An Observer-Egocentric Approach to Real-World Gait Analysis

Alexis Cantaloube[1] Julien Favre[1,2,3]
[1]BioMotion Center at CHUV [2]UNIL [3]The Sense Innovation and Research Center
alexis.cantaloube@chuv.ch

Abstract

Spatiotemporal gait analysis is well established but remains underused in routine practice. We propose an observer-egocentric approach using smart glasses together with monocular 3D human and camera pose estimation to enable gait analysis during routine clinical observation. To support this, we introduce a benchmark dataset of egocentric video captured by a moving observer, with visual-inertial-based reference spatiotemporal parameters, and use it to evaluate a pipeline for real-world gait analysis. Results show the potential and current limitations of observer-egocentric gait analysis.

1. Introduction

Spatiotemporal gait analysis is a proven method for evaluating and studying human movement in clinical and research settings [43, 63]. However, most existing quantitative gait analysis approaches remain impractical for routine use, as they rely on laboratory setups, dedicated instrumentation, and trained personnel [25, 47, 50]. As a result, gait assessment remains largely confined to controlled environments [1], despite growing evidence that real-world measurements may better reflect everyday walking [3, 6, 18, 61]. These limitations highlight the need for gait analysis methods that are easy to use and operable in unconstrained, real-world environments.

Recent advances in computer vision, together with the emergence of smart glasses equipped with cameras and inertial sensors, offer new opportunities for practical gait analysis in real-world settings [14]. Monocular 3D human and camera pose estimation methods have improved substantially, enabling more accurate estimation in unconstrained environments [23, 35]. These developments suggest an observer-egocentric approach in which an external observer, such as a clinician or caregiver, records a walking subject using smart glasses during routine qualitative observation, without the need for complex setups or sensors on the subject

Despite its promise, this approach remains largely unexplored. The closest existing gait analysis approaches are limited to fixed-camera setups in controlled environments [10, 27, 36, 51, 59]. At the same time, most computer vision methods are evaluated almost exclusively using joint-level pose accuracy metrics, such as mean per-joint position error (MPJPE), which do not reflect biomechanically meaningful performance [35]. As a result, it remains unclear whether current methods are suitable for observer-egocentric gait analysis in real-world settings.

The goal of this work is to assess an observer-egocentric approach to spatiotemporal gait analysis. To this end, we introduce a benchmark, including a dataset captured from a moving observer using smart glasses with synchronized inertial reference, as well as a pipeline for spatiotemporal gait parameter extraction and evaluation `https://github.com/Cantalex/nymeria_gait_dataset`.

2. Related Work

2.1. Gait Analysis Systems

Traditionally, gait analysis has relied on marker-based motion capture systems, which provide accurate measurements but require controlled environments, specialized equipment, and extensive setup [7, 11]. More recently, markerless approaches have been proposed to improve accessibility [24], including multi-camera systems [22, 41, 56], monocular methods from static or constrained viewpoints [10, 27, 36, 51, 59], and mobile smartphone-based methods, including gimbal-stabilized and handheld [9, 42], which have shown promising agreement with reference measurements. However, these approaches remain limited to fixed or stabilized camera viewpoints and controlled environments. Other solutions such as pressure-sensitive walkways or wearable inertial sensors also require constrained capture areas, subject instrumentation, or careful calibration [45, 46]. Overall, existing gait analysis systems do not support simple and versatile spatiotemporal gait assessment in unconstrained, hands-free settings.

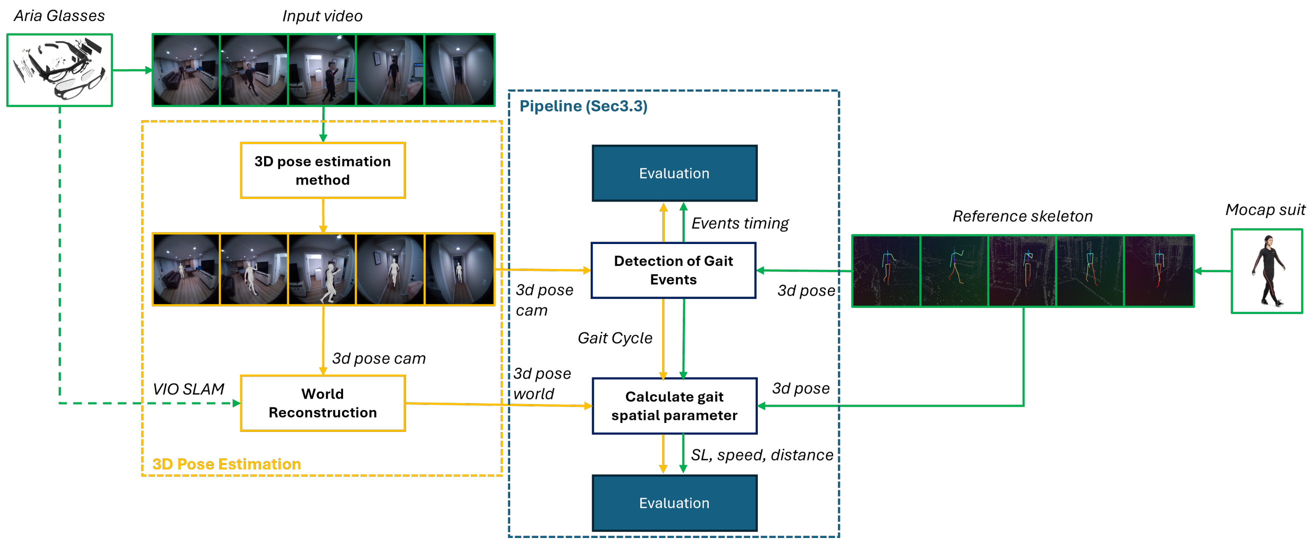

Figure 1. Overview of our evaluation framework. Green represents Nymeria-Gait data, which includes egocentric observer videos, VIO SLAM, and reference skeleton data (Sec. 3.2). Orange represents the 3D pose estimation stage, which takes video input (optionally with VIO SLAM from the glasses) (Sec. 3.4) and returns 3D poses in both camera and world coordinates. Blue represents our gait estimation pipeline and evaluation (Sec. 3.3).

2.2. 3D Pose Estimation

Monocular human motion reconstruction has substantially improved through the use of parametric body models such as SMPL [37], large-scale HMR-style frameworks with vision transformers [20, 29], and methods that couple human motion with camera estimation in world coordinates [32, 60]. At the same time, there is growing interest in biomechanically accurate motion reconstruction [31, 33]. However, most existing approaches are still designed primarily for visual realism and are evaluated using joint-level pose accuracy metrics such as MPJPE. While these metrics are central for training and standardized comparison [39], they do not reflect task-level or biomechanically meaningful performance [35]. As a result, progress reported on existing benchmarks datasets such as Human3.6M [28], 3DPW [58], and EMDB [30] may not necessarily translate to reliable gait analysis.

2.3. Datasets and Benchmarks

Existing datasets for gait and human motion analysis are primarily captured in controlled environments using fixed monocular or multi-camera setups [19, 34], which limits their suitability for observer-egocentric and mobile scenarios [55]. Several datasets provide gait measurements or walking sequences, but either lack synchronized video and rely on static viewpoints [8, 26], or do not include the precise reference information required for biomechanical analysis [8, 44, 54]. More recent real-world motion datasets introduce mobile cameras or egocentric viewpoints and enable the study of human motion from an observer's perspective [28, 30, 38, 58, 66]. However, these datasets are not designed for gait analysis, as they lack structured walking sequences, standardized gait annotations, or evaluation protocols based on spatiotemporal gait parameters (Suppl. Tab. 7). As a result, no existing dataset enables systematic benchmarking of observer-egocentric gait analysis in real-world settings.

