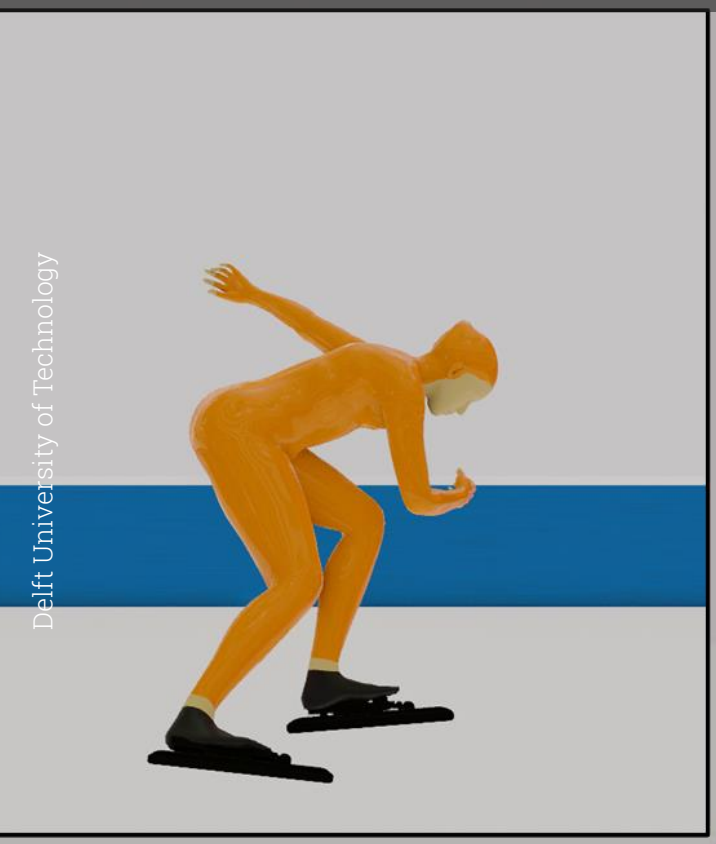
ODAH-SpeedSkater:

Development of a Virtual Video Dataset for Kinematic Analysis in Speed Skating

MSc Thesis

Punitha Devaraja



Delft University of Technology

TU Delft

ODAH-SpeedSkater:

Development of a Virtual Video Dataset for Kinematic Analysis in Speed Skating

by

Punitha Devaraja

to obtain the degree of

MSc in Mechanical Engineering,

at the
Faculty of Mechanical Engineering,
Delft University of Technology
October 2024

Student number:	5757746
Supervisors:	Dr. Ir. E. van der Kruk
	Hassan Osman
Thesis Committee:	Dr. A. Seth
Project Duration:	Apr, 2024 - Oct, 2024

TU Delft

Abstract

Accurate kinematic analysis in speed skating is crucial to understand and improve skaters' unique technique during training. Few research methods have captured joint kinematics in the past, using Inertial Measurement Units (IMUs) or manual annotation of filmed data or marker-based motion capture methods. The large motion capture volume and ice-rink environment hinders these methods to be actively adopted on rink. Thus, with growing accuracy in kinematic estimation algorithms, demand for a biomechanically accurate dataset is high. In this research, we aim to generate a virtual video dataset called ODAH-SpeedSkater, from experimental motion capture data. First, we examine the impact and accuracy of markers during motion capture. After selection of accurate inverse kinematics data, we use the SMPL-X human body model to achieve individual skater body shape and pose throughout the motion, rigging it to a skater-specific scaled OpenSim skeletal model. We render the finalized mesh sequences in an ice rink scene with skater outfits through realistic camera set-ups and configurations. Thus, we successfully create a dataset of 1,326 biomechanically annotated virtual videos of speed skating. Finally, we test our dataset on a pre-trained 3D kinematic estimation algorithm to evaluate its performance on speed skating data. In spite of limited testing, we conclude that training a network exclusively on our dataset may improve its performance, with the ultimate goal of actively implementing such networks in the rink.

Keywords: Speed skating, virtual video dataset, SMPL-X, 3D kinematic estimation algorithm

Contents

1 Introduction

Research on sports biomechanics is growing with technological advancements. Kinematic analysis plays a crucial and fundamental role in this field by providing detailed insights of the motion. In the case of speed skating, kinematic analysis helps to understand the mechanics of skating movements, aiming to reduce injury risk and improving sports performance. Skaters, trainers, coaches, and researchers are particularly interested in factors such as speed, stroke parameters, push-off mechanics, joint kinematics, and power to better analyze and optimize their techniques, ultimately elevating their competitive edge. Currently, although researchers are implementing various methods to analyse kinematics, visual analysis tends to be the most common practice on the rink. However, for accurate and comprehensive kinematic analysis, more advanced method has to be adopted that can be easily integrated on the rink for both training and competitive ends.

Different methods have been adopted for various research studies. One approach involves the use of Inertial Measurement Units (IMU) for both real-time and terminal feedback on parameters such as contact time, stroke frequency and orientations of skates [1]–[3]. Another method used a wind tunnel experiment where influence of air friction on different skating postures were investigated [4]. This allowed for understanding and optimization of the joint angles that determine the skating position. Focusing on joint kinematics, measurement of three-dimensional motion data in the rink has technical difficulties.

Some studies used inertial sensors as a wearable motion analysis system where they are placed on anatomic locations in the wearable suit to capture 3D motion data [5], [6]. Alternatively, the most available and accepted method is the marker-based motion capture system. This method is considered as the gold standard in motion capture in laboratory settings but its accuracy is limited for measuring on the rink mainly due to the large volume of data captured in speed skating [7]. One study implemented the system to capture motion data over a 50 m distance in the straight part of the rink [8]. The accuracy of this system is dependent on the locations of the cameras relative to each other, the distance between the cameras and the markers, the position and number of markers in the field and the marker trajectories within the captured volume. The measurement error of this system compounds due to the dynamic conditions of speed skating, characterized by inclined postures and extended strides.

By contrast, the method most widely practiced by coaches and trainers is a vision based system where the speed skating motion is filmed using high-speed cameras. Camera setups have changed over time as researchers implemented them in their studies. Initially, cameras were placed outside the track, perpendicular to the lane, at a distance of 25-30 m [9]–[12]. Later, a two-camera system was adopted: a panning camera to capture the sagittal view and a static camera placed outside the rink to capture the front view [13]. These studies determined joint kinematics through manual annotation of each frame with the help of a motion analyzer. However, this approach has proven impractical for general use due to the labor-intensive manual processing and potential accuracy issues, especially when dealing with large data. In practice, coaches and trainers often rely primarily on vision to improve skaters' techniques. Filmed videos are used as supplementary tools for visual analysis. Collectively, the methods mentioned are not feasible for obtaining accurate motion data for active use during training and competitions.

The aim of quantitative motion analysis of sports by means of filmed videos is to undertake a detailed analysis of subjects' movements patterns [14]. This approach eliminates the difficulties with motion capture using skin-mounted markers or wearable sensors as well as their potential influence on skater's movements. It also allows for minimal experimental set-up, enabling data collection in the natural environment, i.e. the ice rink. In recent decades, many computer vision algorithms incorporating deep-learning techniques have been developed for video-based motion analysis. Pose estimation algorithms enable markerless motion capture from video inputs, offering high accuracy, fast speed and reduced interference with natural movements [15]. There have been many pose estimation methods published. Some infer 3D poses from 2D joints [16]–[18] while others directly detect the 3D pose [19], [20]. Some are anatomically aware networks [19], [21], [22]. As of October 2024, no 3D pose estimation networks have been specifically applied to speed skating, nor have they been extensively trained on speed skating motion data. A significant challenge exists in the lack of large-scale annotated video datasets needed to train these algorithms. Furthermore, the accuracy of the existing algorithms are yet unknown and their biomechanical fidelity, especially when applied to speed skating, has yet to be thoroughly validated.

In this work, we first address the limited availability of datasets with accurate biomechanical annotations. We generate virtual video datasets based on experimental data collected through motion capture, which will be considered as the ground truth data. The algorithms trained on virtual datasets are more robust, as the virtual video enables automatic and accurate data annotations compared to existing algorithms trained on hand-annotated data, which is prone to human error. To achieve this, we use the synthetic data generation pipeline proposed by Lyn. B. et al (2024) [23] which provides realistic human appearances along with realistic skeleton-mesh registration. We then test the pre-trained algorithm proposed by Lin et al. (2024) [19] which introduces a novel biomechanics-aware network that directly outputs 3D kinematics from two input views with consideration of biomechanical prior and spatio-temporal information.

The aim of this research is to generate biomechanically accurate virtual video data of speed skating, based on measured kinematic data and to test a neural network for video-based 3D kinematic estimation.

The main contributions of this study are the following:

- We generate joint kinematics from raw experiment data.
- We create a virtual speed skating video dataset with accurate kinematic annotations using ground truth motion data for biomechanical models by incorporating the skater's appearance and simulating the camera configurations used in a real ice rink environment.
- We test the existing pre-trained biomechanics-aware 3D kinematics estimation network with our virtual dataset.

2 Methodology

2.1 Overview

The main contributions outlined in the previous section divide into four stages: Stage 1 - Biomechanical model and joint kinematics generation; Stage 2 - SMPL-X mesh generation; Stage 3 - Virtual video generation; and Stage 4 - Testing the pre-trained algorithm. These stages are illustrated in Figure 1. Stage 1 is presented in the sections 2.2 to 2.4, while stages 2 and 3 are discussed in sections 2.5.1 and 2.5.2, respectively. Lastly, stage 4 is introduced in section 2.6.

