OPTIMAL POSITION OF CAMERAS DESIGN IN A 4D FOOT SCANNER

Farzam Tajdari, Christiaan Eijck, Felix Kwa, Christiaan Versteegh, Toon Huysmans, Yu Song

OPTIMAL POSITION OF CAMERAS DESIGN IN A 4D FOOT SCANNER

Farzam Tajdari*
Faculty of
Industrial Design Engineering
Delft University of Technology
f.tajdari@tudelft.nl

Christiaan Eijck
Faculty of
Industrial Design Engineering
Delft University of Technology
C.L.J.Eijck@student.tudelft.nl

Felix Kwa
Faculty of
Industrial Design Engineering
Delft University of Technology
felixkwa@gmail.com

Christiaan Versteegh
R&D Department
BATA Industrials
Best, Netherlands
christiaan.versteegh@bata.com

Toon Huysmans
Faculty of
Industrial Design Engineering
Delft University of Technology
t.huysmans@tudelft.nl

Yu Song
Faculty of
Industrial Design Engineering
Delft University of Technology
y.song@tudelft.nl

ABSTRACT

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

Keywords: Optimal amera network configurations, accuracy, robustness, sensitivity, additive manufacturing, product evaluation;

1 INTRODUCTION

Foot shape is recognized as a very diverse character among the population, containing sex [1–3], and age [4]. This diversity may not be considered in footwear sizing, as available footwear fitting standards require only length and width of foot, and arch length to adjust to standardized shoe sizes [5]. Such diversity challenges the users to define a proper fit, leading to non-suited footwear in terms of comfortable and fitting increase the chance of occurring injury during ambulation [6]. Comfort-based footwear fit has determined direct positive impact on enhancing biomechanical performance [7], decreasing the chance of movement-related injury [8], and practical-wise it

*Address all correspondence to this author.

is the most important factor for users to choose footwear [9]. Thus, Footwear should be fully/semi personalised fit to satisfy any customer in terms of comfort, safety and fitting [10].

Available approaches of designing footwear established on employing static lasts, which supposes the foot includes rigid segments. Due to the fact that the foot is a deformable part of the body [11, 12], the assumption usually leads to failure investigating the dynamic movements in the foot morphology, specifically for a loaded foot in estimation of ankle joint mechanics [13, 14]. Proofs have been studied on articular motion through the loaded foot [15,16], which impacts linear foot measurements, especially during transitioning between postures i.e., from sitting to standing [17] or during the stance phase of gait [18]. The measurements of the dynamic foot recommend morphological variations happening, which are not reachable via static linear and circumferential measurements. Accordingly, characteriing the diverse population of foot shapes would be challenging within individuals in presence of different loading scenarios e.g., gait. One solution to the challenge may be recruiting motion tracking techniques for capturing the foot deformation for shoe penalisation.

Due to the inherent connection between virtual reality systems and naturalistic applications, motion tracking techniques as the key component of the systems should be sufficiently robust and accurate [19]. Partly, motion tracking has been employed to estimate the position and direction of the viewing gate to define the eyepoint to be able to render the most accurate perspective of images. Furthermore, body motion tracking [20, 21] is usually deployed to establish avatars which virtually models the human to the viewer. Meanwhile, optical motion tracking, as a well-known technique, may cover the eyepoint estimation and the dynamic avatar generation. To do so, arrangement of several cameras is an undeniable part of any motion tracking technique. The arrangement should guarantee the free movement of the users to be able to track the targeted features. The feature points (generally detectable via passive or active landmarks) are identified from camera outputs which give the 3D position of the landmarks through triangulation between the multiple camera outputs [22–24]. Apparently, the accuracy and the quality of the 3D reconstructed point-clouds [25, 26] rely on the number of the employed cameras and their configuration. However, the number is competing with the cost of the final product and partly the speed of the process. Thus, this paper focuses on an automatic method to optimally define the position and orientation of a set of cameras in a foot scanner.

