Application of video frame interpolation to markerless, single-camera gait analysis

DUNN, Marcus <http://orcid.org/0000-0003-3368-8131>, KENNERLEY, Adam, MURRELL-SMITH, Zhane, WEBSTER, Kate <http://orcid.org/0000-0002-5573-3688>, MIDDLETON, Kane <http://orcid.org/0000-0002-4914-8570> and WHEAT, Jonathan <http://orcid.org/0000-0002-1107-6452>

Sports Engineering (2023) 26:22
https://doi.org/10.1007/s12283-023-00419-3

TECHNICAL NOTE

Application of video frame interpolation to markerless, single-camera gait analysis

Marcus Dunn[1] · Adam Kennerley[1] · Zhane Murrell-Smith[1] · Kate Webster[2] · Kane Middleton[2] · Jon Wheat[1]

Accepted: 20 April 2023

Abstract

In clinic settings, factors such as time, cost, expertise, and technology feasibility limit the use of instrumented biomechanical analysis. Recent advances in commercial markerless motion capture systems can address patient ease-of-use factors, but are high cost and require specialised equipment, dedicated spaces, and technical expertise. As such, they present similar limitations to biomechanical analyses in clinic settings. Single-camera pose estimation techniques have generated cautious optimism for markerless gait analysis. However, parameters derived using low-cost and low-sample rate cameras commonly used in clinic settings are not yet accurate enough to detect change in complex movement systems. Video frame interpolation is a single-step process that artificially increases the sample rate of videos. This study applied video frame interpolation to videos of walking and demonstrates improved precision for step, stance, swing and double support times, as well as marginal improvements to the precision of ankle and knee joint angles, derived by single-camera pose estimation. Video frame interpolation potentially represents a delimiting factor for gait analysis in clinic settings, as limiting factors such as time, cost, technology feasibility and patient ease-of-use can be minimised.

1 Introduction

The early identification of change within an individual's walking pattern is an important factor in movement health monitoring. In clinic settings, factors such as time, cost, expertise, and technology feasibility limit the use of instrumented biomechanical analysis (e.g., patient-worn markers or sensors). Therefore, patient self-reporting or practitioner observations or both are often relied upon. However, self-reporting bias as well as intra- and inter-clinician reliability can limit understanding [1]. Pose estimation is a computer vision technique that estimates body landmark locations in camera images without markers [2]. Recent advances in commercial markerless motion capture systems, which adopt pose estimation, can address patient ease-of-use [3] and measurement precision concerns [4]. However, such systems require multiple synchronised and precisely calibrated video streams. As such, they are high cost, require specialised equipment, dedicated spaces, and technical expertise. Therefore, commercial markerless systems present similar limitations to traditional biomechanical analyses. Recently, single-camera markerless pose estimation techniques [5, 6] have been used to quantify various parameters of human movement [2], including spatiotemporal parameters of gait [7]. For a detailed review of popular pose estimation

This article is a part of a Topical Collection in Sports Engineering on the Engineering of Sport 14 Conference held at Purdue University USA, edited by Dr Hugo Espinosa, Steven Shade, Dr Kim Blair and Professor Jan-Anders Månsson.

✉ Marcus Dunn
M.Dunn@shu.ac.uk

Adam Kennerley
A.Kennerley@shu.ac.uk

Zhane Murrell-Smith
Z.Murrell-Smith@shu.ac.uk

Kate Webster
K.Webster@latrobe.edu.au

Kane Middleton
K.Middleton@latrobe.edu.au

Jon Wheat
J.Wheat@shu.ac.uk

[1] Academy of Sport and Physical Activity, Sheffield Hallam University, Sheffield, UK

[2] School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, Australia

Published online: 28 April 2023

methods, readers are directed to several recent reviews [8–10].

The accuracy of two-dimensional joint positions derived by pose estimation has generated cautious optimism for human movement analysis applications [1–3]. Recent work [7] demonstrated the potential clinical utility of temporal gait parameters such as stance, swing and step time. Stenum et al. [7] reported mean absolute step time differences of 0.02 s (i.e., pose estimation.vs motion capture). When compared to minimal detectable change (MDC) values reported elsewhere (e.g., 0.02 – 0.08 s [11, 12]), Stenum et al. [7] suggested that pose estimation could detect change in healthy walking. Analysis of the structure of movement variability–afforded by nonlinear analyses–can provide insight into change within complex movement systems [13]. Whilst the accuracy of temporal gait parameters derived using single-camera pose estimation is encouraging, estimates are not precise enough to detect change in complex movement systems using nonlinear analyses. For example, whilst 25 Hz video was used by Stenum et al. [7], research has demonstrated that a sample rate of 120 Hz would be necessary to maintain sufficient precision to calculate sample entropy for walking parameters [14]. Efforts to advance the precision and application of two-dimensional pose estimation have considered the validity of joint location estimates and development of biomechanical models [2]. However, the application of two-dimensional pose estimation for nonlinear analyses of gait particularly in clinic settings has not been addressed. Clinic settings are often limited by feasibility factors such as cost, time, expertise, and dedicated spaces; cameras commonly used are flexible but low-cost and low-sample rate (e.g., 25–30 Hz). Gait parameters derived by single-camera pose estimation might make movement health monitoring in clinic settings more feasible; however, measurement precision must first be addressed.