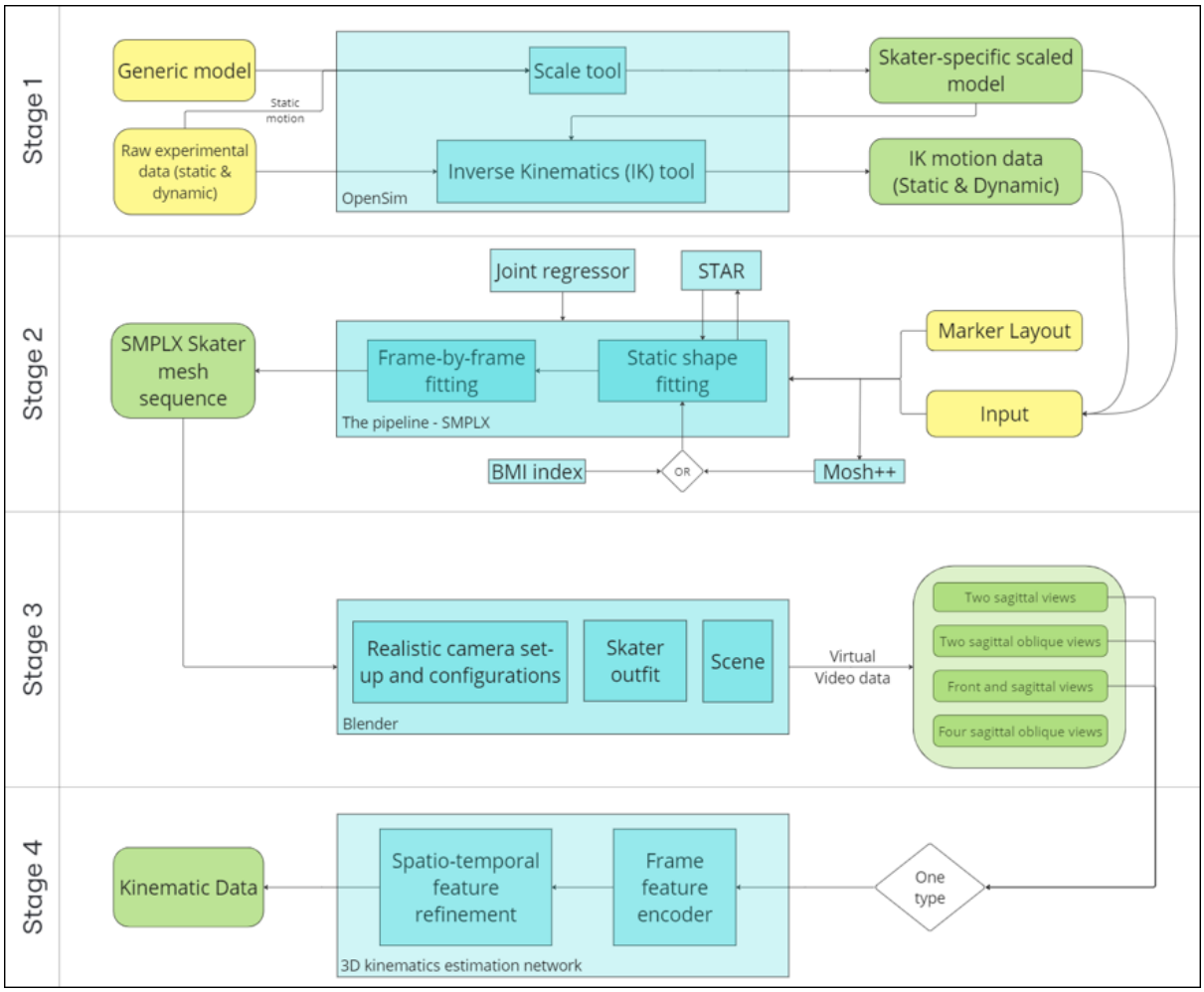


Figure 1: Overview of the four stages of this study

2.2 Experimental Data

This study uses experimental motion capture (MoCap) data, which was collected in a prior study [8] in 2015, which focused on 3D reconstruction of long-track speed skating kinematics. The data was gathered at the indoor ice rink Thialf, in Heerenveen, The Netherlands. Motion data of six elite Dutch skaters (4 men, 2 women; 21.6 $\pm$ 1.02 years; 81 $\pm$ 5.3 kg; 182.83 $\pm$ 4.8 cm) are considered for this research.

For measurements, skaters were equipped with a full-body marker set consisting of 29 markers highlighted in Figure 2(a), with six markers (colored in blue) used only during the static trials. The skating data were recorded using twenty Qualisys cameras (300 Hz), placed on both sides along a 50 m straight section of a 400 m long-track skating rink as shown in the Figure 7 in Appendix A. The dataset contains data of skaters using instrumented Kalpskates while skating at self-selected speeds (normal, slow, medium, and fast) under various movement conditions, such as no arm movement, one arm movement, and both arms moving.

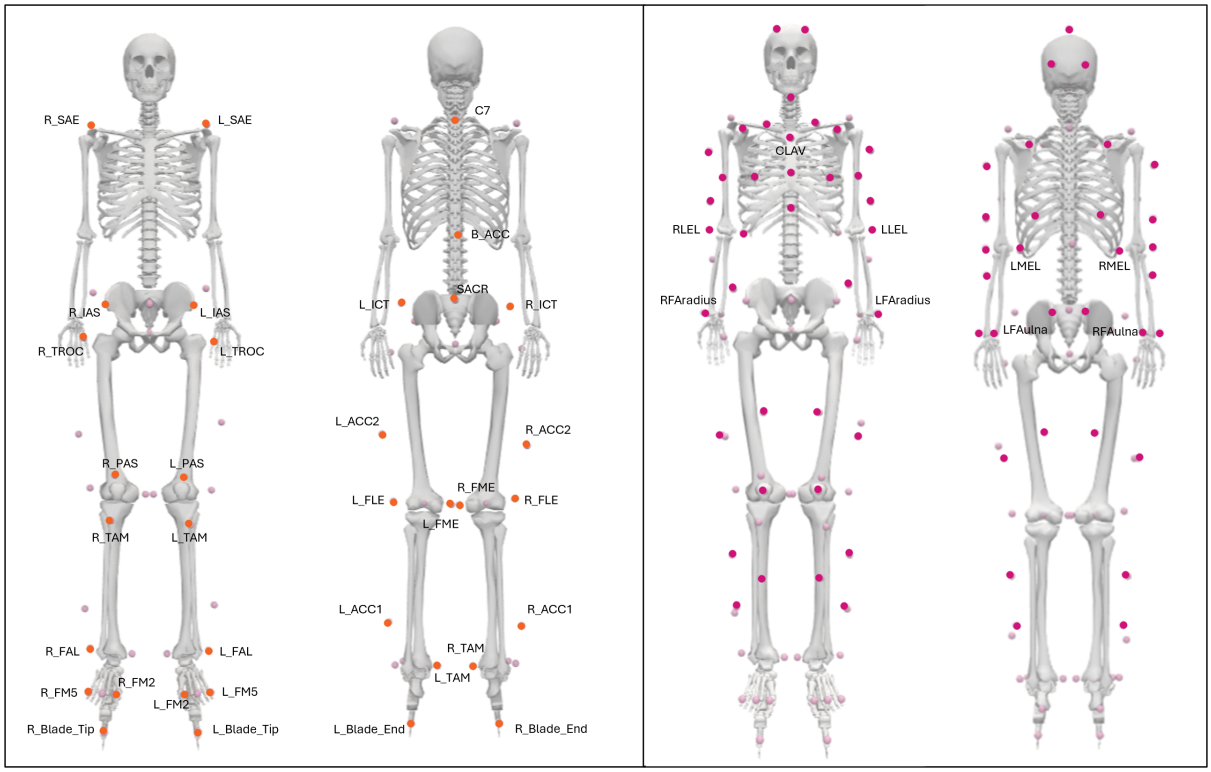


Figure 2: OpenSim skeletal model: a. Experimental markers (left); b. Additional markers for SMPL-X mesh generation (right)

2.3 Biomechanical Model - OpenSim

A full-body OpenSim skeletal model was used to simulate the skating motion. The model consisted of 22 articulated rigid bodies with 20 degrees of freedom in the lower body and 17 degrees of freedom in the upper body. This model was created and validated by Rajagopal et al. (2016) [24]. Skates, consisting of bridges and blades, were added to the generic model using the software OpenSim Creator [25] to represent realistic skating appearance and motion. They were designed in the software Solidworks (2023), taking 'Viking Nagano Gold 2005 Sprint' skating shoes as a reference for design and dimensions. Further scaling and simulations were done using the software OpenSim 4.5 [26]. A set of markers (experimental and joint markers) were placed on the generic model to generate skater-specific skeletal models (Figure 2(a)) and to assist movements. With the help of the Inverse Kinematics (IK) tool and experimental tracking data, skating motions were generated.

2.4 Selection of accurate motion data

Raw experimental MoCap data contain noise, gaps, spikes, jitter, and other artifacts that need to be cleaned and processed [27]. All the tracking files were first trimmed to exclude missing or NaN data at the start and end of the motion. They were further trimmed to eliminate the effect of spikes and jitters in IK output. As mentioned in the reference study [8], the dynamic posture of skaters, compared to the static trial, led to significant shifts of marker positions relative to their anatomical landmarks. This systematic error was evaluated by analyzing the marker error between the experimental data and modal marker data from OpenSim. If the modal marker had a constant offset with the experimental marker throughout all trials of the skater, then that marker was moved to match the experimental marker position.

With the IK results, the maximum marker error and RMS (Root Mean Square) error of our data were not within the limits suggested by OpenSim. Thus, we proceeded further once the systematic error was corrected. These limits depend on the nature of the skeletal model and the motion being examined, which in this case had high instrumented error in the motion due to the large volume of data being captured [8]. Further, motion files (IK results) were selected for next stage if the systematic marker error was reduced, the IK motion was accurate (including visually) regardless of noise in the experimental data, and the RMS errors of each anatomical marker were within a reasonable limit. This limit was subject to the marker's location on the body.

