2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

Optimizing Interaction Space: Enlarging the Capture Volume for Multiple Portable Motion Capture Devices

Muhammad Hilman Fatoni[1,2**], Christopher Herneth[1**], Junnan Li[1], Fajar Budiman[1,2], Amartya Ganguly[1], and Sami Haddadin[1]

***Abstract*— Markerless motion capture devices such as the Leap Motion Controller (LMC) have been extensively used for tracking hand, wrist, and forearm positions as an alternative to Marker-based Motion Capture (MMC). However, previous studies have highlighted the subpar performance of LMC in reliably recording hand kinematics. In this study, we employ four LMC devices to optimize their collective tracking volume, aiming to enhance the accuracy and precision of hand kinematics. Through Monte Carlo simulation, we determine an optimized layout for the four LMC devices and subsequently conduct reliability and validity experiments encompassing 1560 trials across ten subjects. The combined tracking volume is validated against an MMC system, particularly for kinematic movements involving wrist, index, and thumb flexion. Utilizing calculation resources in one computer, our result of the optimized configuration has a better visibility rate with a value of 0.05 $\pm$ 0.55 compared to the initial configuration with -0.07 $\pm$ 0.40. Multiple Leap Motion Controllers (LMCs) have proven to increase the interaction space of capture volume but are still unable to give agreeable measurements from dynamic movement.**

I. Introduction

Marker-based motion capture (MMC) serves as the gold standard for recording human kinematics [1]. In clinical settings, it proves invaluable for assessing movement accuracy, identifying injury risk factors [2], and personalizing diagnostic approaches for conditions like stroke, Parkinson's disease, cerebral palsy, spinal cord injury, and multiple sclerosis [3].

However, MMC systems are expensive, time-consuming to set up and calibrate, and not commonly available while requiring trained personnel [4]. In clinical settings, where patient comfort and convenience are paramount, portable and low-cost markerless motion capture provides a more sensitive tool for research and rehabilitation [5]. Leap Motion controller™ (LMC) represents one such markerless motion capture device that utilizes infrared cameras and a proprietary internal model for tracking hand and finger movements. Despite its low cost and ease of use, a single LMC device did not demonstrate acceptable agreement with marker-based motion capture [6], [7].

To overcome this limitation, Novacek et al. [8] fused angle estimations of multiple LMCs. The Kabsch algorithm [9] was used with weights assigned to recordings of individual LMCs based on confidence levels computed from palm orientation angles. However, their method only gauged confidence levels for the entire hand, while credence in individual fingers was not considered. In this study, a ray-tracing algorithm was proposed, estimating the visibility of individual finger phalanges. Occlusions of virtual markers in LMC measurement frames were computed based on simplified geometric hand models and estimated hand configurations. Resulting finger occlusion metrics and measurements of hand properties from multiple LMCs' marker fused by a Kalman filter [10].

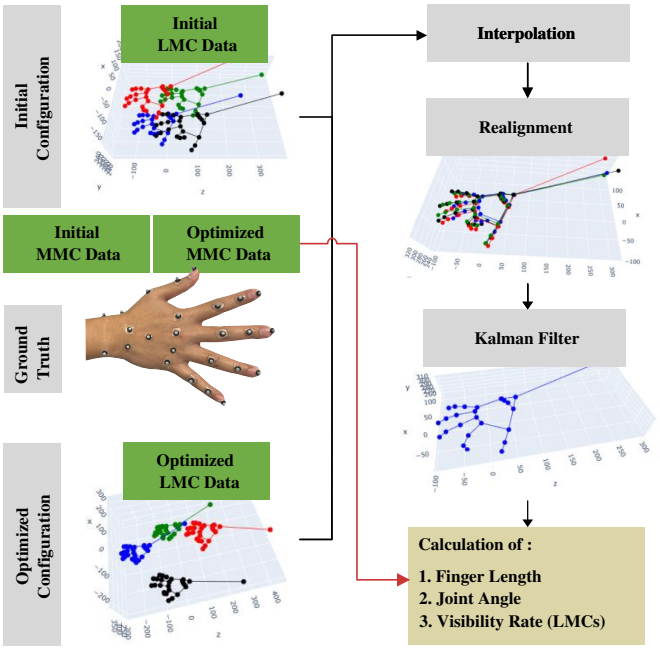

Fig. 1. Multiple LMCs validation pipeline for initial and optimized LMC placement. LMC data is captured - interpolated - realigned - Kalman filtered. Comparisons are made to ground truth optical marker trajectories from MMC.

However, challenges persist, particularly regarding the reliability of capture volume and tracking performance inherent to the LMC. Consequently, optimized LMC positions were computed for multiple LMCs, based on Monte Carlo Simulation [11] of hand trajectories expected in trials. LMCs were located and oriented such that the visibility of as many fingers by as many LMC as possible was ensured in each trial frame.

Figure 1 shows the complete pipeline of this study. Our contributions are as follows: (1) an online multi-LMC framework, fusing measurements of multiple devices where occlusions metrics calculated in the ray-tracing algorithm,

** equal contribution

[1]The authors are with the Munich Institute of Robotics and Machine Intelligence (MIRMI), Technische Universität München (TUM), Germany. [2]The authors are also with Institut Teknologi Sepuluh Nopember (ITS), Indonesia. {muhammadhilman.fatoni, christopher.herneth, junnan.li, fajar.budiman, amartya.ganguly, and haddadin}@tum.de

