Heliyon 10 (2024) e32078

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Calibrationless monocular vision musculoskeletal simulation during gait

Ryo Ueno

Department of Research and Development, ORGO, 2-7 Odori W18, Chuo-ku, Sapporo, 061-1136, Japan

A R T I C L E I N F O

Keywords:
Computer vision
Pose estimation
Musculoskeletal simulation

A B S T R A C T

With computer vision technology and prediction of ground reaction forces (GRF), a previous study performed markerless motion capture and musculoskeletal simulation with two smartphones (OpenCap). A recent approach can reconstruct 3D human motion from a single video without calibration and it may further simplify the motion capture process. However it has not been combined with musculoskeletal simulation and the validity is unclear. Therefore, the purpose of this study was to determine the validity of the musculoskeletal simulation using a monocular vision approach. An open-source dataset that contains motion capture and video data during gait from 10 healthy participants was used. Human motion reconstruction with the skinned human (SMPL) model was performed on each video. Virtual marker data was generated by extracting the position data from the SMPL skin vertices. Inverse kinematics, GRF prediction (only for monocular vision approach), inverse dynamics and static optimization were performed using a musculoskeletal model for experimental motion capture data and the generated virtual markers from videos. Mean absolute errors (MAE) between motion capture based and monocular vision based simulation outcomes were calculated. The MAE were 8.4° for joint angles, 5.0 % bodyweight for GRF, 1.1 % bodyweight*height for joint moments and 0.11 for estimated muscle activations from 16 muscles. The entire MAE was larger but some were comparable to OpenCap. Using the monocular vision approach, motion capture and musculoskeletal simulation can be done with no preparations and is beneficial for clinicians to quantify the daily gait assessment.

1. Introduction

Human movement evaluation, especially gait analysis is a required process for rehabilitation. Kinematic and kinetic research have been reported to study the features of pathological gait in medicine. The kinematic data is commonly obtained with a marker-based motion capture system whereas the kinetic data is measured with force plates in a laboratory space [1]. In addition, neuromuscular assessment is performed using surface electromyography (EMG) sensors. A concern is that the equipment for those measurements are expensive and the measurements require time-consuming preparation that includes sensor placement and calibration. Because of this, most of the hospitals cannot operate motion captures and only the qualitative assessment is conducted without quantitative data for human movement evaluation.

After collecting the data, joint kinematics and kinetics can be calculated using a link segment model. Musculoskeletal models can further simulate neuromuscular activities without EMG sensors. Previous studies reported musculoskeletal modeling approach to evaluate various human movements [2–8]. However, this step requires time-consuming manual software operations and/or

E-mail address: ueno@orgo.co.jp.

https://doi.org/10.1016/j.heliyon.2024.e32078
Received 17 May 2024; Received in revised form 28 May 2024; Accepted 28 May 2024
Available online 31 May 2024

programming techniques to automate the process [9]. This is another barrier for clinical assessments.

To address this problem, a recent study reported a video-based markerless motion capture system using smartphones, called OpenCap [9]. OpenCap uses two to five smartphones to capture the participant's motion and dynamics. The authors automated most of the processes and allowed the users to complete the setup and data collection by 5 and 2 min respectively. After the data collection, users can obtain the kinematics data processed on a server without labeling or gap filling of markers. The kinematics data can be used to calculate further dynamics with predicted ground reaction forces (GRF). Although OpenCap decreased the set up time and total costs for the equipment, it could be decreased further if the motion capture can be done with only a single camera. The OpenCap starts from calibrating the camera extrinsic parameters by capturing a checkerboard image. This is necessary to triangulate the 2D keypoints data from the multiple cameras. If the camera calibration is not necessary and if motion capture is possible only with a single camera, users can start the data collection right away using a smartphone.

Recently, a lot of artificial intelligence (AI) pose estimation research have been reported and the accuracy is improving rapidly [10–13]. Monocular vision pose estimation is studied popularly so that an application with easier use can be developed [14–17]. Many of them still have limitations for biomechanical evaluations in global space. However, some studies reported a way to reconstruct 3D motions with root joint global translations and orientations [18,19] which is critical to evaluate gait biomechanics. Rempe et al. [20] reported a SMPL model based single-camera pose estimation method, HuMoR that takes into account the interaction between foot and ground. When the captured motion and foot-ground interaction are accurate enough, GRF can also be predicted using contact models in either inverse (e.g. CusToM) or forward dynamics based approach (e.g. OpenCap) [21–24]. Then musculoskeletal modeling techniques can further compute the muscle dynamics with those estimated motion and force data [5]. SMPL model is also useful to obtain any set of virtual markers from the skin vertices for the inverse kinematics in musculoskeletal simulation which is not supported by classical pose estimations that only provide joint positions. However, the prediction of GRF and muscle dynamics using the monocular vision pose estimation have not been reported and validated. If the prediction of GRF and muscle dynamics can be performed from monocular vision, this study provides a new insight that facilitate more convenient motion analysis system which can be used in clinical place or sports field. Therefore, the purpose of this study was to report the validity of the joint kinematics, GRF, joint moments and muscle activations predicted using the motion reconstructed from monocular vision approach and musculoskeletal simulation.

