GaitQuest: Markerless Gait Analysis from an Observer-Worn Head-Mounted Display

Silvia Zaccardi[1,2,3], Giulio Rigoni[5], David Beckwée[2,4], Eva Swinnen[2], Bart Jansen[1,3]⋆, and Zimi Sawacha[5]⋆

[1] ETRO, Vrije Universiteit Brussel, Brussels, Belgium
[2] RERE, Vrije Universiteit Brussel, Brussels, Belgium
[3] IMEC, Leuven, Belgium
[4] MOVANT, University of Antwerp, Antwerp, Belgium
[5] DEI, University of Padova, Padova, Italy

Abstract. This work presents GaitQuest, a portable markerless motion capture system. Gait analysis is performed from monocular video captured solely by an RGB camera of a Meta Quest 3 headset, worn by an operator walking freely alongside the participant. To obtain 3D human pose estimates from a moving camera, we introduce CameraHMR$_{\mathrm{cal}}$, an adaptation of CameraHMR that exploits known camera intrinsics and headset odometry to express body pose in a fixed world frame. We evaluated the system on eight healthy participants across posterior and frontal operator viewpoints, with each trial recorded simultaneously with a Vicon reference system. Accuracy was assessed for eight spatiotemporal gait parameters and six lower-limb joint angles. CameraHMR$_{\mathrm{cal}}$ improves spatiotemporal accuracy (best results in posterior view: MAEs $\leq$ 15 ms for all temporal parameters, $\leq$ 2.7 cm for all spatial parameters, and 0.01 m/s for walking speed). For joint kinematics, hip and knee flexion exhibited high waveform agreement with Vicon in both posterior and frontal views ($r > 0.90$), although all joint-angle MAEs exceeded 5°.

Keywords: Markerless motion capture · Gait analysis · Head-mounted display · Human mesh recovery · Egocentric video

1 Introduction

Gait analysis plays an important role in clinical assessment by providing objective movement parameters for diagnosis, treatment planning, and rehabilitation monitoring [1]. Marker-based Motion Capture (MoCap) represents the established gold standard for quantitative gait analysis. However, while accurate, it remains confined to specialised laboratories as it depends on expensive, dedicated infrastructures and highly trained personnel [8,19]. Markerless video-based MoCap eases some of this burden [7,9,21], but often still relies on environment-mounted cameras, imposing fixed acquisition setups that limit use where space, setup time, or camera placement are constrained.

⋆ These authors contributed equally.

Consumer Mixed Reality (MR) head-mounted displays (HMDs) offer an alternative acquisition paradigm [3, 23]. Rather than arranging cameras in the environment, an observer wearing the headset can follow the participant and record gait from an observer-egocentric viewpoint. In this way, the capture device moves with the clinician, requires no environmental camera setup, and allows the viewpoint to be adapted naturally to the task being observed. Unlike common handheld cameras [15], MR HMDs can provide onboard visual-inertial tracking, estimating the camera trajectory over time in a fixed world coordinate frame.

However, headset odometry does not remove the image-level difficulties introduced by a moving observer, including changing viewpoint, scale, perspective, and self-occlusion. Human Mesh Recovery (HMR) methods, which reconstruct a full parametric body model from monocular RGB images, have demonstrated increasing ability to handle these challenges but, so far, they have been mainly validated in static camera setups [7, 20]. To the best of our knowledge, only Cantaloube et al. [3] have demonstrated observer-egocentric markerless gait analysis from smart glasses; however, the validation of their methods is limited to spatiotemporal gait parameter against an IMU-based reference system.

This work presents **GaitQuest**, a portable markerless system that uses the right RGB camera of a Meta Quest 3 headset, worn by an operator walking with the participant, as the sole video source for gait analysis. To reduce scale ambiguity inherent in monocular MoCap, we introduce CameraHMR$_{\text{cal}}$, an adaptation of CameraHMR [14] that replaces estimated camera parameters with known values. Although GaitQuest exploits both camera intrinsics and the HMD 6DoF pose, it does not require any calibration step, markers, or environment preparation, supporting accessible gait analysis outside the laboratory. We assess GaitQuest against marker-based MoCap for both spatiotemporal gait parameters and lower-limb joint kinematics, under realistic observer-egocentric acquisition with natural variation in viewpoint and occlusion.