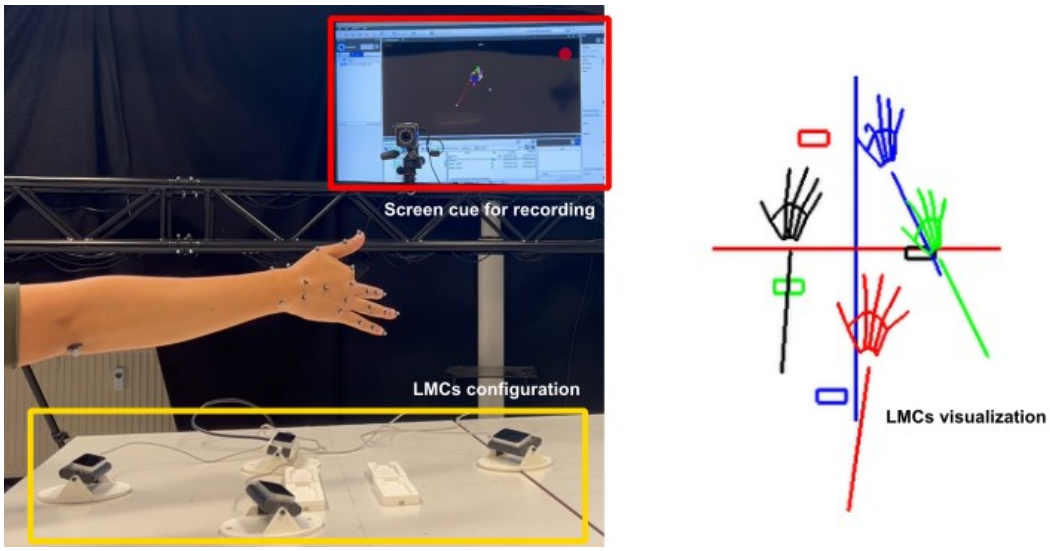

Fig. 2. Experiment setup. Left: Simultaneous recording of MMC and LMCs system. Markers were placed on the hand while an experiment of index flexion movement in the vertical pose of optimized LMCs configuration was running. The cue recording is a big red circle shape, informing the subject to start the trial. Right: Visualization of LMCs marker reading using OpenGL in developed custom program.

(2) an LMC placement optimization, computing ideal LMC placement of multiple devices based on LMC frame samples expected during trials, (3) validation against gold standard MMC, demonstrating the effectiveness of our approach. In summary, this study alleviates the limitations of singular LMC devices and posits a solution for markerless motion capture relevant for rehabilitation and clinical measurements to facilitate reliable diagnosis for hand impairments.

II. Materials and Methods

A. Subject Information

Ten right-handed male subjects with a mean age of 28.90±4.95 years participated in this study. The subjects gave their written informed consent. The study is registered with ethics serial number 2022-588-S-KH. It was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of the Technical University of Munich.

B. Markerless Motion Capture System

The LMC is a markerless portable motion capture device (Ultraleap, Bristol, UK). LMC consists of two 640x240-pixel near-infrared cameras separated 40 mm apart and three LEDs spaced on either side and between the cameras. The dimensions of LMC are 13 mm in height, 80 mm in width, and 30 mm in depth, weighing around 32 grams. Although LMC is extensively used in gaming, several studies have utilised LMC and robots simultaneously [12], [13]. LMC typically operates at 120 Hz [14], but the actual sampling operation is a variable sampling [6], [15], [16]. LMC has an interaction zone of about 60 cm, extending from the device in a 120x150° field of view. It can track 28 marker positions of the hand from the elbow to the tip of the fingers. Previous research [6], [17] used these marker positions to measure hand anthropometrics.

Recent developments of LMC have made it possible to read multiple LMCs simultaneously. The marker data read from LMC following a structured format. They consisted of point coordinates, which are the elbow, wrist, palm, and bones of the fingers. The LMC employed a right-handed Cartesian coordinate system [18]. The local origin of LMC was located at the centre top. The x- and z-axes lie in the plane of the camera sensors, with the x-axis running along the camera baseline. The y-axis is vertical, and the z-axis points towards the user. This study utilised four LMCs in two configurations. The first configuration was decided by placing all four LMCs on a flat surface where all LMCs' centre positions are placed in each corner of a box shape. This initial arrangement was placed assuming it would increase the capture volume. The second configuration was decided from the result of the optimization algorithm, which is shown in Figure 3 (blue LMC configuration). During the pilot testing phase of this study, LMC hand-tracking capability was tested to determine whether the performance was within the datasheet specification. Captured marker data served as indicators to determine whether hand position within LMC interaction space. However, it was found that at a distance of 250 mm from the origin, the value from this axis distance was cut. Because of this, an offset was introduced in the realignment step. Therefore, a second transformation matrix for realignment is deployed to correct this difference.

C. Marker Based Motion Capture System

Sixteen Vero v2.2 cameras (Vicon Motion Systems Ltd, Oxford, UK) were used for validation in this study. The system has a resolution of 2.2 MP and a maximum frame rate of 330 Hz, with a sampling frequency of 100 Hz. This frame rate was set to match the final sampling rate of four LMCs used in this study. Twenty-six markers were carefully placed onto bony landmarks on the arm and hand of the subject. This placement is similar to [6]. However, this study also placed additional markers on the elbow and wrist. The purpose of this placement is to mimic the marker position of the LMC. On the elbow joint, the marker used was 9 mm in diameter, while the other placements were 4 mm. The placement of the marker for the recording of the MMC as ground truth is shown in Figure 1.

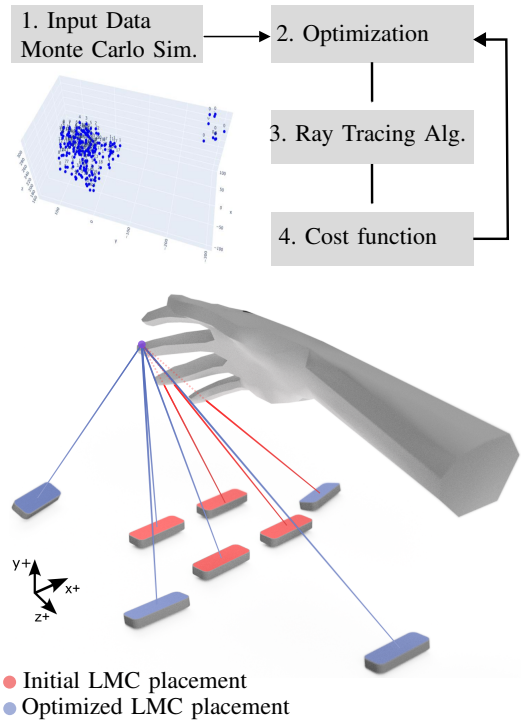

Fig. 3. Optimization procedure: 1: generation of input data; 2 - 4: optimization with ray-tracing algorithm. The bottom shows exemplary LMCs ray-tracing for the tip of the middle finger from optimized positions (blue) and initial positions (red).