2. Methods

2.1. Experimental data and processing

The experimental dataset that was shared by the OpenCap project was used in this study [9]. The dataset included motion capture data and video data for 10 participants (sex: 6 female and 4 male; age = 27.7 ± 3.8 years; body mass = 69.2 ± 11.6 kg; height = 1.74 ± 0.12 m). In this study, only the natural gait trials were used as gait trial is the most common task for patient's motion analysis. The motion capture data contained marker, GRF and EMG data. Marker data included trajectories of 51 markers recorded with 8 cameras at 100 Hz (Motion Analysis Corp., Santa Rosa, CA, USA). GRF was measured with three force plates at 2000 Hz (Bertec Corp., Column-bus, OH, USA). EMG was measured at 2000 Hz using wireless electromyography electrodes (Delsys Corp., Natick, MA, USA)

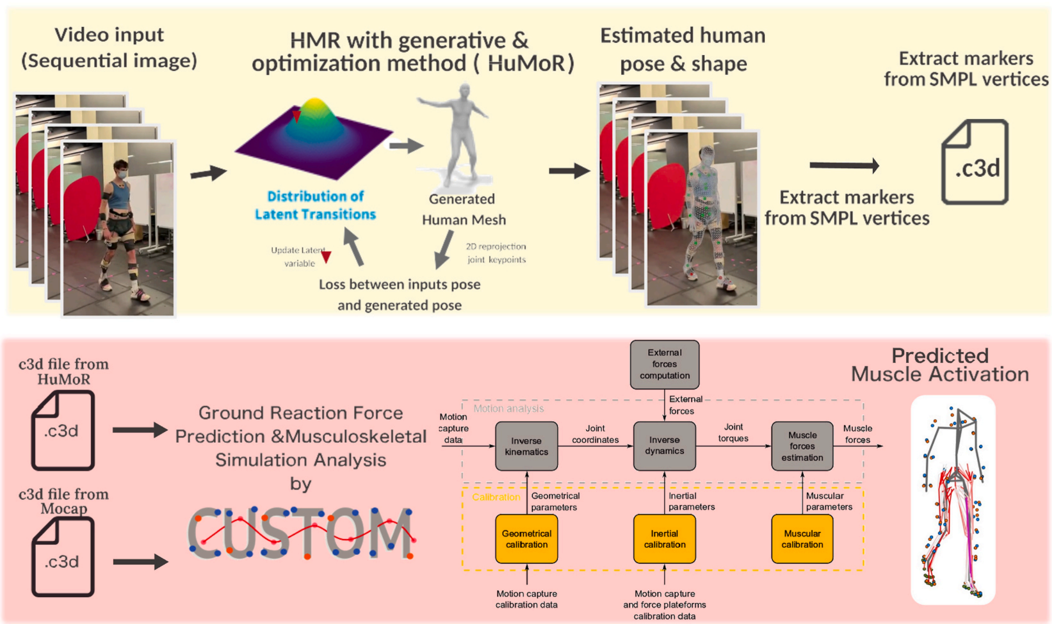

Fig. 1. A summary of the monocular vision musculoskeletal simulation pipeline.

from both side of soleus, gastrocnemius medialis, vastus medialis, vastus lateralis, semitendinosus, biceps femoris long head, gluteus medius, and left tibialis anterior, rectus femoris. Videos were recorded with five smartphones (iPhone 12 Pro, Apple Inc., Cupertino, CA, USA) from different perspectives (at $\pm 70^{\circ}$, $\pm 45^{\circ}$, and 0°, where 0° faces the participant).

Distributed TRC, MOT and STO files that stored the pre-processed marker trajectories, force plate data and EMG data, respectively, were used for this study. Marker and force plate data were low pass filtered at 6 Hz cut-off frequency. EMG data was band pass filtered at 30–500 Hz, rectified and low pass filtered at 6 Hz. EMG data was used only for the visualization in the figure of this study and normalized to the peak value of the estimated muscle activation using motion capture data. Marker data and force plate data were saved in C3D file format using custom code of ezc3d, a python library [25] to prepare for the musculoskeletal simulation step.