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

The accumulative error was analyzed by Sanders-Reed [27] in 3D position estimation via two-camera triangulation. According to the analysis, the configuration of cameras plays a key role on the level of the accumulative error which defined a 90 degree angle between the sensors as the optimal configuration. The accumulative error is maximised when the sensors have the same orientation (either 0 or 180 degree). Olague and Mohr [28] studied the properties of the accumulative error in the process of the triangulation and validated the outputs practically. They proposed a metric of the accumulative error via extracting the component with maximum value through the diagonal of covariance matrix of the 3D point reconstruction. The maximum value indicates the highest accumulative error regarding a certain configuration (position and orientation). To achieve a global optimal solution to the optimal problem, a genetic algorithm is applied. To simplify the problem, Olague et al. considered a constraint that cameras must be located on a plane with a constant distance from the target point. Cowan and Kovesi [29] specified a group of constraints impacting the quality of the cameras' network to circumvent the time-consuming trial-and-error method for designing the configuration of the cameras' network (locations). They mapped the captured object into its surfaces. Then, the solution regarding each surface was extracted individually and independently according to a group of predefined constraints and next, each of the local solutions were integrated. The integration presented a global solution to the optimal problem. For each constraint e.g., resolution loss, sparsity, and occlusion, they determined a valid 3D volume that guarantees any point in that area fulfil the related constraint. Then, they considered the overlap area of those individual 3D volumes as checking area(s) where all constraints are satisfied. Mason [30] continued the idea of Cowan and Kovesi [29] with extended constraints and offered a novel model to determine the optimal placement. The model introduced by Cowan and Kovesi only covers sedentary occlusion known as self-occlusion. Mason's heuristic approach defines a solution for a specific target, however, the result can be a non-optimal output regarding too much simplifications of the target. Moreover, moving targets are big restrictions for Mason's approach. To avoid facing the object surfaces, Wu, Sharma, and Huang [31] contemplate the overlap of the mapped pyramids instead of dealing with a computationally expensive numerical measurement to tackle with the unknown bounds. Chen and Davis [32] assumed a probabilistic model for the occlusion, where a vertical plane was supposed as the occluder next to the target point. The model anticipates the visibility of a target point by at least two cameras considering all valid orientations of the occluder around the vertical axis. Though the Chen and Davis' model perfectly tackles the dynamic occlusion regarding the error metric, the convergence angles of the cameras are not addressed. As a result, the method proceeds to define configurations leading to comparatively large errors in target points alignment outperformed by poor triangulation regarding locating nearly parallel cameras.

The limited space, the number of cameras and the intrinsic and extrinsic properties of the cameras restrict the positioning of

the tracking cameras especially for foot scanning as the bottom of the foot is very complicated to be scanned. Generally, the cameras configuration considering the space constraint and having maximum overlap between cameras and maximum visibility of the foot to the cameras is very challenging. This paper proposes an optimal approach to locate a group of depth cameras D435i - Intel RealSense for a pre-defined box-shape work-space to maximize the percentage of observed points of a target object with minimum accuracy error of the 3D reconstruction in the existence of dynamic occlusion. Generally, work-space location of the target objects is known in many augmented and virtual reality systems. For instance, Rahimian et al. requested the participants to walk across a single-lane road [33], and after some iterations the valid work-space would be identified for a single step of walking. The approach we present estimates the optimal positions and orientations of a group of cameras for a predefined set of target points. Where the best set considers, first, all possible locations of occluders and specifies the cameras' locations in order to have the maximum visibility of the target points between at least two triangulable views i.e., the views' that have the maximum distance with each other and the maximum overlap of capturing the target points. Secondly, the placements should consider that the targets are visible from the widest range of viewpoints to have an acceptable level of overlap for alignment.

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

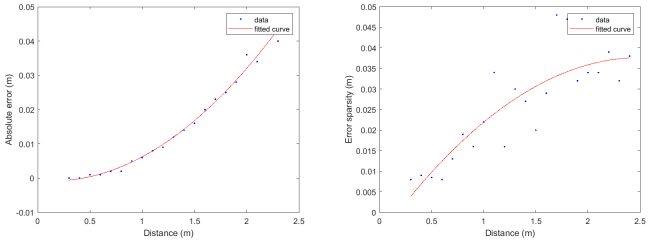
2 METHODOLOGY

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

2.1 Assumptions

To model the camera, we convey the camera parameters based on the distance to a detected point. The parameters that are considered in this paper are the spatial error, the resolution, and the acceptable viewing angle for Depth Camera D435i - Intel RealSense. The error of a 3D point location is the distance between the real measurement and the location estimation by the camera. Practically, the fixed resolution of the camera might lead to sparsity when the distance between the object and the camera becomes large and the error also increases by the distance. According to [34], the error has a parabolic relation with the distance as shown in FIGURE 1(a), where the error has the following equation.

$$Error_{resolution}(dis) = p_1\ dis^2 + p_2\ dis + p_3; \qquad (1)$$



(a) (b)

FIGURE 1: Camera parameter.

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

$$p_1 = 0.009458;\ \ p_2 = -0.002578;\ \ p_3 = -0.000655; \qquad (2)$$

Moreover, the viewing angle of the camera is assumed as a constant and equal to 60° based on [34].

