POSE ESTIMATION OR MANUAL DIGITISING: CAN AUTOMATING TECHNOLOGIES CHANGE THE CURRENT IN-FIELD ASSESSMENT OF HIGH JUMP?

M.Goldacre[1], Z. Born[1], M. Mundt[1], E. Millett[3], E. Phillips[2], J. Alderson[1]
[1]UWA Tech & Policy Lab, University of Western Australia, Perth, Western Australia
[2]Australian Institute of Sport, Canberra, Australia
[3]New South Wales Institute of Sport, New South Wales, Australia

Biomechanists spend significant time completing the time-consuming task of manually digitising 2D videos to derive kinematic spatiotemporal parameters. Recent advances in 2D pose estimation models (PEMs) hold promise for automating the determination of parameters in sport. This study developed an automated PEM digitising and analysis pipeline (AAP) for high jump. We investigated differences in four spatiotemporal and joint angle outputs from traditional manual processing pipelines (MAP) and the AAP using paired t-tests, intra-class correlations and effect size analysis. Statistical analysis revealed that knee angles derived from the MAP and AAP were not different, whereas penultimate foot contact time and both body angle "lean" measures were different. The custom AAP considerably reduced processing time for the selected high jump execution parameters.
KEYWORDS: Biomechanics, high-jump, automation, pose estimation model

INTRODUCTION: The ability to reliably and efficiently assess human movement in ecologically valid settings is an ongoing goal for sports biomechanists. A substantial portion of an applied sports biomechanist's time is spent completing laborious manual annotation of two dimensional (2D) videos captured during training and competition to assess key kinematic and spatiotemporal performance variables. An athlete's coaching and performance team relies on parameters being processed swiftly, however the reliance on manual digitising leaves practitioners dependent on highly time-consuming manual methods that are susceptible to subjective error (Needham et al., 2021). Recent research to reduce reliance on these methods in sports biomechanics, rehabilitation and clinical settings has driven considerable advances in measurement technologies and methods (Colyer et al., 2018; Cronin, 2021). Computer vision-based human pose estimation models (PEMs) automatically identify 'keypoints', labels that sports biomechanists refer to as key anatomical landmarks, unlocking the potential of labour-saving methods for the discipline (Mundt et al., 2022). The reliability of publicly available 2D PEMs has been assessed with promising results in varying motion types, from running (Cronin, 2021) to more specific sporting contexts like boxing (Lahkar et al., 2022) and cycling (Serrancoli et al., 2020). However, further scrutiny and assessment of the reliability and notably, the practicality, of implementing these systems is warranted (Colyer et al., 2018). This study aims to investigate the reliability of a custom PEM and analysis pipeline (AAP) for deriving coach selected kinematic and spatiotemporal performance parameters of national level high jumpers. We compare the outputs to those derived from manual analysis and reporting pipelines (MAP) undertaken by well-trained practitioners in a high-performance sporting environment.

METHODS: This study was approved by the University of Western Australia's Human Ethics Committee. A total of 43 high jump attempts completed during training by two national level athletes (one male and one female) were used for analysis. An a priori power analysis determined a minimum sample size of 34 trials were required to achieve 80% power with an effect size of 0.5 ($\alpha = 0.05$, G*Power version 3.1.9.7, Faul et al., 2007). Two cameras (50Hz) positioned posterior and sagittal to the bar and mat captured 43 independent jumps, resulting in a total of 86 videos. All videos were previously manually digitised and analysed by a single experienced biomechanist. A standard high-jump biomechanical report comprises 14 coach-informed parameters used in the assessment of high jump performance. Recent findings from our research group found that of these 14 current high-performance parameters, four were moderately to strongly correlated with jump success (Born et al., 2023). One of these parameters was jump height (r_s = -0.67) which was not an assessable parameter for this study. The other three performance parameters, penultimate foot contact time (r_s = 0.42), body "lean" measure through hip (r_s = -0.42), knee angle at full foot contact for take-off (r_s = 0.37) were selected for comparison in this study (Born et al., 2023). To increase the number of assessment items between the methods an additional body "lean" measure through ankle (r_s = -0.3) was also included (Table 1).

Table 1: Manual processing definitions and measurement examples for the four selected coach-informed performance parameters.