2.2. Monocular vision motion reconstruction

The pipeline of monocular vision musculoskeletal simulation is illustrated in Fig. 1. The motion reconstruction was performed separately on single videos taken via five different smartphones. A deep learning based approach, HuMoR [20], reported previously was used to reconstruct 3D pose from a single video. HuMoR was chosen because the reconstructed motion was in agreement with the ground without sliding or floating of the foot on the ground. This point was important to combine with the musculoskeletal simulation and to perform the estimation of ground reaction force. HuMoR learned a large number of human motion sequences derived from motion captures. When a RGB video is passed to HuMoR, this approach returns a sequence of 3D human poses. The human pose is based on the SMPL model, a skinned multi-person linear model [26]. The SMPL model consists of a 3D human skeleton and skin, represented with a generalized coordinate system and vertices on the skin surface. The SMPL model is parametrically scalable to fit most of the human shape. In the approach of HuMoR, it first estimates the 2D joint positions from each frame of the video using a 2D keypoint detector. Then HuMoR tries to find the best motion in which the SMPL joint positions (reprojected to 2D image coordinates) match the 2D joint positions out of the learned motions with an optimization. To reproject SMPL joint positions in 2D frames, the camera intrinsic parameters distributed with the dataset were used. Note that calibration of intrinsic parameter is necessary if unknown camera is used. Same intrinsic parameter can be reused if it is already known. Digital zoom or optical zoom should not be changed to use pre-calibrated intrinsic parameters. When HuMoR looks for the best motion, it predicts if there is a foot contact to the ground. In the optimization step, foot height and speed with respect to the predicted ground plane are minimized when the foot was in contact with high probability. Further details should be referred to the original study [20]. MMpose was used for 2D pose estimation in this study [27, preprint]. Distributed pre-trained deep learning model parameters were used for MMpose and HuMoR. No training of the deep learning was performed in this study.

Some technical arrangement was necessary to connect HuMoR data to CusToM musculoskeletal simulation. The output of HuMoR contains the sequence of joint positions, skin vertices and joint poses as well as the predicted ground plane and contact states on heels and toes. To perform musculoskeletal modeling using the HuMoR data, skin vertices were utilized to mimic the marker data. Vertices to replicate the OpenCap marker set were selected. The selection and verification of the marker was performed on 3D visualization tool, Mokka [28]. The sequence of selected markers were saved in C3D file format for the musculoskeletal simulation step as similar to the experimental data. Note that the joint positions and poses derived from HuMoR were not used in this study. The kinematics were recalculated using inverse kinematics with the musculoskeletal model since the degrees of freedom (DOF) were different from the SMPL model. Predicted ground plane data was used to transform the vertex data to the Z-up coordinate system. Contact states data was used to find when the foot was in contact, then subtract the offset between the ground and lowest vertex in Z coordinates. This was to ensure that the feet were close enough to the ground when it should be in contact for the external force prediction step.

2.3. Musculoskeletal modeling

CusToM, a MATLAB toolbox was used for the musculoskeletal simulations [29]. CusToM was chosen since it included a function to estimate external forces. Full body model that consists of 42 DOF, 82 lower extremity muscles and 14 contact points on each foot was used as a generic model. Subject models were calibrated by optimizing the segment length and marker placements of the generic model [30]. Then inverse kinematics, external force prediction, inverse dynamics and static optimization were performed. The external force prediction in CusToM was designed to perform the approach proposed by the previous study [23]. It first detected the contact points that were in contact state with the ground. When the height and speed of a contact point were below 0.05 m and 0.8 m/s, respectively, the contact point was considered to be in contact with the ground. Then the contact points for each foot in contact state were optimized by minimizing forces on the contact points while satisfying full body multibody dynamics equations. The maximum force for each contact point was limited to 40 % of the subject's body weight according to the pilot test of previous study [22]. The predicted GRF was used for the calculation of inverse dynamics. Inverse Dynamics derived the joint torques and residual forces and moments. Finally static optimization was performed to simulate muscle activations and forces. The sum of cubed muscle activations was minimized in static optimization as it was the default setting in CusToM.

2.4. Statistical analysis

Similarly to OpenCap [9], the validation of our approach was performed calculating mean absolute errors (MAE) of joint kinematics, GRF and joint moments. MAE for estimated muscle activations comparing between motion capture based and video based simulations was reported additionally. MAE calculation was performed after pooling all participants and trials data. GRF and joint moments were normalized to percent body weight (%BW) and percent body weight times body height (%BW*BH), respectively.

Following the OpenCap study, twenty-one DOF for joint kinematics, three force values for GRF and 15 DOF for joint moments were included in the analysis showing lower body results that were important for gait analysis. Analysis for muscle activations were performed on all 16 muscles collected experimentally. The statistical results were reported using each camera separately. Only the synchronized and cropped videos distributed in OpenCap dataset were used in the analysis. Since the shared dataset does not include full gait cycle, only the left stance phase was included in the analysis as same as the OpenCap study. Pelvis translations from both motion capture based and video based approach were zeroed at 0 % of left stance phase to match since the monocular vision approach was not designed to estimate in the motion capture space.

3. Results

The best MAE of 8.5° (range:3.7–21.6°) averaged across all joint angles were obtained using the videos captured from 45° right side (Fig. 2). Following results were derived using the same videos unless otherwise noted. Lumbar flexion-extension (16.8°) and both sides of subtalar supination-pronation (right: 21.6°, left: 15.7°) exceeded 10°. When these DOFs were excluded, the average MAE was 6.6°. The MAE averaged across three DOF pelvis translations was 0.020 m (range:0.007–0.034 m). The MAE averaged across all GRFs was 5.0 %BW (range:1.3–9.5 %BW) (Fig. 3). The averaged MAE for all joint moments was 1.1 %BW*BH (range: 0.22–2.7 %BW*BH) (Fig. 4). MAE averaged across all 16 muscles were 0.11 (range: 0.04–0.26) (Fig. 5). Results for specific DOFs and the different smartphones (i.e. different perspectives) were reported in the supplemental table. The reported results are from Cam1 in supplemental tables.