Table 2 in Appendix B provides information about all the skaters and their trials, along with the RMS error for each marker. Data from six skaters, comprising a total of 52 files, were reviewed according to the described procedure. After selection of accurate motion files, the dataset was extended with additional 10 motion files by locking the arms of a skater with arm movements. The final set of motion files were used for virtual video generation.

2.5 The Pipeline

A previously developed and validated pipeline, proposed by Lyn. B. et al. (2023) [23], was used for generating skater-specific 3D human mesh models (stage 2) and rendering the skating motion with a realistic scene setting (stage 3) as shown in Figure 1. The OpenSim skeletal model and the joint angles were the inputs to this pipeline. The pipeline was modified for the speed skater model and motion data with different camera settings to study the skaters' kinematics. Figure 3 illustrates the transition from a skeletal model to a 3D human mesh and finally to a virtual skater, as described in the following subsections.

2.5.1 3D Human skin generation

To generate a 3D human-like mesh, the pipeline uses the Skinned Multi-Person Linear expressive (SMPL-X) model [28]. This model is a skinned vertex-based model that can produce various natural human shapes in different poses. It is driven by an internal armature with 54 joints. It is defined by two parameters: shape β and pose θ. Optimizing these two parameters will provide an optimal SMPL-X mesh that fits best to the given

Figure 3: Overview of the pipeline: a. OpenSim skeletal model b. Fitted SMPLX mesh c. Rendered image of same pose

subject-specific musculoskeletal model and joint angles. A marker layout was used as a reference for mesh generation, manually defined by identifying corresponding SMPL-X vertices as OpenSim markers. These experimental markers, with additional ones highlighted in Figure 2(b), are considered to better represent the human body shape. The pipeline uses MoSh++ (Motion and Shape capture) [29] to create the initial human mesh. It generates a SMPL-X mesh sequence based on the marker trajectories and the marker layout. Using the initial guess, the shape was first optimized with the help of the Sparse Trained Articulated human body Regressor (STAR) model [30] through the static trial. STAR allows for adjusting the size of the mesh using an independent BMI-related shape term. Following this, the pose parameter was optimized frame by frame during the dynamic trial. This was achieved using a joint regressor, minimising a loss function that focused on marker keypoint loss and joint keypoint loss, to produce the best fit between the SMPL-X and OpenSim model. The result of this stage is shown in Figure 3(b).

2.5.2 Rendering - Virtual video

The final stage of the pipeline involved visualizing the motion within the desired scene setting. This was achieved using the software Blender 3.6 LTS with the BLENDER_EEVEE engine. This stage included scene set-up, skater outfit design and camera set-up and configuration, as illustrated in Stage 3 of Figure 1. The scene for speed skating was kept simple, featuring only the ice rink. The skater's outfit was designed to resemble that of a typical speed skater as shown in Figure 3c. To augment appearances, we used 4 different colors which are randomly assigned for each trial.

Camera set-up:
One of the important aspects of kinematic study through video-based motion analysis is the view. In speed skating, skaters steer their skates, resulting in both forward and lateral movements. This creates forward speed causing the skater to move left and right over the rink [31]. Thus, capturing two views for a single motion will provide more information to the network, improving the movement analysis as a result.

To effectively train the network, additional views were introduced to avoid occlusion. Consequently, four different camera setups were used, as shown in Figure 4, along with their corresponding views:

1. Two horizontal side views
2. Two oblique side views (45° inclination towards the skater's motion)
3. Front and side views
4. Four oblique side views (45° inclination with pairs facing each other)

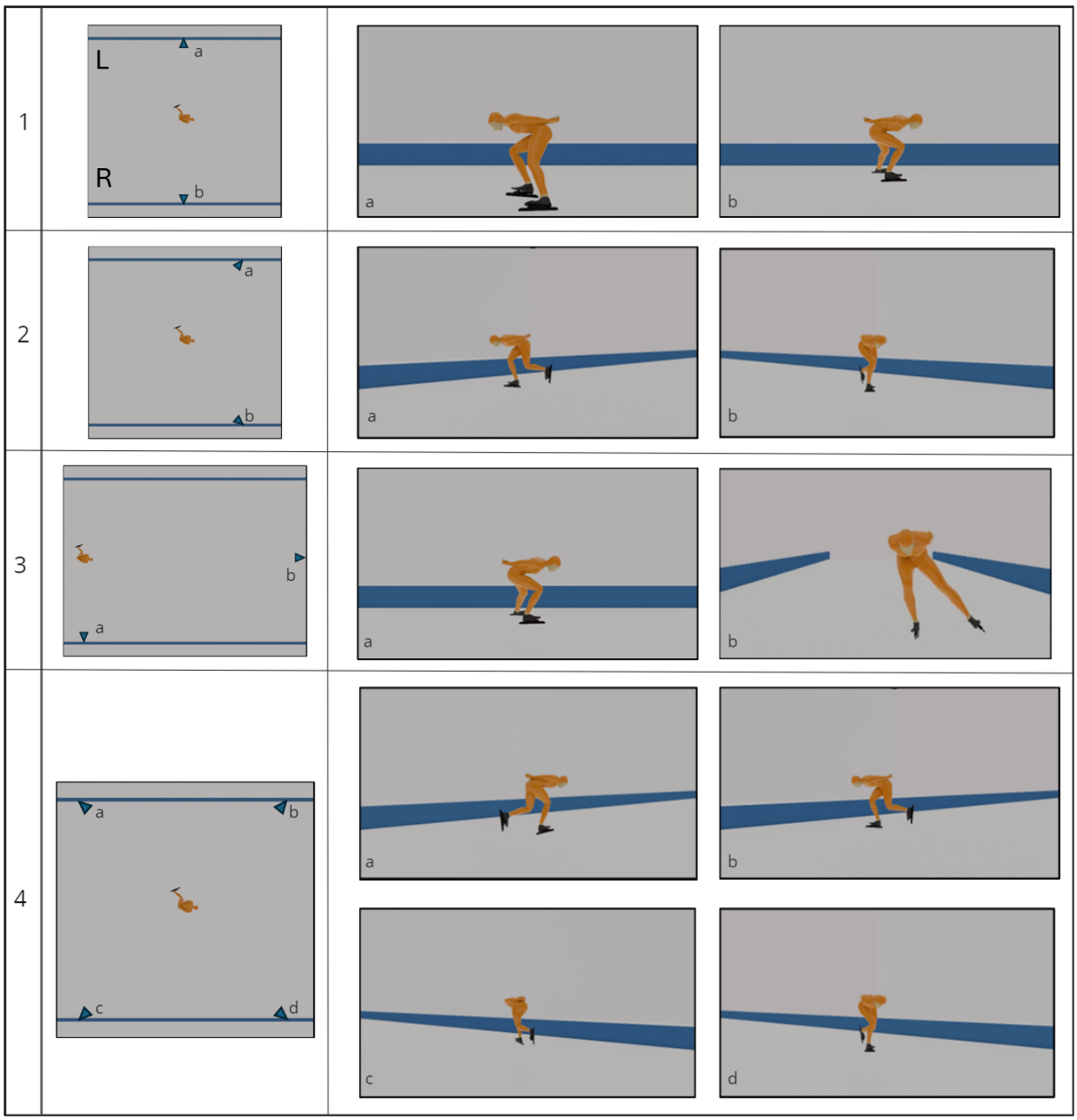


Figure 4: Four different camera setups and their respective views: L (left), R (right), and a, b, c, d correspond to individual camera views.

Camera configuration:
A good video recording dramatically improves the analysis quality and enables networks to produce accurate measurements [32]. Thus, the selection of the most appropriate camera setting is important to undertake high-speed motion analysis with the best quality possible, both spatially and temporally. The camera properties were chosen considering

camera views corresponding to the actual set-up on the rink and the skating speed [33]. We chose to use static cameras positioned along the sides of the rink at a height of 1 m. The side cameras were placed on either sides of the straight section, while the front camera was positioned outside at the curve, facing the straight path. The focal length of the side cameras was set to 35 mm, providing a clear view of the skater while capturing motion dynamics. However, the front camera's focal length was set to 120 mm, considering its position relative to its side camera pair. The sensor fit of the cameras was set to horizontal with the sensor width set to 36 mm. The resolution was set to 1920 by 1080 pixels and collected at a frame rate of 240 fps.

To avoid over-fitting, the network needs to be trained with a large amount of data. Due to limited quantity of accurate experimental MoCap data, we captured a large part of the motion by placing the cameras of the same setup along the track as shown in Figure 5. This greatly expanded the scale of our datasets.