3. Methods

3.1. Overview

Our goal is to evaluate observer-egocentric spatiotemporal gait analysis. To study this setting, we design a benchmark that combines egocentric smart-glasses recordings with synchronized inertial reference data and a modular pipeline for motion reconstruction and gait parameter extraction (Fig. 1). Monocular 3D human and camera pose estimation methods can be evaluated within this framework by comparing estimated gait parameters to reference measurements.

3.2. Nymeria-Gait Dataset

Nymeria-Gait is a dataset designed for observer-egocentric gait analysis which is a subset of the large all-purpose multimodal Nymeria dataset [38]. It consists of egocentric video recordings captured by mobile observers wearing smart glasses [14], while observing a participant equipped with an inertial motion capture suit (Xsens, NL). Video was recorded at 30 Hz and synchronized with reference motion captured at 240 Hz. All devices were time-synchronized, and recordings were spatially registered.

Nymeria-Gait contains 908 walking sequences from 54 subjects (26 males and 28 females; age: 18 to 50 years old, weight: 69±12 kg; height: 1.68±0.09 m), recorded across indoor and outdoor environments and spanning a total walking distance of approximately 7.9 km. The data include straight and turning motion, as well as flat and sloped terrain, captured during unconstrained real-world activities.

Detailed information on dataset construction, sequence selection, annotations, and statistics is provided in the supplementary material (Sec. 5).

3.3. Motion Reconstruction and Gait Evaluation

3.3.1. Pipeline

Given egocentric video captured by a moving observer, our framework reconstructs the subject's 3D motion and evaluates gait parameters derived from the reconstructed trajectories. Monocular 3D pose estimation methods are applied to the observer video for human motion relative to the camera. When required, reconstructed poses are expressed in a global coordinate system using the camera trajectory, enabling spatial gait analysis (Fig. 1).

3.3.2. Evaluation Protocol

The framework is designed to be modular, allowing different monocular pose estimation and camera localization methods to be interchanged and evaluated under a common protocol. Both camera-centric and world-centric reconstructions are supported, depending on the capabilities of the evaluated method.

Gait events are identified from foot trajectories relative to the pelvis along the walking direction for both the evaluated method and the synchronized reference, following prior work [49, 57, 65]. These events are then used to compute cycle, stance, swing, and double support durations. Stride length is also calculated on a cycle basis as the horizontal distance traveled by each foot between consecutive heel strikes. For each sequence, total displacement is computed from pelvis trajectory in the horizontal plane, and average walking speed is obtained by dividing this distance by sequence duration.

The resulting spatiotemporal parameters are compared between the evaluated method and the synchronized reference to assess performance. This design enables systematic benchmarking of observer-egocentric gait analysis pipelines by quantifying how reconstruction errors propagate to interpretable gait measures.

3.4. Method Selection for Experiments

We evaluated multiple monocular 3D human motion reconstruction methods and several variants of camera motion integration within the proposed framework. The evaluation considered both temporal and spatial gait parameters, allowing us to assess how different design choices affect gait estimation accuracy (details in Sec. 6).

Based on this analysis, we selected the best-performing configuration for all subsequent experiments. In particular, we use WHAM [48] combined with camera trajectories provided by the smart-glasses visual-inertial (VIO) simultaneous localization and mapping (SLAM). This configuration achieved the most consistent performance across temporal gait events, spatial gait parameters, and trajectory accuracy in real-world conditions.

4. Experiments

4.1. Experimental Setup

4.1.1. Data Selection and Preprocessing

Experiments are conducted on our Nymeria-Gait dataset (Sec. 3.2) using our pipeline (Sec. 3.3). As Nymeria-Gait is a general-purpose dataset, some recordings revealed to be unsuitable, for example due to persistent foot occlusion that prevents reliable gait event detection. Such sequences were excluded.

The final evaluation set consists of 608 walking sequences, comprising 5,609 gait cycles and covering approximately 4.4 km of walking. These data were collected from 53 subjects across indoor and outdoor environments and correspond to 76 minutes of continuous walking.

4.1.2. Evaluation Metrics

To evaluate the proposed observer-based method, first, the event detection rate was calculated as the percentage of correctly identified heel-strike and toe-off events using the proposed method relative to the total number of these events identified by the reference. Then, based on the identified events, the errors (difference between the proposed and reference methods) were calculated for each cycle, stance, swing, and double support times. Similar error calculations were performed for stride length, gait speed and total distance.

The accuracy was defined as the mean of the individual errors and precision as their standard deviation (SD). In addition, the Pearson correlation coefficient (r) was used to assess the association between parameters obtained with the proposed and reference methods.

4.2. Quantitative Results

4.2.1. Temporal Parameters

The proposed observer-based method detected 96.6% of the 14,337 gait events, with a mean event timing error of 8.7 ms and SD of 37.3 ms. Detailed results for event detection and derived temporal parameters are reported in Table 1.

Overall, events were detected with frame-level accuracy, which propagated to accurate ($\leq 5.7ms$) and precise ($\leq 45.1ms$) temporal parameters. Importantly, 65.4% of

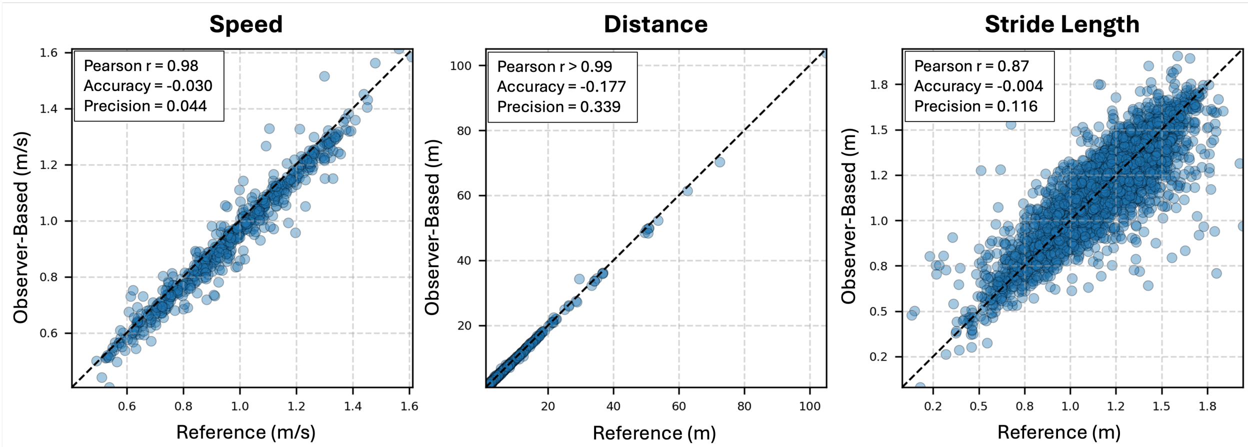

Figure 2. Correlation between spatial gait parameters obtained using the proposed observer-based method and the reference method. Each point for speed and distance represents one of the 608 walking sequences, whereas each point for stride length represents one of the 5,609 gait cycles.

Parameter	Accuracy (ms)	Precision (ms)
Event Detection		
Heel strike	11.4	33.9
Toe-off	5.9	40.3
Gait Timing		
Cycle time	0.2	39.8
Stance time	-5.3	45.1
Swing time	5.5	43.5
Double support	-5.7	41.6

Table 1. Temporal gait parameter accuracy and precision across all analyzed gait cycles. Accuracy is reported as the mean error and precision as the standard deviation (SD).

detected gait events meet the recommended error-tolerance threshold of 20.0 ms [12], indicating that the proposed approach is nearing recommended performance for gait assessment.

4.2.2. Spatial Parameters

Figure 2 reports the accuracy and precision of the estimated spatial gait parameters and illustrates their association with the reference. The mean errors for walking speed and total distance correspond to less than 4.7% of the average values observed in the dataset, while stride length shows a near-zero mean error and a precision corresponding to 9.8% of the average stride length.