4. Discussion

The monocular vision motion capture approach without calibration step makes the motion capture process easier, faster and more user friendly compared to previous approaches. The purpose of this study was to determine the validity of joint kinematics, GRF, joint moments and muscle activations simulated using monocular human motion reconstruction. Although it had large errors in some parameters, many parameters were estimated with accuracies close to the previous approach (e.g. OpenCap [9]). Therefore, this study showed some potentiality that the monocular vision approach reported in this study could be one way to develop user-friendly motion capture systems for future applications. With the current procedure, it takes 20–30 min to process all the processing pipeline dependent on the computational power. Human motion assessment in rehabilitation would be possible when the practitioners are aware of the potential errors. Errors for specific variables and different camera perspectives were provided in Supplemental data. The specific errors

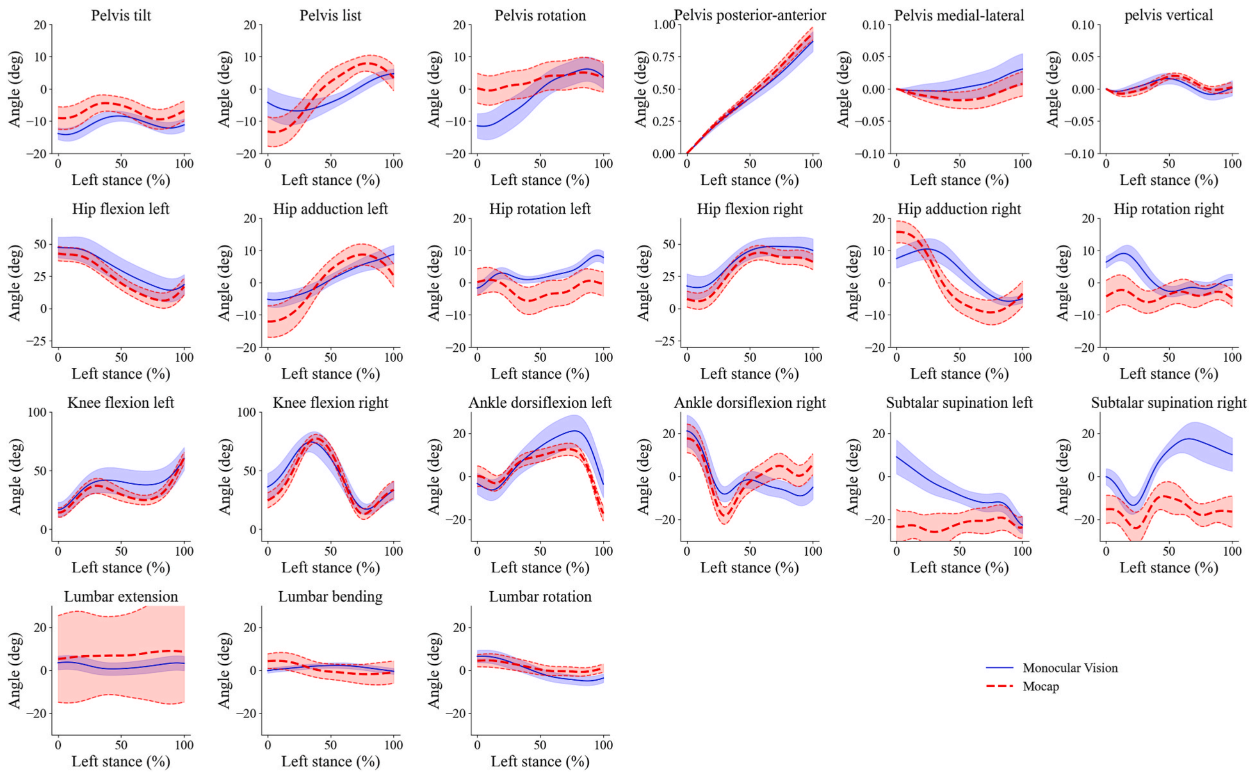

Fig. 2. Joint kinematics comparison. Joint kinematics calculated using experimental motion capture data (red) and monocular vision approach (blue).

