Application of video interpolation to markerless movement analysis

Marcus Dunn[1], Adam Kennerley[1], Kate Webster[2], Kane Middleton[2] and Jon Wheat[1]
[1] Sheffield Hallam University, Sheffield, UK.
[2] La Trobe University, Melbourne, Australia.

Introduction

Markerless video analysis techniques, such as human posture extraction, could address technology complexity limitations of clinic-based movement analysis. Commercial markerless systems often require multiple, precisely calibrated and synchronised video streams. These systems are high cost, require specialised equipment, dedicated spaces, and technical expertise. Single-camera posture extraction[1] has quantified spatiotemporal parameters of gait[2]. However, the identification of events, to determine informative measures such as step time variation, is not precise enough for movement health monitoring. Video quality (e.g., resolution, frame rate) can affect posture extraction accuracy. Video frame interpolation (VFI) artificially increases frame rate by estimating flow between intermediate frames[3]. Whilst VFI does not provide new information, VFI might mitigate factors affecting trajectory post-processing. This study assessed whether VFI can improve markerless step time estimation using a single camera.

Method

The Research Ethics Committee of Sheffield Hallam University approved the secondary analysis of a dataset[4]. Sagittal plane videos (25 Hz) of walking (n = 36) were processed using RIFE video frame interpolation[3] to derive 100 Hz videos. Posture extraction[1] was performed for original (Org) and VFI videos. In addition, trajectories based on Org video posture data were linearly interpolated (LI) to 100 Hz. Concurrent motion capture (MoCap) data and markerless trajectory data were filtered (2nd order low-pass Butterworth filter, 7 Hz cut-off). Foot contact times were identified in MoCap (200 Hz), Org (25 Hz), LI (100 Hz) and VFI (100 Hz) data[5]. Step time was defined as time between contralateral foot contacts. Org, LI and VFI estimates for foot contact time and step time were compared to MoCap using 95% Limits of Agreement (LOA), ratio LOA (dimensionless) was reported in the case of heteroscedasticity (i.e., $|r^2| \geq 0.1$). Root-mean square error (RMSE) was calculated.

Results

LOA for foot contact times were 0.028 ± 0.075 s, -0.004 ± 0.043 s, and -0.003 ± 0.041 s for foot contacts identified using Org, LI and VFI videos, respectively. Foot contact times identified using Org videos were heteroscedastic ($|r^2| = 0.11$); 95%

of ratios were between 2.5% of the mean ratio. RMSE for corresponding foot contact time estimates was 0.047 s, 0.022 s and 0.021 s, respectively. LOA for step time estimates were 0.002 ± 0.092 s, 0.001 ± 0.031 s, and 0.001 ± 0.027 s for step times identified using Org, LI and VFI videos, respectively. RMSE for corresponding step time estimates was 0.021 s, 0.016 s and 0.014 s, respectively.
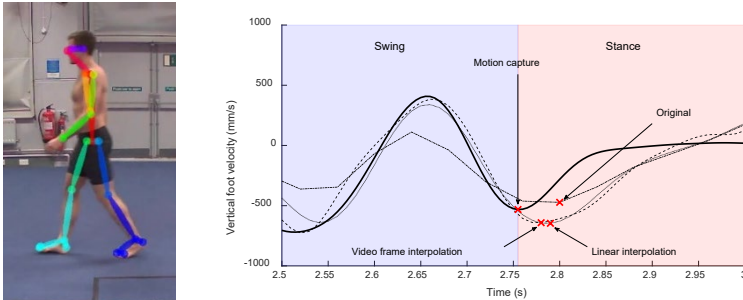

Figure 1. Left: Posture extraction. Right: MoCap (solid), Org (dash-dotted), LI (dotted) and VFI (dashed) derived foot velocities and estimated foot contact event (red crosses).

Discussion and conclusion

LI and VFI markedly improved foot contact and step time estimates, corresponding to a 52.8% and 55.6% (foot contact time), and 24.2% and 33.8% (step time) reduction in RMSE, respectively. Moreover, LI and VFI reduced random errors associated with foot contacts (42.5% and 65.9%, respectively) and step times (45.8% and 70.2%, respectively). Comparing interpolation methods, VFI yields a marginal benefit of 3.2% and 4.3%, when considering random errors of foot contact and step time estimation, respectively. This reflects interframe flow estimation, resulting in different foot trajectories compared to LI. In the context of movement health monitoring, reducing random error within step time estimates is imperative, as step time variability can provide insight into movement health. Research exploring the utility of markerless step time variability monitoring is warranted.

References

1. Cao, Z., Hidalgo, G., Simon, T., Wei, S-E. & Sheikh, Y. (2018). OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *CVPR*, arXiv:1812.08008.
2. Stenum, J., Rossi, C., & Roemmich, R.T. (2020). Two-dimensional video-based analysis of human gait using pose estimation. *PLOS Comp. Biol*. 17(4), e1008935.
3. Huang, Z., Zhang, T., Heng, W., Shi, B. & Zhou, Shuchang. (2020). RIFE: Real-Time Intermediate Flow Estimation for Video Frame Interpolation. *CVPR*, arXiv:2011.06294.
4. Dunn, M., Haake, S., Wheat, J. & Goodwill, S. (2014). Validation of a single camera, spatio-temporal gait analysis system. *Procedia Engineering*, 72, 243-248.
5. O'Connor, C.M., Thorpe, S. O'Malley, M.J. & Vaughan, C.L. (2007). Automatic detection of gait events using kinematic data. *Gait & Posture*, 25, 469-474.