For spatial parameters, acceptable performance thresholds are inherently application-dependent. Nevertheless, the proposed method shows strong association with the reference, with Pearson correlation coefficients >0.87 for walking speed, total distance, and stride length (Figure 2). The low errors for walking speed and total distance support their use in real-world settings. In contrast, a stride length precision close to 10% of the average value highlights a clear area for improvement. Still, these results indicate performance comparable to existing vision-based gait analysis approaches, despite operating in unconstrained, real-world settings with a moving observer [51, 59].

4.2.3. Discussion and Limitations

Overall, these results are encouraging given that they were obtained in an unconstrained, hands-free setting without subject instrumentation or calibration. The observed agreement with prior vision-based gait analysis studies suggests that the proposed approach may similarly meet minimal detectable change thresholds [16, 40]. A detailed analysis of how walking environment, terrain slope, observer viewpoint, and viewing distance influence measurement errors is provided in [5]. This analysis can expose limitations of current monocular motion reconstruction methods that are not captured by joint-level pose errors alone.

Several limitations should be noted. First, the smart glasses operate at 30 Hz, limiting temporal resolution. Second, the use of an inertial motion capture suit as reference limits the evaluation motivating future datasets specifically designed for biomechanical analysis. Third, our pipeline relies on SMPL but more recent body models [2, 15] may offer improved biomechanical fidelity. Evaluating them is a natural direction for future benchmarking on Nymeria-Gait. While only healthy participants were included, the underlying pose estimation and gait event detection methods have been validated on diverse populations, suggesting broader applicability [4, 13, 17, 21, 65]. Finally, although smart glasses were used here, the approach could generalize to any mobile device.

References

[1] Lauren C Benson, Anu M Räisänen, Christian A Clermont, and Reed Ferber. Is this the real life, or is this just laboratory? a scoping review of imu-based running gait analysis. *Sensors*, 22(5):1722, 2022. 1

[2] Romain Brégier, Guénolé Fiche, Laura Bravo-Sánchez, Thomas Lucas, Matthieu Armando, Philippe Weinzaepfel, Grégory Rogez, and Fabien Baradel. Human mesh modeling for anny body. *arXiv preprint arXiv:2511.03589*, 2025. 4

[3] Lorenzo Brognara, Alberto Arceri, Marco Zironi, Francesco Traina, Cesare Faldini, and Antonio Mazzotti. Gait spatio-temporal parameters vary significantly between indoor, outdoor and different surfaces. *Sensors*, 25(5):1314, 2025. 1

[4] Dustin A Bruening and Sarah Trager Ridge. Automated event detection algorithms in pathological gait. *Gait & posture*, 39(1):472–477, 2014. 4

[5] Alexis Cantaloube, Brigitte M. Jolles, and Julien Favre. Spatiotemporal gait analysis using video from a moving observer. SSRN Preprint, 2024. Available at SSRN. 4

[6] Lena Carcreff, Corinna N Gerber, Anisoara Paraschiv-Ionescu, Geraldo De Coulon, Christopher J Newman, Kamiar Aminian, and Stéphane Armand. Comparison of gait characteristics between clinical and daily life settings in children with cerebral palsy. *Scientific reports*, 10(1):2091, 2020. 1

[7] Shanshan Chen, John Lach, Benny Lo, and Guang-Zhong Yang. Toward pervasive gait analysis with wearable sensors: A systematic review. *IEEE journal of biomedical and health informatics*, 20(6):1521–1537, 2016. 1

[8] Roman Chereshnev and Attila Kertész-Farkas. Hugadb: Human gait database for activity recognition from wearable inertial sensor networks. In *Analysis of Images, Social Networks and Texts: 6th International Conference, AIST 2017, Moscow, Russia, July 27–29, 2017, Revised Selected Papers 6*, pages 131–141. Springer, 2018. 2, 3

[9] Anthony Cimorelli, Ankit Patel, Tasos Karakostas, and R James Cotton. Validation of portable in-clinic video-based gait analysis for prosthesis users. *Scientific reports*, 14(1): 3840, 2024. 1

[10] R James Cotton, Emoonah McClerklin, Anthony Cimorelli, Ankit Patel, and Tasos Karakostas. Transforming gait: video-based spatiotemporal gait analysis. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 115–120. IEEE, 2022. 1

[11] Scott L Delp, Frank C Anderson, Allison S Arnold, Peter Loan, Ayman Habib, Chand T John, Eran Guendelman, and Darryl G Thelen. Opensim: open-source software to create and analyze dynamic simulations of movement. *IEEE transactions on biomedical engineering*, 54(11):1940–1950, 2007. 1

[12] Bernhard Dumphart, Djordje Slijepcevic, Fabian Unglaube, Andreas Kranzl, Arnold Baca, and Brian Horsak. The effect of inaccurate initial contact events on kinematics in healthy and pathological gait. *Gait & Posture*, page 110012, 2025. 4

[13] Moataz Eltoukhy, Christopher Kuenze, Michael S Andersen, Jeonghoon Oh, and Joseph Signorile. Prediction of ground reaction forces for parkinson's disease patients using a kinect-driven musculoskeletal gait analysis model. *Medical engineering & physics*, 50:75–82, 2017. 4

[14] Jakob Engel, Kiran Somasundaram, Michael Goesele, Albert Sun, Alexander Gamino, Andrew Turner, Arjang Talattof, Arnie Yuan, Bilal Souti, Brighid Meredith, et al. Project aria: A new tool for egocentric multi-modal ai research. *arXiv preprint arXiv:2308.13561*, 2023. 1, 2

[15] Aaron Ferguson, Ahmed A. A. Osman, Berta Bescos, Carsten Stoll, Chris Twigg, Christoph Lassner, David Otte, Eric Vignola, Fabian Prada, Federica Bogo, Igor Santesteban, Javier Romero, Jenna Zarate, Jeongseok Lee, Jinhyung Park, Jinlong Yang, John Doublestein, Kishore Venkateshan, Kris Kitani, Ladislav Kavan, Marco Dal Farra, Matthew Hu, Matthew Cioffi, Michael Fabris, Michael Ranieri, Mohammad Modarres, Petr Kadlecek, Rawal Khirodkar, Rinat Abdrashitov, Romain Prévost, Roman Rajbhandari, Ronald Mallet, Russell Pearsall, Sandy Kao, Sanjeev Kumar, Scott Parrish, Shoou-I Yu, Shunsuke Saito, Takaaki Shiratori, Te-Li Wang, Tony Tung, Yichen Xu, Yuan Dong, Yuhua Chen, Yuanlu Xu, Yuting Ye, and Zhongshi Jiang. Mhr: Momentum human rig, 2025. 4

[16] Rita Fernandes, Paulo Armada-da Silva, Annelies Pool-Goudaazward, Vera Moniz-Pereira, and António P Veloso. Three dimensional multi-segmental trunk kinematics and kinetics during gait: Test-retest reliability and minimal detectable change. *Gait & posture*, 46:18–25, 2016. 4

[17] Margaret A French, Corey Koller, and Elisa S Arch. Comparison of three kinematic gait event detection methods during overground and treadmill walking for individuals post stroke. *Journal of Biomechanics*, 99:109481, 2020. 4

[18] Irina Galperin, Inbar Hillel, Silvia Del Din, Esther MJ Bekkers, Alice Nieuwboer, Giovanni Abbruzzese, Laura Avanzino, Freek Nieuwhof, Bastiaan R Bloem, Lynn Rochester, et al. Associations between daily-living physical activity and laboratory-based assessments of motor severity in patients with falls and parkinson's disease. *Parkinsonism & related disorders*, 62:85–90, 2019. 1

[19] Saeed Ghorbani, Kimia Mahdaviani, Anne Thaler, Konrad Kording, Douglas James Cook, Gunnar Blohm, and Nikolaus F Troje. Movi: A large multi-purpose human motion and video dataset. *Plos one*, 16(6):e0253157, 2021. 2, 3