2 Methods

2.1 Overview

GaitQuest is a modular pipeline for markerless gait analysis from an observer-worn Meta Quest 3 headset (Fig. 1): a Unity application records the right RGB camera stream of the headset; our proposed HMR method, **CameraHMR$_{\text{cal}}$**, estimates a per-frame SMPL body mesh from the RGB sequence; the reconstructed meshes are then converted into virtual anatomical marker trajectories, representing the input for biomechanical analysis. Resulting spatiotemporal and joint-kinematic outcomes were evaluated against a synchronised Vicon marker-based optical motion-capture system, used as the reference standard.

2.2 Data recording app

Data acquisition is performed by a custom Unity application for Meta Quest 3, adapted from OpenQuestCapture [16]. The app records the right-camera video

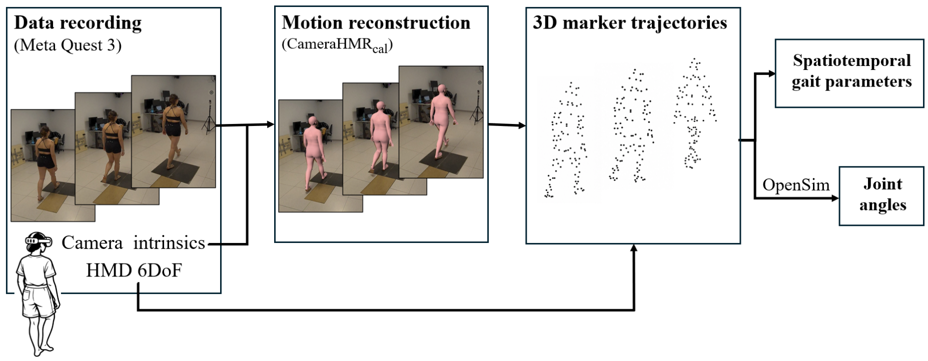

Fig. 1: Overview of the GaitQuest pipeline. A moving observer wearing a Meta Quest 3 records monocular RGB frames, camera intrinsics, and headset 6DoF pose while the participant walks. Our proposed CameraHMR$_{\text{cal}}$ reconstructs the participant's 3D body mesh, from which 3D marker trajectories are extracted in a fixed world coordinate system. These trajectories are used directly to compute spatiotemporal gait parameters and as input to OpenSim to compute joint angles.

stream at approximately 60 fps and logs the calibration and pose metadata required for downstream reconstruction, including the device-calibrated camera intrinsics, the fixed camera-to-headset extrinsic transform, and the per-frame 6DoF headset pose. All data are timestamped in a common Unix-millisecond time base.

2.3 Motion Reconstruction

We introduce **CameraHMR$_{\text{cal}}$**, a calibrated adaptation of CameraHMR [14] for observer-worn headset video. CameraHMR reconstructs SMPL pose parameters θ, shape parameters β, and global translation $\mathbf{t}$ in the camera frame from each RGB image. In its default implementation, the focal length is estimated per frame by an internal focal-length network (FLNet), while the principal point is fixed at the image centre. CameraHMR$_{\text{cal}}$ instead supplies the device-calibrated Meta Quest 3 intrinsics directly to the model, using the calibrated values of f_x, f_y, c_x, and c_y and bypassing FLNet.

2.4 3D Marker Trajectories

The per-frame human body meshes estimated by the HMR methods are converted into virtual anatomical marker trajectories. For each frame, marker positions are obtained from the SMPL mesh using the fixed vertex-to-marker correspondence defined by the Biomechanical Skeleton Model (BSM/SKEL) [10]. The resulting marker trajectories are written to a TRC file in the camera frame, which serves as the common input format for the subsequent coordinate transformation, gait-event detection, and inverse-kinematics steps.

World-Frame Transformation TRC marker positions, initially in the camera frame, are transformed to the world frame. The camera-to-world transform ($\mathbf{T}_{\mathrm{cam}}^{\mathrm{world}}(t)$) is obtained by composing the device-calibrated camera-to-headset extrinsic transform, $\mathbf{T}_{\mathrm{cam}}^{\mathrm{HMD}}$, with the per-frame headset pose, $\mathbf{T}_{\mathrm{HMD}}^{\mathrm{world}}(t)$:

$$\mathbf{T}_{\mathrm{cam}}^{\mathrm{world}}(t) = \mathbf{T}_{\mathrm{HMD}}^{\mathrm{world}}(t)\ \mathbf{T}_{\mathrm{cam}}^{\mathrm{HMD}} \tag{1}$$

PCA Alignment This step defines a common global coordinate system for all trials, compliant with OpenSim [17] conventions. After world-frame transformation, the trajectories shared a fixed vertical axis ($+Y$), but their horizontal heading remained arbitrary. The walking direction was therefore estimated by applying Principal Component Analysis (PCA) to the pelvis-centre trajectory projected onto the horizontal plane. The first principal component defined the forward axis ($+X$), signed according to the net trial displacement, while the lateral axis ($+Z$) was computed as $+X \times +Y$ and oriented toward the participant's right side.