Video frame interpolation (VFI) is a method to artificially increase video sample rate by estimating flow between intermediate video frames [15]. The Real-time Intermediate Flow Estimation (RIFE) VFI algorithm [15] estimates bidirectional optical flow between video frames, then scales and reverses flow to approximate intermediate flow. A neural network iteratively refines flow estimates from coarse-to-fine resolutions, improving processing time and allowing interpolation between arbitrary timesteps [15]. Huang et al. [15] assessed the performance of RIFE VFI with respect to video quality, reporting that it outperformed other methods based on peak signal-to-noise ratio, structural similarity, and computation speed. Regarding gait analysis, VFI was recently applied to videos of walking to assess its effect on temporal gait parameters derived by pose estimation [16]. Improved foot contact and step time estimates were reported (e.g., root-mean square error was ~ 55 and ~ 34% lower than estimates derived using original videos, respectively) and it was noted that VFI might represent a simple approach to improving the precision of gait parameters derived by pose estimation in clinic settings. However, the application of VFI, to derive commonly used gait parameters (e.g., ankle and knee joint angle, stance, step, swing and double support time) using single-camera pose estimation, has not been addressed. The aim of this study was to demonstrate the application of VFI for gait analysis derived by two-dimensional pose estimation, using single-camera video.

2 Method

The Research Ethics Committee of Sheffield Hallam University approved (ER: 43,285,879) the secondary analysis of a publicly available dataset [17]. The dataset comprised synchronised three-dimensional motion capture data and digital video recordings of overground walking for 32 healthy participants (22 men, 10 women). The current study analysed 's1' videos, which comprised straight line walking trials (~ 6 m) and include gait initiation and termination. Data for four participants were excluded owing to labelling inconsistencies; thus, motion capture and video data were analysed for 28 participants (one trial per participant). We aimed to replicate analyses of Stenum et al. [7] to allow comparison.

2.1 Data capture

Three-dimensional motion capture data were obtained using ten Vicon MX-T40 cameras (2,352 × 1,728 pixels, 100 Hz) and two-dimensional digital videos were captured using four Basler Pilot (piA1900-32gc, Ahrensburg, Germany) RGB cameras (960 × 540 pixels, 25 Hz). RGB cameras were mounted on tripods (elevation: ~ 1.3 m) to capture sagittal (left- and right-sides) and frontal (front- and rear-sides) perspectives of walking; only sagittal camera views (left- and right-camera distances: ~ 3.3 m) were analysed in this study (refer to [17] for a schematic). Motion capture and video data were synchronised using Vicon MX Giganet.

2.2 Motion capture data analysis

Motion capture and joint angle data, previously processed as part of the public dataset [17], were imported into Matlab (R2022a, The Mathworks, USA). Gait events were defined as described in Zeni et al. [18], where positive and negative peaks in the horizontal coordinate of ankle-pelvis vectors define heel-strike and toe-off, respectively. Step time, stance time, swing time, and double support time were defined as follows:

– Step time: duration between consecutive contralateral heel-strike events (s).

– Stance time: duration between ipsilateral heel-strike and toe-off events (s).
– Swing time: duration between ipsilateral toe-off and heel-strike events (s).
– Double support time: duration between heel-strike and contralateral toe-off events (s).

Finally, processed ankle and knee angle data were trimmed to identified gait cycles (heel-strike to heel-strike) and normalised to 100 data points.

2.3 Video post-processing, pose estimation and data analysis

Google Colaboratory (four Intel(R) Xeon(R) CPUs @ 2.20 GHz, Tesla T4 GPU, 26 Gb RAM) was used to process original videos using Python notebooks via a web browser. RIFE VFI (code supplied by Huang et al. [15]) was applied to original videos to provide artificially upsampled walking videos (960 × 540 pixels, 100 Hz). Figure 1 illustrates walking images extracted from original

