, -0.020]	-0.292 [-0.320, -0.263]	0.542 [0.513, 0.571]	0.547 [0.519, 0.576]
	Pair 2	-	0.011 [-0.018, 0.039]	-	-0.059 [-0.088, -0.031]	-	-0.102 [-0.130, -0.073]	-	-0.243 [-0.272, -0.215]	-	0.005 [-0.023, 0.034]
Sequential Finger Flexion	**Pair 1**	0.429 [0.405, 0.453]	0.393 [0.369, 0.417]	0.051 [0.027, 0.075]	-0.063 [-0.087, -0.039]	0.061 [0.037, 0.085]	-0.260 [-0.285, -0.236]	0.017 [-0.007, 0.041]	-0.246 [-0.270, -0.222]	0.717 [0.692, 0.741]	0.721 [0.697, 0.745]
	Pair 2	-	-0.036 [-0.060, -0.012]	-	-0.114 [-0.138, -0.090]	-	-0.321 [-0.345, -0.297]	-	-0.263 [-0.287, -0.239]	-	0.005 [-0.019, 0.029]

Table A4: Pairwise differences in means [95% confidence intervals] for the RMSD for each finger and task, across different camera pairs. The values should be read as (row-column). For instance, the RMSD of -5.99° at the finger flexion row, column (Pair 1, Pair 2) is interpreted as a 5.9° higher average for all participants when comparing the Pair 1 RMSD to the Pair 2 RMSD. Pair 1 refers to C/90° to C/L45°, Pair 2 refers to C/90° to L45°/R45°, and Pair 3 refers to C/L45° to L45°/R45°.

		Index		Middle		Ring		Little		Thumb	
		Pair 2	**Pair 3**	**Pair 2**	**Pair 3**	**Pair 2**	**Pair 3**	**Pair 2**	**Pair 3**	**Pair 2**	**Pair 3**
Concurrent Finger Flexion	**Pair 1**	-5.99 [-7.42, -4.57]	-4.84 [-6.26, -3.41]	-0.29 [-1.72, 1.13]	1.06 [-0.37, 2.48]	-0.75 [-2.18, 0.67]	3.80 [2.37, 5.22]	-3.26 [-4.68, -1.83]	3.10 [1.67, 4.52]	-16.38 [-17.81, -14.96]	-15.97 [-17.40, -14.55]
	Pair 2		1.15 [-0.27, 2.58]		1.35 [-0.07, 2.78]		4.55 [3.12, 5.97]		6.35 [4.93, 7.78]		0.41 [-1.02, 1.83]
Index Finger Flexion	**Pair 1**	-20.21 [-21.85, -18.57]	-19.28 [-20.92, -17.64]	0.45 [-1.19, 2.09]	4.23 [2.59, 5.86]	0.78 [-0.86, 2.42]	8.97 [7.33, 10.61]	-1.49 [-3.13, 0.15]	6.25 [4.62, 7.89]	-26.33 [-27.96, -24.69]	-23.86 [-25.50, -22.22]
	Pair 2		0.93 [-0.71, 2.57]		3.78 [2.14, 5.42]		8.19 [6.55, 9.83]		7.74 [6.10, 9.38]		2.47 [0.83, 4.10]
Typing	**Pair 1**	-15.65 [-17.66, -13.64]	-14.43 [-16.44, -12.42]	-0.03 [-2.05, 1.98]	-0.72 [-2.73, 1.29]	1.83 [-0.18, 3.85]	2.48 [0.47, 4.50]	1.07 [-0.94, 3.09]	4.57 [2.56, 6.59]	-28.69 [-30.70, -26.68]	-28.57 [-30.58, -26.56]
	Pair 2		1.22 [-0.79, 3.23]		-0.68 [-2.70, 1.33]		0.65 [-1.36, 2.66]		3.50 [1.49, 5.51]		0.12 [-1.89, 2.13]
Sequential Finger Flexion	**Pair 1**	-6.79 [-8.14, -5.44]	-8.85 [-10.20, -7.51]	0.66 [-0.69, 2.01]	1.21 [-0.14, 2.56]	1.02 [-0.33, 2.37]	10.19 [8.85, 11.54]	0.79 [-0.56, 2.14]	8.92 [7.57, 10.27]	-29.29 [-30.64, -27.94]	-28.16 [-29.51, -26.81]
	Pair 2		-2.06 [-3.41, -0.72]		0.55 [-0.80, 1.90]		9.17 [7.82, 10.52]		8.13 [6.78, 9.48]		1.13 [-0.22, 2.47]