2.5 Gait Parameters Estimation

World-frame marker trajectories (low-pass filtered, 6 Hz), represent the input for both spatiotemporal gait parameters (SGP) and lower-limb joint angles estimation. These outcomes are the gait parameters used to assess GaitQuest against the Vicon reference system.

Spatiotemporal Gait Parameters. We detect heel-strike and toe-off events using the kinematic method of Zeni et al. [24], and derive eight spatiotemporal parameters:

- **Stride length**: anterior–posterior displacement of the heel marker between two consecutive heel strikes of the same limb.
- **Stride time**: time interval between two consecutive heel strikes of the same limb.
- **Step length**: anterior–posterior distance between the heel-marker position of one limb at its heel strike and the heel-marker position of the opposite limb at its preceding heel strike.
- **Step width**: mediolateral distance between the two heel markers at consecutive left and right heel strikes, computed relative to the pelvis position.
- **Stance time**: time interval between heel strike and the subsequent toe-off of the same limb.
- **Swing time**: time interval between toe-off and the subsequent heel strike of the same limb.
- **Double-support time**: time interval between heel strike of one limb and the subsequent toe-off of the opposite limb.
- **Walking speed**: anterior–posterior pelvis displacement divided by the duration of the analysed walking interval.

Joint Angle Estimation Lower-limb joint angles were estimated in OpenSim using the inverse kinematics (IK) tool. Vicon and GaitQuest data were processed with the same generic musculoskeletal model distributed with Pose2Sim [13], with identical joint definitions but system-specific marker sets. The model was scaled to each participant using the static trial and lower-limb markers only.

For Vicon, IK was performed using the anatomical markers placed according to the IORGait protocol [11]. For GaitQuest, IK was performed using BSM/SKEL marker trajectories. Since these markers do not exactly match the IORGait landmarks, participant-specific marker offsets were computed from the GaitQuest static trial and used to define additional OpenSim tracking markers for IK. No Vicon measurements were used at any stage to estimate the GaitQuest joint angles; Vicon data served only as the reference. In addition, GaitQuest pelvic markers were re-referenced to the participant's static neutral posture to reduce residual sagittal pelvic-tilt bias.

The following joint angles are extracted as time-normalised waveforms over the gait cycle: hip flexion/extension, hip abduction/adduction, hip internal/external rotation, knee flexion/extension, ankle dorsiflexion/plantarflexion, and pelvic tilt.

3 Experiments

3.1 Experimental Setup and Data

Laboratory and Instrumentation Data collection was performed at the BioMovLab gait analysis laboratory, Department of Information Engineering, University of Padova, Italy. A Vicon MoCap system (Vicon Motion Systems, Oxford, UK) operating at 250 Hz served as the reference standard. Retroreflective skin markers were placed according to the IORGait protocol [11], which defines anatomical reference frames for the pelvis, thigh, shank, and foot using 26 anatomical landmarks.

Participants Eight healthy adult participants were recruited for this study (6 females, 2 males; Age(years): 25.3 ± 2.2; weight(Kg): 63.10 ± 14.32; height(m): 1.72 ± 0.097). All participants provided written informed consent prior to data collection (the protocol was approved by the local Ethical Committee, Protocol n° 3493/AO/15).

Protocol Participants walked barefoot at their self-selected speed while an observer, wearing the headset, captured them under two viewpoints (Fig. 2):

- Posterior. The observer followed the participant from behind and to the left, in a manner analogous to a clinician accompanying a patient during a corridor walk; since the operator always walked on the participant's left, the right leg was consistently the more occluded limb.