With the same approach for the resolution error and from [34], the sparsity variation has also a parabolic relation with the distance as shown in FIGURE 1(b), where the function has the following equation.

$$Error_{sparsity}(dis) = p_4\ dis^2 + p_5\ dis + p_6; \qquad (3)$$

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

$$p_4 = -0.006931;\ \ p_5 = 0.03473;\ \ p_6 = -0.005851; \qquad (4)$$

2.2 Work-space definition

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

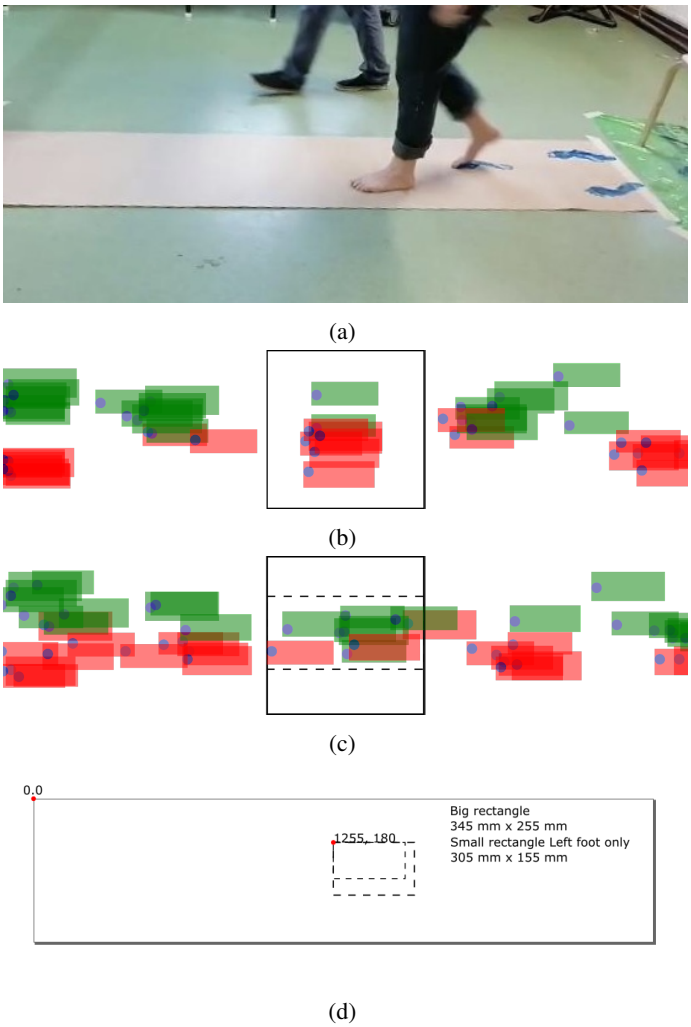


FIGURE 2: (a) Participants were instructed to try to land their feet in the middle without further instructions. (b) Participants were instructed to where to place their feet and to walk to the other side. (c) Average foot position participants. (d) the practical test-bed.

place their feet in a certain position from the edge and were given the command to just walk to the other side.

Both results of the test were digitized and overlaid to show where the participants placed their feet. These images can be seen in FIGURE 2(b) and (c). It was found that the users, with the proper instruction and foot placement, were able to land their feet naturally in the same area. The area is shown as a bonding box in FIGURE 2(d). Considering the right foot, the left foot and both feet these three scenarios of scanning, we extracted three bounding boxes reported in TABLE 1 where we considered the biggest box for our calculation regarding finding the optimal camera positioning. It worth mentioning that the measured maximum height of the foot/ankle during walking belongs to toe-off posture. See the reported height in the table.

TABLE 1: The choice of options.

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

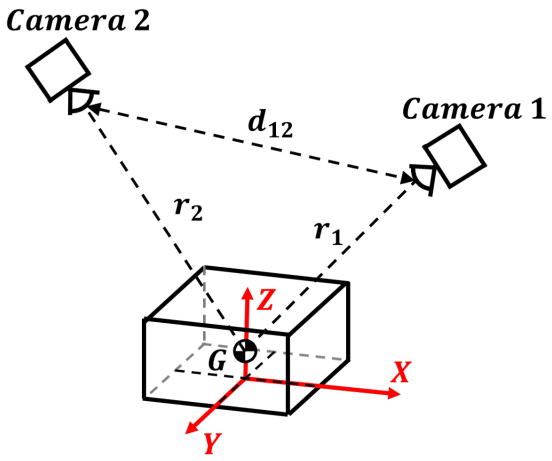


FIGURE 3: Camera parameter.

2.3 Optimal design of the foot scanner

As the future application of the scanner is to scan feet during walking, the work-space of any dynamic foot is considered as predefined as discussed in the previous section and equal to a rectangular box. According to FIGURE 3, r_i is the distance between i^{th} camera to the center of the box, d_{ij} belongs to the distance between i^{th} camera and j^{th} camera, and G is the center of the object. In addition, the (x_i, y_i, z_i) corresponded to the location of i^{th} camera are optimized through the following cost function.

$$J_{tot} = \begin{cases} Minimizing: & J_r, J_e, J_s \\ Maximizing: & J_d, J_v, J_o \end{cases} \quad (5)$$

where J_{tot} is the total cost function including both minimization and maximization terms. Regarding the minimization, J_r explains a term regarding the distance of each camera to the center

of the work-space, denoted as r_i in FIGURE 3, and is considered as follows where n is the number of cameras.