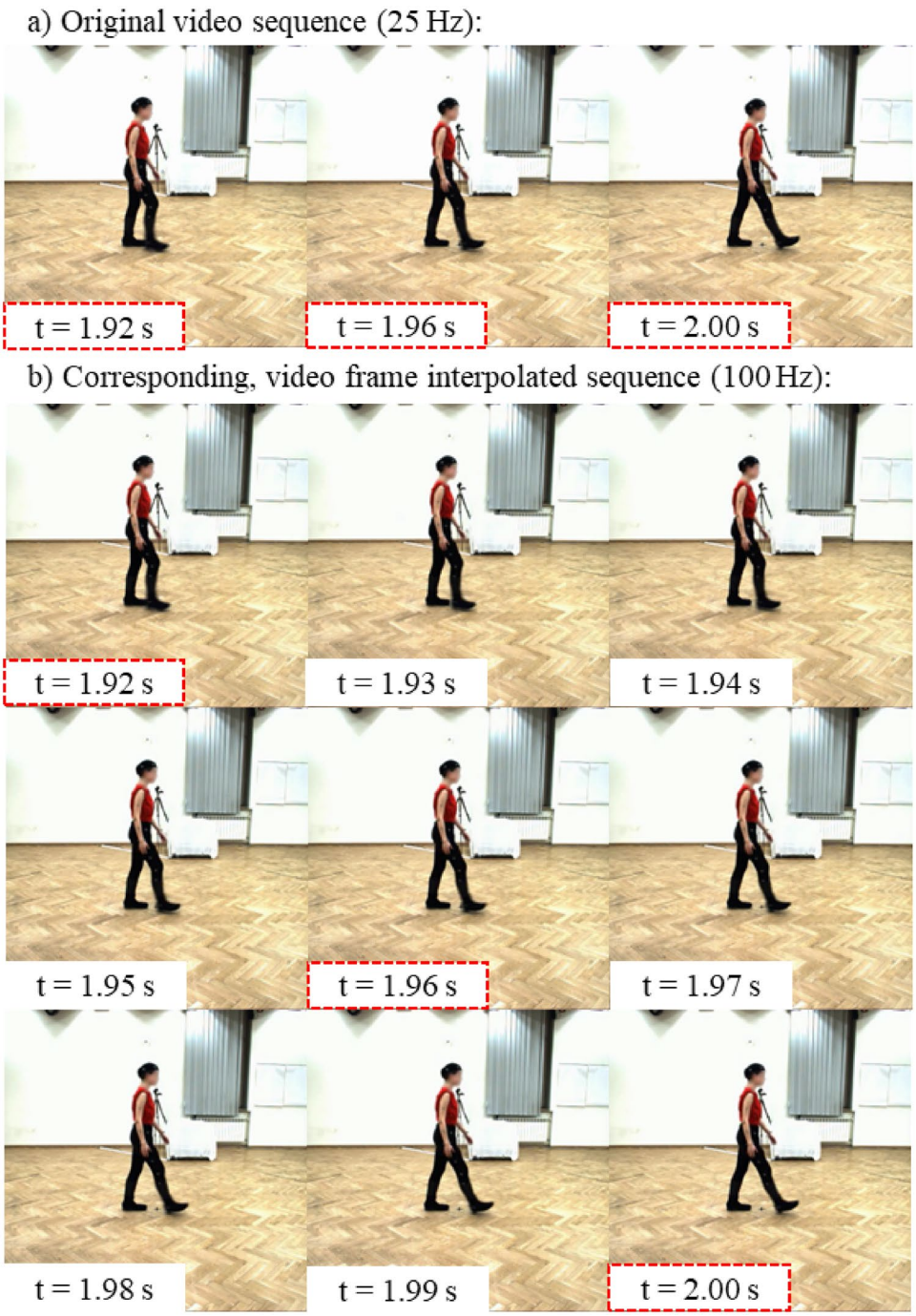

Fig. 1 Walking sequence extracted from original video at 25 Hz (**a**, top row) and the corresponding sequence extracted from VFI video at 100 Hz (**b**, bottom three rows). Timecodes highlighted in red indicate original frames and unhighlighted timecodes indicate interpolated frames

(A) and VFI (B) video at 25 and 100 Hz, respectively, for a 0.08 s time-window.

Subsequently, two-dimensional pose estimation (Body_25 keypoint model) using OpenPose (code supplied by Cao et al. [5]) was performed using Google Colaboratory for the original and VFI videos (example Python notebook for VFI and pose estimation available at: https://github.com/marcusdunn-phd/VideoPostProcessing_PoseEstimation). Subsequently, two-dimensional image coordinates of joint landmarks were imported into Matlab and analysis code supplied by Stenum et al. [7] used to perform the following: (1) manually correct frames with incorrectly identified left and right legs, (2) fill gaps (linear interpolation for gaps spanning up to two video frames) in joint coordinate data (e.g., frames where joints were not detected), and (3) filter joint coordinate data using a zero-lag, 4th order low-pass Butterworth filter (5 Hz cut-off frequency). Videos were not spatially calibrated as in Stenum et al. [7], owing to (1) unmeasured scaling reference points and (2) photogrammetric errors associated with image scaling and lens distortion [19, 20]. The lack of spatial calibration limits the assessment of step length data in this study; however, ankle and knee joint angles were assessed and were defined as follows [7]:

- Knee flexion–extension angle: vectors between the hip and knee joints, and knee and ankle joints, where positive and negative angles represent flexion and extension, respectively.
- Ankle dorsi-plantar flexion: vectors between the knee and ankle joints, and ankle joint and big toe location, where positive and negative angles represent dorsiflexion and plantarflexion, respectively.

Gait events, step time, stance time, swing time, and double support time were defined as described in Sect. 2.2. Ankle and knee angle data were trimmed to the identified gait cycles (heel-strike to heel-strike) and normalised to 100 data points. Finally, 99% confidence intervals were constructed for all normalised ankle and knee angle gait cycles (n = 1102). Gait cycles containing joint angles that exceeded confidence intervals (n = 40) were excluded from further analysis.

2.4 Statistical analyses

For step time, stance time, swing time, and double support time, agreement was assessed using Bland and Altman 95% Limits of Agreement (LOA). In the case of heteroscedasticity (i.e., $|r^2| > 0.1$), ratio LOA (dimensionless) were also reported. Further, root-mean square error (RMSE) was calculated. For ankle and knee angles, 95% Functional Limits of Agreement (FLOA: [21]) and RMSE were calculated for gait cycle normalised joint angles.

3 Results

Table 1 presents LOA, Ratio LOA, and RMSE for temporal gait parameters. Left and right cameras yielded 166 and 164 step times, respectively, and 138 and 136 stance, swing, and double support times, respectively. Systematic errors for VFI videos were between 0.001 and 0.004 s lower than original videos for all temporal gait parameters except double support times derived by the left camera, which were 0.001 s greater. Random errors for VFI videos were between

Table 1 LOA, Ratio LOA, and RMSE for step, stance, swing, and double support times, using original (25 Hz) and VFI (100 Hz) videos of walking