Performance Parameter	Manual Processing Definition	Camera View	Visual Representation
Penultimate foot contact time (seconds)	For the foot contact prior to take-off: the total time between the frame of initial foot contact to the frame of toe-off (i.e. penultimate foot stance time).	Sagittal (relative to the bar and mat)	
Body "lean" measure through hip (°) [refer to blue line in measurement example]	At take-off foot-flat, a vector from the centre of the ankle joint is annotated to intersect the hip joint centre of the same stance leg. "Lean" angle is measured as the global angle of this vector to global vertical.	Sagittal (relative to the bar and mat)	
Body "lean" measure through shoulder (°) [refer to blue line in measurement example]	At take-off foot-flat, a vector from the centre of the ankle joint is annotated to intersect the gleno-humeral joint centre. "Lean" angle is measured as the global angle of this vector to global vertical.	Sagittal (relative to the bar and mat)	
Knee angle at full foot contact (take-off step) (°)	First frame of full foot contact (foot-flat) of the take-off foot. 2D relative (included) angle between visually identified ankle, knee and hip joint centres.	Posterior (relative to the bar and mat)	

OpenPose (Cao, Hidalgo, Simon, Wei, & Sheikh, 2021), a free publicly available 2D PEM, was implemented in Python (Python Software Foundation, v.3.9) and used to estimate 25 keypoints, of which only keypoints of the bilateral hip, foot, knee, ankle, and shoulder joint centres were used in the present comparison. Identification of the key event frames of initial foot contact (IF), foot flat to ground (FF), and toe-off (TO) were defined according to the definitions used in the MAP. IF, FF, and TO were identified in the AAP through threshold implementation utilising the keypoints of the heel and toes. OpenPose keypoints *u*, *v* values were extracted for the relevant frames and the four high jump parameters calculated. Paired t-tests were used to assess for differences between the same four MAP parameters and the AAP calculated parameters. A two-way random effects intraclass correlation model (ICC), with absolute agreement and single rating, was used to measure the reliability of performance parameters between assessment methods. Descriptive data were presented as group means with standard deviation (SD). Cohen's *d* was used to calculate effect sizes, with 0.2, 0.5, and 0.8 representing small, medium, and large effect sizes respectively (Cohen, 2013).

RESULTS: Spatiotemporal and joint angle output differences between the MAP and the AAP are illustrated in Table 2. The MAP consistently reported a longer penultimate foot contact time (0.29s, t = 3.82, p = 0.003, d = 0.93) than the AAP (0.21s), along with a smaller spread (SD = 0.02 and SD = 0.12). The MAP reported a greater body "lean" angle measure for vertical relative to hip (34.8°, t = 10.49, p < 0.001, d = 2.29) when compared to the AAP (28.5°), along with greater manual "lean" angle measures relative to shoulder (32.3°, t = 10.67, p < 0.001, d = 2.33) compared to the AAP (23.7°). The AAP reported slightly smaller distribution dispersions for both hip (SD = 2.2 and SD = 3.1) and shoulder (SD = 1.88 and SD = 2.8) body "lean" measures. For knee angle at full foot contact (take-off step) no significant difference was observed between the MAP (170.9°) and the AAP approaches (169.1°, t = 0.77, p = 0.44, d = 0.17). However, the knee angle did return the largest variation in distribution dispersion between the two pipelines (SD = 14.77 and SD = 3.22). The ICCs were all below 0.51, with 95% confidence intervals between 0.1 - 0.56, indicative of poor reliability between methods (Koo & Li, 2016).

Table 2: Spatiotemporal and joint angle performance parameters (mean ± SD) for manual and automatic processing and analysing pipelines. (*p < 0.05)

	Manual Processing/ Analysis Pipeline (MAP)	Automatic Processing/ Analysis Pipeline (AAP)	Intraclass Correlation	t-value	*P*	Cohen's *d*
Penultimate last contact (seconds)	0.29 ± 0.02	0.21 ± 0.12	0.51	3.82	0.003*	0.93
Body "lean" measure through hip (°)	34.81 ± 3.16	28.53 ± 2.26	0.21	10.49	<0.001*	2.29
Body "lean" measure through shoulder (°)	32.37 ± 2.84	26.77 ± 1.88	0.16	10.67	<0.001*	2.33
Knee angle at foot contact (last step) (°)	170.91 ± 3.22	169.11 ± 14.77	0.07	0.77	0.44	0.17

Notably, the AAP processed a single video in 40 seconds compared to 1,440 seconds, or 24 minutes, for the MAP. In the time taken to manually process one jump the AAP completed analysing 36 videos.

DISCUSSION: This study aimed to investigate select spatiotemporal and joint angle output differences between MAP and AAP methods for high jump 2D video assessment. Three of the four assessed performance parameters returned statistically, and likely functionally, significant differences in mean values (large effect sizes), alongside poor ICC repeatability (Table 2). This finding was unexpected, given previous literature establishing non-significant differences in manual versus automated 2D kinematic analysis methods (Colyer et al., 2018; Cronin, 2021). Temporally dependent parameter differences observed in the present study may be attributed to between-method inconsistency in key event frame identification (e.g., initial foot contact, foot-flat). An additional explanation is the influence of knowledge bias in manual digitising processes as well-trained practitioners anticipate the "normative" or expected values, biasing event detection toward an expected outcome. The smaller AAP standard deviation in body "lean" angles provides support for the standardisation of the AAP method, something that would be beneficial in a high-performance sports system where multiple practitioners across various geographical locations are charged with undertaking MAPs. Implementing AAP methods facilitates the standardisation of performance parameter measurement, an action known to improve the reliability of measurements provided to coaches and athletes (de Oliveira et al., 2019). It was notable that the MAP and AAP achieved similar mean take-off knee angles (differed by < 2°) given the parallax error associated with the posterior camera not being orthogonal to the femur and shank at take-off (Payton & Burden, 2018). The in-field limitations of camera placement supports the ongoing call for further refinement and validation of three-dimensional PEMs from 2D video (Nakano et al., 2020; Pagnon et al., 2022), alongside the call for PEMS to be biomechanically informed for use in sport with complex, fast and multiplanar movements (Mundt et al., 2022). This research highlights that an automated approach analyses 36 jumps with the same time burden as manually processing a single jump, with the biomechanist only required to initiate the automated processing. This dramatic time-cost reduction means sport biomechanists can service a greater absolute number of athletes and increase their availability to the athletes and coaches they already work with.