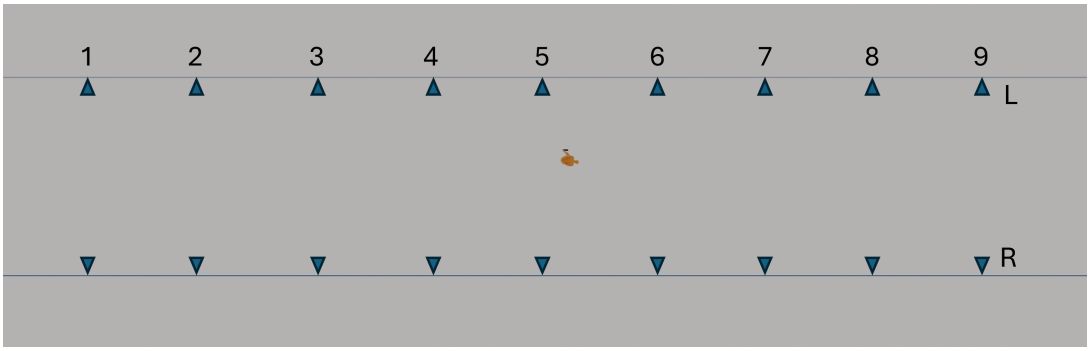


Figure 5: Multiple cameras along the track

2.6 3D Kinematics Estimation Algorithm

The last objective of this research was to test the existing pre-trained biomechanics-aware 3D kinematics estimation algorithm using our virtual video dataset. To achieve this, we selected an end-to-end biomechanics aware network proposed by Lin et al. (2024) [19]. This network was chosen for its ability to predict joint kinematics and body segment scales from video inputs, using an underlying OpenSim skeletal model. It takes two input views to first generate frame features using a frame feature encoder. Later, spatio-temporal feature refinement is applied to refine a sequence of frame features with temporal information. This is illustrated in stage 4 of Figure 1. This network has proven to outperform other state-of-art networks with best generalisation across multiple datasets by solely training on virtual data. Although we were uncertain about the network's performance on sports data, particularly speed skating, we decided to test it on our speed skating video datasets. We evaluated two motion files for each movement condition (no arm movement, one arm movement, both-arm movement) from camera setups 1, 2, and 3, shown in Figure 4. The results were compared with our ground truth kinematics data.

3 Verification

This research followed a multi-stage pipeline with inter-stage data dependency, as shown in the overview in Figure 1. The accuracy of the data at each stage was crucial for ensuring the quality of our dataset and for testing the performance of the algorithm. Therefore, we verified the data at each stage as follows:

3.1 Stage 1 - Experimental data to OpenSim IK data:

As discussed earlier, the marker errors between experimental data and the marker locations from the IK tool were important for establishing the accuracy of the IK data. In this research, the IK data were considered the ground truth. Accordingly, we evaluated the RMS error for each marker so as to understand its movement at its anatomical location during the skating motion, as well as the total RMS error for each trial using the following formulas:

$$\text{RMSE}_{marker3Dj} = \frac{1}{3} \sum_{k=\{x,y,z\}} \sqrt{\frac{1}{f} \sum_{i=1}^{f} \left(\hat{P}_{i,j,k} - P_{i,j,k} \right)^2} \tag{1}$$

$$\text{Total RMSE} = \frac{1}{m} \sum_{j=1}^{m} (\text{RMSE}_{marker3Dj}) \tag{2}$$

where $\hat{P}_{i,j,k}$ and $P_{i,j,k}$ represent the experimental data and IK data for the i-th frame, j-th marker, and k-th component respectively. Additionally, f denotes the total number of frames, and m denotes the total number of markers.

3.2 Stage 2 - OpenSim data to SMPL-X mesh:

After the 3D human skin generation, it is important to check the mean marker spatial error [23] to evaluate how well our fitted mesh can track the input OpenSim motions. This error reflects the quality of the mesh generated by the pipeline by measuring the discrepancy between the virtual marker positions of the OpenSim model and their corresponding vertices on the SMPL-X mesh. This error was calculated using the following formula:

$$E_{\text{marker3D}} = \frac{1}{m \times f} \sum_{j=0}^{n-1} \sum_{i=0}^{f-1} \left\| \hat{P}_{j,i} - P_{j,i} \right\|_2 \tag{3}$$

where $\hat{P}_{j,i}$ is the position of a virtual marker on fitted SMPL-X mesh and $P_{j,i}$ is the position of a virtual marker on OpenSim model

3.3 Stage 3 - OpenSim data to virtual video generation:

After rendering, it was important to assess the motion reconstruction ability of the pipeline in the output virtual videos. Using the cameras' intrinsic and extrinsic matrices, OpenSim virtual markers and SMPL-X virtual markers were projected into the image pixel coordinate system. The discrepancy between them, referred to as the mean marker pixel error [23], was evaluated using the following formula:

$$E_{\text{marker2D}} = \frac{1}{n \times f} \sum_{j=0}^{n-1} \sum_{i=0}^{f-1} \left\| \hat{X}_{j,i} - X_{j,i} \right\|_2 \tag{4}$$

where $\hat{X}_{j,i}$ is the SMPL-X virtual marker position and $X_{j,i}$ is the OpenSim virtual marker position.

4 Network evaluation metric

Our speed skating dataset was used to evaluate the performance of a pre-trained network in its ability to extract 3D kinematics. The metric presented in Lin et al. [19] was used to evaluate our results.

Joint angles were evaluated through **Mean Absolute Error (MAE)**. It was calculated using the following formula:

$$MAE_{\text{angle}} = \frac{1}{f} \sum_{t=1}^{f} \left\| \hat{\theta}_t - \theta_t \right\|_1 \tag{5}$$

where $\hat{\theta}_t$ is the predicted angles and θ_t is the ground truth OpenSim angles for f number of frames in the sequence.

5 Results

5.1 Stage 1 - Experimental data to OpenSim IK data:

In Stage 1, the experimental MoCap data of 5 skaters (3 men and 2 women) over 41 trials was selected from a dataset of 6 skaters (4 men and 2 women) over 51 trials. To evaluate the accuracy of our IK motion files, we calculated the RMSE of each marker and the total RMSE for each trial using Equations 1 and 2. The results are presented in Table 2 (Appendix B). The total RMSE of all the trials is 1.092 cm. We also calculated the RMSE of each marker over all trials, with the results shown in Figure 6. This helped establish the RMSE limits for each marker based on their anatomical location, which are listed below the names of the markers in Table 2 (Appendix B).

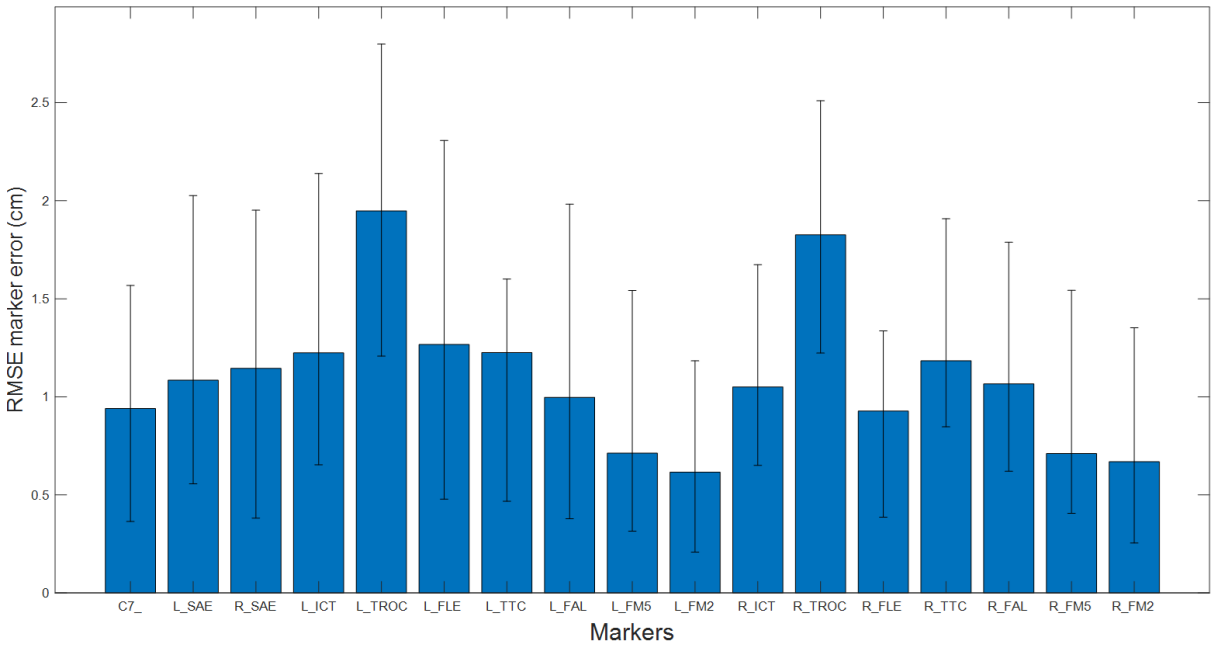


Figure 6: RMS error of each marker over all the trials

The dataset was finalized to include 51 accurate motion data files, consisting of 41 trials and 10 additional motion files, as described in Section 2.4. It comprises various movement conditions: 41 motion files with arms locked behind the back, 8 files with one arm in motion, and 2 files with both arms in motion. These movements were recorded at different self-selected speeds, including 7 files at slow speed, 25 at normal speed, 12 at medium speed, and 7 at fast speed.

5.2 Stage 2 - OpenSim data to SMPL-X mesh:

With the help of the pipeline, we generated skaters' human mesh sequences from the marker trajectories. Since the pipeline optimizes to produce the best fit of the SMPL-X body model with the OpenSim skeletal model, we evaluated the accuracy of this fit using equation 3. The mean marker spatial error of all trials over all participants is 0.033 m. The result of one trial is listed in Table 3 in Appendix C, which includes both experimental markers and some additional markers. It can be observed that the fit of the pelvis, hips and knee is not optimal. However, this varied among different skaters depending on their body shape and motion. Additionally, as the mesh is participant-specific, generated using the static trial and optimized over dynamic poses, the mean marker spatial error for each participant is shown in Table 4 in Appendix C. This suggests that the mesh fitting is influenced by factors beyond gender which will be discussed in section 6.