Camera	N	LOA	r^2	Ratio LOA	RMSE
Origional					
Step time (s)					
Left	166	− 0.001 ± 0.046	0.103	1.00 (×/÷ 1.080)	0.024
Right	164	0.000 ± 0.038	0.092	–	0.019
Stance time (s)					
Left	138	− 0.011 ± 0.052	0.283	0.987 (×/÷ 1.071)	0.029
Right	136	− 0.009 ± 0.047	0.094	–	0.025
Swing time (s)					
Left	138	0.011 ± 0.050	0.034	–	0.027
Right	136	0.009 ± 0.047	0.177	1.020 (×/÷ 1.109)	0.025
Double support time (s)					
Left	138	− 0.009 ± 0.051	0.278	0.947 (×/÷ 1.440)	0.028
Right	136	− 0.009 ± 0.050	0.240	0.954 (×/÷ 1.433)	0.027
Video frame interpolation					
Step time (s)					
Left	166	− 0.001 ± 0.032	0.160	0.999 (×/÷ 1.054)	0.016
Right	164	0.000 ± 0.026	0.287	1.000 (×/÷ 1.041)	0.013
Stance time (s)					
Left	138	− 0.011 ± 0.035	0.212	0.986 (×/÷ 1.046)	0.021
Right	136	− 0.005 ± 0.035	0.181	0.994 (×/÷ 1.049)	0.019
Swing time (s)					
Left	138	0.010 ± 0.036	− 0.001	–	0.021
Right	136	0.005 ± 0.039	0.114	1.011 (×/÷ 1.091)	0.020
Double support time (s)					
Left	138	− 0.010 ± 0.035	0.251	0.933 (×/÷ 1.270)	0.020
Right	136	− 0.005 ± 0.037	0.225	0.967 (×/÷ 1.305)	0.020

0.008 and 0.017 s lower than original videos for all temporal parameters derived by left and right cameras. RMSE for VFI videos were between 0.005 and 0.008 s lower than original videos for all temporal parameters derived by left and right cameras.

Table 2 presents FLOA (systematic and random error components are mean values for the gait cycle) and RMSE for ankle and knee joint angles. Normalised gait cycles included in analyses exist within 99% confidence intervals; Table 2 presents the number of gait cycles (GC) included in each comparison. Systematic errors (mean for gait cycle) for VFI videos were between 0.11 and 1.38° lower than original videos for all joint angles, except right ankle and knee angles derived by the right camera, which were between 0.27 and 1.05° greater. RMSE (mean for gait cycle) for VFI videos were between 0.1 and 2.7° lower than original videos for all joint angles, except left ankle angles estimated by the right camera (0.1° greater; Table 2).

Random errors (mean for gait cycle) for VFI videos were between 0.08 and 3.05° lower than original videos for all joint angles, except left knee angles estimated by the right camera, which were 0.11° greater. Figure 2 and 3 present mean and standard deviation ankle and knee joint angle data and corresponding FLOA throughout the gait cycle. FLOA visualises varying agreement for ankle and knee joint angles derived from the original (Fig. 2) and VFI (Fig. 3) videos throughout the gait cycle.

Table 2 FLOA and RMSE for left and right ankle and knee angles (gait cycle), using original (25 Hz) and VFI (100 Hz) videos of walking (GC is the number of gait cycles included in each comparison)

Camera	GC	FLOA	RMSE
Origional			
Left ankle angle (°)			
Left	63/68	14.72 ± 20.21	19.25 ± 11.30
Right	64/67	18.00 ± 18.40	20.90 ± 8.53
Right ankle angle (°)			
Left	69/71	14.97 ± 20.29	18.29 ± 8.88
Right	66/71	13.14 ± 20.13	17.46 ± 11.42
Left knee angle (°)			
Left	67/68	− 5.60 ± 19.35	10.85 ± 4.99
Right	66/67	− 4.67 ± 18.47	10.30 ± 4.98
Right knee angle (°)			
Left	71/71	− 7.29 ± 19.15	11.23 ± 5.55
Right	71/71	− 8.32 ± 17.84	11.85 ± 5.24
Video frame interpolation			
Left ankle angle (°)			
Left	64/68	14.41 ± 17.87	16.52 ± 6.59
Right	61/65	17.55 ± 15.95	19.53 ± 6.87
Right ankle angle (°)			
Left	69/70	14.36 ± 17.24	16.57 ± 7.19
Right	68/71	14.19 ± 19.53	16.63 ± 8.31
Left knee angle (°)			
Left	62/68	− 5.49 ± 17.12	10.20 ± 5.40
Right	64/65	− 3.29 ± 18.58	10.40 ± 5.94
Right knee angle (°)			
Left	67/70	− 6.84 ± 17.32	10.84 ± 5.80
Right	70/71	− 8.59 ± 17.76	11.76 ± 5.60

4 Discussion

The application of VFI to artificially upsample videos of walking captured at 25 Hz to 100 Hz demonstrated a general improvement to the precision of temporal gait parameters derived by pose estimation (Table 1). For example, RMSE for step, stance, swing, and double support times derived using VFI videos were lower than all original video analyses. This represents an improvement to RMSE ranging between 0.005–0.008 s, or ~ 20–33%. Further, LOA indicated lower systematic errors for step, stance, and swing times, ranging between 0.001–0.004 s (~ 9–44%). An exception to this was systematic errors for double support time, which increased by 0.001 s (~ 11%), but this was only observed in left camera estimates (Table 1). Importantly, random errors for step, stance, swing, and double support time estimates were reduced by between 0.008–0.017 s (~ 17–33%) for VFI videos, when compared to original videos. The reduction of random errors in temporal parameters is an important factor in gait analysis and particularly in health monitoring applications, where the understanding of step-to-step variation can reveal change in complex movement systems [13].