$$J_r = \sum_{i=1}^{n} |r_i|^2 \tag{6}$$

J_e is explaining the error of resolution calculated from (1) where we considered r_i as the dis in (1) for each camera to simplify the calculations. Accordingly, J_e would be extracted as follows

$$J_e = \sum_{i=1}^{n} |Error_{resolution}(r_i)|^2 \tag{7}$$

The last minimization term as J_s defines the sparsity of the points in (3) and is denoted as follows.

$$J_s = \sum_{i=1}^{n} |Error_{sparsity}(r_i)|^2 \tag{8}$$

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

$$J_d = \sum_{i=1}^{n} \sum_{j=i}^{n} |d_{ij}|^2 \tag{9}$$

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

$$J_v = \left(\frac{N_v}{N_{tot}} \right)^2 \tag{10}$$

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

$$J_o = \frac{1}{N_{tot}^2 n^2} \sum_{i=1}^{N_{tot}} n_{o_i}^2 \tag{11}$$

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

$$X = \begin{bmatrix} x_1 \; y_1 \; z_1 \; \dots \; x_{n-1} \; y_{n-1} \; z_{n-1} \; z_n \end{bmatrix} \tag{12}$$

such that

$$\min J_{tot} = \alpha_1 J_r + \alpha_2 J_e + \alpha_3 J_s - \alpha_4 J_d - \alpha_5 J_v - \alpha_6 J_o \tag{13}$$

subject to

$$\begin{aligned}
\text{Camera classification} &: \begin{cases} -z_1 \leq 0 \\ \vdots \\ -z_{n-1} \leq 0 \\ z_n \leq 0 \end{cases} \\
\text{Walking way} &: \begin{cases} -|y_1| + 0.3 \leq 0 \\ \vdots \\ -|y_{n-1}| + 0.3 \leq 0 \end{cases} \\
\text{Allowed work-space of cameras} &: \begin{cases} |y_1| - 0.5 \leq 0 \\ \vdots \\ |y_{n-1}| - 0.5 \leq 0 \end{cases} \\
\text{Bottom camera} &: \begin{cases} y_n = 0 \\ x_n = 0 \end{cases}
\end{aligned} \tag{14}$$

By applying the algorithm in [35], the optimal parameters are designed and for each of the cases the values related to the both feet will be obtained.

3 RESULTS

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

3.1 Solving the optimal problem

In all the experiments one camera is on the bottom of the box (negative values for z axis) and the rest of the cameras have positive value along the z axis as defined in (14). The box includes 3708 points as $N_{tot} = 3708$. In addition the optimization parameters in (13) are considered as $\alpha_1 = 200$, $\alpha_2 = 1.2$, $\alpha_3 = 1.2$,

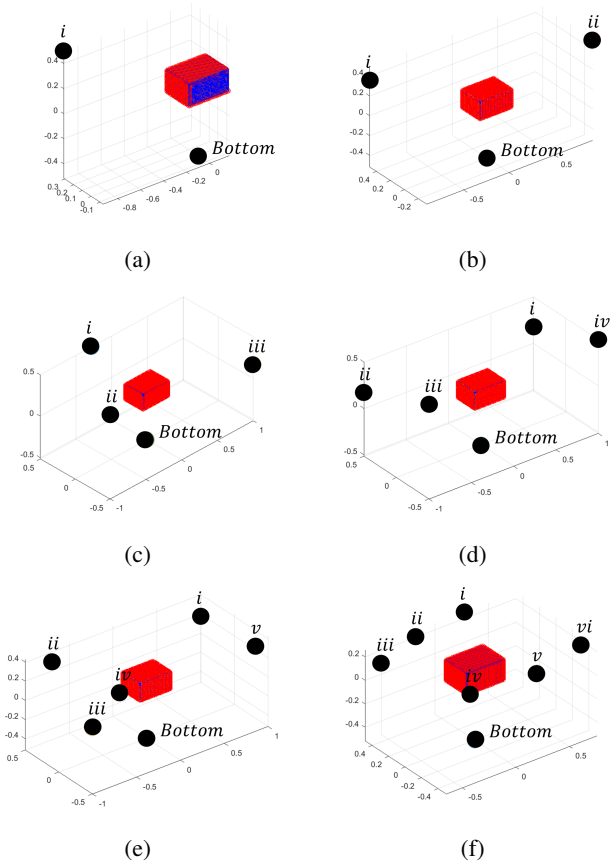


FIGURE 4: The optimal camera positions. (a) One camera on top. (b) Two cameras on top. (c) Three cameras on top. (d) Four cameras on top. (e) Five cameras on top. (f) Six cameras on top.