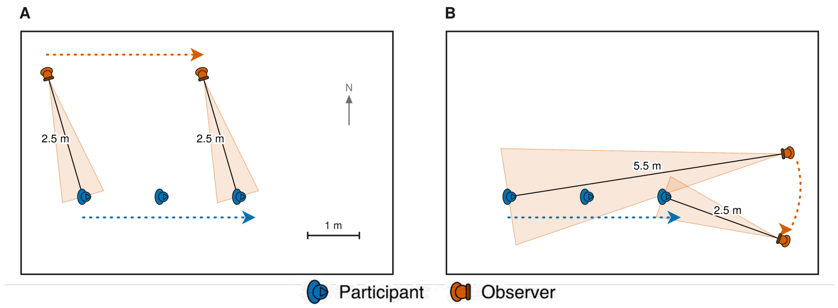

Fig. 2: Top-down schematic of the two operator viewpoints, shown at trial start and end. A: **Posterior** the observer follows behind and to the side of the participant B: **Frontal** the observer moves perpendicular to the participant's path while facing them.

– Frontal. The observer moved perpendicular to the participant's path while facing them.

For each viewpoint, we collected 3 trials comprising 2 complete gait cycles of the left leg (and one of the right), and 3 trials comprising 2 complete gait cycles of the right leg (and one of the left), giving 12 trials per participant. The Vicon system captured all trials simultaneously.

Because the operator moved freely rather than from a fixed position, the field of view and capture angle relative to the participant varied naturally from trial to trial. The realised operator-participant distance varied across trials (posterior: $2.3{\pm}0.2$ m at trial start, $2.5{\pm}0.3$ m at trial end; frontal: $5.2{\pm}0.2$ m to $2.6{\pm}0.2$ m).

3.2 Performance Evaluation

We assess $\mathrm{CameraHMR_{cal}}$ against the default CameraHMR from posterior (post) and frontal (frt) viewpoints, yielding four conditions: post, post-cal, frt, and frt-cal, where *-cal* denotes $\mathrm{CameraHMR_{cal}}$. All conditions are compared with the Vicon reference for spatiotemporal parameters and lower-limb joint angles using mean absolute error (MAE) and Pearson correlation coefficient r. To assess performance under partial occlusion in the post-cal condition, right–left differences in per-participant MAE were calculated for each parameter and reported with 95% bootstrap confidence intervals (CI) based on 20,000 participant-level resamples.

4 Results

4.1 Spatiotemporal Gait Parameters

Table 1 reports Mean Absolute Error (MAE) for eight spatiotemporal parameters under all four conditions (mean ± SD across participants). Fig. 3 shows the

Table 1: MAE (mean ± SD across participants) for eight spatiotemporal parameters, all four GaitQuest conditions vs. Vicon. Bold indicates the lowest MAE per row.

Parameter	post	post-cal	frt	frt-cal
Walking speed ($\mathrm{m\,s^{-1}}$)	0.04 ± 0.02	$\mathbf{0.01 \pm 0.00}$	0.17 ± 0.05	0.05 ± 0.03
Stride length (cm)	13.5 ± 1.9	$\mathbf{1.5 \pm 0.3}$	40.9 ± 16.6	8.6 ± 3.8
Step length (cm)	12.7 ± 1.2	$\mathbf{2.3 \pm 0.6}$	33.5 ± 7.0	6.5 ± 1.6
Step width (cm)	2.6 ± 1.4	2.7 ± 1.5	$\mathbf{1.6 \pm 0.7}$	$\mathbf{1.6 \pm 0.8}$
Stride time (ms)	10 ± 1	$\mathbf{9 \pm 2}$	36 ± 9	37 ± 8
Stance time (ms)	16 ± 4	$\mathbf{15 \pm 3}$	52 ± 13	55 ± 14
Swing time (ms)	$\mathbf{14 \pm 4}$	$\mathbf{14 \pm 4}$	35 ± 15	37 ± 15
Double-support time (ms)	$\mathbf{15 \pm 4}$	$\mathbf{15 \pm 4}$	46 ± 14	48 ± 14

Table 2: MAE (mean across participants) for six spatiotemporal parameters in the post-cal condition for the left and right sides. Right–left differences are reported with 95% bootstrap CIs.

Parameter	Left	Right	Δ (right − left) [95% CI]
Stride length (cm)	1.42	1.67	+0.25 [−0.09, +0.59]
Step length (cm)	2.26	2.36	+0.10 [−0.23, +0.44]
Step width (cm)	2.84	2.65	−0.18 [−0.27, −0.08]
Stride time (ms)	9.0	9.5	+0.5 [−1.6, +3.0]
Stance time (ms)	16.2	13.0	−3.1 [−5.6, +0.4]
Swing time (ms)	14.4	13.7	−0.7 [−2.9, +1.7]

corresponding GaitQuest vs. Vicon scatter plots for walking speed, step length, step width, and stance time under the post-cal condition, with Pearson correlation coefficient r computed separately for each limb and reported in the panel legends.