[20] Shubham Goel, Georgios Pavlakos, Jathushan Rajasegaran, Angjoo Kanazawa, and Jitendra Malik. Humans in 4d: Reconstructing and tracking humans with transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 14783–14794, 2023. 2

[21] Rejane Vale Gonçalves, Sérgio Teixeira Fonseca, Priscila Albuquerque Araújo, Vanessa Lara Araújo, Tais Martins Barboza, Gabriela Andrade Martins, and Marisa Cotta Mancini. Identification of gait events in children with spastic cerebral palsy: comparison between the force plate and algorithms. *Brazilian Journal of Physical Therapy*, 24(5):392–398, 2020. 4

[22] Yoni Gozlan, Antoine Falisse, Scott Uhlrich, Anthony Gatti, Michael Black, and Akshay Chaudhari. Opencapbench: A benchmark to bridge pose estimation and biomechanics. *arXiv preprint arXiv:2406.09788*, 2024. 1

[23] Yan Guo, Tianhan Gao, Aoshuang Dong, Xinbei Jiang, Zichen Zhu, and Fuxin Wang. A survey of the state of the art in monocular 3d human pose estimation: Methods, benchmarks, and challenges. *Sensors (Basel, Switzerland)*, 25(8): 2409, 2025. 1

[24] Thomas Hellsten, Jonny Karlsson, Muhammed Shamsuzzaman, and Göran Pulkkis. The potential of computer vision-based marker-less human motion analysis for rehabilitation. *Rehabilitation Process and Outcome*, 10: 11795727211022330, 2021. 1

[25] Salhah Hobani, Anas Mohammed Alhakami, Shadab Uddin, Fuzail Ahmad, and Hana Alsobayel. Perceived application and barriers for gait assessment in physical therapy practice in saudi arabia. *Life*, 13(1):50, 2022. 1

[26] Martin Hofmann, Jürgen Geiger, Sebastian Bachmann, Björn Schuller, and Gerhard Rigoll. The tum gait from audio, image and depth (gaid) database: Multimodal recognition of subjects and traits. *Journal of Visual Communication and Image Representation*, 25(1):195–206, 2014. 2, 3

[27] Brian Horsak, Mark Simonlehner, Viktoria Quehenberger, Bernhard Dumphart, Philipp Wegscheider, Andreas Kranzl, and Djordje Slijepcevic. Validity and reliability of monocular 3d markerless gait analysis in simulated pathological gait: A comparative study with opencap. *Journal of Biomechanics*, page 112986, 2025. 1

[28] Catalin Ionescu, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu. Human3. 6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. *IEEE transactions on pattern analysis and machine intelligence*, 36(7):1325–1339, 2013. 2, 3

[29] Angjoo Kanazawa, Michael J Black, David W Jacobs, and Jitendra Malik. End-to-end recovery of human shape and pose. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7122–7131, 2018. 2

[30] Manuel Kaufmann, Jie Song, Chen Guo, Kaiyue Shen, Tianjian Jiang, Chengcheng Tang, Juan José Zárate, and Otmar Hilliges. Emdb: The electromagnetic database of global 3d human pose and shape in the wild. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 14632–14643, 2023. 2, 3

[31] Marilyn Keller, Keenon Werling, Soyong Shin, Scott Delp, Sergi Pujades, C Karen Liu, and Michael J Black. From skin to skeleton: Towards biomechanically accurate 3d digital humans. *ACM Transactions on Graphics (TOG)*, 42(6): 1–12, 2023. 2

[32] Muhammed Kocabas, Ye Yuan, Pavlo Molchanov, Yunrong Guo, Michael J Black, Otmar Hilliges, Jan Kautz, and Umar Iqbal. Pace: Human and camera motion estimation from in-the-wild videos. In *2024 International Conference on 3D Vision (3DV)*, pages 397–408. IEEE, 2024. 2

[33] Farnoosh Koleini, Muhammad Usama Saleem, Pu Wang, Hongfei Xue, Ahmed Helmy, and Abbey Fenwick. Biopose: Biomechanically-accurate 3d pose estimation from monocular videos. *arXiv preprint arXiv:2501.07800*, 2025. 2

[34] Bogdan Kwolek, Agnieszka Michalczuk, Tomasz Krzeszowski, Adam Switonski, Henryk Josinski, and Konrad Wojciechowski. Calibrated and synchronized multi-view video and motion capture dataset for evaluation of gait recognition. *Multimedia Tools and Applications*, 78(22):32437–32465, 2019. 2, 3

[35] Yang Liu, Changzhen Qiu, and Zhiyong Zhang. Deep learning for 3d human pose estimation and mesh recovery: A survey. *Neurocomputing*, page 128049, 2024. 1, 2

[36] Luca Lonini, Yaejin Moon, Kyle Embry, R James Cotton, Kelly McKenzie, Sophia Jenz, and Arun Jayaraman. Video-based pose estimation for gait analysis in stroke survivors during clinical assessments: a proof-of-concept study. *Digital Biomarkers*, 6(1):9–18, 2022. 1

[37] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J Black. Smpl: A skinned multi-person linear model. In *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, pages 851–866. 2023. 2

[38] Lingni Ma, Yuting Ye, Fangzhou Hong, Vladimir Guzov, Yifeng Jiang, Rowan Postyeni, Luis Pesqueira, Alexander Gamino, Vijay Baiyya, Hyo Jin Kim, et al. Nymeria: A massive collection of multimodal egocentric daily motion in the wild. In *European Conference on Computer Vision*, pages 445–465. Springer, 2024. 2, 1

[39] Julieta Martinez, Rayat Hossain, Javier Romero, and James J Little. A simple yet effective baseline for 3d human pose estimation. In *Proceedings of the IEEE international conference on computer vision*, pages 2640–2649, 2017. 2

[40] Dara Meldrum, Ciara Shouldice, Ronan Conroy, Kim Jones, and Malcolm Forward. Test–retest reliability of three dimensional gait analysis: Including a novel approach to visualising agreement of gait cycle waveforms with bland and altman plots. *Gait & posture*, 39(1):265–271, 2014. 4

[41] David Pagnon, Mathieu Domalain, and Lionel Reveret. Pose2sim: an end-to-end workflow for 3d markerless sports kinematics—part 1: robustness. *Sensors*, 21(19):6530, 2021. 1

[42] JD Peiffer, Kunal Shah, Irina Djuraskovic, Shawana Anarwala, Kayan Abdou, Rujvee Patel, Prakash Jayabalan, Brenton Pennicooke, and R James Cotton. Portable biomechanics laboratory: Clinically accessible movement analysis from a handheld smartphone. *arXiv preprint arXiv:2507.08268*, 2025. 1

[43] Jacquelin Perry and Judith Burnfield. *Gait analysis: normal and pathological function*. CRC Press, 2024. 1

[44] Rahm Ranjan, David Ahmedt-Aristizabal, Mohammad Ali Armin, and Juno Kim. Computer vision for clinical gait analysis: A gait abnormality video dataset. *arXiv preprint arXiv:2407.04190*, 2024. 2, 3

[45] Abdul Saboor, Triin Kask, Alar Kuusik, Muhammad Mahtab Alam, Yannick Le Moullec, Imran Khan Niazi, Ahmed Zoha, and Rizwan Ahmad. Latest research trends in gait analysis using wearable sensors and machine learning: A systematic review. *Ieee Access*, 8:167830–167864, 2020. 1

[46] Ozell Sanders, Bin Wang, and Kimberly Kontson. Concurrent validity evidence for pressure-sensing walkways measuring spatiotemporal features of gait: A systematic review and meta-analysis. *Sensors*, 24(14):4537, 2024. 1

[47] Yashoda Sharma, Lovisa Cheung, Kara K Patterson, Andrea Iaboni, et al. Factors influencing the clinical adoption of

quantitative gait analysis technologies for adult patient populations with a focus on clinical efficacy and clinician perspectives: protocol for a scoping review. *JMIR research protocols*, 12(1):e39767, 2023. 1