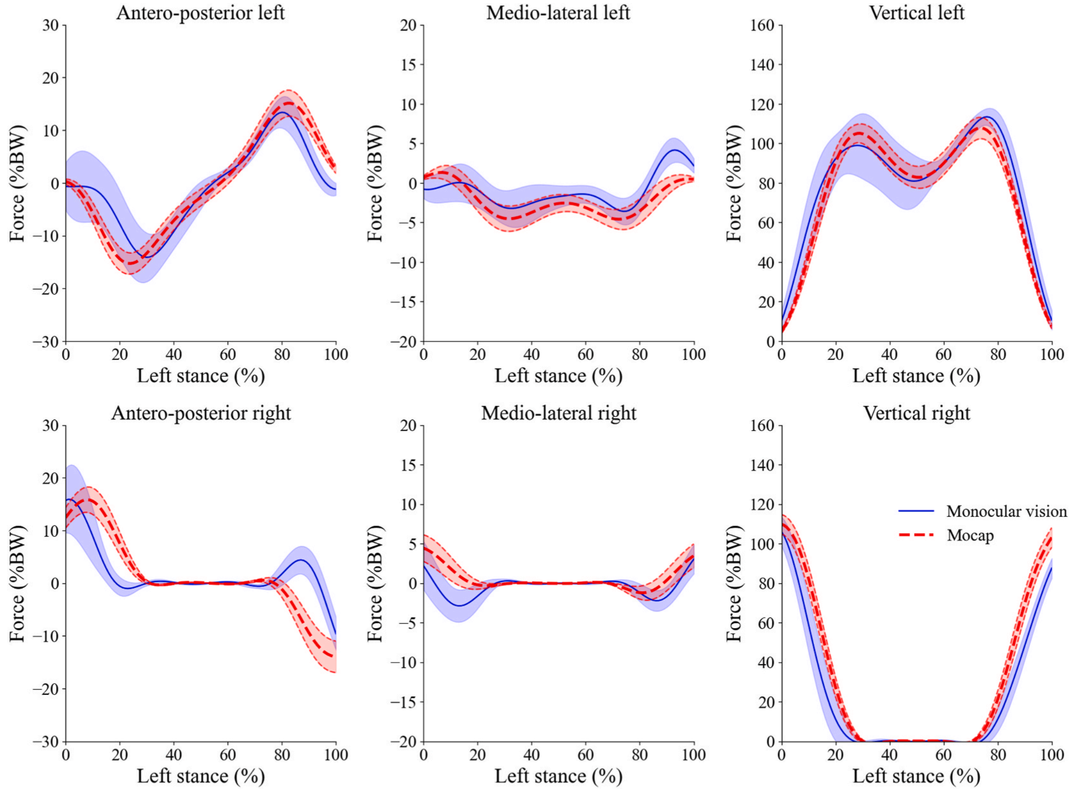

Fig. 3. Ground reaction forces comparison. Experimentally measured ground reaction forces (red) and predicted ground reaction forces using monocular vision approach (blue).

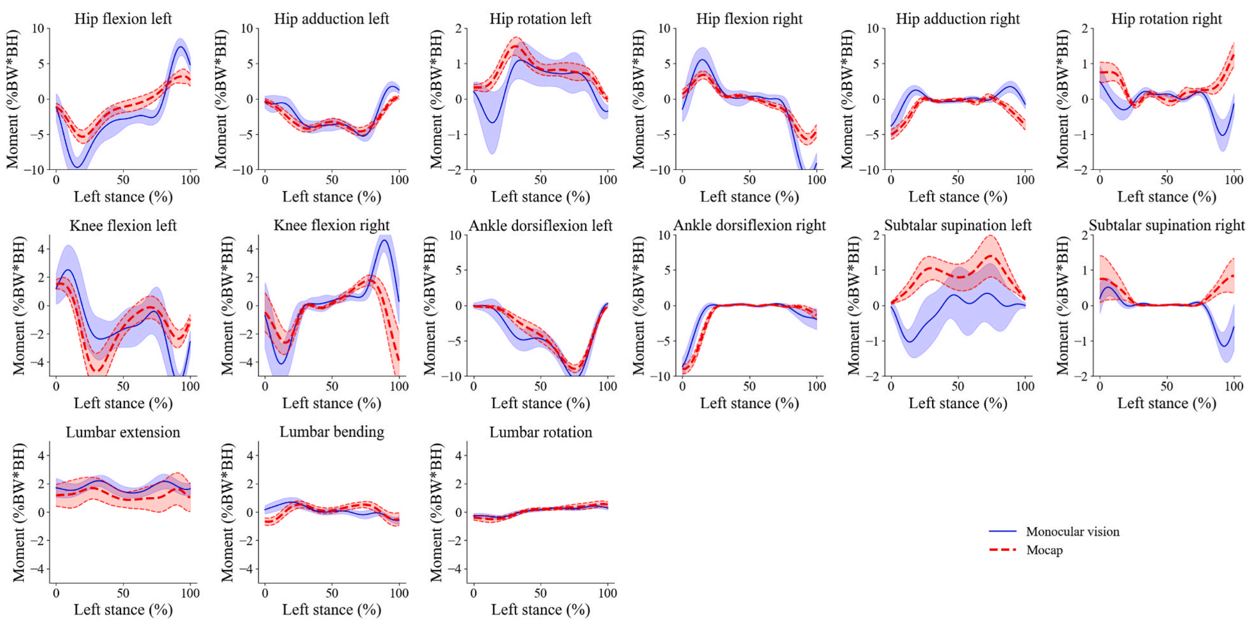

Fig. 4. Joint moments comparison.Joint moments calculated using experimentally measured motion capture data (red) and monocular vision approach (blue).

