Human Pose Estimation Based Biomechanical Feature Extraction for Long Jumps

Qi Gan
LTCI
Télécom Paris & IP Paris
Palaiseau, France
qi.gan@telecom-paris.fr

Sao Mai Nguyen
FLOWERS team, U2IS
ENSTA, IP Paris & Inria
Palaiseau, France
nguyensmai@gmail.com

Mounîm A. El-Yacoubi
SAMOVAR
Télécom SudParis, CNRS & IP Paris
Palaiseau, France
mounim.el_yacoubi@telecom-sudparis.eu

Eric Fenaux
Ef-e-science
Paris, France
eric.fenaux@gmail.com

Stéphan Clémençon
LTCI
Télécom Paris & IP Paris
Palaiseau, France
stephan.clemencon@telecom-paristech.fr

***Abstract*—Biomechanical features describing movements and poses of athletes have been proposed by experts to help study athletic performances, but the traditional way of measuring those features are high-cost, time-consuming and intrusive. In this paper, we propose a deep learning-based method that can estimate athletic biomechanical features from typical broadcast competition videos, i.e. single-camera-shot moving videos. This method involves state-of-the-art human pose estimation models and a biomechanical analysis to reconstruct the trajectory. We then leverage the reconstructed trajectory to estimate the target features. To evaluate the method, we gathered a dataset from the long jump World Championships of 2017 and 2018, comprising 22 expert-proposed long-jump biomechanical features about the trajectories, taking-off and landing characteristics. Our experiments show the effectiveness of the pipeline in automatically estimating the biomechanical features. By analysing the results, we identify the challenges towards high-accuracy athletes' feature estimations from monocular broadcast competition videos. Code is available at *https://github.com/QGAN2019/Long_Jump_Feature_Estimation*.**

***Index Terms*—human pose estimation, athletes' biomechanical feature extraction, long jump**

I. Introduction

Athlete techniques analysis can help evaluate athletes' movements and provide guidance for training sessions [1], [2]. To better quantify these techniques, experts have proposed biomechanical features as metrics [3]–[9], which constitute physical quantities of athletes' poses and trajectories. However, the collection of these features requires sensor systems that can be costly and time-consuming. Automating this process is desirable to make it more cost-effective and efficient, thus applicable in a wider range of scenarios. A possible solution to this problem is to leverage modern artificial intelligence models to estimating human poses from videos, and recovering athletes' 3D trajectories from the estimated poses. The biomechanical features can then be calculated from these 3D trajectories. While some studies have successfully extracted basic features from estimated 2D athlete poses in videos [1], [2], to the best of our knowledge, there has been no research dedicated to extracting comprehensive features from 3D trajectories reconstructed from videos, which could offer a more nuanced understanding of athlete movements. To bridge this research gap, we present a method to estimate long-jump-related biomechanical features using world championships competition videos, with the long jump as a primary example. Specifically, we focus on monocular (i.e. single-camera-shot), dynamic (i.e. moving-camera-shot), low-frame-rate (i.e. 25 frames per second) videos containing a single athlete in each frame. The single-athlete condition mitigates complexity arising from human pose occlusion, while the other settings are prevalent in online competition videos. The proposed method follows a pipeline that takes advantage of off-the-shelf deep neural network models for athlete pose estimation and utilizes biomechanical analysis to estimate pose global positions. To evaluate the performance of the method, we collected a dataset consisting of recordings of 26 jumps by 22 athletes from online videos [10]–[13], along with ground truths of 22 long-jump-related biomechanical features provided by online reports [3]–[6]. We conducted experiments under different settings to explore the influence of different components in the pipeline. The contributions of this work are summarized as follows:

- We propose a method that can automatically estimate athletes' biomechanical features from online competition videos which are monocular, dynamic and low frame rate.
- We introduce a biomenchanical analysis method that can place 3D estimated athlete poses to their global position, in order to reconstruct athletes' 3D trajectories.
- We perform extensive experiments to analyze the influence of different components on feature estimation performance. The results provide guidance toward higher accuracy estimation of athletes' biomechanical features.