APPENDIX B – SUPPLEMENTARY FIGURES

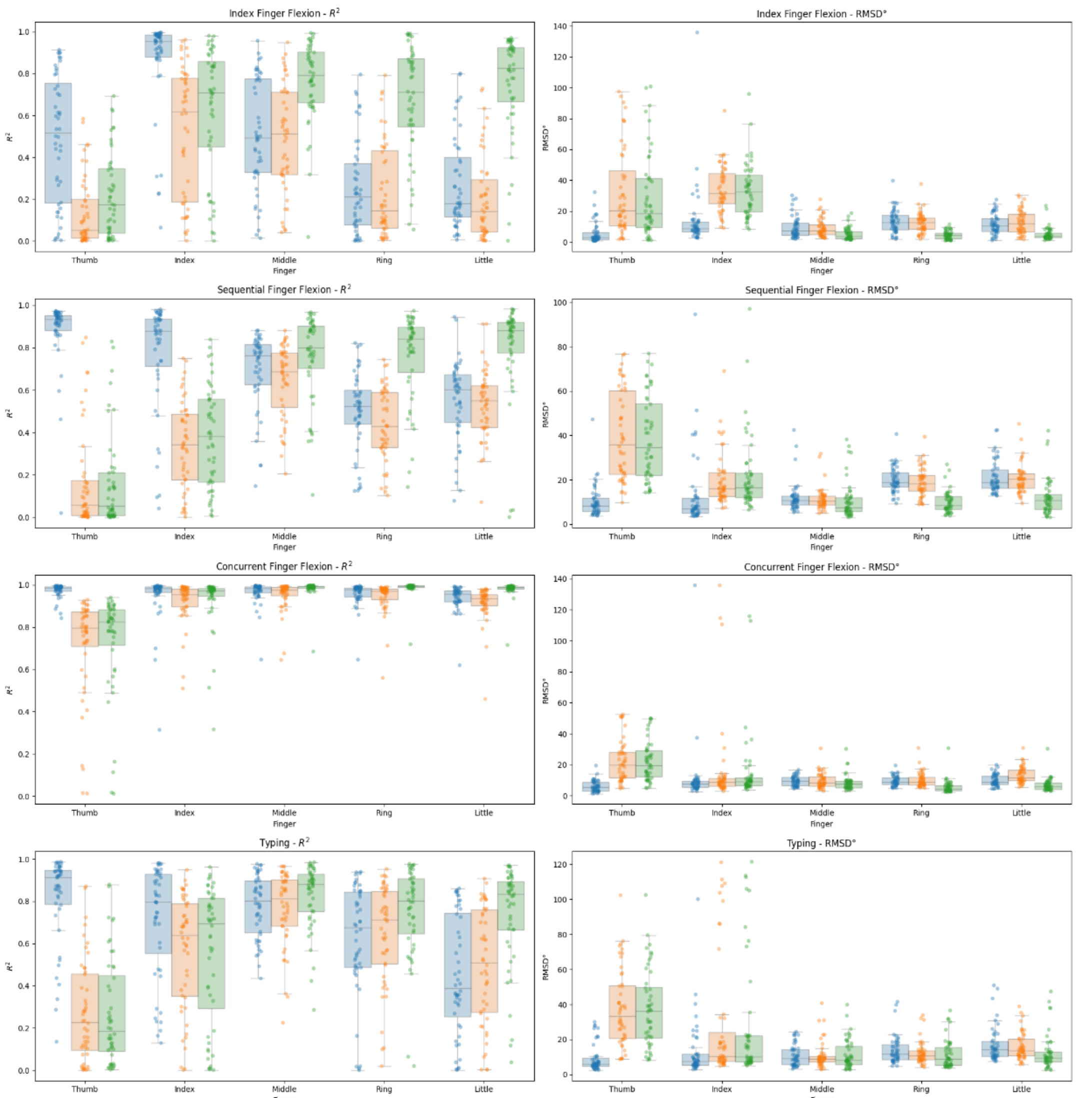

Figure A1: Box plot comparison of camera pair R² and RMSD across multiple fingers (thumb, index, middle, ring, and little) and tasks (index finger flexion, finger roll flexion, finger flexion, and typing) with the markers off (non-concurrent). The left column presents the R² for all tasks and fingers, indicating the correlation between camera pairs, while the right column shows the corresponding RMSD, reflecting the differences in angles measured between camera pairs. The **(blue) box plots** represent the comparison between C/90° and C/L45°, the **(orange) box plots** represent the comparison between C/90° and L45°/R45°, and the **(green) box plots** represent the comparison between C/L45° and L45°/R45°. Each point within the box plots denotes the R² or RMSD for each trial for each participant.