D. Experimental Protocol

The subject sat comfortably on a chair behind a desk. On the top of the desk, four LMCs were configured in two settings: initial configuration or optimized configuration. The subject put their right hand on top of the LMCs (around 15 cm) and produced one static trial and three dynamic trials in two hand poses for each configuration. Each trial was performed five times. The motions were flexion of the index finger, the thumb, and the wrist. The subject performed two open-hand poses, one with the palm facing downward (horizontal) and the other with the palm facing to the left (vertical). These poses were chosen to assess the configuration efficacy in ensuring the visibility of at least one LMC on the finger. Data was recorded simultaneously for both LMC and the MMC, as shown in Figure 2.

E. LMC Placement Optimization

The study comprises two steps. The first step involves determining the optimized configuration for four LMCs, shown in Figure 3. The second step, shown in Figure 1, entails validating the optimization results through experimentation. During the first step, ten hand poses serve as reference points to generate virtual markers. These reference hand poses encompass static configurations during index, thumb, and wrist flexion in horizontal and vertical palm positions. Hand marker datasets required for optimization were generated using Monte Carlo simulation [11], which simulated various wrist and finger configurations along expected trial trajectories. Virtual markers, representing the LMC marker-set, were extracted from each configuration. Ray-tracing against these virtual markers of multiple LMCs frames was used to optimize the experimental location and orientation of the four LMCs setup. An optimal LMCs placement was defined by the metric $\mathfrak{M}$ described in (1), where F_i corresponds to the minimum of LMCs seeing a particular finger across all fingers in marker frame i. At the same time, N represents the number of frames in the dataset. The metric ensures prioritization of frames with invisible fingers ($F_i = 0$) while providing a transparent gradient towards increasing the number of LMCs with unobstructed views on finger markers. This metric is used as a cost function mentioned in the optimization process in Figure 3.

$$\mathfrak{M} = \sum_{i=1}^{\#\text{LMC}s} \sum_{j=1}^{N} \frac{F_j}{N^i} \tag{1}$$

The visibility of individual fingers for a particular LMC frame was determined by testing the following scenarios on the virtual marker set M of the frame:

1) Does the vector pointing from the subject elbow to the wrist $\vec{f_a}$ form an acute angle with the LMC forward direction?
2) Are all LMC virtual markers with location vectors $\vec{m}$ within the LMC Field of View (FoV)
$\forall m \in M, k \in 1,2,3,4 : |\vec{m}| <= L \ \wedge \ \vec{m} \cdot \vec{n_k} > 0$
Here the FoV was outlined as the set of normal vectors $\{\vec{n_0}, \vec{n_1}, \vec{n_2}, \vec{n_3}\}$ describing the planes forming the inverted LMC visibility pyramid as well as the maximal distance to the LMC camera centre L.
3) Does a ray pointing from the LMC camera centre to any virtual finger marker intersect the palm of the same hand?
4) Does the same ray intersect another finger before reaching the finger marker tested for visibility?

Conditions 3 and 4 were tested for each frame and each marker belonging to a given finger. Occlusions of a single finger marker labelled the entire finger as invisible. Condition 3 was assessed by computing the regression plane (2) formed by the palm marker set $\{m_{P0}, m_{P1}, ..., m_{P5}\}$ and calculating the intersection point p_1 of the marker (m) - LMC ray $\vec{r}$ with that plane (3). An occlusion occurred when the intersection point lived within the convex hull formed by the palm markers ($\exists \ x \ s.t. \ |m_{P0}, m_{P1}, ..., m_{P5}| \cdot x \leq p_1 \ \& \ x_i > 0$), was in the LMC FoV: ($\arccos(p1 \cdot m) = 0$) and closer to the LMC than the marker ($|p_1| < |m|$).

$$U\Sigma[\vec{v_1}, \vec{v_2}, \vec{v_3}]^T = SVD(\boldsymbol{P_i} - \overline{\boldsymbol{P_i}}) \tag{2}$$

$$p_1 = m - \vec{r} \cdot \frac{\vec{v_3} \cdot (p - m_{P0})}{\vec{v_3} \cdot \vec{r}} \tag{3}$$

Condition 4 was formalized by constructing cylinders of 1 cm diameter along the three phalanges of each finger. A marker was invisible if the marker-LMC ray intersected any phalange cylinder of another finger, and that intersection point was located between the LMC and the marker. Say p_{h1} and p_{h2} are markers at a finger's PIP and DIP joints, and $\vec{p_h}$ is the unit vector pointing from p_{h1} to p_{h2}. Then the marker m - LMC ray $\vec{r}$ intersects the phalange if the shortest distance d (4) between $\vec{p_h}$ and $\vec{r}$ was smaller than the

cylinder radius ($d < 0.5$ cm), the point of shortest distance p_2 (5) on $\vec{\boldsymbol{p_h}}$ lay between p_{h1} and p_{h2} ($|p_2 - p_{h1}| < |\vec{\boldsymbol{ph}}|$)and p_2 was closer to the LMC than m ($|p_2| < |m|$).

$$d = \frac{abs((\vec{\boldsymbol{p_h}} \times \vec{\boldsymbol{r}}) \cdot (p_{h1} - m))}{|\vec{\boldsymbol{p_h}} \times \vec{\boldsymbol{r}}|} \tag{4}$$

$$p_2 = p_{h1} + \vec{\boldsymbol{p_h}} \cdot \frac{(\vec{\boldsymbol{r}} \times (\vec{\boldsymbol{p_h}} \times \vec{\boldsymbol{r}})) \cdot (m - p_{h1})}{((\vec{\boldsymbol{p_h}} \times \vec{\boldsymbol{r}}) \cdot (\vec{\boldsymbol{p_h}} \times \vec{\boldsymbol{r}}))} \tag{5}$$

Metric (1) was minimized by Particle Swarm Optimization (PSO) [19], optimizing the placement and orientation of a total of four LMCs. LMCs coordinates $\{x, y, \phi, \theta\}$ represent their 2D location as well as their rotation along the LMC long and vertical axes.