The organization of the paper is summarized as follows: Section II briefly reviews related research works on the topics of human pose estimation and athletes' biomechanical feature estimation. Section III introduces the proposed method to estimate athletes' biomechanical features. Section IV describes the dataset that we collect to evaluate the proposed method. Section V introduces the experiments and presents both the quantitative and qualitative, together with discussions on the influences of each component in the proposed method. Finally, section VI concludes the paper.

II. Related works

Human pose estimation: In the realm of human pose estimation from monocular videos, a key focus of research involves the accurate inference of single-person 3D poses. A general way to achieve this involves first estimating 2D poses from frames and then estimating 3D poses from the estimated 2D poses. This methodology capitalizes on recent advancements in 2D human pose estimation models, but lifting 3D poses from 2D representations still presents a significant challenge, primarily due to the loss of depth information. Prior studies have addressed this challenge by incorporating information from human body constraints (kinematics) [14], [15], temporal dynamics [16], [17], physical laws [18], [19], or their combinations [20]. In this context, we adopt YOLOv6 [21], ViTPose [22] and MHFormer [17] for, respectively, object detection, 2D human pose estimation and 3D human pose lifting, due to their proven state-of-the-art performance and superior efficacy demonstrated on our dataset.

Biomechanical feature estimation: Biomechanical features have been used to analyze athlete techniques in many works [7]–[9], but have been collected manually. So far, not much work has been done to automate the estimation process. Most studies relevant to this topic focus on estimating athlete poses instead of features, which mainly transfers deep neural network models to sports poses [23]–[25]. Only a handful studies have investigated the feature estimation problem. Wang et al. [1] have applied pose estimation to skiing videos to help evaluate the movements of skaters by estimating the 2D poses of athletes and their boards. Van der Meijden [2] used OpenPose [26] to estimate 2D poses to extract four features from relatively high frame rate (60 Frame/s) monocular sprint videos. However, to our best knowledge, no work has yet attempted to estimate the biomechanical features that requires the reconstruction of athletes' global 3D poses (or trajectories) from broadcast competition videos. In this work, we take the long jump as an example to study the feasibility and challenges of estimating athletes' biomechanical features with modern human pose estimation models.

III. Feature extraction pipeline

To calculate the biomechanical features of an athlete from his/her competition video, we proposed a method consisting of the pipeline shown in Fig. 1. The pipeline takes a frame sequence from competition videos as input, and extracts target features through three steps: 1) Human detection and pose estimation, 2) Trajectory reconstruction, 3) Biomechanical feature calculation. In the first step, YOLOv6 [21] and ViTPose [22] are used to perform athlete detection and 2D pose estimation, respectively. The estimated poses are corrected by the annotation tool. Then 3D poses are estimated using MHFormer [17]. In the second step, the estimated 3D poses are placed to their global positions according to biomechanical analysis. In the third step, the reconstructed trajectory from the previous step is used to calculate the target features of the athlete.

A. Human pose estimation

The athletes are firstly detected by YOLOv6 [21], with the bounding boxes of class 'person' selected. The bounding box of the largest area in each frame is regarded as the athlete's area, as the athletes are always the objects focused on in broadcast competition videos. Then, the 2D poses of the athletes are estimated by ViTPose [22]. The Coco-whole body keypoints are estimated and transformed into Human3.6M [27] format (including keypoints for big toes). Despite the high performance of the state-of-the-art models, some highly-blurred or self-occluded frames still cause considerable detection errors, e.g. the swap of left and right legs, the erroneously estimated arms when occluded by athlete trunk and mixture between a ponytail and an arm, etc. We have thus manually corrected those errors in estimated 2D key points using an annotation tool developed by ourselves. With the tool, one can visualize the 2D key points in image frames and drag key points to correct them. With the new annotation, we retrain the 2D human pose estimation model.

The 3D poses are estimated from the corrected 2D poses by MHFormer [17]. To include the keypoints for big toes, we re-trained the MHFormer model with Human3.6M [27].

B. Athlete trajectory reconstruction

Premises: The above estimated 3D poses are in relative position, which cannot be directly used to extract features related to athlete's displacements. To place the 3D poses back to their global positions, we propose a method by introducing a biomechanical analysis. This analysis is based on two premises: 1) The athletes run on hard horizontal ground; 2) There is no foot-sliding when the athletes run. As the two premises are valid for the process before landing in long jump, we focus mainly on the process before (and including) athletes' feet touching the sand pit.