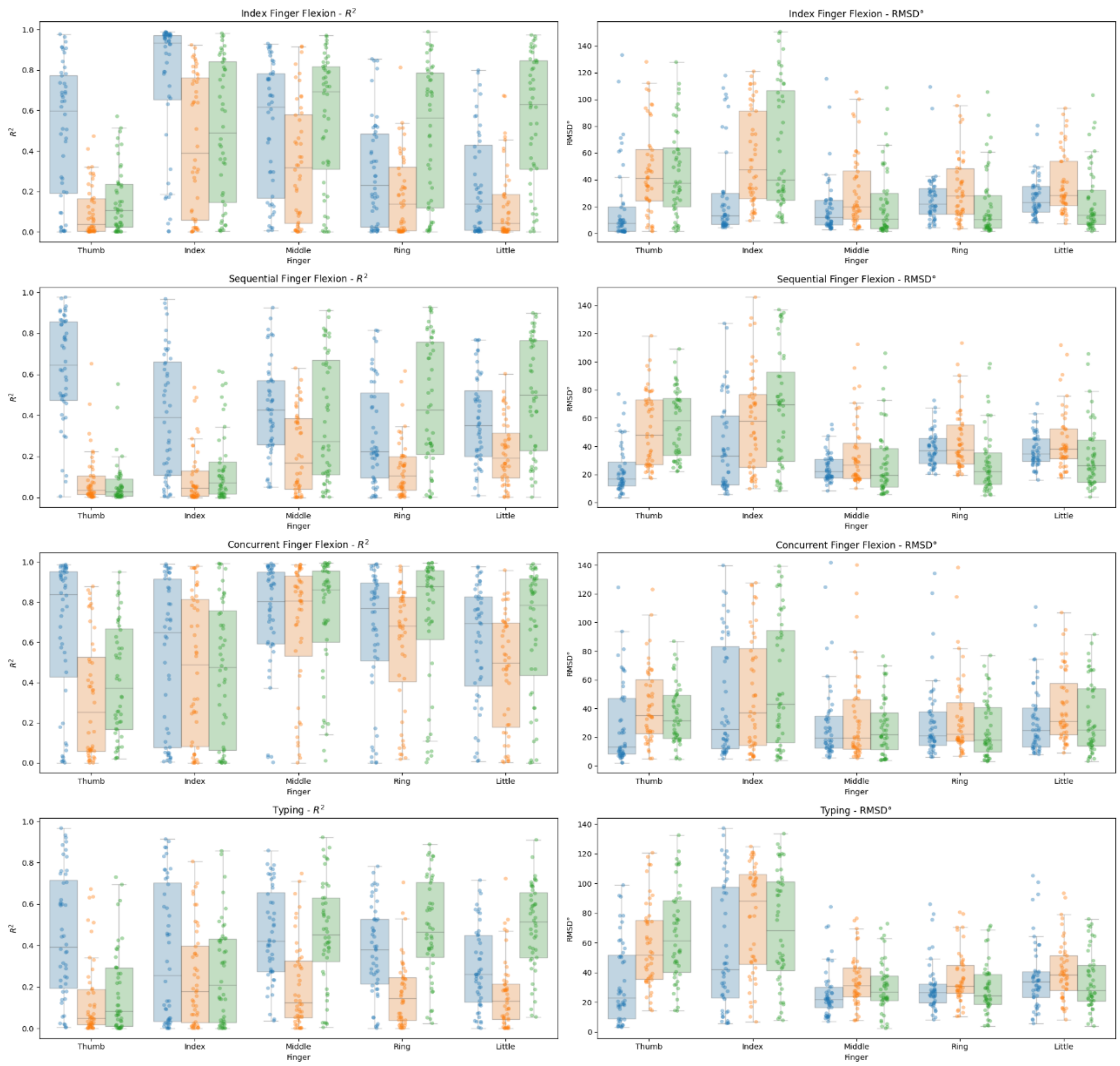

Figure A2: Box plot comparison of camera pair R^2 and RMSD across multiple fingers (thumb, index, middle, ring, and little) and tasks (index finger flexion, finger roll flexion, finger flexion, and typing) with the markers on (concurrent). The left column presents the R^2 for all tasks and fingers, indicating the correlation between camera pairs, while the right column shows the corresponding RMSD values, reflecting the differences in angles measured between camera pairs. The **(blue) box plots** represent the comparison between C/90° and C/L45°, the **(orange) box plots** represent the comparison between C/90° and L45°/R45°, and the **(green) box plots** represent the comparison between C/L45° and L45°/R45°. Each point within the box plots denotes the R^2 or RMSD for each trial for each participant.

APPENDIX C – PILOT STUDY NACOB ABSTRACT

MARKERLESS MOTION CAPTURE OF THE HAND AND FINGERS

Nigel Majoni, Daanish M. Mulla, & Peter J. Keir*
Department of Kinesiology, McMaster University, Hamilton, ON, Canada, L8S 4K1
email: pjkeir@mcmaster.ca

Introduction

Hand and finger movements are poorly captured in most workplace assessments but play important roles in the development of hand-related musculoskeletal disorders [1]. All manual tasks require the use of the hands and loading of the upper extremity typically occurs through the hands. However, the hands are often poorly represented in most biomechanical and ergonomic research studies. Marker-based systems have been used to evaluate the kinematics of the hand but are typically limited to laboratory settings [2]. Marker-based motion capture systems may provide the desired accuracy but are not practical or feasible in the workplace. Computer vision is a field of artificial intelligence that allows computers to extract information from images and videos. Recent developments in computer vision have influenced the production of programs with the ability to track motion by recording a simple video of the desired motion, without the use of markers. This work aims to use an open source markerless motion capture program and train a network to evaluate and track hand and finger movements and postures, allowing easier collection in the workplace.

Methods