F. Data Collection

The proposed framework employed a single computer to collect data from four LMCs, in contrast with other studies that utilize multiple computers for this purpose [20]–[22]. The computer specifications included an 11th Gen Intel(R) Core(TM) i9 processor with 32GB RAM. Preliminary investigations revealed that each LMC requires a specific USB bandwidth, allowing for connectivity of two LMCs per USB port. Consequently, in this study, the LMCs were connected to the computer through two USB ports, with two LMCs sharing the same port via a USB extender. Moreover, external power was provided to the extender to ensure sufficient current supply of all LMCs.

A custom program[1] was developed to acquire and process data from these four LMCs on a single computer. The program utilizes Software Development Kit (SDK) 5.6.1 provided by Ultraleap, capable of simultaneously reading data from multiple LMCs. Developed in C++, it facilitates the use of the C-based SDK. Each LMC was identified by its serial number and named LMC1, LMC2, LMC3, and LMC4. Timestamped data from each LMC was interpolated to ensure synchronization at 100 Hz using B-spline interpolation, with the initial timestamp defined when the LMC first detects the hand. As noted in prior studies [6], [17], [20], the LMC system lacks a fixed frame rate, employing variable sampling instead. Standardizing this sampling was the initial data processing step. Ganguly et al. [6] resampled data at 150 Hz to align with MMC, while Houston resampled LMC data at 100 Hz [17]. Wang et al. also employed interpolation with a sampling interval of 0.033s [20]. Our study found the original sampling rate, when all four LMCs were read, to be approximately 11 - 34 Hz, lower than the specified maximum of 120 Hz. Data visualization was employed using OpenGL, and Universal Datagram Protocol (UDP) communication synchronized LMC and MMC system recording times. Each trial commenced when all LMCs began recording, triggering the MMC system recording protocol simultaneously.

The experiment for each subject comprised 16 trial types conducted across two different LMCs configurations and two distinct hand poses, resulting in a total of 80 trials per subject. Trial specifics are outlined in Table I. Each trial involved the recording of two datasets: LMC data and MMC data. All subjects completed all experiment trials except for Subject 7, who did not participate in wrist movement trials. The total dataset comprised 1560 records from both the LMC and the MMC System.

[1]https://github.com/hilmanfatoni/Multi-LMC_Optimization

TABLE I
LIST OF MOTION TYPE AND NUMBER OF TRIALS IN EXPERIMENT

LMC Configuration	Number of Trial
Initial Configuration	
Pose: Open hand poses horizontal	
1. Static	5 times
2. Index flexion	5 times
3. Thumb flexion	5 times
4. Wrist flexion	5 times
Pose: Open hand poses vertical	
5. Static	5 times
6. Index flexion	5 times
7. Thumb flexion	5 times
8. Wrist flexion	5 times
Optimized Configuration	
Pose: Open hand poses horizontal	
9. Static	5 times
10. Index flexion	5 times
11. Thumb flexion	5 times
12. Wrist flexion	5 times
Pose: Open hand poses vertical	
13. Static	5 times
14. Index flexion	5 times
15. Thumb flexion	5 times
16. Wrist flexion	5 times

G. Data Processing

As shown in Figure 1, four types of marker data were recorded: Initial LMC Data, Initial MMC Data, Optimized LMC Data, and Optimized MMC Data. LMC data went through the stages of interpolation, realignment, and Kalman filtering. For MMC data, these processes were not carried out; instead, the process of calculating hand properties such as finger length and joint angle was calculated based on marker data. MMC's marker data was preprocessed by verifying marker data from the MMC system to prevent potential mislabeling or missing markers. During preprocessing, It was found that the simultaneous recording process between MMC and LMCs would only be problematic if the marker were too close to the LMCs. Markers that are close to the LMCs, for example, when performing a wrist flexion movement, will experience mislabeling because the LMC light is recognized as a marker. For LMCs data, they went through an interpolation process after being acquired. This process resulted in all LMCs data sampled exactly every 10 ms and recorded in the file where markers from LMC1, LMC2, LMC3, and LMC4 were listed sequentially.

LMCs data then underwent a realignment process. Different from other studies [8], [17] where using the Kabsch Algorithm to reorient other sensors to the same reference sensor, this study uses a transformation matrix that transforms all markers relative to each LMC to the world coordinate which located in the centre of experiment. The Kabsch Algorithm only works when both the reference sensor and the sensor to be transformed can "see" the same corresponding marker. However, in this study, such conditions are not consistently met, as there are instances where one or more LMCs fail to detect the hand. To address this, an algorithm to handle undetected hand or a Not a Number (NaN) value was carried

out. After realignment, all four markers from LMC were fused using Kalman Filter [10]. This study uses the Kalman Model from Houston [17]. This study assumes that markers read by LMC follow Gaussian distribution and are linearly related to the state. A different part from the Kalman Model was modified to handle NaN value. In the update step of Kalman Filter, NaN marker value that comes from one of the LMCs would be skipped and not used as part of updating the state.

Fused LMC data and MMC data were then calculated together to measure finger length and joint angle. The purpose of the experiment is to validate the optimized configuration against the gold standard MMC system. The metrics used were finger length and joint angle. Additionally, the origin LMC data underwent calculation of visibility rate. The finger length was calculated by measuring the distance of markers placed between finger segments from static trial data. The markers placement of MMC system was configured in the exact location of LMC markers. By doing so, the calculation result between the MMC system and LMC (both in initial and optimized configurations) should be the same. The joint angle was calculated by using four markers. From these four markers, every two markers will make a line. This line represents the finger segment. Then, there will be two lines formed by these four markers, and one of their ends meets. By utilizing the cosine function from those two lines, the angle of the joint was calculated. This angle calculation is adapted from [6]. In this study, the joint angle was measured from the dynamic trial that focused on the movement of the index finger, thumb, and wrist. Based on that movement, the resulting joint angles are Index Metacarpophalangeal (MCP), Index Proximal Interphalangeal (PIP), Index Distal Interphalangeal (DIP), Thumb MCP, Thumb DIP, and Wrist Joint. The last metric used is the visibility rate. Each LMC from each trial would have visibility values per frame, which was calculated using a ray-tracing algorithm.