Biomechanical analysis: We propose to split the long jump process periodically into two phases in order to apply biomechanical analysis (see Fig. 2). Phase I is the process when one of the athlete's feet continuously contacts the ground, while Phase II is the process between two steps when the athlete's whole body is in the air. During Phase I, as the toe key point continuously contacts the ground without sliding, we can align the 3D poses during Phase I by translation according to the toe key point. During Phase II, because the body of the athlete can be approximated as being only affected by gravity, neglecting other effects like winds, the trajectory of the center of mass of the athlete follows the laws of free fall motion. Thus, if the velocity of the athlete's center of mass at the beginning of Phase II is known, the following positions of the athlete's 3D poses can be derived by aligning the center of mass to

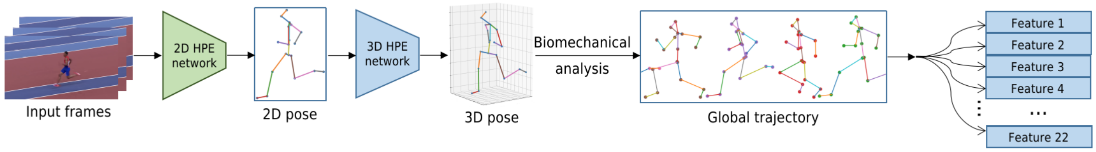

Fig. 1. Overview of the proposed pipeline. With input frames of athletes, 2D poses and 3D poses are estimates subsequently. Then the 3D trajectory is recovered from 3D poses through biomechanical analysis. Finally, different features are calculated from the trajectory.

its trajectory. Noticing that the velocity at the beginning of Phase II is the velocity at the end of Phase I, the subsequent positions, starting from a specific Phase I, can be alternatively calculated according to the above two phases.

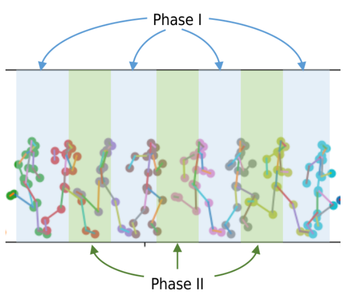

Fig. 2. Phase split. Phase I: When at least one foot touches the ground. Phase II: When both feet are in the air

Phase splitting: We split the athlete's movement into phases by estimating whether their foot contacts the ground. The foot-ground contacting label is obtained by aligning the 2D positions of the key points of left and right toes and setting a threshold to determine whether one foot touches the ground. However, as the above biomechanical analysis is sensitive to errors in phase splitting, we manually corrected the estimated foot-ground contact label using the interface shown in Fig. 3.

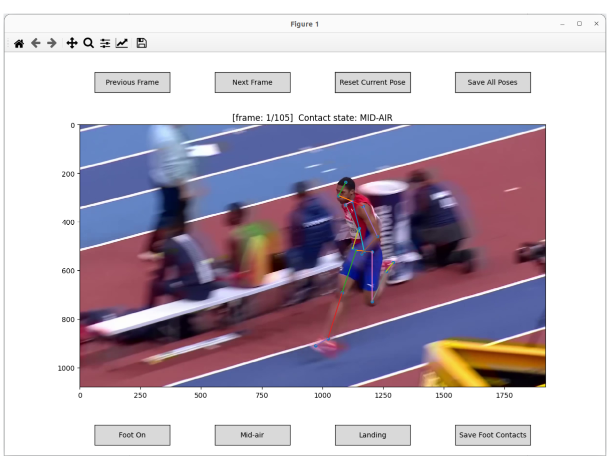

Fig. 3. The annotation tool. The above four buttons are used to correct 2D poses. The below four buttons are used to correct foot-ground contact labels. The keypoints of the 2D pose displayed in the frame can be dragged by cursor to correct the positions.