Table 2 reports the right–left difference in MAE for the six limb-specific spatiotemporal parameters in the post-cal condition, with positive values indicating greater error for the more frequently occluded right limb.

4.2 Lower-Limb Joint Angles

Table 3 reports MAE and Pearson r (mean ± SD across participants) for each joint angle and condition. Fig. 4 shows the corresponding mean ± SD waveforms for the post-cal condition, with per-limb Pearson r reported in the panel legends.

5 Discussion

CameraHMR$_{\mathbf{cal}}$ yields accurate spatiotemporal parameters. CameraHMR$_{\mathrm{cal}}$ consistently reduced spatiotemporal errors compared with the default CameraHMR (Table 1). The posterior viewpoint yielded the most accurate spatiotemporal estimates, with temporal MAE below 20 ms and spatial MAE below 3 cm.

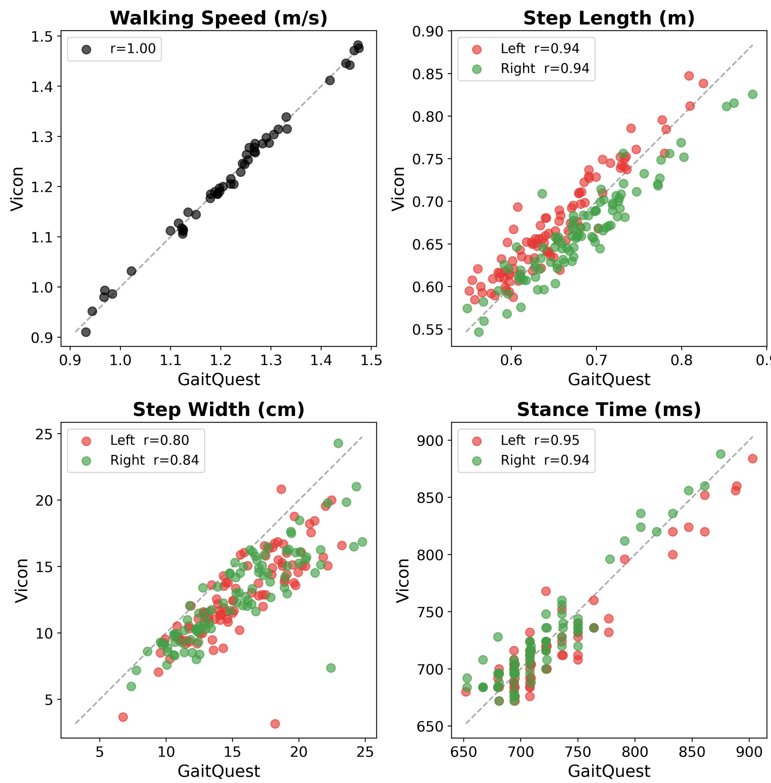

Fig. 3: GaitQuest vs. Vicon scatter plots (post-cal) for walking speed, step length, step width, and stance time, coloured by limb (left: red; right: green, the occluded limb in this viewpoint). Pearson r per limb (pooled, or overall for walking speed) is shown in each panel's legend.

The largest gains were observed for spatial parameters, while temporal errors improved only marginally. This pattern is consistent with the role of calibrated intrinsics improving 3D spatial reconstruction.

Viewpoint affects spatiotemporal accuracy. Frontal-view errors were approximately 2–4$\times$ larger for most parameters (Table 1). This degradation is likely related to the larger and more variable operator–participant distance in the frontal condition ($\sim$5.2–2.6 m within each trial), compared with the more stable posterior condition ($\sim$2.3–2.5 m). Moreover, the posterior setup was not a purely rear view but closer to a sagittal/oblique view, making displacements along the walking direction easier to observe. Step width was the main exception, with lower MAE in the frontal condition (1.6 cm) than in the posterior condition (2.6–2.7 cm), consistent with frontal views better resolving mediolateral foot separation.