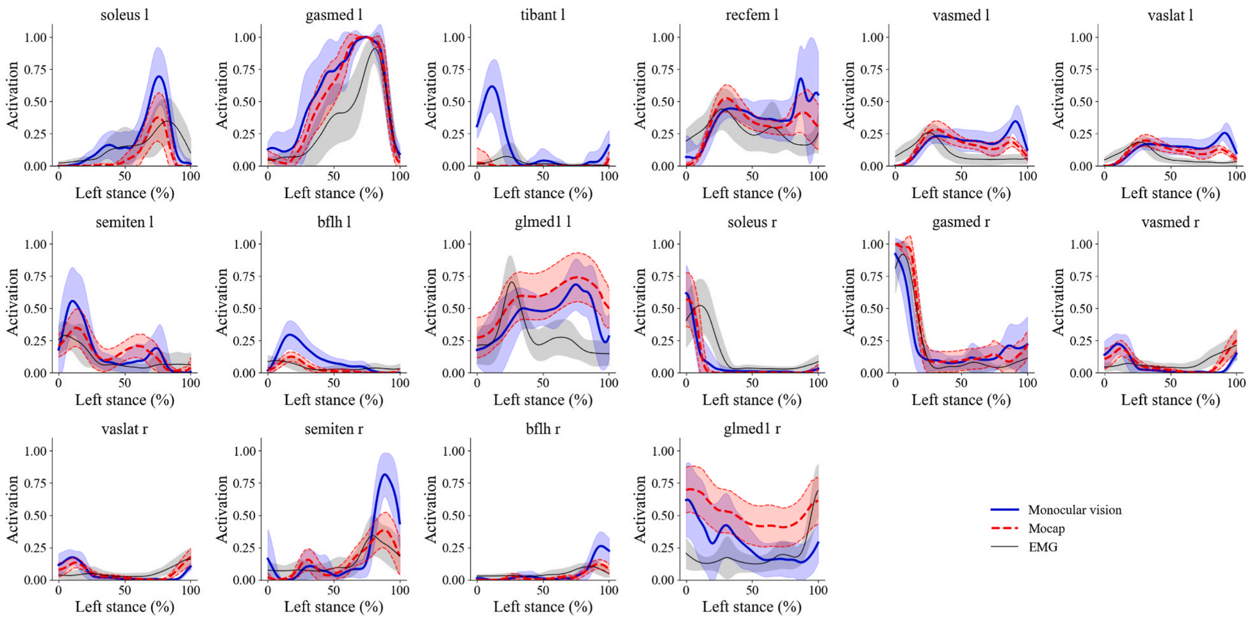

Fig. 5. Muscle activations comparison. Muscle activations estimated using experimentally measured motion capture data (red) and monocular vision approach (blue) compared to electromyography (EMG) data (brack). EMG was normalized to peak value of estimated muscle activation from motion capture based simulation in each trial.

should be referred to consider the future use of this study's approach based on the user's needs.

Calibration-less single smartphone motion capture was possible because the HuMoR was designed to robustly generate smooth and natural sequential motion that fit to an observation (i.e. 2D keypoints for video input). Conditional variational autoencoder (CVAE) was trained using sequential motion data of AMASS [31], a large dataset of SMPL motion. Therefore, the generated motion was natural even if only the limited and noisy input from a single smartphone was available.

The average MAE in joint angles was 8.5° while it was 4.1° in OpenCap. The larger MAE were found in the subtalar angles and the lumbar joints in this study. Since the other joints were estimated with better accuracy, improving these joint angles helps to achieve a better entire score. HuMoR fitted the learned motion to the 2D keypoints by reprojecting 3D joints to 2D space in an optimization process. However, 2D keypoints information was limited to determine pelvis-trunk relative rotations from one perspective. The 2D keypoints detection on the feet was qualitatively less accurate to reconstruct subtalar supination-pronation, especially when a blur exists on the foot. From a single perspective, it would be very difficult to detect sufficient 2D keypoints on the foot to determine the subtalar angle. OpenCap addressed these problems by augmenting the 3D marker positions using a long short-term memory (LSTM) model. This type of augmentation may help improve the accuracy. In addition, SMPL motion data was used to train HuMoR while CusToM was used to evaluate the kinematics. This modality difference might induce some errors.

With regards to the kinetics estimation, the MAE in GRF averaged across both sides of three DOF forces was 5.0 %BW while it was 3.8 %BW in OpenCap. The average MAE for joint moments were 1.1 %BW*BH and it was 0.75 %BW*BH in OpenCap. Although the MAE was larger compared to OpenCap, these results were encouraging that our approach could reasonably assess the kinetics outcomes during gait only with a single smartphone. Our approach to estimate kinetics outcomes was fully dependent on the kinematics quality. As described above, there remains room to improve joint kinematics estimation. The accuracy of kinetics outcomes should become better when the kinematics estimation is improved. It should also be noted that our approach of the prediction of kinetics was robust since it was an inverse dynamics based approach, and task-dependent user inputs are not required unlike the muscle-driven simulation as reported in OpenCap study.