$\alpha_4 = 1$, $\alpha_5 = 15$, and $\alpha_6 = 1$. The values are defined via trial and error method emphasising on having as small as possible scanner dimensions (high $\alpha_1 = 200$). The results of solving the optimal problem are numerically reported in TABLE 2. According to the table, except the cases with having one or two cameras on top, the other cases have J_v equal to one meaning all the points on the box were visible to at least one camera. The position of the cameras are also visually shown in FIGURE 4, where the points with red color on the box are the captured points with at least one camera and the blue ones are not captured with any cameras.

3.2 Percent of overlap between pair of cameras

A crucial step to reconstruct a scanned object through multiple depth cameras is using surface registration methods. Surface registration is capable to assemble multiple 3D point clouds in a common coordinate system via aligning the overlapping parts of the point clouds [36, 37]. The point clouds may contain characteristics of a single shape as a mesh structure to explain the surface of the 3D objects or scenes. To form a complete 3D shape, different 3D captured datasets of an object from different viewpoints are needed which should cover all the targeted areas of the object with acceptable level of overlap between each pair of the views. This facilitates a well-performed registration to recombine all the datasets and reconstruct as accurate as possible the 3D mesh of the original scanned object or scenes [38]. In the registration process of a multiple-camera scanner, these datasets must be registered on each other and the quality of the registration is very depending on the level of overlap between each two pair of cameras from the object. Thus, here we investigate the degree of overlap based on the number of used cameras with optimal positions.

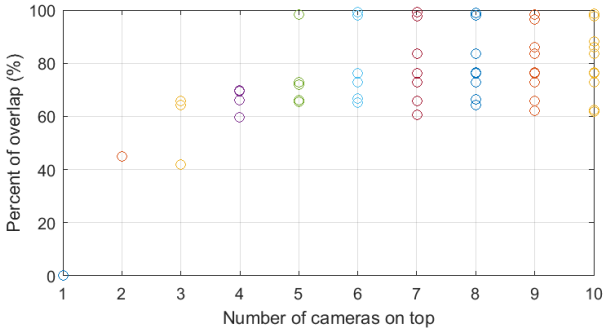


FIGURE 5: Percent of overlap variation based on the number of cameras on top.

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

$$P_o^i = \frac{N_o^{i,j}}{N_o^i} \times 100 \tag{15}$$

where, $N_o^{i,j}$ is the number of points on the scanned object captured with both camera i^{th} and camera j^{th}, and N_o^i is the total number of points captured with camera i^{th}. To study the effect, we designed the optimal camera locations through minimizing the cost function in (13) for different number of cameras on top and reported the overlap percent for each pair of cameras in FIGURE 5. According to the figure, the overlap range increases by increasing the number of top cameras up to 5. After the 5 cameras on top the range is more or less constant. Thus, 5 cameras on top should be enough to have acceptable registration. However, as we would like to have a symmetric design we prefer 6 cameras on top to have three cameras on each sides.

3.3 Scanner prototype

The scanner used in this study is the first prototype of a 4D foot scanner at TUDelft [39]. As shown in FIGURE 6, the scanner utilizes seven RealSense D435i depth cameras to capture a

TABLE 2: The optimal positions as [x,y,z] in cm and percent of captured area of the box ($\sqrt{J_v}$).

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

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

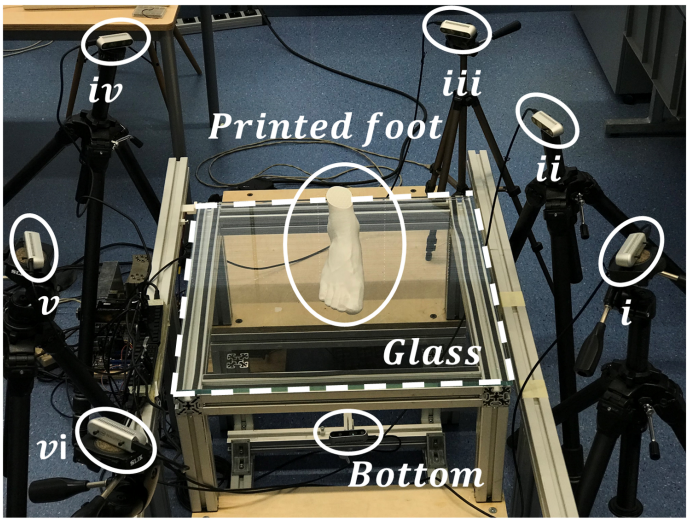


FIGURE 6: Scanner.

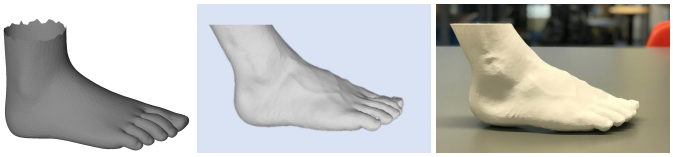
3.4 Accuracy

Through the explained 3D registration method in [40] which uses a probability function [41–44], we use seven cleaned frames captured by each camera to reconstruct the foot. According to the method each cleaned data-set is rigidly registered on a reference model to build a rough point-cloud of the foot, then the reference model is non-rigidly registered on the point-cloud to have a meaningful mesh. Due to the larger overlap between each data-set captured by each camera in the proposed method than the study in [40], their registration method shows better performance though our camera arrangement comparing to their non-optimal camera arrangement. Accordingly, the used reference model (source mesh) is shown in FIGURE 7(a).