DeepLabCut is an open-source Python package used for 2D and 3D markerless animal pose estimation [3, 4]. Currently, a 2D model was developed to track hand motion using DeepLabCut. Preparing to train the network began with collecting a sample video at 30 Hz and 1080p. 10 participants (five males and five females) were recruited to complete three hand tasks: wrist flexion\extension and forearm pronation\supination, grasping, and finger manipulation. A total of 1400 frames were extracted from the videos collected by way of Kmeans clustering and were then manually digitized to give the model a training dataset to be trained identify joint locations. Twenty-one anatomical landmarks were identified. Specifically, the model identified the wrist, the carpometacarpal joint of the thumb, the metacarpal joints of digits 1-5, the proximal interphalangeal joints of digits 2-5, the distal interphalangeal joints of digits 1-5, and the endpoints of digits 1-5 The network was trained with 500,000 iterations using a training dataset to enable 2D motion capture. Evaluation of the trained network determined the performance of the trained neural network. Figure 1 shows the results of the trained network in identifying joint locations of the hand.

Results and Discussion

The current model is limited to analyzing movements in two dimensions using a single camera. Following successful model development in 2D, a second camera will be included to move to three-dimensional tracking of hand motion. The multicamera system will be developed with mobility in mind such that non-laboratory settings may be used for video capture. Training will also include the industrial use of gloves and tools to ensure capability. This will allow us to develop a model capable of markerless capture of hand and finger movements in the workplace that can be integrated with other force-sensing equipment. These kinematic and external force data may then be used as input to our hand simulations in OpenSim. Ultimately, this data will be used to improve the representation of the hand in biomechanical and ergonomics assessments to improve our assessment of the risk of developing musculoskeletal disorders in the workplace

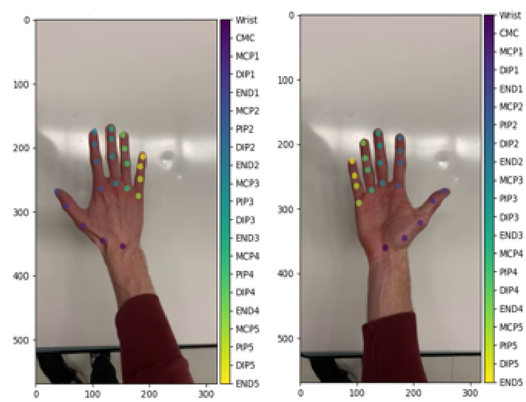

Figure 1: (a) Joint locations of the hand modelled in pronation within DeepLabCut. (b) Joint locations of the hand modelled in supination within DeepLabCut. Marker colours represent each joint or anatomical location based on the scale at the