[48] Soyong Shin, Juyong Kim, Eni Halilaj, and Michael J Black. Wham: Reconstructing world-grounded humans with accurate 3d motion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2070–2080, 2024. 3, 5

[49] Emily H Sinitski, Edward D Lemaire, Natalie Baddour, Markus Besemann, Nancy L Dudek, and Jacqueline S Hebert. Fixed and self-paced treadmill walking for able-bodied and transtibial amputees in a multi-terrain virtual environment. *Gait & posture*, 41(2):568–573, 2015. 3

[50] Julie Stebbins, Marian Harrington, and Caroline Stewart. Clinical gait analysis 1973–2023: Evaluating progress to guide the future. *Journal of biomechanics*, 160:111827, 2023. 1

[51] Jan Stenum, Cristina Rossi, and Ryan T Roemmich. Two-dimensional video-based analysis of human gait using pose estimation. *PLoS computational biology*, 17(4):e1008935, 2021. 1, 4

[52] Zachary Teed and Jia Deng. Droid-slam: Deep visual slam for monocular, stereo, and rgb-d cameras. *Advances in neural information processing systems*, 34:16558–16569, 2021. 5

[53] Zachary Teed, Lahav Lipson, and Jia Deng. Deep patch visual odometry. *Advances in Neural Information Processing Systems*, 36:39033–39051, 2023. 5

[54] Yonghong Tian, Lan Wei, Shijian Lu, and Tiejun Huang. Free-view gait recognition. *PloS one*, 14(4):e0214389, 2019. 2

[55] Luke K Topham, Wasiq Khan, Dhiya Al-Jumeily, and Abir Hussain. Human body pose estimation for gait identification: A comprehensive survey of datasets and models. *ACM Computing Surveys*, 55(6):1–42, 2022. 2

[56] Scott D Uhlrich, Antoine Falisse, Łukasz Kidziński, Julie Muccini, Michael Ko, Akshay S Chaudhari, Jennifer L Hicks, and Scott L Delp. Opencap: Human movement dynamics from smartphone videos. *PLoS computational biology*, 19(10):e1011462, 2023. 1

[57] Baptiste Ulrich, Alejandro N Santos, Brigitte M Jolles, David H Benninger, and Julien Favre. Gait events during turning can be detected using kinematic features originally proposed for the analysis of straight-line walking. *Journal of biomechanics*, 91:69–78, 2019. 3

[58] Timo Von Marcard, Roberto Henschel, Michael J Black, Bodo Rosenhahn, and Gerard Pons-Moll. Recovering accurate 3d human pose in the wild using imus and a moving camera. In *Proceedings of the European conference on computer vision (ECCV)*, pages 601–617, 2018. 2, 3

[59] Hanwen Wang, Bingyi Su, Lu Lu, Sehee Jung, Liwei Qing, Ziyang Xie, and Xu Xu. Markerless gait analysis through a single camera and computer vision. *Journal of Biomechanics*, 165:112027, 2024. 1, 4

[60] Yufu Wang, Ziyun Wang, Lingjie Liu, and Kostas Daniilidis. Tram: Global trajectory and motion of 3d humans from in-the-wild videos. In *European Conference on Computer Vision*, pages 467–487. Springer, 2024. 2, 5

[61] Elke Warmerdam, Jeffrey M Hausdorff, Arash Atrsaei, Yuhan Zhou, Anat Mirelman, Kamiar Aminian, Alberto J Espay, Clint Hansen, Luc JW Evers, Andreas Keller, et al. Long-term unsupervised mobility assessment in movement disorders. *The Lancet Neurology*, 19(5):462–470, 2020. 1

[62] Lan Wei, Yonghong Tian, Yaowei Wang, and Tiejun Huang. Swiss-system based cascade ranking for gait-based person re-identification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2015. 3

[63] Michael W Whittle. *Gait analysis: an introduction*. Butterworth-Heinemann, 2014. 1

[64] Yufei Xu, Jing Zhang, Qiming Zhang, and Dacheng Tao. Vitpose: Simple vision transformer baselines for human pose estimation. *Advances in neural information processing systems*, 35:38571–38584, 2022. 5

[65] JA Zeni Jr, JG Richards, and JS2384115 Higginson. Two simple methods for determining gait events during treadmill and overground walking using kinematic data. *Gait & posture*, 27(4):710–714, 2008. 3, 4

[66] Siwei Zhang, Qianli Ma, Yan Zhang, Zhiyin Qian, Taein Kwon, Marc Pollefeys, Federica Bogo, and Siyu Tang. Egobody: Human body shape and motion of interacting people from head-mounted devices. In *European conference on computer vision*, pages 180–200. Springer, 2022. 2, 3

An Observer-Egocentric Approach to Real-World Gait Analysis

Supplementary Material

5. Nymeria-Gait Dataset Construction and Statistics

Overview: Figure 3 illustrates the Nymeria-Gait setup. It shows example egocentric observer views across diverse environments and lighting conditions, as well as synchronized 3D skeletal motion from the motion capture suit.

Dataset: Table 5 summarizes the data modalities available in the Nymeria dataset [38], distinguishing between those captured by the observer, participant, and other sources of information. The modalities highlighted in red are those specifically used in Nymeria-Gait.

While Nymeria-Gait is primarily concerned with gait analysis from the observer's perspective, the dataset includes additional modalities that may enable alternative gait-related applications, such as participant-centric gait analysis.

Nymeria-Gait consists of 908 walking sequences from 54 participants in both indoor and outdoor environments, covering a total of 7,890 meters and spanning 142 minutes.

Table 2 summarizes participant demographics and anthropometric measurements. Table 3 summarizes the dataset's composition, such as sequence count, duration, and distance covered. Finally, Table 4 shows gait characteristics, environmental conditions, and scripted interactions.

Statistic	**Value**
Total Participants	54
Total Locations	10
Demographic	
Male	26
Female	28
Age 18-24	10
Age 25-30	10
Age 31-35	12
Age 36-40	11
Age 41-45	7
Age 46-50	4
Anthropometric	
Avg Height (cm)	168.26
Avg Weight (kg)	68.89

Table 2. Summary of participant statistics.

Statistic	**Value**
Total Sequences	908
Total Frames	256,413
Avg Frames per Sequences	282.39
Time	
Total (minutes)	142
Indoor (minutes)	39
Outdoor (minutes)	102
Distance	
Total (m)	7,890
Indoor (m)	2,397
Outdoor (m)	5,492

Table 3. Summary of dataset sequence characteristics.

Statistic	**Value**
Gait	
Events	26,685
Cycles	11,437
Environment	
Indoor Sequences	377
Outdoor Sequences	531
Slope	
Flat Sequences	771
Up Sequences	86
Down Sequences	51
Scripts	
S2 Where is X	22
S16 Simon says	34
S18 Hike	16
S19 Fresh air	36

Table 4. Summary of motion and environmental factors.

Participant Selection: The Nymeria dataset includes 20 scripts designed to capture a wide range of daily activities in both indoor and outdoor settings. We conducted a thorough review of the entire dataset and identified participants who had at least two recordings, one in an outdoor script (*Fresh Air* or *Hike*) and one in an indoor script (*Simon Says*, *Where is X*). These scripts were chosen because they naturally include walking sequences, such as navigation, object search, and verbal instructions (details in Tab. 6).