III. Results

A. Optimized LMCs Configuration

The result of the first step in this study method was an optimized LMCs configuration. Figure 3 depicts the visualization of initial and optimized configurations. Optimized figuration has a unique position where LMC2 was located rather far forward in the Z+ axis direction: 342,57 mm. The orientation of three LMCs: LMC2, LMC3, and LMC4, which face in negative rotation of X axis: -7.57°, -8.72°, and -12.06° respectively, had the purpose of seeing the hand in the vertical pose. Table II shows the coordinates of initial and optimized configurations in detail.

B. Validity and Repeatability

The MMC system was employed alongside the initial and optimized configurations of LMCs to validate the effectiveness of using multiple LMCs for assessing hand properties. The Vicon system, serving as the MMC, is considered the gold standard for validating the measurement outcomes of the LMC. In this study, measurements of finger length and the Range of Motion (RoM) for specific parts of the hand and fingers were calculated and compared.

TABLE II
CENTER POINT COORDINATES AND AXIS ROTATION OF LMCS

	Coordinate (mm)			Rotation against axis (degree)	
	x	y	z	x	y
Initial Configuration					
LMC1	-60	0	60	0	0
LMC2	60	0	60	0	0
LMC3	-60	0	-60	0	0
LMC4	60	0	-60	0	0
Optimized Configuration					
LMC1	-120.37	0	256.29	29.90	16.50
LMC2	-69.97	0	342.57	-7.57	-7.04
LMC3	-190.60	0	88.70	-8.72	-4.93
LMC4	178.88	0	100.90	-12.06	12.38

Figure 4a shows the average inter-subject finger length across all trials. The graph shows that the result has the same trend between initial and optimized configurations. The measurement of finger length from MMC is always more significant than from LMCs, with a difference to ground-truth for all fingers is 24.23±6.35 mm. The vertical hand pose also has a more substantial difference compared to the horizontal hand pose in both LMCs configurations, where the differences to ground truth are 22.93±7.11 mm and 25.53±5.18 mm, respectively. This indicates that vertical pose had presented challenges in obtaining accurate readings from the LMC. Due to the distribution and unique positioning of the LMCs, this study hypothesized that the proposed optimized configuration would yield improved marker readings, potentially leading to more accurate finger length measurements. In fact, as illustrated in Figure 4a, the optimized configuration generally resulted in less significant finger length values than the initial configuration. In the optimized configuration, the finger length differences from the ground truth are 25.82±5.91 mm. This is more significant than the initial configuration differences, 22.64±6.39 mm.

The result of RoM calculation is shown in Figure 4b. The mean difference of all joint angles to ground truth is 29.49°±27.85. With this difference, the result of RoM measurement from LMC is far from agreeable. Especially the Index DIP, the ground-truth value of Index DIP only has a maximum value of 12.06°, where the LMC reading is 64.61°. A comparison between the initial and optimized configurations shows no improvement; both exhibit significant deviations from the ground truth values, with differences of 25.63°±22.62 and 35.84°±33.73, respectively.

C. Visibility Rate

The ray-tracing algorithm implemented in this study ensures that all fingers are in the line of sight of LMCs. LMC uses its internal models to predict all the markers of the finger. When certain parts of the hand are detected, such as the palm marker, LMC can predict all the finger markers even though the finger is occluded. This case is defined as detected from LMC's internal model. Detected from the internal model can lead to a misplacement of the finger joint position. Therefore, the proposed optimized configuration aims to detect the hand in true definition. The ray-tracing algorithm could prove this detection. This study uses the term visibility rate to define whether the finger is truly detected or comes from LMC's internal model. Visibility rate with

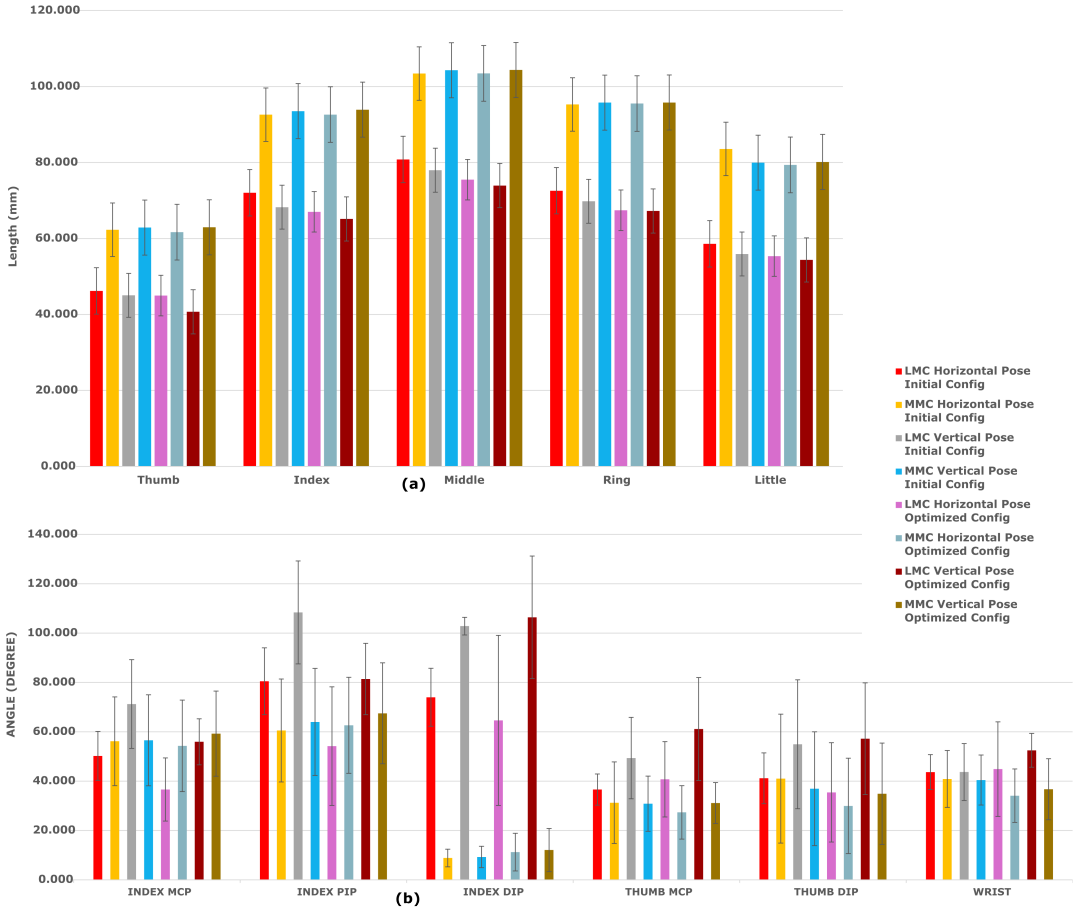

Fig. 4. (a) Average inter-subject finger length for all static trial of each LMC and MMC in two configurations and hand poses. (b) Average inter-subject joint angle for all dynamic trials of each LMC and MMC in two configurations and hand poses.

value -1 means that all the hand parts are undetected. This could be because the hand is outside the detection range of LMC or full oclusion of the hand. Value 0 means that finger marker data is acquired, but comes from an internal model and is not truly detected. A visibility rate with a value of 1 means that the finger is truly detected: the finger is in the line of sight of LMC.