Body mass center estimation: To calculate the center of mass of a human body, we first calculate the center of mass of different segments in the 3D pose skeleton, and approximate the center of mass of the whole body according to the human body weight distribution. Referring to the human skeleton definition of Human3.6M [27], the center of mass of 'upper arm', 'forearm', 'thigh', 'shank' and 'foot' are approximated by the mid-point of the corresponding links; the center of mass of head is approximated by the mid-point between key point 'Head' and 'Thorax'; the center of mass of torso is approximated by the key point 'Spine'; and the center of mass of each hand is approximated by the key point of related 'Wrist'. Finally, the center of mass of the whole body is estimated by the weighted sum of the above centers, with the values of the weights provided in [28]. In order to map the 3D pose to real athlete body size, we took the official height of each athlete as a reference, while the height of the 3D skeleton is approximated by the sum of a half foot length, calf, thigh, torso and the distance between 'Thorax' and 'Head'.

IV. Datasets

We collected a dataset consisting of clips of long jumps with ground truths of features. The jumps are from the IAAF World Indoor Championships 2018 Long Jump Men & Women and the IAAF World Championships London 2017 Long Jump Men & Women. We referred to the biomechanical reports of IAAF [3]–[6] for ground truth of the 22 biomechanical features. The reports provided biomechanical features of 52 jumps of 40 athletes. Among them, there are 12 men and 12 women in the report of 2017's competition, and 15 men and 13 women in that of 2018's competition. Features are provided for one jump for each athlete in each competition. The videos are taken from the IAAF official channel of Youtube [10]–[13]. Because the competition videos do not record all the jumps, we only obtained 26 jumps of 22 athletes (11 men and 11 women) with biomechanical features' ground truth.

There are in total 43 features provided for each jump. Excluding irrelevant features (e.g. personal best), a not quantifiable feature (center of mass trajectory), and contact/flight/step time (because the frame rate is 25 (Frame/s), which is too low to estimate these features), only 22 features remain, which are explained in Table I (for more details see [3]–[6]). In this work, we divide the 22 features into 3 groups:

- **The trajectory features:** The features related to taking-off velocity and jumping distance. These features are directly related to the trajectories of long jump athletes from taking off to landing. They are thus highly correlated with the performance of long jump. They reveal the dynamics of the center of mass of an athlete's body

around and just before the moment of taking off. The accuracy of estimation of these features depends on both the pose estimation models and the biomechanical analysis.
- **The taking-off pose features:** The features related to athletes' poses when they take off from the board. These features reveal the information of athlete poses around the taking off moment. The performance of estimating them is only related to pose estimation performance.
- **The landing pose features:** The features related to athletes' poses when they land into the sand pit. Although the dynamics of the center of mass of an athlete are determined after taking off, a better landing pose can help avoid extra distance loss due to unnecessary body parts contacts with the sand pit. The estimation of these poses relies on the performance of pose estimation models, and of rare poses in particular. Notice that the ground truths of the landing pose features are not provided in the reports of the IAAF World Championships London 2017 Long Jump Men & Women [3], [5].

TABLE I
EXPLANATIONS TO THE BIOMECHANICAL FEATURES [3]–[6]

Features	Explanations
EF_dist	Effective distance
LS_len	Last step length
LS_vel	Mean velocity during last step
TO_vel_H	Horizontal velocity at take-off
TO_vel_V	Vertical velocity at take-off
TO_vel_loss_H	Loss in horizontal velocity during foot-on-board
TO_vel	Magnitude of the velocity at take-off
TO_ang	Angle of the velocity at take-off
CM_lower	Center of mass height difference between the lowest and the beginning during foot-on-board
TD_body_inclin	Body inclination angle at touchdown
TO_body_inclin	Body inclination angle at take-off
TD_trunk_inclin	Trunk inclination at touchdown
TO_trunk_inclin	Trunk inclination at take-off
TO_thigh_ang	Leading thigh angle at take-off
TO_thigh_ang_vel	Mean leading thigh angular velocity during foot-on-board
TD_knee_ang	Knee angle at touchdown
TO_knee_ang_min	Minimum knee angle during foot-on-board
TO_knee_ang_range	Knee angle range during foot-on-board
LD_knee_ang	Knee angle when landing
LD_trunk_ang	Trunk angle when landing
LD_dist	Distance between center of mass and the first landing point

V. RESULTS

We evaluated the performance of the pipeline in Fig. 1 on the estimation of features introduced in the previous section. The mean absolute error (MAE) was used as the metric to evaluate the estimation accuracy of each feature. To simplify the discussion, we examined the performance on the 3 groups of features separately. Furthermore, we designed four settings to explore the influence of different elements of the pipeline and thus to address the challenges in applying computer vision models to biomechanical feature estimation in sports (see II), which are:

- **Setting 1 (S1):** Following the pipeline in Fig. 1, firstly, the athletes in input frames were detected by YOLOv6-Large [21]. The largest bounding box of the 'person' class of each frame was selected as the one for the athlete. Then, ViTPose-Large [22] was used to perform 2D human pose estimation. The foot-ground contacts were estimated by a signal-processing based method. Both the 2D human pose and foot-ground contact labels were corrected with the annotation tool. Then the 3D poses were estimated using MHFormer [17] with sequence length set to 351, which we retrained on Human3.6M [27] to include 2 keypoints for big toes in order to calculate more accurate features. Afterwards, the trajectory of the athlete was reconstructed by estimating the global positions of the 3D poses through our biomechanical analysis. Finally, we estimated the features from the reconstructed trajectory.
- **Setting 2 (S2):** The same process of Setting 1 is followed, except that for the 3D poses, we excluded the keypoints for big toes, which are thus not used in feature calculation. This is because most of the models for 2D/3D human pose estimation focus on human skeletons without toe keypoints [17], [16]. By comparing Setting 2 with Setting 1, we would like to examine the influence of the toe keypoints to the estimation of features.
- **Setting 3 (S3):** The same process of Setting 1 is followed, except that we retrained MHFormer with sequence length of 81. According to [17], the performance with length 81 performed worse than with length 351. Thus, by comparing Setting 3 with Setting 1, we could examine the influence of 3D lifting models to the final feature estimation.
- **Setting 4 (S4):** The same process of Setting 1 is followed, with the only difference being that the 2D human poses were not corrected by the annotation tool. By comparing Setting 4 with Setting 1, we could investigate the effects of the quality of 2D keypoints.

TABLE II
SUMMARY OF THE FOUR EXPERIMENTAL SETTINGS

Setting	Toe keypoints included?	2D pose corrected?	MHFormer sequence length
S1	Y	Y	351
S2	N	Y	351
S3	Y	Y	81
S4	Y	N	351

A. The trajectory features

The trajectory features are directly related to the athlete's final performance (the distance). The results are summarized in Table III. Among the four settings, S1 outperforms the others significantly. Under S1, the pipeline provided the most accurate estimation of the features dominated by taking off horizontal velocity, i.e. the *effective distance* ('EF_dist'), the *taking-off horizontal velocity* ('TO_vel_H') and the *taking-off velocity* ('TO_vel'). This shows that the pipeline successfully recovered the horizontal positions of the 3D poses during taking off. Regarding the *loss in horizontal velocity during foot-on-board* ('TO_vel_loss_H'), the error should be within approximately twice that of 'TO_vel_H', since it represents the difference between two horizontal velocities. On the other hand, the features closely related to vertical velocity, i.e. the *taking-off vertical velocity* ('TO_vel_V') and the *taking-off velocity angle* ('TO_ang'), were poorly estimated. We claim that the low frame rate is the major cause of this difference. Due to the fact that the *vertical velocity at takeoff* undergoes rapid changes compared to the *horizontal velocity at takeoff*, a slight delay in identifying the precise moment of takeoff results in greater errors when estimating vertical velocities. Moreover, the *last-step length* ('LS_Len') is notably impacted by the low frame rate. This is primarily due to the brevity of the last steps, typically lasting only 1 or 2 frames, which complicates the calculation of step length based on the product of velocity and time. Finally, the *center of mass lowering* ('CM_lower') is too small to estimate accurately because it is around the limit of the accuracy of MHFormer for 3D human pose estimation.

Compared to S1, S2 did not include keypoints of toes. The decrease in the performance of S2 shows that the length of foot is big enough to cause significant difference in trajectory feature estimations. The poor performance of S3 highlights the importance of a good 3D lifting model. The performance of S4 is between the best (S1) and the others, which indicates that the toe keypoints and the performance of 3D lifting models are more important than the quality of 2D poses.