Table 3: MAE (°) and Pearson r (mean ± SD across participants) for lower-limb joint angles, all four GaitQuest conditions vs. Vicon. Bold indicates the lowest MAE and the highest r per row.

	post		post-cal		frt		frt-cal	
Joint angle	MAE	r	MAE	r	MAE	r	MAE	r
Hip flex/ext	**5.4±2.7**	**0.97±0.03**	5.6±2.7	**0.97±0.03**	7.0±2.5	0.92±0.07	7.0±2.5	0.92±0.07
Hip abd/add	5.3±0.9	0.65±0.28	5.1±1.0	0.65±0.27	5.0±1.1	0.66±0.35	**4.9±1.1**	**0.68±0.34**
Hip int/ext rot	8.6±1.8	0.11±0.23	**8.5±1.9**	0.11±0.25	8.9±2.2	**0.16±0.25**	9.0±2.3	0.14±0.28
Knee flex/ext	7.7±2.5	**0.97±0.02**	**7.6±2.6**	**0.97±0.02**	8.6±1.9	0.90±0.05	8.4±1.9	0.90±0.05
Ankle dorsi/plantar	11.4±1.9	0.45±0.08	11.6±2.0	0.45±0.08	**7.4±1.7**	**0.55±0.09**	7.5±1.7	0.54±0.09
Pelvis tilt	3.0±1.4	0.18±0.31	2.9±1.2	0.16±0.31	**2.1±0.9**	0.24±0.13	2.2±1.0	**0.26±0.13**

CameraHMR$_{\mathbf{cal}}$ is robust to partial occlusion. In the posterior viewpoint, the operator walked on the participant's left, making the right leg the more frequently occluded limb. Despite this asymmetry, Fig. 3 shows comparable agreement for both legs, and Table 2 shows small right–left differences in MAE across all spatiotemporal parameters, with no consistent increase in error for the more frequently occluded limb. These results suggest that CameraHMR$_{\text{cal}}$ is robust to the partial self-occlusions induced by an observer-worn viewpoint.

Joint-angle magnitude errors exceed 5°. Joint-angle MAE was generally above 5° (Table 3), exceeding the threshold commonly regarded as clinically acceptable [12], but in line with errors reported for markerless HMR-based pipelines with static cameras [7, 20]. Differently from the spatiotemporal results, CameraHMR and CameraHMR$_{\text{cal}}$ yielded very similar joint-angle errors. This is expected: joint angles are computed from the relative configuration of body segments and are therefore largely insensitive to a global scale error. Joint-angle errors were also broadly comparable between the two viewpoints. It is nonetheless worth noting that Pearson correlations were higher for the posterior view for hip and knee flexion ($r = 0.97$ vs. 0.90–0.92).

5.1 Limitations and future work

This study has several limitations. First, the evaluation was conducted on a small cohort of healthy young participants ($N = 8$), and performance on older adults and individuals with pathological gait remains to be established. The Zeni et al. [24] method assumes that peaks in heel and toe motion relative to the pelvis correspond to initial contact and toe-off; this assumption may not hold for pathological gait patterns with atypical foot contact or highly irregular stepping [2]. This is particularly relevant in Parkinsonian gait, where forefoot-first contacts and highly irregular stepping patterns, including freezing of gait, can occur [5,6]. In addition, GaitQuest and Vicon were not temporally synchronised through a shared hardware trigger; instead, the same gait-event detection algorithm was applied independently to each modality to derive spatiotemporal gait parameters. Future validation in clinical populations should therefore employ gait-event