Heteroscedasticity was present for temporal gait parameters derived using both original and VFI video sequences (Table 1). Previous, comparable work [16] only reported heteroscedasticity for foot contact time, when using 25 Hz video sequences, and did not report heteroscedasticity for any parameters derived using VFI videos. There are several factors that might cause heteroscedasticity in this study, including the temporal resolution of comparator motion capture data, the camera perspective of participants, and details of the foot trajectory post-processing. For example, the motion capture data in the current study have a temporal resolution of 0.01 s, in contrast to 0.005 s in Dunn et al. [16]. Low temporal resolution limits the precision to which the parameters can be assessed and is a known limitation of publicly available datasets of this type [1]. Further, walking distance in the current dataset was 6.3 m and 4 m in Dunn et al. [16]. This potentially yields a wider range of participant-camera perspectives in the current study than in previous research [16]. Linked to this, Stenum et al. [7] demonstrated heteroscedasticity in step length estimates (i.e., step length error related to perpendicular distance between camera and participant). Heteroscedasticity reflects the

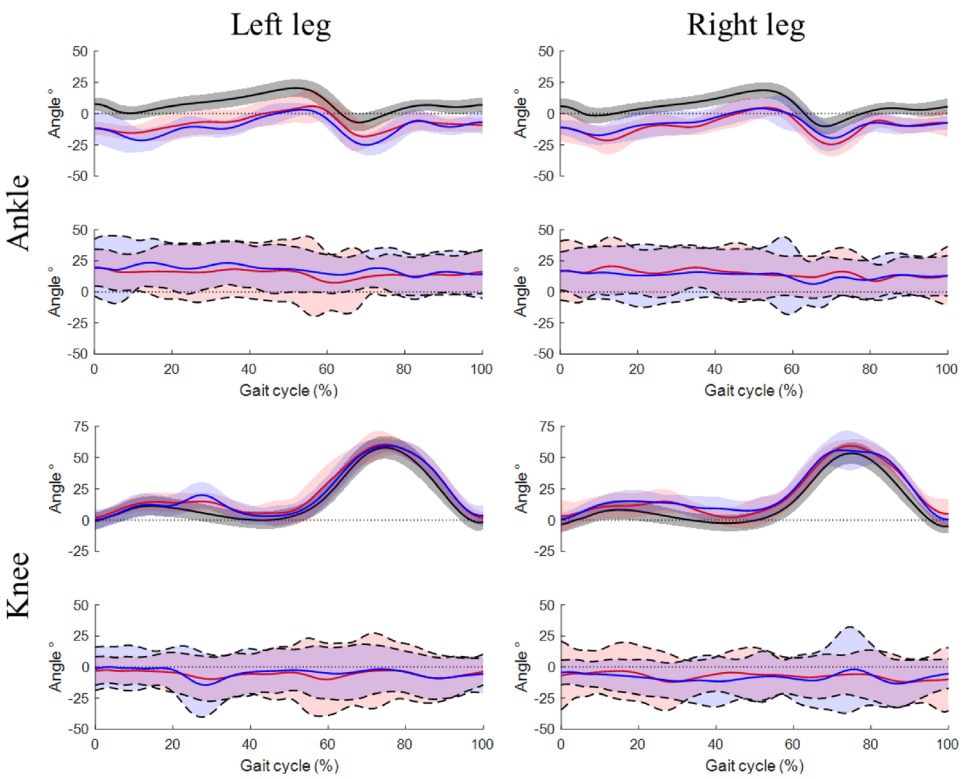

Fig. 2 Ankle and knee joint angle estimates (gait cycle) derived using original videos. Top rows (graph pairs) are mean and standard deviation for motion capture (grey), left camera (red) and right camera (blue) joint angles. Bottom rows (graph pairs) are FLOA for left camera (red) and right camera (blue) joint angles

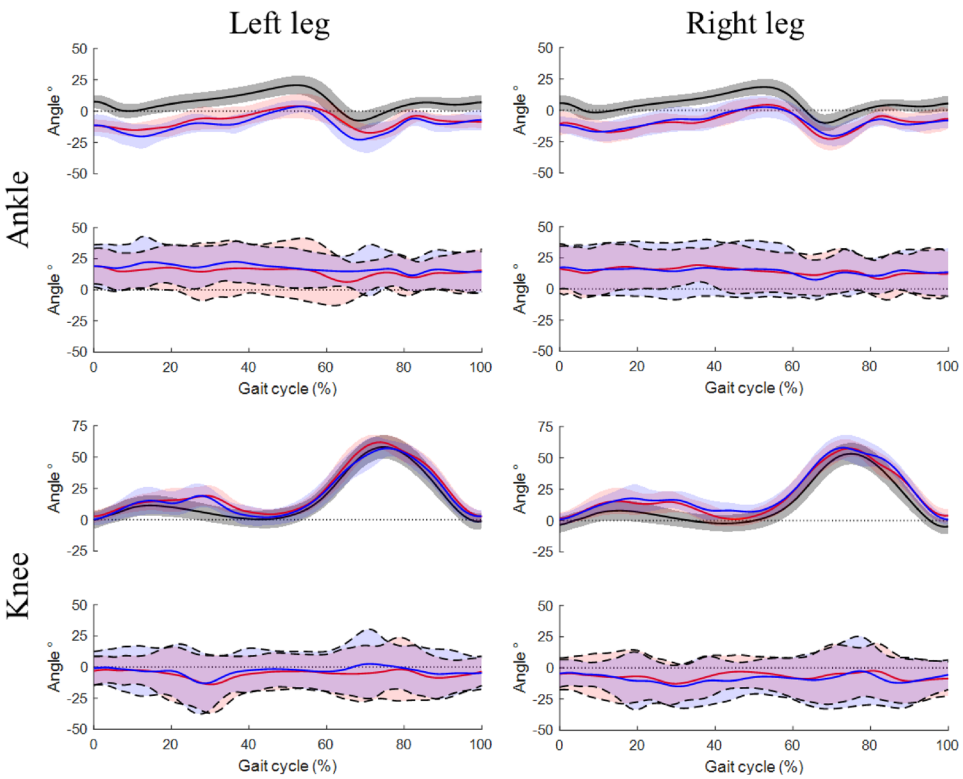

Fig. 3 Ankle and knee joint angle estimates (gait cycle) derived using VFI videos. Top rows (graph pairs) are mean and standard deviation for motion capture (grey), left camera (red) and right camera (blue) joint angles. Bottom rows (graph pairs) are FLOA for left camera (red) and right camera (blue) joint angles