Significance

The hand and fingers are currently poorly represented, or even neglected, in most evaluations of external loads acting on the body. An accurate markerless system to capture the kinematics of the hand and fingers in the workplace would immensely improve our ability to assess the injury risk in the upper extremity. Concurrent to the 2D model, we are also working on using multiple cameras to track 3D bilateral hand movements with the open-source Python toolkit Anipose [5]. Ultimately, this will be incorporated with real-time force inputs for the rapid output of assessments.

Acknowledgments

Funding was provided by NSERC (RGPIN-2016-06460).

References

[1] Amell et al., 2001. *J Occupational rehabilitation.* 11:255-265
[2] Cocchiarella et al., 2016. *Comput Method Biomech Biomed Engin.* 19:639-647
[3] Mathis et al., 2018. *Nat Neurosci.* 21:1281-1289.
[4] Nath et al., 2019. *Nat Protoc.* 14:2152-2176.
[5] Karashcuk et al. 2021. *Cell Rep.* 36:1

APPENDIX D – ETHICS CONSENT FORM

McMaster University

Inspiring Innovation and Discovery

LETTER OF INFORMATION AND CONSENT

Validity of markerless motion capture for the hands and fingers

Principal Investigator:

Dr. Peter Keir, PhD

Department of Kinesiology

McMaster University

Hamilton, Ontario, Canada

(905) 525-9140 ext. 23543

E-mail: pjkeir@mcmaster.ca

Student Investigator:

Nigel Majoni, MSc Candidate

Department of Kinesiology

McMaster University

Hamilton, Ontario, Canada

(905) 525-9140

E-mail: majonit@mcmaster.ca

Purpose of the Study:

We use our hands to interact with the external environment every day. Hand and finger postures and movement play important roles in the development of hand-related musculoskeletal disorders and injuries. Nevertheless, these postures and movements are often poorly represented in many biomechanical and ergonomic studies. As such, a broader understanding of how the hands and fingers interact with the world is needed to expand the avenue in assessing the risk of developing hand-related musculoskeletal disorders in the workplace and at home. Movement varies between individuals, and assessing hand activity is difficult. Tracking upper limb movements, particularly in the hands and fingers, is challenging. Simple hand movements such as grasping depend on several factors ranging from the size and shape of the object to the postures being adopted by the hand and fingers. Furthermore, accurate assessment of hand and finger kinematics can be limited by technology. Historically, human movement is tracked by putting markers on specific points of the body. These are bulky and intrusive and will likely change how people move. In recent years, there has been an advancement in the evaluation of markerless motion capture. There is a growing interest in investigating the feasibility of markerless motion capture systems as a viable alternative to marker-based systems and exploring their potential for advancing research in motion analysis. Currently, no markerless motion capture system is available that can provide accurate tracking of hand and finger kinematics. Consequently, we need to determine the accuracy of markerless motion capture systems for hand and finger kinematics compared to marker-based motion capture systems.

Procedures involved in the Research:

The study will involve a single laboratory session taking approximately 2 hours to complete. All procedures will be completed by the researchers in the study.

1) An informed consent form with details of the experiment will be explained, and all questions will be answered before signing.
2) Participants will be seated at the assessment table, and markers will be placed on the participants by the researcher to collect motion capture data. Anatomical landmarks will be palpated, and participants will be asked to flex certain joints to ensure the location of the specific body being collected. Markers will be placed at landmarks on the right hand. The markers will be taped down using tape. These procedures are required to obtain a high-quality signal.
3) Participants will be told that video recording will begin for each task performed using the Logitech c920 webcam. These cameras will only be recording their hands and finger up until the elbow and upper arm.
4) Once the markers and video is recording are secure, the participant will be asked to complete a series of six (6) functional hand tasks. The tasks are:
    I. Typing on a keyboard. Participants will be asked to type out a phrase three (3) times "The quick brown fox jumps over the lazy dog." The sentence contains all the letters of the alphabet, and all the keys on the keyboard will be used.
    II. Stacking a Jenga tower. Participants will be asked to stack half a Jenga tower, and then unstack the tower. They will be instructed to do so by taking three (3) building blocks and stacking them perpendicular to the other blocks and unstacking them the exact same way.
    III. Perform the peg test. Participants will have the pegboard in front of them with pegs beside them. They will be instructed to insert all the pegs available into the pegboard

in sequential order and then remove the pegs in sequential order. They will perform three (3) rounds of this to complete the task.