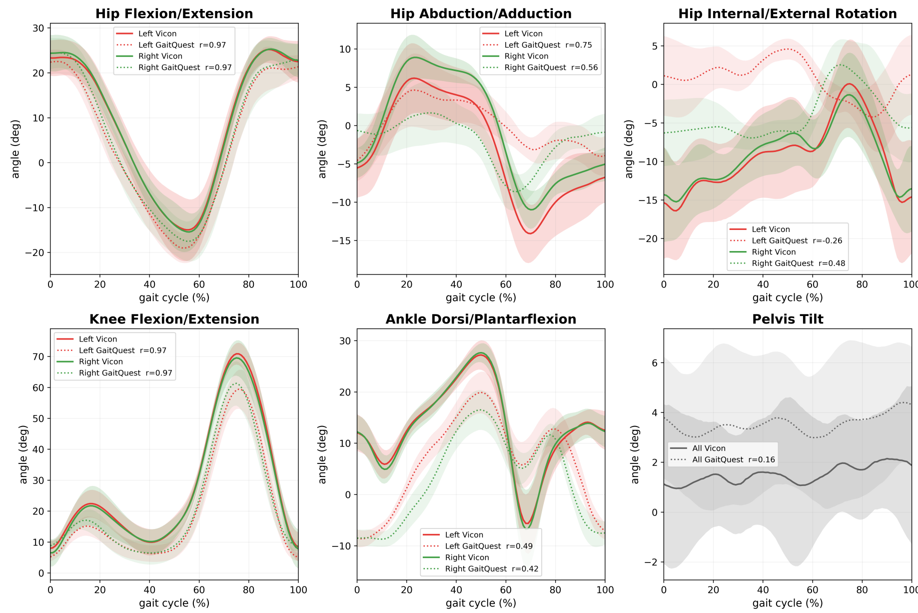

Fig. 4: Mean (line) ± SD (shaded band) joint-angle waveforms over the normalised gait cycle for Vicon (solid) and GaitQuest (dotted) in the post-cal condition, coloured by limb (left: red; right: green; pelvis tilt pooled across limbs, grey). Pearson r per limb is shown in each panel's legend.

detection methods validated for the target pathology, or manually labelled gait events, and may additionally include direct temporal synchronisation between systems.

Second, the pipeline operates offline and therefore does not yet support real-time gait analysis. Future work should investigate whether combining known camera intrinsics and headset odometry with alternative HMR methods, such as WHAM [18], Multi-HMR 2 [4], and SAM 3D Body [22], could further improve reconstruction accuracy under observer motion.

6 Conclusion

We introduce GaitQuest, a portable markerless MoCap system for quantitative gait analysis from an observer-worn RGB camera. The proposed methods exploit both camera known intrinsics and headset odometry, enabling accurate spatio-temporal gait parameters estimation. Joint-angle estimation remains more challenging: while waveforms are preserved, absolute angular errors are still too large for reliable lower-limb kinematic analysis. These results position GaitQuest as a benchmark and baseline for studying observer-worn, calibration-aware human motion recovery in realistic walking scenarios.

Acknowledgements

S.Z. is funded from Research Foundation Flanders (FWO) with project number FWOSB139.

References

1. Baker, R., Esquenazi, A., Benedetti, M.G., Desloovere, K.: Gait analysis: clinical facts. Eur J Phys Rehabil Med. **52**(4), 560–574 (2016)
2. Bruening, D.A., Ridge, S.T.: Automated event detection algorithms in pathological gait. Gait & Posture **39**(1), 472–477 (2014). https://doi.org/10.1016/j.gaitpost.2013.08.023
3. Cantaloube, A., Jolles, B.M., Favre, J.: Spatiotemporal gait analysis using video from a moving observer. Journal of Biomechanics **205**, 113402 (2026). https://doi.org/10.1016/j.jbiomech.2026.113402
4. Fiche, G., Weinzaepfel, P., Brégier, R., Baradel, F.: Multi-HMR 2: Multi-person camera-centric human detection, mesh recovery and tracking (2026), https://arxiv.org/abs/2606.14841
5. Filtjens, B., Nieuwboer, A., D'Cruz, N., Spildooren, J., Slaets, P., Vanrumste, B.: A data-driven approach for detecting gait events during turning in people with parkinson's disease and freezing of gait. Gait & Posture **80**, 130–136 (2020). https://doi.org/10.1016/j.gaitpost.2020.05.026
6. Ghislieri, M., Agostini, V., Rizzi, L., Knaflitz, M., Lanotte, M.: Atypical gait cycles in parkinson's disease. Sensors **21**(15), 5079 (2021). https://doi.org/10.3390/s21155079
7. Horsak, B., Simonlehner, M., Quehenberger, V., Dumphart, B., Wegscheider, P., Kranzl, A., Slijepčević, D.: Validity and reliability of monocular 3D markerless gait analysis in simulated pathological gait: A comparative study with OpenCap. J. Biomech. **193**, 112986 (2025). https://doi.org/10.1016/j.jbiomech.2025.112986
8. Hulleck, A.A., Mohan, D.M., Abdallah, N., El Rich, M., Khalaf, K.: Present and future of gait assessment in clinical practice: Towards the application of novel trends and technologies. Front Med Technol. **4**, 901331 (2022)
9. Kanko, R.M., Laende, E.K., Davis, E.M., Selbie, W.S., Deluzio, K.J.: Concurrent assessment of gait kinematics using marker-based and markerless motion capture. J. Biomech. **127**, 110665 (2021). https://doi.org/10.1016/j.jbiomech.2021.110665
10. Keller, M., Werling, K., Shin, S., Delp, S., Pujades, S., Liu, C.K., Black, M.J.: From skin to skeleton: Towards biomechanically accurate 3D digital humans. In: ACM ToG, Proc. SIGGRAPH Asia. vol. 42 (Dec 2023)
11. Leardini, A., Sawacha, Z., Paolini, G., Ingrosso, S., Nativo, R., Benedetti, M.G.: A new anatomically based protocol for gait analysis in children. Gait & Posture **26**(4), 560–571 (2007). https://doi.org/10.1016/j.gaitpost.2006.12.018
12. McGinley, J.L., Baker, R., Wolfe, R., Morris, M.E.: The reliability of three-dimensional kinematic gait measurements: A systematic review. Gait & Posture **29**(3), 360–369 (2009). https://doi.org/10.1016/j.gaitpost.2008.09.003
13. Pagnon, D., Domalain, M., Reveret, L.: Pose2sim: An open-source python package for multiview markerless kinematics. Journal of Open Source Software (2022). https://doi.org/10.21105/joss.04362, https://joss.theoj.org/papers/10.21105/joss.04362