location of participants and their steps in camera images, as well as image scaling techniques used to estimate real-world measurements [19]. Oblique camera perspectives of participants might also affect the precision of gait event detection. This study follows gait event detection in Stenum et al. [7], where foot contact events were defined as the peak

horizontal coordinate in ankle-pelvis vectors [18], whereas the foot (heel and toe midpoint) velocity algorithm [22] was used by Dunn et al. [16]. Further, trajectories were filtered using 7 Hz (2nd order) and 5 Hz (4th order) Butterworth filters in Dunn et al. [16] and the current study, respectively. Thus, foot trajectories in the current study would likely exhibit smoother profiles than in Dunn et al. [16]. Whilst smoother trajectories might aid robustness of automated event detection, the precision of identified events might also be affected. Moreover, the relative performance of trajectory- and velocity-based approaches to gait event detection has been reported to vary for different cohorts in different studies (e.g., [18, 23]). Thus, different gait event detection methods and lower temporal resolution of comparator data might explain inconsistencies between temporal gait parameter error distributions in previous research [16] and the current study.

Regardless, the application of VFI in this study demonstrated a consistent improvement to temporal gait parameters and importantly, reduced random errors when estimating temporal gait parameters. When considering clinical applications of temporal gait metrics, MDC values for step time in healthy walking have been reported to range between 0.02–0.08 s [11, 12]. Stenum et al. [7] suggested that step times derived using pose estimation (mean and maximum errors of 10 ms and 100 ms, respectively) were sensitive to change. The current study demonstrates that VFI further improves the precision of temporal gait parameters such as step time, reducing both systematic and random error components by up to 44 and 33%, respectively (Table 1). The reduction of random errors in particular (i.e., narrower LOA) is an important factor in gait analysis and movement health monitoring. This is because an understanding of the magnitude and structure of step time variation can reveal changes in complex movement systems [9]. Further research into the role of markerless gait analysis and VFI, to understand change in complex movement systems, is necessary and warranted.

Estimating ankle and knee joint angles from VFI videos yielded marginal improvements when compared to original videos (Table 2). For left and right ankle and knee joint angles, FLOA indicated marginally lower systematic errors (0.11–1.38°, or ~ 2–29%), except for higher systematic errors for right ankle and knee angles, derived by the right camera (1.05° and 0.27°, or 8 and 3%, respectively). Random errors for ankle and knee angles were reduced by 0.08–3.05° (0.4–15%), except for marginally increased random errors for left knee angles derived by the right camera (0.11° or 0.6%). Stenum et al. [7] presented mean and standard deviation joint angles at fixed intervals (~ 3%) throughout the gait cycle, as well as mean absolute error (MAE); thus, assessing the magnitude and impact of outlying joint angle data points is difficult. Our analysis of pose estimated joint angles first excludes trials outside 99% confidence intervals (n = 40 or ~ 3.6%). Whilst this yields fewer trials for analysis (Table 2), retaining a confident estimate of 'real' joint angles is a pragmatic approach that can be applied to automatic tools and analyses in clinic settings. Moreover, FLOA aid interpretation of the varying precision of joint angles derived by pose estimation across the gait cycle, using original and VFI videos. For example, systematic differences between motion capture and pose estimated ankle angles reflect foot segment definition. Further, differences within knee joint angles emerge during stance (~ 30%) and swing (~ 80%) phases for ~ 10% of the gait cycle and exist in both original and VFI videos (e.g., Fig. 2, 3). Differences in knee angles might reflect the precision of markerless knee joint estimates during contralateral swing and stance phases (e.g., limb obfuscation or occlusion, or manual post-processing when recreating joint trajectories), as comparable differences were not observed in ankle joint angles. The application of VFI was not anticipated to improve the precision of joint angle estimates, since the spatial resolution of VFI videos remain unchanged. The improvements for VFI video estimates (Table 2) likely reflect trajectory post-processing, where additional video frames and thus pose estimated data (inclusive of different frame-by-frame estimation errors), result in different trajectories to those derived using original videos.

Video quality factors might affect the precision of pose estimation [24]. Huang et al. [15] assessed VFI image quality using peak signal-to-noise ratio and structural similarity. Huang et al. [15] reported that their approach outperformed other VFI methods, except for one closed-sourced software. Results of our study indicate that VFI improved the temporal precision of walking gait analyses derived by pose estimation. This suggests that VFI might benefit other gait analysis applications using pose estimation (e.g., running, sprinting), or other image-based object tracking tasks. However, any assumptions of underlying movements or analyses should be carefully considered. Regarding clinical applications of joint angle data in healthy walking, Wilken et al. [25] reported MDC values of 4–6° and 4° for knee and ankle angles, respectively. Stenum et al. [7] reported MAE for knee and ankle angles derived by pose estimation as 5.6 and 7.4°, respectively. Therefore, and based on [25], Stenum et al. [7] suggested that only knee angle data were sensitive to change in healthy walking. In this study, gait cycle averaged errors were greater than those reported by Stenum et al. [7]. This greater error might reflect the greater sensitivity of RMSE to outlying data points than MAE and might also reflect manual approaches to the post-processing of obfuscated or occluded joint locations. However, when considering joint angles derived using VFI videos, FLOA (Table 2) illustrate a general improvement when compared to joint angles derived using standard videos. This indicates that VFI might improve the clinical application of knee and

ankle angle data derived by pose estimation in healthy walking. Further work is required to clarify this, the effects of interventions to reconstruct joint trajectories, and any subsequent impact in clinical applications.