This study experiment has eight combinations of pose and motion that were done in initial and optimized configurations. Figure 5 shows the mean visibility rate of each LMC across trials and subjects specific to each motion in each configuration. The result indicates that in optimized configuration, one LMC always truly detects any movement. In the initial configuration, the value of the visibility rate is mostly less than 0.25, which means either the marker is not detected or the marker comes from the internal model of LMC. LMC3 was the LMC in the optimized configuration that 83.33% can read the marker with a visibility rate of more than 0.75. It means that in optimized configuration, at least "one LMC can truly detect" policy is fulfilled. This policy is not satisfied in the initial configuration when doing vertical pose.

IV. Discussion

The use of multiple LMCs in previous research has been studied by [8], [17], [20]–[22]. They employed multiple LMCs and arranged them in specific configurations to ensure accurate hand tracking. Albeit the placement is working in their study, there is no analysis or discussion regarding the optimality of the configuration. From those studies, only Houston et al. [17] use multiple LMCs for measuring hand anthropometrics. Houston [17] placed each LMC angled by 15° toward the hand in triangle configuration. In this study, we introduced a method to find the optimized placement for multiple LMCs for measuring hand anthropometrics. Our result agrees with previous research carried out by Ganguly et al. [6] regarding the measurement of finger length in a static pose. Although Ganguly [6] uses only one LMC, we have a similar trend where the value of finger length was always less than MMC as the ground truth value in both initial and optimized configurations. The RoM result significantly deviated from the ground truth, particularly for the Index DIP joint. The LMCs consistently recorded more significant angles for this joint, likely due to the substantial changes occurring during index flexion. The rapid movement in this joint posed challenges for our Kalman model, leading to inaccuracies in state prediction.

Employing ray-tracing for the confidence estimation in a LMC measurement was motivated by a pilot study on a single LMC, where the effects of occluding different hand parts were investigated. The results showed that blocking the line of sight to a single virtual marker location of a particular finger leads to immediate changes in the LMC marker output of the same finger. The nature of the occlusion via palm,

		INITIAL CONFIGURATION					OPTIMIZED CONFIGURATION				
		Thumb	Finger	Middle	Ring	Little	Thumb	Finger	Middle	Ring	Little
STATIC HORIZONTAL POSE	LMC1	0.8000	0.8000	0.8000	0.8000	0.8000	0.1364	0.1364	0.1364	0.1364	0.1364
	LMC2	0.8000	0.8000	0.8000	0.8000	0.8000	0.0000	0.0000	-0.0004	0.0000	0.0000
	LMC3	1.0000	1.0000	1.0000	1.0000	1.0000	0.6000	0.6000	0.5970	0.6000	0.6000
	LMC4	1.0000	1.0000	1.0000	1.0000	1.0000	0.1462	0.1462	0.1462	0.1044	0.1462
INDEX FLEXION HORIZONTAL POSE	LMC1	1.0000	1.0000	0.9458	1.0000	1.0000	0.1500	0.1500	0.1448	0.1472	0.1500
	LMC2	0.9200	0.9200	0.9200	0.9200	0.9200	0.3206	0.2798	0.1188	0.2872	0.3206
	LMC3	1.0000	1.0000	0.9830	1.0000	1.0000	0.7984	0.7956	0.4364	0.6716	0.7738
	LMC4	0.9960	1.0000	0.9996	0.9998	1.0000	0.3434	0.3440	0.3198	0.1648	0.3448
THUMB FLEXION HORIZONTAL POSE	LMC1	1.0000	0.8878	0.9202	0.9920	1.0000	0.3450	0.3450	0.3448	0.3448	0.3448
	LMC2	1.0000	0.9792	1.0000	1.0000	1.0000	0.2580	0.1286	0.1372	0.2716	0.2794
	LMC3	0.9600	0.9568	0.9600	0.9600	0.9600	0.8800	0.7466	0.6890	0.6852	0.7666
	LMC4	1.0000	1.0000	1.0000	1.0000	1.0000	0.3424	0.2600	0.2218	0.0364	0.3452
WRIST FLEXION HORIZONTAL POSE	LMC1	0.9556	0.9948	0.9832	0.9448	0.9456	0.3094	0.3122	0.3138	0.3128	0.3126
	LMC2	0.8930	0.8346	0.9044	0.9130	0.9174	0.6850	0.6832	0.6318	0.6784	0.6832
	LMC3	0.8684	0.9554	0.9530	0.8908	0.8532	0.9946	0.9966	0.7740	0.7894	0.8876
	LMC4	0.8808	0.8704	0.9708	0.9902	0.9652	0.4446	0.3736	0.3324	0.2566	0.4446
STATIC VERTICAL POSE	LMC1	0.8996	0.1100	0.1000	0.0000	1.0000	-0.4160	-0.4160	-0.4160	-0.4160	-0.4960
	LMC2	-0.6800	-0.6800	-0.6800	-0.4600	-0.3600	-0.2072	-0.2914	-0.3200	-0.4164	-0.2072
	LMC3	-0.1176	-0.5200	-0.5200	-0.5200	-0.0400	1.0000	1.0000	1.0000	0.7986	1.0000
	LMC4	-0.2684	-0.3888	-0.0458	0.0096	0.0818	-0.6000	-0.6000	-0.6000	-0.6000	-0.6000
INDEX FLEXION VERTICAL POSE	LMC1	0.5130	0.3542	0.0510	-0.0958	0.8000	-0.7074	-0.7590	-0.7570	-0.7218	-0.7000
	LMC2	-0.6700	-0.6930	-0.6572	-0.4670	-0.4000	-0.1674	-0.2390	-0.2938	-0.4438	-0.1708
	LMC3	0.1668	-0.3356	-0.3056	-0.1992	0.2800	0.8812	0.9478	0.9514	0.8446	1.0000
	LMC4	-0.1808	-0.4270	-0.2456	-0.2114	0.0400	0.1106	-0.1066	0.0604	-0.1772	0.1200
THUMB FLEXION VERTICAL POSE	LMC1	0.7668	-0.0114	0.0390	-0.0302	0.8800	-0.0278	-0.1200	-0.0774	-0.0504	-0.0478
	LMC2	-0.3134	-0.4678	-0.2084	-0.1032	-0.0002	-0.0446	-0.2308	-0.2688	-0.3210	-0.0876
	LMC3	0.4220	-0.1676	-0.1766	-0.1928	0.5600	0.9996	0.8448	0.9090	0.6520	0.9990
	LMC4	-0.3042	-0.2170	-0.0150	-0.0222	0.0362	-0.1918	0.1060	0.0678	0.0348	0.0304
WRIST FLEXION VERTICAL POSE	LMC1	0.5650	0.0176	-0.0140	0.0218	0.9092	-0.2686	-0.3342	-0.3580	-0.3124	-0.2816
	LMC2	-0.6286	-0.6534	-0.6020	-0.4802	-0.3200	0.5158	0.3742	0.3592	0.1318	0.5174
	LMC3	0.1366	-0.2120	-0.1586	-0.0998	0.4400	0.9494	0.8010	0.8264	0.6280	0.9970
	LMC4	0.0078	-0.1706	0.0894	0.1578	0.3524	-0.2880	-0.3524	-0.3190	-0.4056	-0.2204