14. Patel, P., Black, M.J.: CameraHMR: Aligning people with perspective (2024), https://arxiv.org/abs/2411.08128
15. Peiffer, J.D., Shah, K., Djuraskovic, I., Anarwala, S., Abdou, K., Patel, R., Jayabalan, P., Pennicooke, B., Cotton, R.J.: Portable biomechanics laboratory: Clinically accessible movement analysis from a handheld smartphone (2025), https://arxiv.org/abs/2507.08268
16. samuelm2: OpenQuestCapture: Capture and store real-world sensor data on Meta Quest 3. https://github.com/samuelm2/OpenQuestCapture (2024)
17. Seth, A., Hicks, J.L., Uchida, T.K., Habib, A., Dembia, C.L., Dunne, J.J., Ong, C.F., DeMers, M.S., Rajagopal, A., Millard, M., Hamner, S.R., Arnold, E.M., Yong, J.R., Lakshmikanth, S.K., Sherman, M.A., Ku, J.P., Delp, S.L.: Opensim: Simulating musculoskeletal dynamics and neuromuscular control to study human and animal movement. PLoS Computational Biology **14**(7), e1006223 (2018). https://doi.org/10.1371/journal.pcbi.1006223
18. Shin, S., Kim, J., Halilaj, E., Black, M.J.: WHAM: Reconstructing world-grounded humans with accurate 3D motion. In: Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR). pp. 2070–2080 (2024)
19. Simon, S.R.: Quantification of human motion: gait analysis—benefits and limitations to its application to clinical problems. J Biomech. **37**(12), 1869–1880 (2004)
20. Souri, A., Sikdar, S., Raffegeau, T.E.: Validity of monocular human mesh reconstruction for estimating lower-extremity joint kinematics: A comparison of SAM3D and CameraHMR. In: Computer Vision for Biomechanics Workshop (CVBW), CVPR Workshops (2026), https://cvbw2026.github.io/assets/papers/14.pdf
21. Uhlrich, S.D., Falisse, A., Kidziński, Ł., Muccini, J., Ko, M., Chaudhari, A.S., Hicks, J.L., Delp, S.L.: OpenCap: Human movement dynamics from smartphone videos. PLOS Comput. Biol. **19**(10), 1–26 (2023)
22. Yang, X., Kukreja, D., Pinkus, D., Sagar, A., Fan, T., Park, J., Shin, S., Cao, J., Liu, J., Ugrinovic, N., Feiszli, M., Malik, J., Dollár, P.: SAM 3D Body: Robust full-body human mesh recovery. Technical Report, Meta FAIR (2025)
23. Zaccardi, S., Arapi, A., Frantz, T., Brahimetaj, R., Debeuf, R., Beckwée, D., Swinnen, E., Jansen, B.: HoloMoCap: Real-time clinical motion capture with HoloLens 2. In: 2025 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR). pp. 218–222 (2025). https://doi.org/10.1109/AIxVR63409.2025.00042
24. Zeni, J.A., Richards, J.G., Higginson, J.S.: Two simple methods for determining gait events during treadmill and overground walking using kinematic data. Gait & Posture **27**(4), 710–714 (2008). https://doi.org/10.1016/j.gaitpost.2007.07.007