MAE for estimated muscle activations was reported in this study which was not reported in OpenCap. The average MAE for sixteen muscles was 0.11 and the waveforms of estimated muscle activations from monocular vision approach were similar to that of optical motion capture based approach. The results indicated that the monocular vision approach can be used to estimate muscle activations to obtain similar outcomes that are based on the optical motion capture and force plate data. Similarly to kinetic outcomes, the accuracy becomes better with improved kinematics estimation.

Some limitations should be considered for the procedure used in this study. Tasks such as drop jump can not be analyzed using the current version of our approach because HuMoR was applicable only for the motions on the flat ground [20]. To apply for the motion on a step, a different approach should be used to reconstruct motions instead of HuMoR. Second, HuMoR was a dataset driven model in which the reconstruction of the motion was possible only if the motion was learned. Most likely it fails to reconstruct a reasonable motion for which is not included in the training dataset. This point should be solved by retraining the HuMoR with a specific motion dataset. Third, HuMoR is not flexible to different sampling rates of videos. HuMoR was trained with 30 fps so it sequentially generates a pose in 1/30 s future for each image frame of the video in an optimization. If the video is 120 fps, for example, motion progress in each frame is too small (one fourth) and not learned in HuMoR. To use it for high speed videos such as for sports, HuMoR should be retrained with the higher fps as same as the sampling rate of the high speed videos. With regard to the analysis of this study, the monocular vision approach was compared to optical motion capture approach. However, the results of optical motion capture approach also contained errors from the measurement. It is not possible to differentiate measurement errors and inaccuracies from the processing including many estimation steps. Comparing monocular vision approach to inertial sensor based approach was not performed and would also be interesting for the future study as inertial sensor can be used outside of the laboratory.

5. Conclusion

Validity of monocular vision musculoskeletal simulation was reported for gait trial in this study. Although some errors exist, the monocular vision approach can estimate joint kinematics, GRF, joint moments and muscle activation similarly to the optical motion capture system and OpenCap. Monocular vision musculoskeletal simulation makes the motion capture further user friendly by eliminating most of the preparation step. This will help clinicians to operate quantitative motion assessment in daily works.

CRediT authorship contribution statement

Ryo Ueno: Writing – original draft, Software, Investigation, Conceptualization.

Declaration of competing interest

Ryo Ueno is employee of ORGO Inc.

Acknowledgements

No external funding was used for this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e32078.