To evaluate the accuracy of the results with the introduced algorithm, we scanned a foot with Artec Eva scanner in a no-load condition as shown in FIGURE 7(b), then the scanned foot was printed out via Ulimaker S5 scanner shown in FIGURE 7(c). The printed foot is scanned with our scanner in static mode which is comparable with the scan data of the Artec Eva. Next, the output of the Artec Eva scanner dimensions are compared with the corresponded dimensions of our method output data depicted in FIGURE 8. We first find the scaling ratio which maps the length of the reconstructed foot to the length of the foot collected with the Artec Eva scanner and then calculate the errors for the

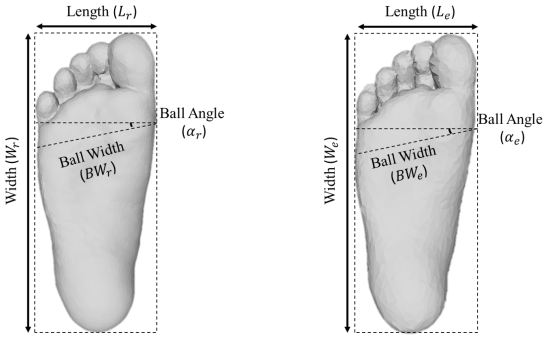
TABLE 3: Components of the final product in FIGURE 6.

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



(a) (b) (c)

FIGURE 7: No-load foot. (a) The source mesh. (b) Scanned through Artec Eva scanner. (c) Printed Artec Eva output via Ultimaker S5 printer.



(a) (b)

FIGURE 8: Dimension definition: (a) Artec Eva. (b) Our method.

other dimensions as width, ball width, and ball angle according to [45]. Based on FIGURE 8, L_r, W_r, BW_r, and α_r explain the length, width, ball width, and ball angle of the real scanned foot in the no-load case with the Artec Eva scanner respectively. And, L_e, W_e, BW_e, and α_e define the length and width of the estimated foot through our method respectively. The scaling factor of r is considered as follows

$$r = \frac{L_r}{L_e} \tag{16}$$

TABLE 4: Error results of the parameters introduced in FIGURE 8 with $r = 600.4$.

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

where in our experiment, $L_r = 222.2$ mm, $L_e = 0.3701$ mm which results in $r = 600.4$. Thus any dimension extracted from our method is multiplied to the r and compared with the corresponded values on the Artec Eva scanner output summarised in TABLE 4. According to the table, the errors for the width is 1.8 mm, for the ball width is 2.5 mm, and for the ball angle is 0.3°, which shows the average percentage error based on the Artec Eva scanner is about 2.5%.

4 CONCLUSION

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

ACKNOWLEDGMENT

This work is supported by Dutch NWO Next UPPS - Integrated design methodology for Ultra Personalised Products and Services project under Grant 15470.

REFERENCES

[1] Wunderlich, R. E., and Cavanagh, P. R., 2001. "Gender differences in adult foot shape: implications for shoe design.". *Medicine and science in sports and exercise,* ***33***(4), pp. 605–611.

[2] Krauss, I., Grau, S., Mauch, M., Maiwald, C., and Horstmann, T., 2008. "Sex-related differences in foot shape". *Ergonomics,* ***51***(11), pp. 1693–1709.

[3] Krauss, I., Valiant, G., Horstmann, T., and Grau, S., 2010. "Comparison of female foot morphology and last design in athletic footwear—are men's lasts appropriate for women?". *Research in Sports Medicine,* ***18***(2), pp. 140–156.

[4] Tomassoni, D., Traini, E., and Amenta, F., 2014. "Gender and age related differences in foot morphology". *Maturitas,* ***79***(4), pp. 421–427.

[5] Boppana, A., and Anderson, A. P., 2021. "Dynamic foot morphology explained through 4D scanning and shape modeling". *Journal of Biomechanics,* ***122***, p. 110465.

[6] Dobson, J. A., Riddiford-Harland, D. L., Bell, A. F., and Steele, J. R., 2018. "The three-dimensional shapes of underground coal miners' feet do not match the internal dimensions of their work boots". *Ergonomics,* ***61***(4), pp. 588–602.

[7] Pryhoda, M. K., Wathen, R. J., Dicharry, J., Shelburne, K. B., Feeney, D., Harrison, K., and Davidson, B. S., 2021. "Alternative upper configurations during agility-based movements: part 1, biomechanical performance". *Footwear Science,* ***13***(1), pp. 91–103.