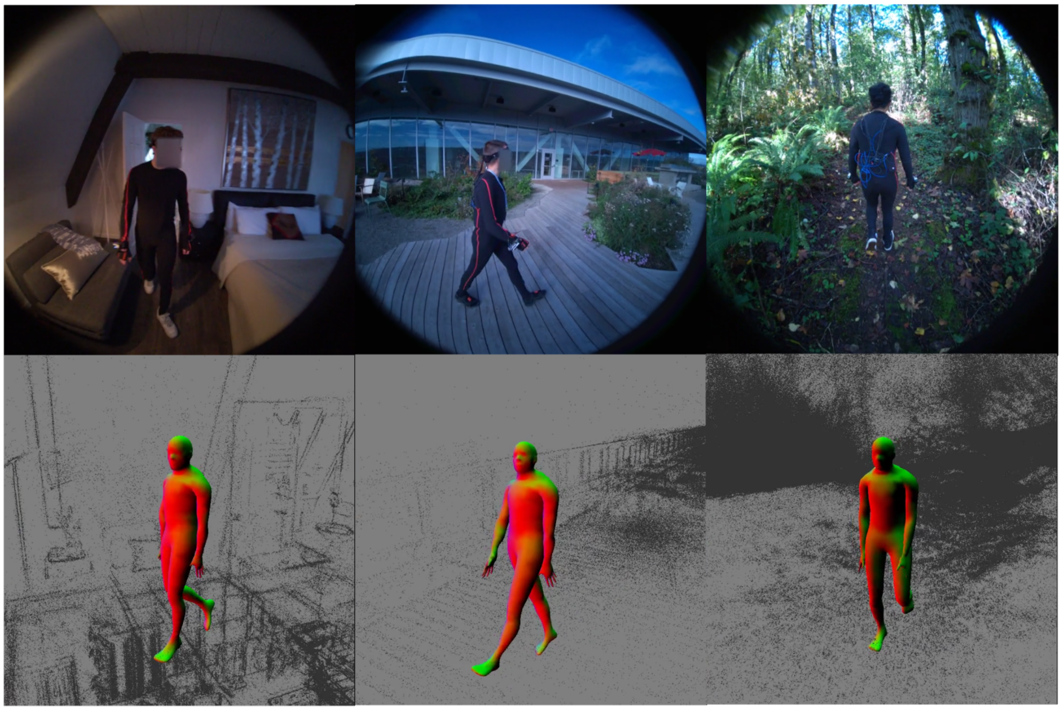

Figure 3. Overview of Nymeria-Gait: example egocentric observer images (top) illustrating environmental and viewpoint diversity, and synchronized 3D skeletal motion from the participant (bottom) captured using an inertial-based motion capture suit.

Data Type	**Observer**	**Participant**	**Additional**
RGB Video (30 FPS)	✓	✓	
Grayscale Video (30 FPS)	✓	✓	(Used for Global Alignment)
Eye Tracking (10 FPS)	✓	✓	
IMU Motion Data (1 kHz)	✓	✓	(Right: 1 kHz, Left: 800 Hz)
MiniAria Wristbands (10 FPS)		✓	
3D Skeletal Motion (240 Hz)		✓	(XSens Motion Capture)
SLAM Trajectory (6DoF)	✓	✓	(Wristband + Glasses)
Parametric Human Model		✓	(Derived from XSens)
Magnetometer (10 Hz)	✓	✓	Not in wristbands
Scene Point Clouds (Global Alignment)	✓	✓	(Generated via SLAM)
Demographic data		✓	

Table 5. Available Data Modalities in the Nymeria Dataset. Red are the one we use in the framework

Scenario ID	**Name**	**Description**
S2	Where is X	Searching for misplaced objects in a home environment; includes direction changes, varied speeds, and environmental interaction.
S16	Simon Says	Task-driven walking based on verbal instructions (e.g., carrying groceries, measuring furniture); includes turning, stopping, and accelerating.
S18	Hike	Walking on both flat and hilly outdoor trails; includes varied terrain slopes.
S19	Fresh Air	Recreational outdoor activities including walking, jogging, and light sports.

Table 6. Scenario descriptions

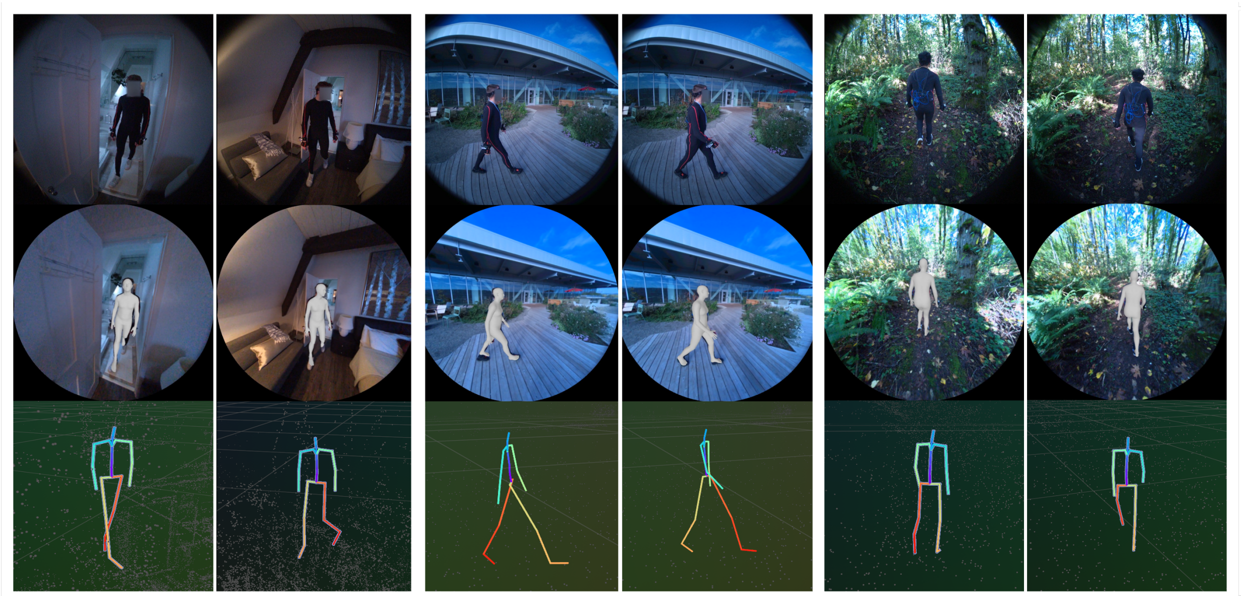

Figure 4. Nymeria-Gait

Dataset	Camera Type	3D Skeleton	Gait Dataset	Gait Pipeline	Indoor	Outdoor	Participants
TUM-GAID [26]	Fixed Monocular		✓		✓		305
CASIA [8]	Fixed Multiview		✓		✓		124
PKU-HumanID [62]	Fixed Multiview		✓			✓	18
GAVD [44]	Fixed Monocular		✓		✓		300+
Human3.6M [28]	Fixed Multiview	✓			✓		11
GPJATK [34]	Fixed Multiview	✓	✓		✓		32
MoVi [19]	Fixed Multiview	✓	✓		✓		90
3DPW [58]	Monocular Mobile	✓			✓	✓	7
EMDB [30]	Monocular Mobile	✓			✓	✓	10
EgoBody [66]	Egocentric	✓			✓		36
Nymeria-Gait (ours)	**Egocentric**	✓	✓	✓	✓	✓	**54**

Table 7. Comparison of motion datasets.

Table 8 provides participant metadata, including gender, age group, weight, height, and recording location, along with the total indoor and outdoor video duration for each participant.

Sequences: Each selected participant's observer video was manually inspected to identify sequences in which the participant walked while remaining entirely visible with minimal occlusion allowed. Walking sequences were carefully segmented to ensure that they included at least two full gait cycles (mid-stance to mid-stance). Each sequence was then annotated with the observer's perspective (back, side, front, or combined), environment type (indoor or outdoor), terrain slope (flat, up, or down), and trajectory pattern (straight, turn, backward or combined).

Selected frames were then undistorted and devignetted using Project Aria Tools, ensuring a consistent pinhole projection. Additionally, synchronized observer SLAM trajectories and 3D skeletal reference motion were extracted (Figure 4)

To account for potential occlusions caused by undistortion, all processed sequences were reassessed, and an annotation was added to indicate whether the participant's feet were still fully visible.