5.3 Stage 3 - OpenSim data to virtual video generation:

The generated SMPL-X human mesh sequences were rendered in a virtual set-up to produce virtual skating videos that accurately represent the captured skating motion. The 2D representation of 3D motion data is crucial and should be evaluated. The cameras' location, rotation, and projections significantly influence this representation, especially

when aiming for accurate kinematic estimation by networks from 2D data. Using Equation 4, the mean marker pixel errors were evaluated across four different camera set-ups for three distinct movement conditions: no arm movement, one-arm movement, and both-arm movement, as shown in Table 1. Considering the resolution (1920 x 1080 pixels), one pixel error in the image coordinate system corresponds to an approximate error of 0.267 cm in 3D space. From Table 1, it can be inferred that the discrepancy between SMPL-X virtual markers and OpenSim virtual markers is least in the oblique views. This reflects the effectiveness of the camera set-up and its configurations to simulate realistic conditions on the rink.

Table 1: Mean pixel error of different camera set-ups for different skating conditions

Condition	**Mean marker pixel error (pixels) – camera set-ups**			
	Two horizontal side views	Two oblique side views	Front & side views	Four oblique side views
No arm movement	9.05	5.03	8.83	5.43
One arm movement	11.12	5.65	10.18	5.92
Both-arm movement	10.85	5.82	10.22	5.97
Mean	10.34 pixels ∼2.76 cm	5.5 pixels ∼1.47 cm	9.74 pixels ∼2.6 cm	5.77 pixels ∼1.54 cm

At the end of this stage, we created a virtual video dataset of speed skating consisting of 1,326 videos based on 51 motion capture trials from 5 skaters. Following its parent dataset, ODAH (OpenSim Driven Animated Human) [23], it is named ODAH-SpeedSkater.

5.4 Stage 4 - Network evaluation:

To evaluate the performance of the chosen pre-trained 3D kinematics estimation network for speed skating, we tested it on our ODAH-Speedskater dataset as described in section 2.6. As this network inputs two views, we tested it on each movement condition for our camera set-ups with two views and the results were evaluated using the equation 5 as shown in Table 5 and 6 in Appendix D. The results are derived from limited testing as a result of time constraints. However, it is evident that the MAE is high for our dataset compared to the network's performance on its trained ODAH dataset.

The tables list the results of two camera set-ups: two side views and two oblique views. The network's results on the third camera setup, which included front and side views, were incomplete. It attempted to detect the motion at the beginning, specifically during the first 20-30% of the frame time, but ultimately failed to capture the kinematics in the subsequent frames.

6 Discussions, Limitations and Future Work

Stage 1 - Experimental trc data to OpenSim IK data:

The accuracy and reproducibility of experimental MoCap data was crucial for obtaining realistic speed skating movement in the virtual videos. Our results on marker error, as described in Section 5.1, show that the selected experimental motion data, obtained through a carefully described selection process, are biomechanically accurate for speed skating motion, with a total RMSE of less than 1.5 cm. The rejected motion files, after cleaning the raw experimental data, exhibited high noise, mainly in pelvis and hip markers, and in some cases certain markers were not properly tracked. These files were rejected based on observing the OpenSim marker trajectories, as IK tools can handle some noise in experimental data. Our observation on each rejected file is listed in the end of the table 2 (Appendix B), with numbered footnotes.

Of course, our model is not entirely without limitations. The skeletal model used in this research was designed for accurate gait simulations with main focus on lower extremity [24]. This simplified generic model poses some challenges for a full representation of skating motion, mainly regarding joints. Further, the scaling of the generic model was based on marker positions which was limiting for some parts of the body like the foot and the pelvis. Thus, it might be beneficial to use imaging techniques such as Magnetic Resonance Imaging (MRI) and 3D body scans that can help with the anthropometry of the skaters, providing accurate measurements of segment lengths including joint centers. These measurements can then be used to scale the dimensions and inertial properties of a generic model to fit an individual subject [34].

Stage 2 - OpenSim data to SMPL-X mesh:

As described in section 5.2, the accuracy of mesh fitting with our ground truth data is 3.3 cm, with multiple factors contributing to this value. First, the pipeline uses the SMPL-X body model for mesh generation which offers identity-dependent shape and non-rigid pose-dependent shape. The model was trained on 3D scans of 40 individuals (20 males and 20 females) capturing wide variety of body shapes in natural human poses [28]. However, generating accurate body shapes of speed skaters, mainly muscular lower body and toned upper body, proved challenging.

Further, the initial guess of shape and pose was done by MoSh++ which uses marker layout to estimate body shape [29]. As seen in Figure 2, the additional markers were added to help define the overall body shape. However, these markers do not accurately represent the specific body shape of individual skaters. During optimisation, if the MoSh++ results were not accurate, then the independent BMI-related term in STAR was used to manually iterate and produce a similar shape as shown through the shape-factor parameter in Table 4 (Appendix C). Additionally, the joint regressor used was trained on a small dataset of manually fitted templates of walking motion. When the joint fit was not good, the joint offsets of STAR and SMPL-X with respect to OpenSim were manually iterated. Thus, to avoid the manual iterations to produce best shape and pose, it is advantageous to use the 3D body scans of the skaters where both the representation of surface mesh as well as its relation with underlining skeleton structure, mainly joint position, are accurate. This will

facilitate accurate representation of skaters and their MoCap data in virtual video. Given the limitations, the mean marker spatial error of our data can be considered accurate as it represents the most accurate result we could achieve using our methodology.

Stage 3 - OpenSim data to virtual video generation:

One of our aims was to represent the speed skating MoCap data as a video through the lens of a realistic on-rink camera set-up. The mean pixel error of our mesh fit in skating motion varied for different camera positions and set-ups. The error for direct views was high compared to oblique views for our camera set-up and configuration. This value is the 2D representation of our mean marker spatial error which on reducing can help all views to be more accurate. Although this grants insight in which views might be better for kinematic estimation and on-rink implementation, it is still to be evaluated through a network's performance.

Our dataset has been partially validated through our verification metric and visual evaluation. However, for complete validation, it is recommended to capture real videos along with MoCap data using our camera set-up and configuration. As the camera captures a small area, a MoCap set-up can be adopted. It is recommended to also capture static data in the skating pose to ensure the accuracy and relevance of marker placements to the motion [8]. The virtual video can be generated using our methodology, and tested on the network trained on our dataset.

Stage 4 - Network evaluation:

The final goal was to test the network trained on synthetic data generated through the same pipeline along with its promising performance against other state-of-art. It can be inferred from our results in Table 5 and Table 6 (Appendix D) that the performance of the network is not accurate when compared with our ground truth data. This is possibly because the network was trained on simple motions like sitting down, throwing and catching. Its understanding of speed skating poses and movements may yet be limited. Although our testing was limited, it is noteworthy that the angles of the right hip, knee, and ankle are significantly lower than those of the left leg in the side view camera setup. As the skater moves forward and engages in lateral movements, the accuracy of capturing the full body becomes critical. This highlights the necessity for the network to be trained on a dataset specifically designed for such conditions. For the third camera set-up, associating front view with side view was challenging to the network. The issue of the network detecting only the initial few frames while failing to generalize to subsequent frames may be indicative of overfitting. Thus, an important part of the future work should be to train the network with our dataset and then further evaluate its performance.

7 Conclusion

In this research, we developed 'ODAH-SpeedSkater', the first virtual video dataset based on experimental data for speed skating. The impact and accuracy of markers during motion capture were highlighted during selection of accurate IK data for virtual video generation. We successfully generated an accurate virtual skater mesh and corresponding

motion sequences for five skaters over 51 MoCap trials, using ground truth kinematics. This was achieved by rigging an SMPL-X human body model to a skater-specific scaled OpenSim skeletal model and its associated joint angles. Further, the finalized mesh sequences were rendered in an ice rink scene with skater outfits through realistic camera set-ups and configurations. Thus, we created a large dataset of 1,326 biomechanically accurate annotated speed skating videos. The dataset has been verified and partially validated through our metrics. Yet, it is beneficial to first further validate the dataset with our proposed camera set-ups and configurations. Finally, we tested the dataset on a pre-trained 3D kinematic estimation algorithm to evaluate its performance on speed skating data. Although we were unable to test the network extensively,we conclude that training the network on our dataset may improve its performance and avoid overfitting. For future research, we recommend to continue testing our dataset with other pre-trained algorithms. We also recommended to train and test networks with our dataset and real videos. We are confident that these networks will provide accurate kinematic analysis for speed skating and will soon make their active contribution on the ice rink.

References

[1] J. v. d. Eb, S. Gereats, and A. Knobbe, "Enhancing the performance of elite speed skaters using skateview: A new device to measure performance in speed skating," in *Proceedings*, MDPI, vol. 49, 2020, p. 133.

[2] J. van der Eb, H. Mossink, E. Kiel, S. Geraets, D. Veeger, and P. J. Beek, "Analysis of in competition speed skating using imu's," in *36th Conference of the*, 2018.

[3] E. Van der Kruk, A. Schwab, F. Van Der Helm, and H. Veeger, "Getting the angles straight in speed skating: A validation study on an imu filter design to measure the lean angle of the skate on the straights," *Procedia engineering*, vol. 147, pp. 590–595, 2016.

[4] G. J. van Ingen Schenau, "The influence of air friction in speed skating," *Journal of Biomechanics*, vol. 15, no. 6, pp. 449–458, 1982.

[5] T. Purevsuren, B. Khuyagbaatar, K. Kim, and Y. H. Kim, "Investigation of knee joint forces and moments during short-track speed skating using wearable motion analysis system," *International Journal of Precision Engineering and Manufacturing*, vol. 19, pp. 1055–1060, 2018.

[6] Y. Tomita, T. Iizuka, K. Irisawa, and S. Imura, "Detection of movement events of long-track speed skating using wearable inertial sensors," *Sensors*, vol. 21, no. 11, p. 3649, 2021.

[7] E. Van der Kruk and M. M. Reijne, "Accuracy of human motion capture systems for sport applications; state-of-the-art review," *European journal of sport science*, vol. 18, no. 6, pp. 806–819, 2018.