The use of an existing dataset [17] presents several limitations when understanding the application of VFI to gait analysis. Filming factors (e.g., fixed camera height, distance, and perspective), as well as the temporal precision of motion capture data (e.g., 100 Hz) limit the exploration of VFI for use in the calculation of gait metrics. Moreover, insufficient and inconsistent camera calibration information prohibited the understanding of VFI on the calculation of some spatial gait metrics (e.g., step length), owing to photogrammetric errors associated with image scaling and lens distortion [19, 20]. The dataset presented by [17] was used owing to the wide range of participants, multi-view colour video, and synchronised, three-dimensional motion capture data. Whilst such datasets are extremely valuable, the aforementioned factors are known limitations of datasets of this type [1]. Further, findings in this study are relevant to single-camera pose estimation techniques using RGB images. This is because standard videos are necessary to allow the application of VFI used in this study [15]; other single-camera, markerless motion capture tools, such as the Microsoft Kinect [26–28] are dependent on depth data, limiting the application of VFI in this case [15]. Finally, and whilst this study aimed to replicate processes in Stenum et al. [7], knee angle errors that emerge mid-cycle likely indicate different approaches during manual post-processing of joint trajectories for obfuscated or occluded limbs. Further research is necessary to identify systematic approaches and best practices, for the interrogation and reconstruction of such trajectories. This highlights the necessity for well-defined, systematic solutions to markerless, single-camera gait analysis particularly in clinic settings and reaffirms a key benefit of VFI, as its application might complement automatic gait analysis solutions.

Traditional biomechanical tools have not been successfully implemented in clinic-based practice; this reflects complexities related to technology and data collection protocols [29]. Advances in commercial markerless systems can address patient ease-of-use concerns [3]; however, the space and technological requirements of these systems present similar limitations to traditional biomechanical tools in clinic settings. Single-camera pose estimation represents a promising advance to the feasibility of clinic-based gait analysis [3]. The use of high-speed cameras is a logical route to address sample rate considerations of advanced gait analysis techniques but the feasibility of using high-speed cameras in a clinic is questionable [29]. The current study demonstrated that VFI yields marked improvements to temporal gait parameter estimates, and markedly improves random error components in particular. Moreover, VFI did not detrimentally effect image quality and yields marginal improvements in ankle and knee joint angle estimates. In the context of movement health monitoring, reducing random error within gait parameters is imperative, as the magnitude and structure of movement variability can provide insight into movement health [14].

5 Conclusion

Single-camera pose estimation techniques have generated cautious optimism for markerless gait analysis in clinic settings. However, parameters derived using low-cost and low-sample rate cameras are not yet sensitive to detect change in complex movement systems. This study demonstrated that by upsampling single-camera videos of walking with VFI (from 25 to 100 Hz), the precision of gait parameters derived by pose estimation can be markedly improved. RMSE were improved by up to 33% for step, stance, swing, and double support times, and by up to 8% for knee and ankle joint angles. Our findings represent a novel contribution to markerless gait analysis derived by single-camera pose estimation, as VFI can improve derived gait metrics and be systematically applied to videos captured in clinic settings. VFI therefore represents a delimiting factor for clinic-based gait analysis using pose estimation, as limiting factors associated with traditional analysis techniques (e.g., time, cost, dedicated space, patient ease-of-use, and clinician expertise) can be minimised. However, acceptable precision to monitor change in complex movement systems is not easily defined; research addressing the reliability and sensitivity of markerless single-camera gait analysis is warranted.

Author contribution MD Conceptualisation, Data curation, Formal analysis and investigation, Writing–original draft preparation, Writing–review and editing. AK: Writing–review and editing. ZM-S Formal analysis and investigation, Writing–review and editing. KW, KM, JW, Writing–review and editing.

Funding No funding was received for conducting this study.

Data availability Example Python notebook (code) for video frame interpolation and pose estimation available at: https://github.com/marcusdunn-phd/VideoPostProcessing_PoseEstimation.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethical approval The Research Ethics Committee of Sheffield Hallam University approved (ER: 43285879) the secondary analysis of a publicly available dataset.

Consent for publication All authors have given their consent for publication.