Visibility rate less than 0.25 | Visibility rate more than 0.75

Fig. 5. Map of mean visibility rate from ten subjects for each LMC. Visibility rate value: -1: Hand not detected, 0: Finger marker detected from LMC internal model, 1: Finger marker truly detected from ray tracing.

other fingers, or external objects did not influence that behaviour. Since the inner workings of the LMC model are proprietary, the performance measure based on the visibility of virtual markers forms a lower bound on the confidence threshold. Furthermore, a LMC might not register a hand at the borders of its FoV, even though the ray-tracing algorithm predicted good visibility. During optimization, the maximal LMC visibility range and the FoV angles were thus reduced from 0.6 m to 0.4 m and from 120°/150° to 100°/100°, respectively.

This study identified several limitations associated with the LMC. Previous research has highlighted the ongoing issue of variable sampling with LMC [6], [17], [20]. When multiple LMCs are employed, the variable sampling rate decreases to the maximum frame rate divided by the number of LMCs. For instance, with four LMCs utilized, the variable sampling rate for each LMC is capped at 30 Hz (calculated as 120 Hz divided by 4). Experimental results confirmed this limitation, with LMC sampling rates ranging from 11 to 34 Hz.

Another constraint observed with LMC is axis-cutting. Although the LMC specifications indicate an interaction zone of approximately 60 cm within a field of view (FoV) of 120°x150°, markers beyond 250 mm from the LMC centre are affected. This axis-cutting phenomenon was demonstrated in the initial phase of our study. It aimed to determine the edge of the LMC capture volume. During this experiment, the hand was moved along the x, y, and z-axes to identify areas where it was not detected. While the hand remained normal-detected within the 60 cm interaction zone, it exhibited abnormal shapes in the visualization when positioned around the 250 mm mark along any axis, despite all markers being detected. Through data inspection, the coordinate value of this axis was cut when the hand distance was more than 250 mm. This will introduce an offset if LMCs configuration places one of the LMC away than 250 mm. This offset needs to be included in the realignment process to solve this. On the other hand, if the hand moves further than 350 mm, all markers are still detected, but the unusual shape of the hand becomes more prominent. An error in marker reading may occur if the subject places their hands too far away. To prevent any further error introduced by this limitation, this study suggests the placement of multiple LMCs should be around a maximum of 250 mm at any axis. The LMC internal model of marker prediction for elbow marker was mostly incorrect to actual ground truth. Although all the hand markers were detected, the elbow marker data is not correctly predicted by the internal model of LMC. This leads to differences in elbow marker position coming from each LMC. In this study, the reading of the elbow marker from LMC1 has a difference of 10 - 20 mm in the y-axis after the realignment process compared to other LMCs. This difference leads to unstable state prediction for the elbow marker in the Kalman filter process. Although LMC has limitations, the marker read from LMC still can be used well for hand gesture recognition. Using multiple LMC in the optimized configuration can ensure that all the markers have better visibility than in the initial configuration.

From the optimization result of LMCs configuration, three LMCs are located on the left side of origin (x-) and one LMC on the right side (x+). This placement is logical because the reference hand pose used for optimization involves the motion of the index and thumb fingers in a vertical pose that can be seen properly from the left side. To satisfy the correct marker reading, the optimized configuration gave results that LMC1, LMC2, and LMC3 were placed on the left side of the world origin coordinates. The LMC1 also has rotation

in the x-axis, which ensures the fingers are fully detected when making wrist motion. In conclusion, the ray tracing algorithm will place all LMCs in multiple configurations in a way that all fingers are in the line of sight of at least one LMC.

For future work, the optimization component will incorporate a refined LMC model along with additional metric terms aimed at aligning LMC orientations to configure where trial markers are centred in the FoV. To address bandwidth limitations and variable sampling drops in a single computation for multiple LMCs, several strategies could be considered for future research: (1) Implementation of a multi-threaded process for computation and online Kalman filtering. Currently, the computation in this study's program occurs within the same thread as data reading from LMCs. Separating these tasks into different threads can reduce the load on marker reading. (2) Utilization of a computer equipped with more USB ports. Assigning each LMC its own USB port ensures that bandwidth requirements are adequately met. (3) Exploration of a Kalman Filter model tailored specifically for LMCs, where the relationships between marker positions are precisely defined. This Kalman Filter model should also be capable of handling missing or NaN value to ensure that it does not adversely affect the prediction state.