[8] E. van der Kruk, A. Schwab, F. van der Helm, and H. Veeger, "Getting in shape: Reconstructing three-dimensional long-track speed skating kinematics by comparing several body pose reconstruction techniques," *Journal of biomechanics*, vol. 69, pp. 103–112, 2018.

[9] R. De Boer, G. Ettema, H. Van Gorkum, G. De Groot, and G. van Ingen Schenau, "A geometrical model of speed skating the curves," *Journal of biomechanics*, vol. 21, no. 6, pp. 445–450, 1988.

[10] G. v. van Ingen Schenau, G. De Groot, and R. De Boer, "The control of speed in elite female speed skaters," *Journal of biomechanics*, vol. 18, no. 2, pp. 91–96, 1985.

[11] G. van Ingen Schenau, G. De Groot, and A. Hollander, "Some technical, physiological and anthropometrical aspects of speed skating," *European journal of applied physiology and occupational physiology*, vol. 50, no. 3, pp. 343–354, 1983.

[12] G. van Ingen Schenau and G. De Groot, "On the origin of differences in performance level between elite male and female speed skaters," *Human Movement Science*, vol. 2, no. 3, pp. 151–159, 1983.

[13] H. Houdijk *et al.*, "Push-off mechanics in speed skating with conventio...," 2000.

[14] C. J. Payton and R. Bartlett, *Biomechanical evaluation of movement in sport and exercise*. Routledge Abingdon, Oxon, UK, 2007.

[15] X. Suo, W. Tang, and Z. Li, "Motion capture technology in sports scenarios: A survey," *Sensors*, vol. 24, no. 9, p. 2947, 2024.

[16] D. Pagnon, M. Domalain, and L. Reveret, "Pose2sim: An open-source python package for multiview markerless kinematics," *Journal of Open Source Software*, vol. 7, no. 77, p. 4362, 2022.

[17] S. D. Uhlrich, A. Falisse, Ł. Kidziński, *et al.*, "Opencap: Human movement dynamics from smartphone videos," *PLoS computational biology*, vol. 19, no. 10, e1011462, 2023.

[18] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7291–7299.

[19] Z.-Y. Lin, B. Lyu, J. C. Fernandez, E. Van Der Kruk, A. Seth, and X. Zhang, "3d kinematics estimation from video with a biomechanical model and synthetic training data," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 1441–1450.

[20] M. Bittner, W.-T. Yang, X. Zhang, A. Seth, J. van Gemert, and F. C. van der Helm, "Towards single camera human 3d-kinematics," *Sensors*, vol. 23, no. 1, p. 341, 2022.

[21] T. Chen, C. Fang, X. Shen, Y. Zhu, Z. Chen, and J. Luo, "Anatomy-aware 3d human pose estimation with bone-based pose decomposition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 1, pp. 198–209, 2021.

[22] L. Sigal, A. O. Balan, and M. J. Black, "Humaneva: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion," *International journal of computer vision*, vol. 87, no. 1, pp. 4–27, 2010.

[23] L. Bofan, "Synthetic human motion video generation based on biomechanical model," *Master's thesis, Delft University of Technology*, 2023.

[24] A. Rajagopal, C. L. Dembia, M. S. DeMers, D. D. Delp, J. L. Hicks, and S. L. Delp, "Full-body musculoskeletal model for muscle-driven simulation of human gait," *IEEE transactions on biomedical engineering*, vol. 63, no. 10, pp. 2068–2079, 2016.

[25] A. Kewley, J. van Beesel, and A. Seth, *OpenSim Creator*, version 0.5.14, Sep. 2024. [Online]. Available: `https://github.com/ComputationalBiomechanicsLab/opensim-creator`.

[26] S. L. Delp, F. C. Anderson, A. S. Arnold, *et al.*, "Opensim: Open-source software to create and analyze dynamic simulations of movement," *IEEE transactions on biomedical engineering*, vol. 54, no. 11, pp. 1940–1950, 2007.

[27] K. Chen, Y. Wang, S.-H. Zhang, S.-Z. Xu, W. Zhang, and S.-M. Hu, "Mocap-solver: A neural solver for optical motion capture data," *ACM Transactions on Graphics (TOG)*, vol. 40, no. 4, pp. 1–11, 2021.

[28] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black, "Smpl: A skinned multi-person linear model," in *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, 2023, pp. 851–866.

[29] N. Mahmood, N. Ghorbani, N. F. Troje, G. Pons-Moll, and M. J. Black, "Amass: Archive of motion capture as surface shapes," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 5442–5451.

[30] A. A. Osman, T. Bolkart, and M. J. Black, "Star: Sparse trained articulated human body regressor," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VI 16*, Springer, 2020, pp. 598–613.

[31] E. van der Kruk, "Parameter analysis for speed skating performance," 2018.

[32] B. Pueo, "High speed cameras for motion analysis in sports science," *Journal of Human Sport and Exercise*, vol. 11, no. 1, pp. 53–73, 2016.

[33] A. J. Roete, I. K. Stoter, R. P. Lamberts, M. T. Elferink-Gemser, and R. T. Otter, "Introducing a method to quantify the specificity of training for races in speed skating," *The Journal of Strength & Conditioning Research*, pp. 10–1519, 2022.

[34] J. L. Hicks, T. K. Uchida, A. Seth, A. Rajagopal, and S. L. Delp, "Is my model good enough? best practices for verification and validation of musculoskeletal models and simulations of movement," *Journal of biomechanical engineering*, vol. 137, no. 2, p. 020 905, 2015.

A Appendix A

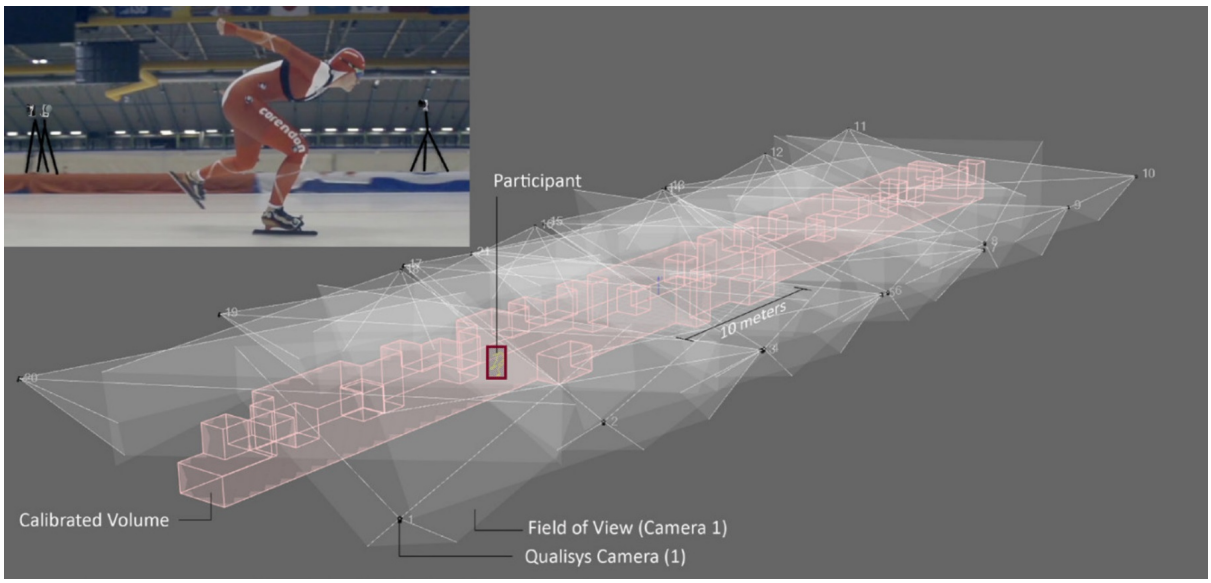


Figure 7: Experimental set-up of twenty Qualisys cameras located along the 50 m straight part of the rink that was the calibrated volume. The skaters were equipped with a full body marker set as described. The pink areas indicate the calibrated volumes. The field of view of each camera is shown for interpretation [8]

B Appendix B

Table 2: RMSE of each marker and total RMSE for each trial of the six skaters are listed. RMS error limits for each marker are indicated below its name. Footnotes for rejected files are listed at the end of the table.