TABLE III
MAE OF ESTIMATING TRAJECTORY FEATURES

Features (Unit)	GT Mean (Std)	S1	S2	S3	S4
EF_dist (m)	7.64 (0.67)	**0.44**	1.16	1.70	<u>0.84</u>
LS_len (m)	2.13 (0.15)	**0.82**	1.52	1.17	<u>0.86</u>
LS_vel (m/s)	9.35 (0.47)	<u>1.01</u>	2.45	2.99	**0.94**
TO_vel_H (m/s)	8.50 (0.56)	**0.79**	1.82	2.23	<u>1.75</u>
TO_vel_V (m/s)	3.57 (0.30)	<u>3.10</u>	4.14	4.67	**1.97**
TO_vel_loss_H (m/s)	-1.48 (0.38)	<u>2.00</u>	**1.25**	2.14	4.15
TO_vel (m/s)	9.22 (0.52)	**0.82**	2.47	2.79	<u>1.66</u>
TO_ang (°)	23 (2)	<u>20</u>	29	34	**15**
CM_lower (cm)	3.4 (1.8)	<u>3.7</u>	<u>3.7</u>	**2.8**	12.4

B. The taking-off pose features

The results on the taking-off pose features are shown in Table IV. The overall performance of estimating those pose-angle features was satisfactory, with errors around or below 15° under the best settings (S1 and S2). On the contrary, the angular-velocity-related features ('TO_thigh_ang_vel' and 'TO_knee_ang_vel') had large errors, as they were limited by the low time resolution.

Comparing the different settings, S1 and S2 had similar performance because of the exclusion of foot-related factors in these features. On the contrary, S3 and S4 performed much worse than S1 and S2 on knee-related features, which indicates the importance of human pose estimation accuracy to the estimation of these features.

TABLE IV
MAE OF ESTIMATING TAKING-OFF POSE FEATURES

Features (Unit)	GT Mean (Std)	S1	S2	S3	S4
TD_body_inclin (°)	-35 (3)	13	11	**5**	<u>10</u>
TO_body_inclin (°)	18 (3)	14	18	**8**	<u>13</u>
TD_trunk_inclin (°)	-7 (6)	<u>7</u>	<u>7</u>	<u>7</u>	**6**
TO_trunk_inclin (°)	0 (5)	<u>11</u>	<u>11</u>	**10**	**10**
TO_thigh_ang (°)	-12 (7)	**6**	**6**	<u>11</u>	<u>11</u>
TO_thigh_ang_vel (°/s)	588 (80)	<u>216</u>	<u>216</u>	**143**	273
TD_knee_ang (°)	167 (4)	**21**	**21**	43	<u>27</u>
TO_knee_ang_min (°)	140 (7)	**13**	**13**	<u>19</u>	40
TO_knee_ang_range (°)	27 (5)	**13**	**13**	<u>26</u>	27
TO_knee_ang_vel (°/s)	-492 (94)	**208**	**208**	<u>426</u>	530

C. The landing pose features

The landing poses are more different from daily poses appearing in public dataset [27], [29] than taking-off poses are. Therefore, estimating these rare poses was more difficult, which explained that the errors about angular features were slightly higher than those of taking-off pose features (see Table V). Nonetheless, the overall absolute errors of these features were relatively small, which indicates that human pose estimation performance was still acceptable for landing poses.

TABLE V
MAE OF ESTIMATING LANDING POSE FEATURES

Features (Unit)	GT Mean (Std)	S1	S2	S3	S4
LD_knee_ang (°)	133 (13)	<u>20</u>	<u>20</u>	31	**19**
LD_trunk_ang (°)	7 (16)	<u>21</u>	<u>21</u>	**19**	<u>21</u>
LD_dist (m)	0.63 (0.12)	**0.10**	<u>0.11</u>	**0.10**	**0.10**

D. Qualitative results

To qualitatively evaluate the performances of the pipeline under different settings, we plotted the side-view and bird-view of one reconstructed trajectory (see Fig. 4). This trajectory was reconstructed under setting 1 from the video of Juan Miguel Echevarría [13], which is his best jump in the 2018 competition. From the two views of the reconstructed trajectory, the movement details during the whole jump process are recognizable, e.g. the adjustment of running step before taking off, the swing of the arms during flight, etc. Meanwhile, from the bird-view plot, the reconstructed trajectory is not perfectly straight, which is caused by the rotation of the camera. This issue is not considered in the proposed pipeline, because it has little influence on the target biomechanical features.