[8] Mundermann, A., Stefanyshyn, D. J., Nigg, B. M., and others, 2001. "Relationship between footwear comfort of shoe inserts and anthropometric and sensory factors". *Medicine and science in sports and exercise,* ***33***(11), pp. 1939–1945.

[9] Mart\'\inez-Mart\'\inez, J. M., Mart\'\in-Guerrero, J. D., Soria-Olivas, E., Bernabeu, J. A., Escandell-Montero, P., Stark, R. H., Serrano-López, A. J., and Montiel, E., 2017. "Use of SOMs for footwear comfort evaluation". *Neural Computing and Applications,* ***28***(7), pp. 1763–1773.

[10] Minnoye, A., Tajdari, F., Doubrovski, E., Wu, J., Kwa, F., Elkhuizen, W., Huysmans, T., and Song, Y., 2022. "PERSONALIZED PRODUCT DESIGN THROUGH DIGITAL FABRICATION". *International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*.

[11] Tajdari, F., Tajdari, M., and Rezaei, A., 2021. "Discrete time delay feedback control of stewart platform with intelligent optimizer weight tuner". In 2021 IEEE International Conference on Robotics and Automation (ICRA), pp. 12701–12707.

[12] Tajdari, F., and Ebrahimi Toulkani, N., 2021. "Implementation and intelligent gain tuning feedback–based optimal torque control of a rotary parallel robot". *Journal of Vibration and Control*, p. 10775463211019177.

[13] Zelik, K. E., and Honert, E. C., 2018. "Ankle and foot power in gait analysis: Implications for science, technology and clinical assessment". *Journal of Biomechanics,* ***75***, pp. 1–12.

[14] Vickers, N. J., 2017. "Animal communication: when i'm calling you, will you answer too?". *Current biology,* ***27***(14), p. R713–R715.

[15] Lundgren, P., Nester, C., Liu, A., Arndt, A., Jones, R., Stacoff, A., Wolf, P., and Lundberg, A., 2008. "Invasive in vivo measurement of rear-, mid-and forefoot motion during walking". *Gait & posture,* ***28***(1), pp. 93–100.

[16] Wolf, P., Stacoff, A., Liu, A., Nester, C., Arndt, A., Lundberg, A., and Stuessi, E., 2008. "Functional units of the human foot". *Gait & posture,* ***28***(3), pp. 434–441.

[17] Xiong, S., Goonetilleke, R. S., Zhao, J., Li, W., and Witana, C. P., 2009. "Foot deformations under different load-bearing conditions and their relationships to stature and body weight". *Anthropological Science,* ***117***(2), pp. 77–88.

[18] Grau, S., and Barisch-Fritz, B., 2018. "Improvement of safety shoe fit-evaluation of dynamic foot structure". *Footwear Science,* ***10***(3), pp. 179–187.

[19] Lanzoni, D., Vitali, A., Regazzoni, D., and Rizzi, C., 2022. "Design of Customized Virtual Reality Serious Games for the Cognitive Rehabilitation of Retrograde Amnesia After Brain Stroke". *Journal of Computing and Information Science in Engineering,* ***22***(3).

[20] Tajdari, F., Toulkani, N. E., and Zhilakzadeh, N., 2020. "Semi-Real Evaluation, and Adaptive Control of a 6DOF Surgical Robot". In 2020 11th Power Electronics, Drive Systems, and Technologies Conference (PEDSTC), pp. 1–6.

[21] Tajdari, F., Toulkani, N. E., and Zhilakzadeh, N., 2020. "Intelligent Optimal Feed-Back Torque Control of a 6DOF Surgical Rotary Robot". In 2020 11th Power Electronics, Drive Systems, and Technologies Conference (PEDSTC), pp. 1–6.

[22] Elkhuizen, W., Essers, T., Song, Y., Geraedts, J., Weijkamp, C., Dik, J., and Pont, S., 2019. "Gloss, color, and topography scanning for reproducing a Painting's appearance using 3D printing". *Journal on Computing and Cultural Heritage (JOCCH),* ***12***(4), pp. 1–22.

[23] Malhan, R., Jomy Joseph, R., Bhatt, P. M., Shah, B., and Gupta, S. K., 2022. "Algorithms for Improving Speed and Accuracy of Automated Three-Dimensional Reconstruction With a Depth Camera Mounted on An Industrial Robot". *Journal of Computing and Information Science in Engineering,* ***22***(3), p. 31012.

[24] Tajdari, F., Huysmans, T., Yang, Y., and Song, Y., 2022. "Feature preserving non-rigid iterative weighted closest point and semi-curvature registration". *IEEE Transactions on Image Processing,* ***31***, pp. 1841–1856.

[25] Goto, L., Lee, W., Molenbroek, J. F. M., and Goossens, R., 2015. "Analysis of a 3D anthropometric data set of children for design applications". In Proceedings 19th Trien-

nial Congress of the IEA, Vol. 9, p. 14.