Name	Trial	Selection	Total RMSE	RMSE of each marker (cm)															
				C7	L_SAE	R_SAE	L_ICT	L_TROC	L_FLE	L_TTC	L_FAL	L_FM5	L_FM2	R_ICT	R_TROC	R_FLE	R_TTC	R_FAL	R_FM5	R_FM2
				<2cm	<2cm	<2cm	<3cm	<3cm	<3cm	<3cm	<2cm	<2cm	<2cm	<3cm	<3cm	<3cm	<= 3cm	<2cm	<2cm	<2cm
P1	T1	✓	0.92	0.61	0.87	0.79	1.51	2.19	0.81	0.62	0.62	0.42	0.25	1.3	1.89	0.89	0.87	0.94	0.56	0.57
	T2	✓	0.94	0.65	0.79	0.67	1.4	1.77	0.9	0.69	0.61	0.49	0.43	1.41	2.2	1	1.13	0.8	0.55	0.46
	T3	✓	0.86	0.51	0.66	0.49	1.7	1.71	0.76	0.47	1.03	0.34	0.22	1.46	1.95	0.66	0.94	0.98	0.41	0.28
	T4	✓	1	0.46	0.9	0.56	1.46	2.08	1	0.71	0.68	0.91	0.82	1.01	1.53	1.16	1.23	1.23	0.7	0.59
	T5	✓	0.85	0.45	0.74	0.56	1.17	1.9	1.06	0.74	0.38	0.31	0.36	0.9	1.62	0.9	1.07	0.99	0.79	0.53
	T6	✓	0.91	0.36	0.67	0.57	1.24	1.83	1.05	0.74	0.51	0.46	0.36	1.2	1.79	0.96	1.24	1.09	0.77	0.65
	T7	✓	0.92	0.59	0.84	0.38	1.74	2.11	0.89	0.56	0.83	0.36	0.21	1.34	1.87	0.78	1.35	1.03	0.45	0.25
	T8	✓	1.04	0.47	1.22	1.09	1.94	2.21	1.07	0.72	0.63	0.41	0.39	1.2	1.58	1.04	1.33	1.09	0.72	0.49
	T9	✓	1.03	0.61	0.84	1.04	1.31	1.83	1.14	0.77	0.68	0.46	0.46	0.93	2.03	1.1	1.52	1.19	0.85	0.67
	T10	✗[1]	1.07	0.47	0.92	1.12	1.28	1.75	1.13	0.79	1.13	0.47	0.61	0.93	1.9	1.14	1.49	1.29	0.84	0.89
P2	T1	✗[2]	9.79	2.29	4.22	3.97	16.23	79.29	5.53	4.83	4.28	1.64	2.05	9.69	11.25	6.37	4.83	5.73	2.18	1.97
	T2	✗[3]	1.96	1.34	3.85	2.45	2.31	3.97	1.37	1.47	0.88	0.61	0.5	2.16	5.8	1.14	1.6	0.75	1.88	1.3
	T3	✗[4]	1.87	1.16	1.59	1.26	2.98	4.37	1.08	1.24	1.04	0.5	0.75	3.43	5.03	1.98	2.13	1.13	1.07	1.09
	T4	✗[5]	1.58	1.05	1.52	1	3.79	2.87	0.97	1.3	1.4	0.54	0.7	2.02	3.78	1.04	1.42	1.32	1.13	1.1
	T5	✗[4]	2.19	1.1	1.62	1.34	4.72	7.86	2.01	2.06	1.63	0.57	0.78	2.87	4.02	1.7	1.99	1.12	1.06	0.85
P3	T1	✓	0.88	0.78	0.86	1.07	0.91	1.43	0.52	1.36	0.87	0.57	0.53	0.7	1.36	0.48	1.37	0.82	0.64	0.63
	T2	✓	0.94	0.85	0.91	1.3	0.92	1.43	0.48	1.46	0.78	0.66	0.5	0.65	1.39	0.84	1.91	0.75	0.59	0.47
	T3	✓	0.99	0.99	0.96	1.19	1.5	1.47	0.6	NaN	0.85	0.66	0.48	1.38	1.48	1.15	1.86	1.23	0.54	0.54
	T4	✓	0.94	1.08	1.06	1.14	1.36	1.61	0.53	1.46	0.92	0.61	0.42	1.05	1.22	0.39	1.28	0.64	0.61	0.63
P4	T1	✓	1.15	1.09	0.76	0.98	1.78	2.15	1.41	1.19	1.22	1.21	0.77	1.03	1.61	0.75	0.85	1.14	0.95	0.6
	T2	✓	1.08	0.77	0.75	0.95	1.17	1.88	1.48	1.13	0.67	1.54	1.02	1.26	1.59	0.97	0.86	0.9	0.92	0.47
	T3	✗[6]	1.22	1.25	2.28	1.9	1.62	2.33	1.75	0.97	0.65	1.08	0.64	0.75	2.09	0.6	0.85	1.17	0.66	0.2
	T4	✗[7]	152.28	234.68	174.96	159.07	294.12	189.67	78.72	108.96	180.86	123.96	128.49	207.47	144.73	165.22	102.58	82.69	124.3	88.21
	T5	✗[8]	1.55	1.06	2.16	2.06	1.72	2.33	1.77	1.06	1.1	1.07	0.7	1.17	2.37	3.97	1.49	1.24	0.67	0.32

Table 2: RMSE of each marker and total RMSE for each trial of the six skaters are listed. RMS error limits for each marker are indicated below its name. Footnotes for rejected files are listed at the end of the table.

Name	Trial	Selection	Total RMSE	RMSE of each marker (cm)															
				C7	L_SAE	R_SAE	L_ICT	L_TROC	L_FLE	L_TTC	L_FAL	L_FM5	L_FM2	R_ICT	R_TROC	R_FLE	R_TTC	R_FAL	R_FM5	R_FM2
				<2cm	<2cm	<2cm	<3cm	<3cm	<3cm	<3cm	<2cm	<2cm	<2cm	<3cm	<3cm	<3cm	<= 3cm	<2cm	<2cm	<2cm
	T6	✓	1.12	1.01	0.72	0.95	1.66	2.13	1.6	1.23	0.8	1.15	0.79	1.27	1.59	0.81	0.87	0.76	0.99	0.64
	T7	✓	1.09	1	0.71	0.89	1.45	2.19	1.44	1.24	0.75	1.09	0.72	1.46	1.5	0.9	0.94	0.83	0.98	0.49
	T8	✓	1.09	1	0.61	0.93	1.5	2.12	1.47	1.22	0.84	1.24	0.94	1.03	1.59	0.83	0.87	0.97	0.9	0.55
	T9	✓	1.3	0.87	1.97	1.79	1.13	2.68	1.61	1.22	0.63	1.18	0.72	0.71	2.51	1.03	1.02	1.21	0.89	0.88
	T10	✓	1.39	1.05	2.03	1.83	1.49	2.8	1.52	1.12	1.34	1.3	0.86	1.14	2.44	0.97	1.18	1.16	0.97	0.51
	T11	✓	1.32	0.89	1.92	1.91	1.37	2.58	1.61	1.47	0.87	1.23	0.78	0.82	2.26	1.06	1.15	1.13	0.88	0.48
P5	T1	✓	1.11	1.05	1.19	1.21	1.29	1.91	0.68	1.29	0.88	0.55	0.43	1.06	1.53	1.13	1.26	1.25	0.87	1.22
	T2	✗[9]	1.12	1.07	1.32	1.46	0.96	1.39	0.66	1.11	2.07	0.9	0.51	0.88	1.72	0.87	1.3	1.41	0.59	0.87
	T3	✓	0.96	0.99	1.43	1.39	1.07	1.21	0.64	1.07	1	0.43	0.24	0.65	1.68	0.87	1.43	0.62	0.6	0.93
	T4	✓	1.16	1	1.09	1.41	1.99	2.22	0.74	1.28	1.45	0.71	0.6	0.87	1.67	0.99	1.32	0.93	0.6	0.85
	T5	✓	1.11	1.08	1.07	1.39	1.49	1.77	0.8	1.22	1.07	0.74	0.63	0.8	1.65	1.13	1.25	1.12	0.69	1.02
	T6	✓	1.2	0.93	1.21	1.17	2.14	1.91	0.68	1.28	1.4	0.74	0.6	1.33	1.96	1.01	1.33	1.07	0.72	0.93
	T7	✓	1.21	1.18	1.14	1.45	1.41	1.79	1.29	1.23	1.24	0.68	0.5	0.99	1.79	1.06	1.45	1.25	0.91	1.14
	T8	✓	1.03	1.27	1.35	1.39	0.94	1.7	0.6	1.23	0.92	0.72	0.54	0.87	1.37	1.02	1.15	0.77	0.67	0.94
	T9	✓	1.21	1.3	1.36	1.27	1.7	1.91	0.86	1.57	0.97	0.73	0.54	1.12	1.33	1.21	1.45	0.99	0.95	1.35
	T10	✓	0.99	1.47	1.19	1.66	0.77	1.26	0.54	1.18	0.74	0.66	0.52	0.87	1.74	0.86	1.14	0.81	0.55	0.87
	T11	✓	1.07	1.17	1.35	1.45	1.15	1.62	0.75	1.36	0.92	0.7	0.51	0.8	1.62	0.98	1.23	0.97	0.68	0.93
P6	T1	✓	1.04	0.82	0.93	0.86	0.86	2.01	1.35	1.39	0.87	0.64	0.94	0.95	1.96	0.79	1.23	1.07	0.55	0.51
	T2	✓	1.05	1.04	1.29	0.99	0.69	1.53	1.31	1.35	1.06	0.64	0.59	1.01	1.75	0.77	1.38	1.22	0.61	0.62
	T3	✓	1.28	1.46	1.53	1.55	1.14	2.01	1.63	1.58	1.14	0.71	0.82	1.67	1.97	0.87	1.11	1.33	0.66	0.66
	T4	✓	1.13	1.39	1.64	1.3	0.75	1.82	1.49	1.39	1.13	0.55	0.47	1.28	1.94	0.82	1.12	0.9	0.58	0.57
	T5	✓	1.16	1.36	1.48	1.38	0.71	1.92	1.47	1.59	1.12	0.6	0.59	1.28	1.91	0.83	1.17	1.12	0.58	0.67
	T6	✓	1.51	1.57	1.85	1.47	1.27	2.21	2.13	1.51	1.53	0.8	0.86	1.38	2.33	1.34	1.21	1.79	1.54	0.93
	T7	✓	1.4	1.43	1.4	1.95	1.11	2.07	2.31	1.6	1.98	0.95	1.18	1.18	2.01	1.07	1.01	1.29	0.57	0.61
	T8	✓	1.1	1.21	1.34	1.25	0.9	2.18	1.9	1.21	1	0.56	0.53	0.97	2.03	0.89	0.97	0.83	0.51	0.5
	T9	✓	1.15	1.49	1.44	1.49	0.85	1.87	1.83	1.28	0.99	0.6	0.66	1.13	1.94	0.93	1	0.87	0.52	0.59
	T10	✓	1.24	1.54	1.44	1.46	1.19	1.96	1.65	1.41	1.14	0.6	0.62	1.33	1.91	0.9	1.07	1.64	0.62	0.67

Table 2: RMSE of each marker and total RMSE for each trial of the six skaters are listed. RMS error limits for each marker are indicated below its name. Footnotes for rejected files are listed at the end of the table.