VI. CONCLUSIONS

This work proposed a pipeline that applies modern human pose estimation techniques and biomechanical analysis to long jump biomechanical features' estimation. The experimental results revealed that the majority of features were

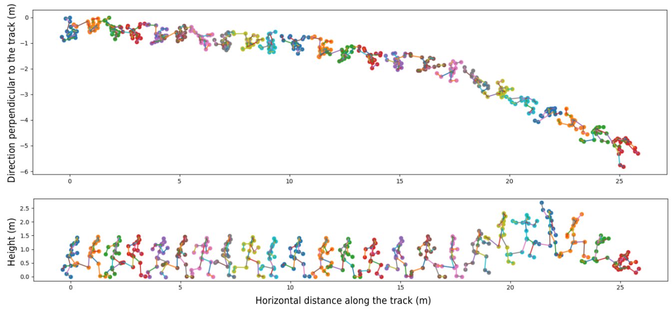

Fig. 4. The birdview (above) and sideview (below) of reconstructed trajectory. The trajectory is reconstructed from the monocular competition video of Juan Miguel Echevarrıa's best jump in the 2018 competition. The units of the coordinates are all in meters.

accurately estimated when employing high-quality human pose estimations, except for fast-changing-rate features which presented difficulty due to the low time resolution of broadcast competition videos, suggesting the potential utility of pose interpolation to enhance time resolution of pose sequences. Additionally, including toe keypoints could improve the estimation of both trajectory-related features and pose-related features, highlighting the need to adapt and retrain models lacking toe keypoints. Overall, the proposed method demonstrates feasibility in automating the feature estimating process using normal-quality broadcast competition videos, with implications for simplifying athlete feature estimation, enabling biomechanical-feature-based analysis for daily training, and facilitating large-scale biomechanical feature analysis of online competition videos.

References

[1] J. Wang, K. Qiu, H. Peng, J. Fu, and J. Zhu, "AI Coach: Deep Human Pose Estimation and Analysis for Personalized Athletic Training Assistance," in *Proceedings of the 27th ACM International Conference on Multimedia*. Nice France: ACM, Oct. 2019, pp. 374–382.

[2] K. van der Meijden, "Analyzing sprint features with 2d human pose estimation," Master's thesis, Tilburg University, Tilburg, Netherlands, 2019.

[3] C. Tucker, G. Nicholson, M. Cooke, A. Bissas, and S. Merlino, "Biomechanical report for the iaaf world championships london 2017: Long jump women's," World Athletics, Tech. Rep., 2018.

[4] C. Tucker, A. Bissas, and S. Merlino, "Biomechanical report for the iaaf world indoor championships 2018: Long jump women," World Athletics, Tech. Rep., 2019.

[5] C. Tucker, G. Nicholson, M. Cooke, A. Bissas, and S. Merlino, "Biomechanical report for the iaaf world championships london 2017: Long jump men's," World Athletics, Tech. Rep., 2018.

[6] C. Tucker, A. Bissas, and S. Merlino, "Biomechanical report for the iaaf world indoor championships 2018: Long jump men," World Athletics, Tech. Rep., 2019.

[7] J. Schiffer, "The horizontal jumps," *New Studies in Athletics*, no. 3, 2011.

[8] L. Mendoza and E. Nixdorf, "Biomechanical Analysis of the Horizontal Jumping Events at the 2009 IAAF World Championships in Athletics," *New Studies in Athletics*, no. 3, 2011.

[9] N. P. Linthorne, M. S. Guzman, and L. A. Bridgett, "Optimum take-off angle in the long jump," *Journal of Sports Sciences*, vol. 23, no. 7, pp. 703–712, Jul. 2005.

[10] World Athletics, "Women's Long Jump Final — IAAF World Championships London 2017," https://www.youtube.com/watch?v=5zyPJe6tJBU, 2019.

[11] ——, "Men's Long Jump Final — IAAF World Championships London 2017," https://www.youtube.com/watch?v=wNpA02SNgUg, 2019.

[12] ——, "Women's Long Jump — World Indoor Athletics Championships Birmingham 2018," https://www.youtube.com/watch?v=3pWvc059tAc, 2020.