V. Conclusion

This study introduces a comprehensive approach to markerless motion capture using an online multi-LMC framework with a single computation system. It incorporates a placement optimization algorithm to determine ideal device placement based on expected LMC frame samples during trials. The methodology is validated against the gold standard MMC, demonstrating its effectiveness and reliability. By addressing the limitations of singular LMC devices, our approach provides a robust solution for markerless motion capture, offering clinicians a valuable tool for assessing grasping movements and informing rehabilitation strategies effectively. In terms of ratios, the finger length measurements obtained from the initial and optimized configurations are consistent with those of the MMC system. The ratios between fingers in both configurations match those of the MMC across all poses. But value-wise, there is no significant improvement from the initial and the optimized configuration to the validation method using MMC system. Both configurations also could not accurately measure RoM from dynamics movement and have a significant deviation from the ground truth value. Although neither of the multiple LMCs configurations fully matched the performance of the standard MMC system, the proposed optimized LMCs configuration demonstrated better detection of fingers during vertical pose motion. It achieved a mean visibility rate of 0.05 ± 0.55, compared to -0.07 ± 0.40 for the initial configuration. This improved detection suggests the potential for further optimization of configurations tailored to specific hand poses.

ACKNOWLEDGMENT

This work was supported by the Federal Ministry of Education and Research of the Federal Republic of Germany (BMBF) by funding the project AI.D under Project Number 16ME0539K. The first author completed this work while receiving a scholarship funded by Indonesia Endowment Fund for Education (LPDP).

References

[1] G. E. Sheiko, A. N. Belova, N. N. Rukina, and N. L. Korotkova, "Possibilities of using biomechanical human motion capture systems in medical rehabilitation (review)," *Physical and rehabilitation medicine, medical rehabilitation*, vol. 4, p. 181–196, Oct. 2022.

[2] S. Hogue, X. Guo, R. A. Morrison, S. McDowell, and A. C. Shembel, "Use of motion capture technology to study extrinsic laryngeal muscle tension and hyperfunction," *The Laryngoscope*, vol. 133, p. 3472–3481, June 2023.

[3] A. C. Vincent, H. Furman, R. C. Slepian, K. R. Ammann, C. Di Maria, J. H. Chien, K.-C. Siu, and M. J. Slepian, "Smart phone-based motion capture and analysis: Importance of operating envelope definition and application to clinical use," *Applied Sciences*, vol. 12, p. 6173, June 2022.

[4] R. Cotton, A. DeLillo, A. Cimorelli, K. Shah, J. Peiffer, S. Anarwala, K. Abdou, and T. Karakostas, "Markerless motion capture and biomechanical analysis pipeline," *Arxiv*, 03 2023.

[5] Y. Meng, I. Bíró, and J. Sárosi, "Markerless measurement techniques for motion analysis in sports science," *Analecta Technica Szegedinensia*, vol. 17, p. 24–31, Apr. 2023.

[6] A. Ganguly, G. Rashidi, and K. Mombaur, "Comparison of the performance of the leap motion controllertm with a standard marker-based motion capture system," *Sensors*, vol. 21, no. 5, p. 1750, 2021.

[7] A. Ganguly, G. Rashidi, and K. Mombaur, "Can leap motion controller replace conventional marker-based motion capture systems?," *XXVIII Congress of the International Society of Biomechanics*, 2021.

[8] T. Novacek, C. Marty, and M. Jirina, "Project multileap: Fusing data from multiple leap motion sensors," in *2021 IEEE 7th International Conference on Virtual Reality (ICVR)*, pp. 19–25, 2021.

[9] W. Kabsch, "A solution for the best rotation to relate two sets of vectors," *Acta Crystallographica Section A*, vol. 32, p. 922–923, Sept. 1976.

[10] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, vol. 82, pp. 35–45, 03 1960.

[11] N. Metropolis and S. Ulam, "The monte carlo method," *J. Am. Stat. Assoc.*, vol. 44, p. 335, 1949.

[12] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of the leap motion controller," *Sensors (Basel)*, vol. 13, pp. 6380–6393, May 2013.

[13] G. Du and P. Zhang, "A markerless human–robot interface using particle filter and kalman filter for dual robots," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2257–2264, 2015.

[14] Ultraleap, *Leap Motion Controller Data Sheet*, 10 2022. Issue 6.

[15] A. H. Smeragliuolo, N. J. Hill, L. Disla, and D. Putrino, "Validation of the leap motion controller using markered motion capture technology," *Journal of Biomechanics*, vol. 49, no. 9, pp. 1742–1750, 2016.

[16] E. Niechwiej-Szwedo, D. Gonzalez, M. Nouredanesh, and J. Tung, "Evaluation of the leap motion controller during the performance of visually-guided upper limb movements," *PLoS One*, vol. 13, p. e0193639, Mar. 2018.

[17] A. Houston, V. Walters, T. Corbett, and R. Coppack, "Evaluation of a multi-sensor leap motion setup for biomechanical motion capture of the hand," *Journal of Biomechanics*, vol. 127, p. 110713, 2021.

[18] Ultraleap, "Leap Concepts - Ultraleap documentation — docs.ultraleap.com." `https://docs.ultraleap.com/api-reference/tracking-api/leapc-guide/leap-concepts.html`. [Accessed 11-01-2024].

[19] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4, pp. 1942–1948 vol.4, 1995.

[20] Y. Wang, Y. Wu, S. Jung, S. Hoermann, S. Yao, and R. W. Lindeman, "Enlarging the usable hand tracking area by using multiple leap motion controllers in vr," *IEEE Sensors Journal*, vol. 21, no. 16, pp. 17947–17961, 2021.

[21] V. Kiselev, M. Khlamov, and K. Chuvilin, "Hand gesture recognition with multiple leap motion devices," in *2019 24th Conference of Open Innovations Association (FRUCT)*, pp. 163–169, 2019.

[22] T. Hu, X. Zhu, X. Wang, T. Wang, J. Li, and W. Qian, "Human stochastic closed-loop behavior for master-slave teleoperation using multi-leap-motion sensor," *Sci. China Technol. Sci.*, vol. 60, pp. 374–384, Mar. 2017.