Name	Trial	Selection	Total RMSE	RMSE of each marker (cm)															
				C7	L_SAE	R_SAE	L_ICT	L_TROC	L_FLE	L_TTC	L_FAL	L_FM5	L_FM2	R_ICT	R_TROC	R_FLE	R_TTC	R_FAL	R_FM5	R_FM2
				<2cm	<2cm	<2cm	<3cm	<3cm	<3cm	<3cm	<2cm	<2cm	<2cm	<3cm	<3cm	<3cm	<= 3cm	<2cm	<2cm	<2cm
	T11	✓	0.99	0.53	0.56	0.75	0.95	2.05	1.35	1.39	0.88	0.64	0.95	0.74	1.96	0.78	1.24	1.07	0.54	0.5
	T12	✓	0.92	0.59	0.89	0.69	0.71	1.64	1.34	1.21	0.95	0.47	0.36	0.77	1.8	0.79	1.4	0.88	0.57	0.58
	T13	✓	1.2	0.96	1.08	1.39	0.91	1.99	1.64	1.58	1.13	0.71	0.82	1.59	1.99	0.88	1.1	1.33	0.66	0.66
	T14	✓	0.95	0.6	0.61	0.65	0.65	1.95	1.46	1.4	1.13	0.54	0.47	0.78	1.97	0.79	1.14	0.9	0.57	0.56
	T15	✓	1.01	0.57	0.76	0.77	0.68	2.02	1.44	1.6	1.11	0.59	0.58	0.74	1.92	0.82	1.18	1.12	0.58	0.67
	T16	✓	1.4	1.02	1.13	1.38	1.08	2.25	2.13	1.48	1.51	0.79	0.86	1.06	2.44	1.31	1.13	1.74	1.52	0.94
	T17	✓	1.31	0.79	0.77	1.87	1.18	2.15	2.3	1.58	1.97	0.95	1.18	0.97	2	1.06	1.02	1.29	0.56	0.61
	T18	✓	1.02	0.78	0.88	0.79	1	2.2	1.91	1.2	1	0.55	0.53	0.73	2.04	0.91	0.97	0.82	0.5	0.5
	T19	✓	0.99	0.68	0.57	0.85	0.96	2.01	1.81	1.25	0.99	0.59	0.66	0.69	1.95	0.91	1.02	0.87	0.51	0.59
	T20	✓	1.06	0.63	0.58	0.64	1.21	2.12	1.62	1.39	1.13	0.59	0.61	0.73	1.94	0.88	1.08	1.64	0.61	0.67

[1]Foot and shoulder markers are unstable visually
[2]High noise in the middle of the trial mainly pelvis and hip markers
[3]Pelvis and hip markers are not tracked properly
[4]Pelvis and hip markers are unstable
[5]Pelvis markers are highly unstable
[6]C7 and pevlis markers are bad at different sections
[7]Pelvis and hip markers have very high noise throughout trial
[8]Knee and ankle markers are bad at different sections
[9]Pelvis and foot markers are unstable

C Appendix C

Table 3: Mean marker spatial error of one trial

Marker name	**Mean marker spacial error (m)**	**Marker name**	**Mean marker spacial error (m)**
C7	0.0241	R_TROC	0.0611
R_SAE	0.0286	L_TROC	0.0646
L_SAE	0.0323	R_FLE	0.0572
CLAV	0.0319	R_FME	0.0419
RFAradius	0.0175	L_FLE	0.0483
RFAulna	0.0211	L_FME	0.0271
RLEL	0.0356	R_TTC	0.0339
RMEL	0.0165	L_TTC	0.025
LFAradius	0.0206	R_FAL	0.0283
LFAulna	0.0264	R_TAM	0.0348
LLEL	0.0157	L_FAL	0.0299
LMEL	0.0268	L_TAM	0.0358
R_IAS	0.0336	R_FM2	0.0268
L_IAS	0.0298	R_FM5	0.0435
R_ICT	0.0434	L_FM2	0.0125
L_ICT	0.0516	L_FM5	0.0435

Table 4: Mean marker spatial error for each participant along with their shape-factor. ORI: original (MoSh++ initial guess); BMI: independent Body Mass Index (BMI) for STAR model instead of ORI

Participants	**Gender**	**Shape-factor**	**Mean marker spatial error (m)**
P1	M	BMI	0.02979
P2	F	BMI	0.03137
P3	M	ORI	0.03037
P4	M	ORI	0.03377
P5	F	ORI	0.03356

D Appendix D

Table 5: MAE of joint angles (lower-body) for two types of camera set-ups.
No 1,2 and 3 indicates no arm movement, one arm movement and both-arm movement

Camera set-up	No	pelvis_ tilt	pelvis_ list	pelvis_ rotation	hip_ flexion_r	hip_ adduction_r	hip_ rotation_r	knee_ angle_r	knee_angle _r_beta	ankle_ angle_r	hip_ flexion_l	hip_ adduction_l	hip_ rotation_l	knee_ angle_l	knee_angle _l_beta	ankle_ angle_l
Two side views	1	7.32	3.29	5.99	14.55	9.99	4.75	10.63	9.59	9.59	31.90	28.10	6.51	20.05	19.35	15.03
		28.57	3.73	12.52	24.62	6.59	8.25	19.72	20.70	9.38	8.57	15.71	5.15	8.42	7.80	24.28
	2	30.20	2.02	13.04	43.19	3.07	4.12	6.90	6.64	8.36	9.25	17.22	17.80	30.97	28.65	25.92
		34.52	12.07	18.09	26.76	3.80	13.52	16.91	17.36	6.15	29.17	17.51	2.38	13.41	14.23	14.91
	3	32.37	8.63	15.08	30.44	5.00	9.47	24.65	24.03	9.82	13.41	11.87	6.60	9.02	9.32	16.98
		45.40	5.69	23.76	69.95	11.19	5.85	17.40	17.17	14.90	17.93	21.31	15.04	18.23	18.44	13.74
Two oblique views	1	18.59	10.02	24.38	15.56	8.38	5.49	12.48	12.61	11.89	37.99	6.44	18.51	18.34	18.64	6.68
		40.02	13.01	22.93	21.31	13.79	11.39	17.57	17.29	13.01	41.18	13.47	11.17	12.07	12.34	11.26
	2	55.58	22.29	42.14	26.67	17.90	5.31	21.86	21.75	7.25	55.49	10.48	12.39	7.78	7.63	15.32
		30.43	15.50	7.02	31.81	4.98	10.01	15.05	14.80	13.14	21.13	6.74	10.65	20.08	20.79	11.09
	3	43.53	15.39	45.36	23.00	12.15	10.16	16.51	16.35	8.79	39.79	12.30	11.80	18.84	19.13	16.19
		45.83	16.76	42.25	21.87	14.95	14.94	18.07	17.37	10.39	42.92	11.90	8.95	21.81	21.44	19.58

Table 6: MAE of joint angles (upper body) for two types of camera set-ups.
No 1,2 and 3 indicates no arm movement, one arm movement and both-arm movement

Camera set-up	No	lumbar_ extension	lumbar_ bending	lumbar_ rotation	arm_ flex_r	arm_ add_r	arm_ rot_r	elbow_ flex_r	pro_ sup_r	wrist_ flex_r	wrist_ dev_r	arm_ flex_l	arm_ add_l	arm_ rot_l	'elbow_ flex_l'	'pro_ sup_l'	'wrist_ flex_l'	'wrist_ dev_l'
Two side views	1	7.40	5.86	4.66	76.60	4.93	14.97	19.18	24.55	6.04	29.63	86.16	18.17	3.39	20.00	7.42	18.77	28.73
		34.96	8.12	9.94	91.96	5.85	15.81	24.37	23.07	5.91	29.66	109.78	18.09	29.86	22.30	7.17	18.61	28.56
	2	36.30	4.15	17.33	94.48	27.61	45.65	10.93	25.21	0.42	0.29	88.87	22.05	12.77	52.90	49.00	0.29	1.40
		32.41	9.40	5.12	38.15	39.29	36.33	14.30	27.29	0.43	0.31	92.04	16.98	34.84	36.70	43.91	0.35	1.33
	3	31.40	5.30	7.09	92.98	45.31	51.19	9.88	26.55	0.45	0.26	23.97	3.52	57.78	82.61	40.15	0.38	1.06
		26.09	4.64	14.08	157.06	56.08	69.46	11.79	24.13	0.49	0.37	21.42	9.50	46.34	88.63	43.18	0.60	1.83
Two oblique views	1	7.59	2.71	6.80	72.55	8.88	10.53	7.41	20.66	6.02	29.40	84.90	12.93	9.48	14.01	4.01	18.72	28.59
		32.30	8.36	7.46	78.34	9.03	14.16	15.57	14.79	6.07	29.52	81.79	12.01	15.20	15.04	8.90	18.73	28.67
	2	35.44	5.86	10.46	9.59	14.49	23.93	51.34	31.98	0.47	0.33	63.92	10.63	12.85	37.24	54.90	0.34	1.28
		33.30	3.28	10.23	87.96	16.10	28.48	10.24	30.28	0.33	0.28	94.98	17.05	23.10	41.47	43.12	0.32	1.35
	3	19.60	7.63	9.69	24.52	17.42	21.08	60.07	26.88	0.33	0.26	45.61	10.77	22.36	17.85	48.73	0.33	1.46
		21.50	9.03	6.66	33.62	16.61	22.36	50.39	30.11	0.33	0.23	56.39	18.48	37.42	24.67	49.81	0.32	1.47