[13] ——, "Men's Long Jump — World Indoor Championships Birmingham 2018," https://www.youtube.com/watch?v=bZE4faCuw_I, 2020.

[14] K.-M. Jin, B.-S. Lim, G.-H. Lee, T.-K. Kang, and S.-W. Lee, "Kinematic-Aware Hierarchical Attention Network for Human Pose Estimation in Videos," in *IEEE Winter Conf. on Applications of Computer Vision*, 2023, pp. 5725–5734.

[15] J. N. Kundu, S. Seth, R. M. V, M. Rakesh, V. B. Radhakrishnan, and A. Chakraborty, "Kinematic-Structure-Preserved Representation for Unsupervised 3D Human Pose Estimation," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 07, pp. 11 312–11 319, Apr. 2020, number: 07.

[16] D. Pavllo, C. Feichtenhofer, D. Grangier, and M. Auli, "3D Human Pose Estimation in Video With Temporal Convolutions and Semi-Supervised Training," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA: IEEE, Jun. 2019, pp. 7745–7754.

[17] W. Li, H. Liu, H. Tang, P. Wang, and L. Van Gool, "MHFormer: Multi-Hypothesis Transformer for 3D Human Pose Estimation," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, LA, USA: IEEE, Jun. 2022, pp. 13 137–13 146.

[18] E. Gartner, M. Andriluka, H. Xu, and C. Sminchisescu, "Trajectory Optimization for Physics-Based Reconstruction of 3d Human Pose from Monocular Video," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, LA, USA: IEEE, Jun. 2022, pp. 13 096–13 105.

[19] S. Shimada, V. Golyanik, W. Xu, and C. Theobalt, "PhysCap: Physically plausible monocular 3D motion capture in real time," *ACM Transactions on Graphics*, vol. 39, no. 6, pp. 1–16, Dec. 2020.

[20] J. Zhang, Z. Tu, J. Yang, Y. Chen, and J. Yuan, "MixSTE: Seq2seq Mixed Spatio-Temporal Encoder for 3D Human Pose Estimation in Video," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, LA, USA: IEEE, Jun. 2022, pp. 13 222–13 232.

[21] C. Li, L. Li, H. Jiang, K. Weng, Y. Geng, L. Li, Z. Ke, Q. Li, M. Cheng, W. Nie, Y. Li, B. Zhang, Y. Liang, L. Zhou, X. Xu, X. Chu, X. Wei, and X. Wei, "YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications," Sep. 2022.

[22] Y. Xu, J. Zhang, Q. Zhang, and D. Tao, "ViTPose: Simple vision transformer baselines for human pose estimation," in *Advances in Neural Information Processing Systems*, 2022.

[23] T. Baumgartner and S. Klatt, "Monocular 3D Human Pose Estimation for Sports Broadcasts Using Partial Sports Field Registration," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 5109–5118.

[24] J. Hwang, S. Park, and N. Kwak, "Athlete Pose Estimation by a Global-Local Network," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Honolulu, HI, USA: IEEE, Jul. 2017, pp. 114–121.

[25] M. Fastovets, J.-Y. Guillemaut, and A. Hilton, "Athlete Pose Estimation from Monocular TV Sports Footage," in *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. OR, USA: IEEE, Jun. 2013, pp. 1048–1054.

[26] Z. Cao, G. H. Martinez, T. Simon, S. Wei, and Y. A. Sheikh, "Openpose: Realtime multi-person 2d pose estimation using part affinity fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[27] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 7, pp. 1325–1339, Jul. 2014, conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence.

[28] M. Adolphe, J. Clerval, Z. Kirchof, R. Lacombe-Delpech, and B. Zagrodny, "Center of mass of human's body segments," *Mechanics and Mechanical Engineering*, vol. Vol. 21, no. nr 3, 2017.

[29] D. Mehta, H. Rhodin, D. Casas, P. Fua, O. Sotnychenko, W. Xu, and C. Theobalt, "Monocular 3D Human Pose Estimation in the Wild Using Improved CNN Supervision," in *2017 International Conference on 3D Vision (3DV)*. Qingdao: IEEE, Oct. 2017, pp. 506–516.