Segmented sequences and annotations were stored in a structured JSON format. Additionally, a separate JSON file tracks excluded sequences where foot occlusion occurs after undistortion.

Fake Name	Gender	Age Group	Weight (kg)	Height (cm)	Location	Indoor Time (s)	Outdoor Time (s)
douglas_martin	Male	41-45	68.0	175.0	Loc_BX	155.90	133.50
debra_melton	Male	31-35	59.0	155.0	Loc_BX	177.20	115.30
scott_hutchinson	Male	18-24	73.0	174.0	Loc_BX	38.40	232.70
philip_morales	Female	31-35	67.0	165.0	Loc_BX	34.30	293.70
angela_gomez	Female	31-35	59.0	168.0	Loc_BX	42.90	442.00
ethan_jacobson	Male	25-30	69.0	175.0	Loc_BX	77.80	221.80
stacey_lamb	Male	41-45	83.0	183.0	Loc_BX	97.40	505.40
austin_lopez	Female	36-40	89.0	166.0	Loc_BX	128.70	139.60
tamara_gibbs	Female	18-24	59.0	162.0	Loc_BX	10.40	0.00
stephanie_arnold	Female	36-40	55.0	156.0	Loc_BX	82.50	4.00
richard_gallegos	Female	25-30	75.0	174.0	Loc_BX	95.50	643.80
colleen_drake	Male	36-40	82.0	182.0	Loc_BX	9.10	3.30
kevin_shaw	Male	36-40	68.0	176.0	Loc_BX	6.80	135.00
bobby_griffith	Female	25-30	99.0	161.0	Loc_BX	63.90	213.20
andrew_taylor	Female	36-40	89.0	159.0	Loc_BX	161.50	818.10
jessica_webster	Female	25-30	69.0	167.0	Loc_BX	7.80	359.00
jeremy_lewis	Male	36-40	69.0	166.0	Loc_BX	38.10	0.00
justin_martin	Female	36-40	59.0	162.0	Loc_BX	0.00	260.00
suzanne_romero	Male	18-24	83.0	188.0	Loc_BX	77.40	0.00
xavier_norris	Female	41-45	64.0	155.0	Loc_BX	45.70	0.00
tasha_lee	Female	31-35	52.0	162.0	Loc_BX	9.20	33.40
megan_mejia	Male	18-24	75.0	180.0	Loc_BX	38.60	0.00
holly_keller	Male	18-24	81.0	181.0	Loc_BX	78.70	0.00
diane_williams	Female	31-35	59.0	176.0	Loc_BX	69.90	0.00
brady_pearson	Female	36-40	64.0	163.0	Loc_BX	13.40	273.60
patricia_gutierrez	Male	25-30	97.0	174.0	Loc_BX	7.80	0.00
greg_clark	Male	46-50	88.0	164.0	Loc_BX	19.90	0.00
alec_meza	Female	25-30	59.0	155.0	Loc_BX	32.60	356.20
emily_farmer	Male	18-24	82.0	183.0	Loc_BX	4.00	225.80
kurt_young	Male	46-50	73.0	175.0	Loc_38	20.10	73.20
anthony_chen	Female	18-24	50.0	163.0	Loc_BX	0.00	40.50
michael_vargas	Female	36-40	59.0	168.0	Loc_42	19.70	16.60
steven_vang	Female	18-24	49.0	167.0	Loc_42	30.40	4.80
frank_hayden	Female	25-30	59.0	155.0	Loc_41	19.60	75.90
anna_chambers	Male	31-35	68.0	191.0	Loc_41	14.90	53.40
randy_martin	Female	31-35	58.0	158.0	Loc_40	22.10	65.20
logan_walton	Male	46-50	88.0	164.0	Loc_40	26.30	54.60
dean_krause	Male	41-45	73.0	172.0	Loc_36	51.70	74.20
jason_brown	Female	41-45	67.0	165.0	Loc_36	61.20	19.10
angela_garcia	Female	18-24	46.0	149.0	Loc_36	5.00	14.00
stacie_cross	Male	41-45	62.0	159.0	Loc_35	70.70	3.30
arthur_byrd	Male	46-50	88.2	164.0	Loc_35	6.10	9.50
alicia_drake	Male	31-35	60.0	180.0	Loc_35	62.80	2.00
shelley_jones	Female	36-40	62.0	158.0	Loc_33	28.70	6.80
amy_rosales	Male	31-35	62.0	160.0	Loc_35	55.20	2.40
lucas_flores	Female	31-35	67.0	170.0	Loc_33	29.50	19.80
johnathan_good	Female	31-35	65.0	171.0	Loc_33	10.30	49.90
daniel_kim	Female	31-35	69.0	167.0	Loc_33	107.50	108.20
kirk_flowers	Male	18-24	70.0	168.0	Loc_31	39.80	3.40
corey_coleman	Female	25-30	61.0	157.0	Loc_31	8.90	12.40
ashley_reyes	Male	41-45	69.0	169.0	Loc_31	11.20	15.10
glenn_richardson	Male	25-30	70.0	170.0	Loc_31	19.80	0.00
david_vega	Female	36-40	66.0	181.0	Loc_31	12.80	6.10
clayton_bradley	Male	25-30	64.0	178.0	Loc_32	23.40	21.70

Table 8. Participant metadata including gender, age group, weight, height, recording location, and total indoor/outdoor video time.

6. Method Evaluation and Selection

This appendix reports additional experiments conducted to select a representative state-of-the-art configuration for observer-egocentric gait analysis. We show here two monocular human motion reconstruction methods, WHAM [48] and TRAM [60], and several variants that differ in their use of camera motion estimation. These experiments are used solely to motivate the method choice adopted in the main paper, not to establish a comprehensive benchmark.

6.1. Methods

6.1.1. Preprocessing:

For the comparison of the state-of-the-art monocular 3D pose estimation models, we first selected participants from the Nymeria-Gait dataset who executed the *HIKE* scenario at location *loc_BX*. In this scenario, participants hike through the woods on various trails, including both easy flat paths and moderately hilly terrain. *HIKE* is especially relevant for the model comparisons because this scenario provides continuous gait sequences over long periods of time and it is the only scenario with sequences involving both uphill and downhill walking. These variations contribute to the study's novelty, variety, and importance by considering realistic and diverse gait conditions in addition to simple flat walking.

After selecting the sequences, we used a two-stage filtering process to ensure evaluation consistency. First, we excluded any sequences with foot occlusions in the observer's video because our goal is to evaluate model precision under normal conditions rather than robustness to occlusion. By removing these cases, we ensure that model performance is evaluated without any additional confounding variables. Second, we removed sequences in which any tested method failed to produce outputs for 100% of frames, ensuring that all approaches were tested under the same conditions. These filtering steps remove sources of variation unrelated to model accuracy.

This results in a final evaluation dataset of 9,223 events and 3,952 gait cycles for analysis among the 28 identified participants for the present model comparisons.

6.1.2. Evaluation Setup:

Evaluation Metrics: Temporal Gait Errors: We evaluate the accuracy of heel strike and toe-off detection (ETE, in ms) and the event detection rate (DR, in %), which measures how frequently estimated gait events match reference events. Using these detected events, we calculate errors in stride time (ETE), stance time (StTE), swing time (SwTE), and double stance time (DSE) in milliseconds.

Spatial Gait Errors: We evaluate stride length (SL) and walking speed errors in absolute terms (m and m/s, respectively), while total distance traveled errors are reported as a percentage of the reference total distance to normalize variations across sequences.

Trajectory Accuracy: We use Root Trajectory Error (RTE, in%), as described in previous work ([48, 60]). RTE quantifies the differences between estimated and reference trajectories following rigid alignment without scaling.

Comparison of Methods: Nymeria-Gait offers a framework for evaluating monocular 3D pose estimation in camera and world coordinates. TRAM's **VIMO** model and WHAM using VitPose ([64]) (**WHAM(ViT)**) use the camera coordinate system to assess pose reconstruction accuracy and temporal gait consistency without requiring global positioning.