[26] Tajdari, M., Pawar, A., Li, H., Tajdari, F., Maqsood, A., Cleary, E., Saha, S., Zhang, Y. J., Sarwark, J. F., and Liu, W. K., 2021. "Image-based modelling for adolescent idiopathic scoliosis: mechanistic machine learning analysis and prediction". *Computer methods in applied mechanics and engineering,* ***374***, p. 113590.

[27] Sanders-Reed, J. N., 2001. "Error propagation in two-sensor three-dimensional position estimation". *Optical Engineering,* ***40***(4), pp. 627–636.

[28] Olague, G., and Mohr, R., 2002. "Optimal camera placement for accurate reconstruction". *Pattern recognition,* ***35***(4), pp. 927–944.

[29] Cowan, C. K., and Kovesi, P. D., 1988. "Automatic sensor placement from vision task requirements". *IEEE Transactions on Pattern Analysis and machine intelligence,* ***10***(3), pp. 407–416.

[30] Mason, S., and others, 1997. "Heuristic reasoning strategy for automated sensor placement". *Photogrammetric engineering and remote sensing,* ***63***(9), pp. 1093–1101.

[31] Wu, J. J., Sharma, R., and Huang, T. S., 1998. "Analysis of uncertainty bounds due to quantization for three-dimensional position estimation using multiple cameras". *Optical Engineering,* ***37***(1), pp. 280–292.

[32] Chen, X., and Davis, J., 2008. "An occlusion metric for selecting robust camera configurations". *Machine Vision and Applications,* ***19***(4), pp. 217–222.

[33] Rahimian, P., O'Neal, E. E., Yon, J. P., Franzen, L., Jiang, Y., Plumert, J. M., and Kearney, J. K., 2016. "Using a virtual environment to study the impact of sending traffic alerts to texting pedestrians". In 2016 IEEE Virtual Reality (VR), pp. 141–149.

[34] Ahn, M. S., Chae, H., Noh, D., Nam, H., and Hong, D., 2019. "Analysis and noise modeling of the Intel RealSense D435 for mobile robots". In 2019 16th International Conference on Ubiquitous Robots (UR), pp. 707–711.

[35] Lagarias, J. C., Reeds, J. A., Wright, M. H., and Wright, P. E., 1998. "Convergence properties of the Nelder–Mead simplex method in low dimensions". *SIAM Journal on optimization,* ***9***(1), pp. 112–147.

[36] Yang, Y., Yuan, T., Huysmans, T., Elkhuizen, W. S., Tajdari, F., and Song, Y., 2021. "Posture-invariant three dimensional human hand statistical shape model". *Journal of Computing and Information Science in Engineering,* ***21***(3).

[37] Yang, Y., Zhou, H., Song, Y., and Vink, P., 2021. "Identify dominant dimensions of 3D hand shapes using statistical shape model and deep neural network". *Applied Ergonomics,* ***96***, p. 103462.

[38] Lempitsky, V., and Boykov, Y., 2007. "Global optimization for shape fitting". In 2007 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8.

[39] Kwa, F. S. S., 2021. "Design of an accurate and low cost 4D foot scanner for podiatrists".

[40] Tajdari, F., Kwa, F., Versteegh, C., Huysmans, T., and Song, Y., 2022. "Dynamic 3D foot mesh reconstruction based on Nonrigid iterative closest-farthest points registration". *International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*.

[41] Tajdari, F., Ghaffari, A., Khodayari, A., Kamali, A., Zhilakzadeh, N., and Ebrahimi, N., 2019. "Fuzzy control of anticipation and evaluation behaviour in real traffic flow". In 2019 7th International Conference on Robotics and Mechatronics (ICRoM), IEEE, pp. 248–253.

[42] Tajdari, F., Toulkani, N. E., and Nourimand, M., 2020. "Intelligent architecture for car-following behaviour observing lane-changer: Modeling and control". In 2020 10th International Conference on Computer and Knowledge Engineering (ICCKE), IEEE, pp. 579–584.

[43] Tajdari, F., Golgouneh, A., Ghaffari, A., Khodayari, A., Kamali, A., and Hosseinkhani, N., 2021. "Simultaneous intelligent anticipation and control of follower vehicle observing exiting lane changer". *IEEE Transactions on Vehicular Technology,* ***70***(9), pp. 8567–8577.

[44] Tajdari, F., Kabganian, M., Rad, N. F., and Khodabakhshi, E., 2017. "Robust control of a 3-dof parallel cable robot using an adaptive neuro-fuzzy inference system". In 2017 Artificial Intelligence and Robotics (IRANOPEN), IEEE, pp. 97–101.

[45] Tang, U. H., Siegenthaler, J., Hagberg, K., Karlsson, J., and Tranberg, R., 2017. "Foot anthropometrics in individuals with diabetes compared with the general Swedish population: Implications for shoe design". *Foot Ankle Online J,* ***10***(3), p. 1.