References

1. Wade L, Needham L, McGuigan P, Bilzon J (2022) Applications and limitations of current markerless motion capture methods for clinical gait biomechanics. Peer J 10:12995. https://doi.org/10.7717/peerj.12995
2. Cronin NJ (2021) Using deep neural networks for kinematic analysis: challenges and opportunities. J Biomech 123:110460. https://doi.org/10.1016/j.jbiomech.2021.110460
3. Stenum J, Cherry-Allen KM, Pyles CO, Reetzke RD, Vignos MF, Roemmich RT (2021) Applications of pose estimation in human health and performance across the lifespan. Sensors (Basel) 21(21):7315. https://doi.org/10.3390/s21217315
4. Kanko RM, Laende E, Selbie WS, Deluzio KJ (2021) Inter-session repeatability of markerless motion capture gait kinematics. J Biomec 121:110422. https://doi.org/10.1016/j.jbiomech.2021.110422
5. Cao Z, Hidalgo G, Simon T, Wei SE, Sheikh Y (2021) Openpose: realtime multi-person 2D pose estimation using part affinity fields. IEEE Trans Pattern Anal Mach Intell 43(1):172–186. https://doi.org/10.1109/TPAMI.2019.2929257
6. Mathis A, Mamidanna P, Cury K, Abe T, Murthy V, Mathis M, Bethge M (2018) DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci 21:1281–1289. https://doi.org/10.1038/s41593-018-0209-y
7. Stenum J, Rossi C, Roemmich RT (2021) Two-dimensional video-based analysis of human gait using pose estimation. PLoS Comput Biol 17(4):e1008935. https://doi.org/10.1371/journal.pcbi.1008935
8. Chen Y, Tian Y, He M (2020) Monocular human pose estimation: a survey of deep learning-based methods. Computer Vis Image Underst 192:102897. https://doi.org/10.1016/j.cviu.2019.102897
9. Colyer S, Evans M, Cosker D, Salo A (2018) A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. Sports Med-Open 4:24. https://doi.org/10.1186/s40798-018-0139-y
10 Dang Q, Yin J, Wang B, Zheng W (2019) Deep learning based 2D human pose estimation: a survey. Tsinghua Sci Tech 24:663–676. https://doi.org/10.26599/TST.2018.9010100
11 Meldrum D, Shouldice C, Conroy R, Jones K, Forward M (2014) 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. https://doi.org/10.1016/j.gaitpost.2013.07.130
12. Fernandes R, Armada-da-Silva P, Pool-Goudzwaard AL, Moniz-Pereira V, Veloso AP (2016) Three dimensional multi-segmental trunk kinematics and kinetics during gait: Test-retest reliability and minimal detectable change. Gait Posture 46:18–25. https://doi.org/10.1016/j.gaitpost.2016.02.007
13. Fuller JT, Amado A, Emmerik RE, Hamill J, Buckley JD, Tsiros MD, Thewlis D (2016) The effect of footwear and footfall pattern on running stride interval long-range correlations and distributional variability. Gait Posture 44:137–142. https://doi.org/10.1016/j.gaitpost.2015.12.006
14. Fallahtafti F, Wurdeman SR, Yentes JM (2021) Sampling rate influences the regularity analysis of temporal domain measures of walking more than spatial domain measures. Gait Posture 88:216–220. https://doi.org/10.1016/j.gaitpost.2021.05.031
15. Huang Z, Zhang T, Heng W, Shi B, Zhou S (2020) RIFE: real-time intermediate flow estimation for video frame interpolation. Arxiv preprint arXiv. 2011:06294. https://doi.org/10.48550/arXiv.2011.06294
16. Dunn M, Kennerley A, Webster K, Middleton K, Wheat J (2022) Application of video interpolation to markerless movement analysis. In: Proceedings of the ISEA Engineering of Sport 14 Conference. https://docs.lib.purdue.edu/resec-isea/2022/session07/3/.
17. Kwolek B, Michalczuk A, Krzeszowski T, Switonski A, Josinski H, Wojciechowski K (2019) Calibrated and synchronized multi-view video and motion capture dataset for evaluation of gait recognition. Multimed Tools Appl 78:32437–32465. https://doi.org/10.1007/s11042-019-07945-y
18. Zeni JA Jr, Richards JG, Higginson JS (2008) Two simple methods for determining gait events during treadmill and overground walking using kinematic data. Gait Posture 27(4):710–714. https://doi.org/10.1016/j.gaitpost.2007.07.007
19. Brewin MA, Kerwin DG (2003) Accuracy of scaling and DLT reconstruction techniques for planar motion analyses. J Appl Biomech 19:79–88. https://doi.org/10.1123/jab.19.1.79
20. Dunn M, Wheat J, Miller S, Haake S, Goodwill S (2012) Reconstructing 2D planar coordinates using linear and nonlinear techniques. In: Proceedings of the 30th Annual Conference of Biomechanics in Sports. https://ojs.ub.uni-konstanz.de/cpa/article/view/5310.
21. Pini A, Markström JL, Schelin L (2022) Test-retest reliability measures for curve data: an overview with recommendations and supplementary code. Sports Biomech 21(2):179–200. https://doi.org/10.1080/14763141.2019.1655089
22 O'Connor CM, Thorpe S, O'Malley MJ, Vaughan CL (2007) Automatic detection of gait events using kinematic data. Gait Posture 25(3):469–474. https://doi.org/10.1016/j.gaitpost.2006.05.016
23 Zahradka N, Verma K, Behboodi A, Bodt B, Wright H, Lee SCK (2020) An evaluation of three kinematic methods for gait event detection compared to the kinetic-based gold standard. Sensors (Basel). 20(18):5272
24. Hardy P, Dasmahapatra S, Kim H (2021) Can super resolution improve human pose estimation in low resolution scenarios? Arxiv preprint arXiv. 2107:02108. https://doi.org/10.5220/0010863700003124
25. Wilken JM, Rodriguez K, Brawner M, Darter BJ (2012) Reliability and minimal detectible change values for gait kinematics and kinetics in healthy adults. Gait Posture 35(2):301–307. https://doi.org/10.1016/j.gaitpost.2011.09.105
26. Ng K-D, Mehdizadeh S, Iaboni A, Mansfield A, Flint A, Taati B (2020) Measuring gait variables using computer vision to assess mobility and fall risk in older adults with dementia. IEEE J Translat Eng Health Med 8:1–9. https://doi.org/10.1109/JTEHM.2020.2998326
27. Kastaniotis D, Theodorakopoulos I, Theoharatos C, Economou G, Fotopoulos S (2015) A framework for gaitbased recognition using Kinect. Pattern Recognit Letters 68(2):327–335. https://doi.org/10.1016/j.patrec.2015.06.020
28. Mentiplay B, Perraton L, Bower K, Pua Y-H, McGaw R, Heywood S, Clark R (2015) Gait assessment using the microsoft Xbox one kinect: concurrent validity and inter-day reliability of

spatiotemporal and kinematic variables. J Biomech 48(10):2166–2170. https://doi.org/10.1016/j.jbiomech.2015.05.021
29. Gokeler A, Dingenen B, Mouton C, Seil R (2017) Clinical course and recommendations for patients after anterior cruciate ligament injury and subsequent reconstruction: A narrative review. EFORT Open Rev 2(10):410–420. https://doi.org/10.1302/2058-5241.2.170011