References

[1] A.D. Sylvester, S.G. Lautzenheiser, P.A. Kramer, A review of musculoskeletal modelling of human locomotion, Interface Focus 11 (2021) 20200060.
[2] S.L. Delp, J.P. Loan, A computational framework for simulating and analyzing human and animal movement, Comput. Sci. Eng. 2 (2000) 46–55.
[3] C.L. Dembia, N.A. Bianco, A. Falisse, J.L. Hicks, S.L. Delp, OpenSim Moco: musculoskeletal optimal control, PLoS Comput. Biol. 16 (2020) 1–21.
[4] H. Hatze, The complete optimization of a human motion, Math. Biosci. 28 (1976) 99–135.
[5] A. Falisse, G. Serrancolí, C.L. Dembia, J. Gillis, I. Jonkers, F. De Groote, Rapid predictive simulations with complex musculoskeletal models suggest that diverse healthy and pathological human gaits can emerge from similar control strategies, J. R. Soc. Interface 16 (2019), https://doi.org/10.1098/rsif.2019.0402.
[6] D. Heinrich, A.J. Van den Bogert, W. Nachbauer, Estimation of joint moments during Turning maneuvers in alpine skiing using a three dimensional musculoskeletal skier model and a forward dynamics optimization framework, Front. Bioeng. Biotechnol. 10 (2022) 894568.
[7] A. Navacchia, R. Ueno, K.R. Ford, C.A. DiCesare, G.D. Myer, T.E. Hewett, EMG-informed musculoskeletal modeling to estimate realistic knee anterior shear force during drop vertical jump in female athletes, Ann. Biomed. Eng. (2019), https://doi.org/10.1007/s10439-019-02318-w.
[8] T.K. Uchida, J.L. Hicks, C.L. Dembia, S.L. Delp, Stretching your energetic budget: how tendon compliance affects the metabolic cost of running, PLoS One 11 (2016) e0150378.
[9] S.D. Uhlrich, A. Falisse, Ł. Kidziński, J. Muccini, M. Ko, A.S. Chaudhari, J.L. Hicks, S.L. Delp, OpenCap: human movement dynamics from smartphone videos, PLoS Comput. Biol. 19 (2023) e1011462.
[10] A. Kanazawa, M.J. Black, D.W. Jacobs, J. Malik, End-to-end recovery of human shape and pose, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, 2018, https://doi.org/10.1109/cvpr.2018.00744.
[11] C. Malleson, J. Collomosse, A. Hilton, Real-time multi-person motion capture from multi-view video and IMUs, Int. J. Comput. Vis. 128 (2020) 1594–1611.
[12] D. Pavllo, C. Feichtenhofer, D. Grangier, M. Auli, 3D human pose estimation in video with temporal convolutions and semi-supervised training, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 7745–7754. IEEE.
[13] H. Rhodin, F. Meyer, J. Spörri, E. Muller, V. Constantin, P. Fua, I. Katircioglu, M. Salzmann, Learning monocular 3D human pose estimation from multi-view images, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018, pp. 8437–8446.
[14] Y. Feng, J. Yang, M. Pollefeys, M.J. Black, T. Bolkart, Capturing and animation of body and clothing from monocular video, in: SIGGRAPH Asia 2022 Conference Papers (SA '22 Conference Papers), December 6â•fi9, 2022, Daegu, Republic of Korea 1, 2022, https://doi.org/10.1145/3550469.3555423.
[15] S. Peng, J. Hu, 3D human pose estimation in video with temporal and spatial transformer, in: International Conference on Image, Signal Processing, and Pattern Recognition (ISPP 2023), SPIE, 2023, pp. 89–94.
[16] S. Shimada, V. Golyanik, W. Xu, P. Pérez, C. Theobalt, Neural monocular 3D human motion capture with physical awareness, ACM Trans. Graph. 40 (2021) 1–15.
[17] K. Zhou, X. Han, N. Jiang, K. Jia, J. Lu, HEMlets pose: learning part-centric heatmap triplets for accurate 3D human pose estimation, in: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, 2019, https://doi.org/10.1109/iccv.2019.00243.
[18] J. Li, S. Bian, C. Xu, G. Liu, G. Yu, C. Lu, D &D: learning human dynamics from dynamic camera, in: Computer Vision – ECCV 2022, Springer Nature Switzerland, 2022, pp. 479–496.
[19] Y. Yuan, U. Iqbal, P. Molchanov, K. Kitani, J. Kautz, GLAMR: global occlusion-aware human mesh recovery with dynamic cameras, in: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2022, https://doi.org/10.1109/cvpr52688.2022.01076.
[20] D. Rempe, T. Birdal, A. Hertzmann, J. Yang, S. Sridhar, L.J. Guibas, HuMoR: 3D human motion model for robust pose estimation, Proceedings of the IEEE International Conference on Computer Vision (2021) 11468–11479.

[21] E.J. Dijkstra, E.M. Gutierrez-Farewik, Computation of ground reaction force using Zero Moment Point, J. Biomech. 48 (2015) 3776–3781.

[22] R. Fluit, M.S. Andersen, S. Kolk, N. Verdonschot, H.F.J.M. Koopman, Prediction of ground reaction forces and moments during various activities of daily living, J. Biomech. 47 (2014) 2321–2329.

[23] A. Muller, C. Pontonnier, X. Robert-Lachaine, G. Dumont, A. Plamondon, Motion-based prediction of external forces and moments and back loading during manual material handling tasks, Appl. Ergon. 82 (2020) 102935.

[24] S. Skals, M.K. Jung, M. Damsgaard, M.S. Andersen, Prediction of ground reaction forces and moments during sports-related movements, Multibody Syst. Dyn. 39 (2017) 175–195.

[25] B. Michaud, M. Begon, ezc3d: an easy C3D file I/O cross-platform solution for C++, Python and MATLAB, J. Open Source Softw. 6 (2021) 2911.

[26] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, M.J. Black, SMPL: a skinned multi-person linear model, ACM Trans. Graph. 34 (2015) 1–16.

[27] T. Jiang, P. Lu, L. Zhang, N. Ma, R. Han, C. Lyu, Y. Li, K. Chen, RTMPose: real-time multi-person pose estimation based on MMPose, ArXiv [Cs.CV] (2023). http://arxiv.org/abs/2303.07399 [preprint].

[28] A. Barre, S. Armand, Biomechanical ToolKit: open-source framework to visualize and process biomechanical data, Comput. Methods Progr. Biomed. 114 (2014) 80–87.

[29] A. Muller, C. Pontonnier, P. Puchaud, G. Dumont, CusToM: a Matlab toolbox for musculoskeletal simulation, J. Open Source Softw. 4 (2019) 927.

[30] M.S. Andersen, M. Damsgaard, B. MacWilliams, J. Rasmussen, A computationally efficient optimisation-based method for parameter identification of kinematically determinate and over-determinate biomechanical systems, Comput. Methods Biomech. Biomed. Eng. 13 (2010) 171–183.

[31] N. Mahmood, N. Ghorbani, N.F. Troje, G. Pons-Moll, M. Black, AMASS: archive of motion capture as surface shapes, in: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, 2019, https://doi.org/10.1109/iccv.2019.00554.