IV. Open fist, close fist task. Participants will be asked to perform this task using a metronome (sound cue) and do this for 15 seconds. It will be repeated 3 times.

V. Open hand forearm pronation and supination task. Participants will be asked to perform this task using a metronome (sound cue) and do this for 15 seconds. It will be repeated 3 times.

VI. Open hand wrist flexion and extension task. Participants will be asked to perform this task using a metronome (sound cue) and do this for 15 seconds. It will be repeated 3 times.

Each of these tasks will be completed for three (3) trials.

5) Once the participants have completed data collection, markers will be removed.

Potential Harms, Risks or Discomforts:

Minimal risks are anticipated for this study.

Psychological Risk

There is a potential risk that you might feel embarrassed if you cannot perform the tasks correctly. This will be managed by explaining and demonstrating how to perform each task prior to collection correctly.

Social Risk

There is a potential risk to privacy regarding the video recording. This will be managed by ensuring there are no distinguishable features in the recordings, and all videos will be kept on a password-protected computer.

Skin Sensitivity

You may experience mild skin irritation/redness from the adhesive of the reflective marker and tape. This is similar to the irritation that may be caused by a bandage and typically fades within 2 to 3 days.

Potential Benefits:

The outcome of the study will allow us to validate our markerless motion capture system. The research will not benefit you directly.

Incentive

You will receive $20 for participating in this study as remuneration for your time. Your contact information may be shared with the Kin Grad Admin to ensure your compensation. The amount received is taxable. It is your responsibility to report this amount for income tax purposes.

Confidentiality:

Your identity will be kept confidential, and the data collected will be used for research purposes only. The information directly pertaining to you will be locked in a cabinet or stored electronically on a password-protected computer for 7 years. During the collection, there may be undergraduate research assistants present in the lab space.

Participation:

Your participation in this study is voluntary. If you decide to participate, you can decide to stop at any time, even after signing the consent form or part-way through the study. If you decide to stop participating, there will be no consequences for you. If you choose to withdraw at any time in the study, you will still be compensated for your time. Once the data collection is completed, you will not be able to withdraw from the study, as the data is being collected without your name. .

Information about the Study Results:

You may obtain information about the study results by contacting Dr. Peter Keir at (905) 525-9140 (x 23543) or indicating "Yes" at the bottom of this form.

Questions about the Study:

If you have questions or need more information about the study itself, please contact me at:

Nigel Majoni
majonit@mcmaster.ca

This study has been reviewed by the McMaster University Research Ethics Board and received ethics clearance. If you have concerns or questions about your rights as a participant or about the way the study is conducted, please contact:

McMaster Research Ethics Secretariat

Telephone: (905) 525-9140 ext. 23142

C/o Research Office for Administrative Development and Support

E-mail: ethicsoffice@mcmaster.ca

CONSENT

- I have read the information presented in the information letter about a study being conducted by Nigel Majoni and Dr. Peter J. Keir, of McMaster University.
- I have had the opportunity to ask questions about my involvement in this study and to receive the additional details I requested.
- I understand that if I agree to participate in this study, I may withdraw from the study at any time.

- I understand that if I choose to withdraw from the study after the collection, you cannot withdraw your data since your data is being collected without your name
- I understand I will receive a signed copy of this form via email
- I agree to be recorded via marker-based and markerless motion capture to capture motion data of the upper extremity. Only your hands and fingers and potentially your elbow will be recorded.
- I agree to participate in the study.

Signature: ____________________________________ Date: _________________________

Name of Participant (Printed) __________________________________

1._____ **YES**, I would like to receive a summary of the study's results

Please send them to me at this email address: ____

OR to this mailing address: ____

____

____

______ NO, I do not want to receive a summary of the study's results

Person Obtaining Consent

Signature: ____________________________________ Date: ________________________

Name of Participant (Printed) __________________________________