To analyze trajectory estimation strategies, we evaluate five variations of methods for reconstructing motion in world coordinates (Figure 5). **TRAM** and **VIMO-ARS** use Masked DROID-SLAM [52] with scale estimation and Aria SLAM, respectively, to reconstruct global motions. **WHAM** uses DPVO [53], whereas **WHAM-ARAV** refines angular velocity using Aria SLAM. **WHAM-ARS** replaces WHAM's trajectory refinement network with Aria SLAM camera poses to enable direct world-space alignment.

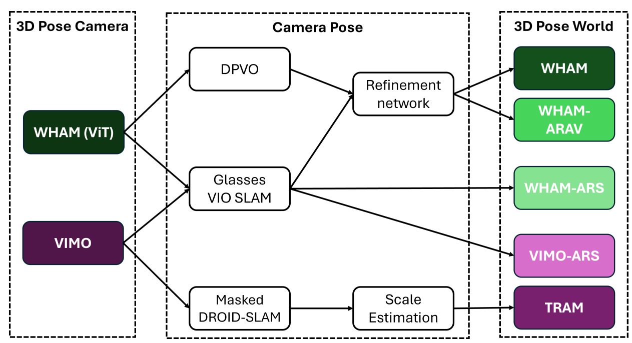

Figure 5. Variations of methods for reconstructing motion in world coordinates.

Method	ETE	DR	STE	StTE	SwTE	DSE
VIMO	10.2 ± 47.9	94.6	-0.9 ± 47.2	-43.0 ± 61.6	42.1 ± 61.2	-44.7 ± 57.1
WHAM (ViT)	8.8 ± 38.7	94.7	-0.3 ± 43.0	-3.8 ± 48.1	3.6 ± 45.9	-4.2 ± 43.7

Table 9. Evaluation of event-related gait analysis metrics. Ordered by Event Timing Error (ETE). Metrics: ETE (Event Timing Error, ms), DR (Detection Rate, %), STE (Stride Time Error, ms), StTE (Stance Time Error, ms), SwTE (Swing Time Error, ms), DSE (Double Support Error, ms). The table presents the mean and standard deviation for each metric.

Method	Speed	Distance	SL	RTE
TRAM	0.38 ± 0.43	37.8 ± 43.4	156 ± 454	12.7
WHAM	-0.01 ± 0.15	-1.9 ± 16.3	-21 ± 232	4.2
WHAM-ARAV	-0.01 ± 0.15	-1.8 ± 16.2	-21 ± 232	4.2
VIMO-ARS	-0.03 ± 0.04	-2.8 ± 4.1	2 ± 97	1.0
WHAM-ARS	-0.03 ± 0.04	-3.2 ± 3.5	0 ± 83	0.8

Table 10. Evaluation of different methods based on gait parameters, including Speed (m/s), Distance (%), Stride Length (SL, in mm), and Relative Trajectory Error (RTE, in %). The table presents the mean and standard deviation for each type of error.

6.2. Results and Ablation Study

Our results are designed as an ablation study to isolate the role of SLAM in gait estimation accuracy. We compare WHAM and TRAM with and without glasses's SLAM integration to determine how it affects trajectory accuracy and gait metrics.

6.2.1. Quantitative Results:

Table 9 compares event-based gait analysis between WHAM(ViT) and TRAM's VIMO models, using the camera coordinate system. Metrics include detection rate (DR), event timing error (ETE), and timing errors for stride, stance, swing, and double support phases.

Table 10 assesses spatial gait parameters errors for each method, including walking speed, total distance, stride length, and Root Trajectory Error (RTE) in world coordinates. Figure 6 focuses specifically on walking speed, illustrating the comparison between TRAM and VIMO-ARS, as well as between WHAM and WHAM-ARS, and highlighting the effect of integrating visual-inertial SLAM from the glasses.

Environments and Slope:

Nymeria-Gait supports condition-specific evaluation of model performance under different walking environments and terrain slopes. Table 11 shows gait parameter errors (distance and stride length) for each model across five conditions: indoor, outdoor, uphill, downhill, and flat terrain.

6.2.2. Conclusion

Across all evaluated configurations, WHAM-ARS achieved the most consistent performance across temporal gait events, spatial gait parameters, and trajectory accuracy. Based on these results, WHAM-ARS was selected for all experiments in the main paper.

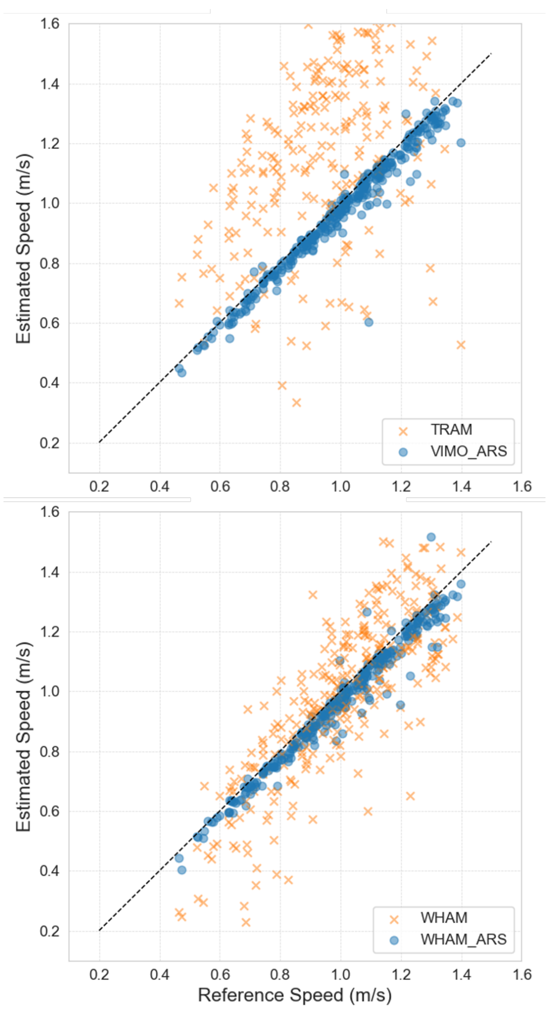

Figure 6. Comparison of speed estimation.

Models	Distance (%)					Stride Length (mm)				
	Indoor	**Outdoor**	**Up**	**Down**	**Flat**	**Indoor**	**Outdoor**	**Up**	**Down**	**Flat**
TRAM	43.8 ± 41.0	36.7 ± 45.8	28.6 ± 31.3	32.7 ± 37.9	41.4 ± 48.3	52 ± 499	158 ± 446	9 ± 570	233 ± 344	134 ± 455
WHAM	5.1 ± 12.7	-5.3 ± 15.8	-3.7 ± 14.9	-6.3 ± 14.1	-2.4 ± 16.2	84 ± 202	-37 ± 219	-18 ± 209	-73 ± 215	13 ± 222
WHAM-ARAV	5.1 ± 12.7	-5.2 ± 15.8	-3.6 ± 14.9	-6.3 ± 14.1	-2.4 ± 16.1	84 ± 202	-37 ± 219	-18 ± 208	-72 ± 214	14 ± 222
VIMO-ARS	-2.1 ± 4.4	-2.8 ± 2.6	-3.5 ± 2.8	-3.4 ± 2.5	-2.4 ± 3.2	3 ± 77	0 ± 94	-2 ± 106	0 ± 90	2 ± 85
WHAM-ARS	-2.5 ± 3.2	-3.8 ± 3.1	-4.4 ± 3.1	-4.3 ± 2.7	-3.2 ± 3.2	0 ± 67	-1 ± 82	-1 ± 92	-1 ± 83	0 ± 74

Table 11. Comparison of different models based on gait parameters across various environments. The table presents the errors (mean and standard deviation) for distance (in percentage) and stride length (in millimeters) under five different conditions: indoor, outdoor, uphill, downhill, and flat terrain.