CONCLUSION: This study investigated selected spatiotemporal and joint angle output differences derived from a traditional MAP and a newly developed AAP for the reporting of high jump performance parameters. The AAP significantly reduced processing and analysis time compared with the MAP. Although this study shows promise for automated 2D video analysis, further research is needed to identify the causal mechanisms of between-method output differences. This research highlights the importance of ensuring consistent key event and parameter definitions to generate implementable automated pipelines. The custom AAP approach in this study offers a substantial saving in the labour costs required to produce high priority training reports for coaches and athletes.

REFERENCES:

Born, Z., Goldacre, M., Mundt, M., Millet, E., Phillips, E., & Alderson, J. (2023). Simplifying the biomechanical assessment of high jump technique. Submitted to *XLI Congress of the International Society of Biomechanics in Sport* (ISBS).

Cao, Z., Hidalgo, G., Simon, T., Wei, S.E., & Sheikh, Y. (2021). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 43*(1), 172-186. doi:10.1109/TPAMI.2019.2929257

Cohen J. (2013). *Statistical power analysis for the behavioral sciences*. 2nd ed. Taylor and Francis.

Colyer, S. L., Evans, M., Cosker, D. P., & Salo, A. I. T.(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 Medicine- Open, 4*(1), 1-15. doi:10.1186/s40798-018-0139-y

Cronin, N. J. (2021). Using deep neural networks for kinematic analysis: Challenges and opportunities. *Journal of Biomechanics, 123*, 110460. doi:10.1016/j.jbiomech.2021.110460

De Oliveira, F.C.L., Fredette, A., Echeverria, S.O., Batcho, C.S., & Roy, J.S.(2019). Validity and Reliability of 2-Dimensional Video-Based assessment to Analyze Foot Strike Pattern and Step Rate During Running: A Systematic Review. *Sports health, 11*(5), 194173811984479- 194173811984415. Doi:10.1177/1941738119844795

Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). GPower 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2), 175-191. doi:10.3758/BF03193146

Koo, T.K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of chiropractic medicine*, *15*(2), 155-163. Doi:10.1016/j.jcm.2016.02.012

Lahkar, B. K., Muller, A., Dumas, R., Reveret, L., & Robert, T. (2022). Accuracy of a markerless motion capture system in estimating upper extremity kinematics during boxing. *Frontiers in sports and active living, 4*, 939980. doi:10.3389/fspor.2022.939980

Mundt, M., Oberlack, H., Goldacre, M., Powles, J., Funken, J., Morris, C., Wolfgang, P., & Alderson, J. (2022). Synthesising 2D Video from 3D Motion Data for Machine Learning Applications. *Sensors (Basel, Switzerland), 22*(17), 6522. doi:10.3390/s22176522

Nakano, N., Sakura, T., Ueda, K., Omura, L., Kimura, A., Lino, Y., Fukashiro, S., & Yoshioka, S. (2020). Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras. *Frontiers in sports and active living, 2*, 50. doi:10.3389/fspor.2020.00050

Needham, L., Evans, M., Cosker, D. P., Wade, L., McGuigan, P. M., Bilzon, J. L., & Colyer, S. L. (2021). The accuracy of several pose estimation methods for 3D joint centre localisation. *Scientific reports, 11*(1), 20673. doi:10.1038/s41598-021-00212-x

Pagnon, D., Domalain, M., & Reveret, L. (2022). Pose2Sim: An End-to-End Workflow for 3D Markerless Sports Kinematics-Part 2: Accuracy. *Sensors (Basel, Switzerland), 22*(7), 2712. doi:10.3390/s22072712

Payton, C. J., & Burden, A. (2018). *Biomechanical Evaluation of Movement in Sport and Exercise: The British Association of Sport and Exercise Sciences Guide* (2 ed.). Florence: Routledge.

Serrancoli, G., Bogatikov, P., Huix, J. P., Barbera, A. F., Egea, A. J. S., Ribe, J. T., Kanaan-Izquierdo, S., & Susin, A. (2020). Marker-Less Monitoring Protocol to Analyze Biomechanical Joint Metrics During Pedaling. *IEEE access, 8*, 122782-122790. doi:10.1109/ACCESS.2020.3006423