Realtime Football Curve Shot Posture Assessment Using Markerless Pose Estimation

Rahul Yadav
Amity Inst. of Information Technology
Amity University Uttar Pradesh
Noida, India
rahul.yadav18@s.amity.edu

Gurpreet Kaur
Amity Inst. of Information Technology
Amity University Uttar Pradesh
Noida, India
gkaur10@amity.edu

Soni Gupta
Amity Innovation and Design Centre
Amity University Uttar Pradesh
Noida, India
sgupta10@amity.edu

Vinayak Majhi
Amity Innovation and Design Centre
Amity University Uttar Pradesh
Noida, India
majhi.vinayak@gmail.com

Sujata Pandey
Amity Innovation and Design Centre
Amity University Uttar Pradesh
Noida, India
spandey@amity.edu

***Abstract*—Evaluating the body posture of players during a football curve shot is essential in enhancing performance and reducing injuries. However, the traditional approach is based mainly on the observation of the coach, which is subjective and not consistent. This paper proposes an efficient system in real-time posture evaluation based on computer vision techniques. The method is based on the use of multiple views of the image, which are synchronized using an audio-based clap detection method. It also makes use of markerless pose estimation to identify the key joints of the body and track the movement of the player. The motion-based approach is used to locate the ball and determine the exact time of contact between the foot and the ball. At the exact time of contact, crucial angles such as knee, trunk, and ankle angles are calculated. These angles are compared with the threshold values using the rule-based approach to check whether the posture is correct or not. It also displays an understandable grading of the shot made by the player. The proposed approach does not demand the availability of huge training data and is computationally inexpensive. In summary, this piece of research provides an approach that is viable and comprehensible for the automated analysis of football posture.**

***Keywords*—Football Biomechanics, Pose Estimation, Computer Vision, Sports Analytics, Performance Evaluation**

I. Introduction

Football is considered one of the top-ranked games in the world, and techniques such as the curve shot are considered vital in determining the outcome of football matches. A good curve shot demands proper body positioning, balance, and synchronization of body joints. A slight variation in body positioning during ball-foot strike is known to greatly influence the direction, speed, and accuracy of the curve shot. In general, football players depend on their coaches to analyze and improve their body positioning through visual inspection. However, such an analysis is subjective, time-consuming, and may vary depending on the coach's experience.

With the advent of computer vision and artificial intelligence, the ability to automatically analyze human motion using video has become possible. Markerless pose estimation techniques enable the estimation of the positions of the body's joints without the need for any special sensor or wearable equipment. This has opened up new possibilities for the objective analysis of sports performance. However, the problems of multi-view synchronization, event detection, and interpretation of the biomechanical data need to be addressed.

In this paper, a real-time posture evaluation system for football curve shots based on multi-view video analysis is proposed. The proposed system can synchronize multi-view images by audio-based synchronization, identify the time of foot-ball contact, and obtain key joint angles through pose estimation. A rule-based thresholding technique is used for posture correctness evaluation and a grading system.

The contributions of this work are: (1) the development of the multi-view synchronization technique based on audio cues, (2) the detection of the contact between the foot and the ball via motion analysis, and (3) the application of the rule-based system in the biomechanical evaluation of posture. The approach is efficient and does not demand huge training data.

II. Related Work

Recent advances in computer vision have greatly enhanced the capabilities of human pose estimation, and this has led to the practical application of this technique in sports biomechanics and performance analysis. Tools like MediaPipe are efficient and can be applied in a real-time scenario. [1]. Prior studies such as OpenPose proposed the idea of multi-person 2D pose estimation with the introduction of part affinity fields. This laid the groundwork for real-time human pose tracking[2]. Later studies such as DeepPose used deep learning to improve the accuracy of the methods with the introduction of convolutional neural networks. [3].

Some works have also extended the pose estimation to video-based and 3D analysis. VNect and VIBE are some of the methods that allow for 3D pose estimation in real-time from monocular video, which enhances motion understanding

[4], [11]. Also, the use of multi-view methods has been investigated to improve the accuracy of the estimation process, as well as to address the issue of occlusions by using 3D poses from multiple camera views [6]. Additionally, Temporal models have also been proposed to capture motion consistency, which can improve the estimation of poses in videos[5], [8].

The domain of sports is also where the concept of pose estimation has been used in the analysis of the performance of sportsmen. For example, computer vision has been used in the analysis of football players and their actions [7]. Additionally, there are new datasets, such as AthletePose3D and SportsPose, that have been proposed for the purpose of conducting studies on sports-specific pose estimation tasks.[9], [10]. These studies show the increasing interest in the use of AI-based techniques for sports analysis.

Despite all of these developments, there are some limitations to the existing methods. Some of the limitations include the fact that most of the existing methods make use of deep learning models, which are computationally expensive. Moreover, most of the methods focus on either pose estimation or action recognition, with limited focus on posture evaluation. Multi-view methods require the data to be synchronized, which is a limitation in most cases. Moreover, most of the existing methods do not offer direct feedback, which can be easily understood by players or coaches.

To address these issues, the present work proposes a novel system that is based on a lightweight and interpretable system for football posture assessment. The proposed system is different from previous systems in that it utilizes audio-based synchronization for multi-view alignment, motion-based football contact detection, and markerless pose estimation without the use of large datasets. The proposed system also utilizes a rule-based feedback mechanism that enables the conversion of biomechanical parameters into significant insights, making the system more practical and interpretable.

III. Proposed Methodology

A. System overview

The proposed system aims at providing automatic evaluation of the posture of the curve shot in the football match using the concept of multi-view video analysis. The entire system is divided into several steps, namely video acquisition, synchronization, event detection, pose estimation, biomechanical evaluation, and feedback. Video acquisition is the first step in which the video is acquired using the multi-view system, and the synchronization is performed using the audio-based synchronization system. Then, the event is detected, the pose estimation is performed, the important angles are calculated, and the evaluation is done using the rule-based system.

B. Multi-View Video Acquisition

To get the complete view of the movements of the player, videos are recorded in multiple angles such as front, side, and rear. This helps in reducing the problem of occlusion and provides a better view of the posture of the player. The videos are recorded individually by each camera.

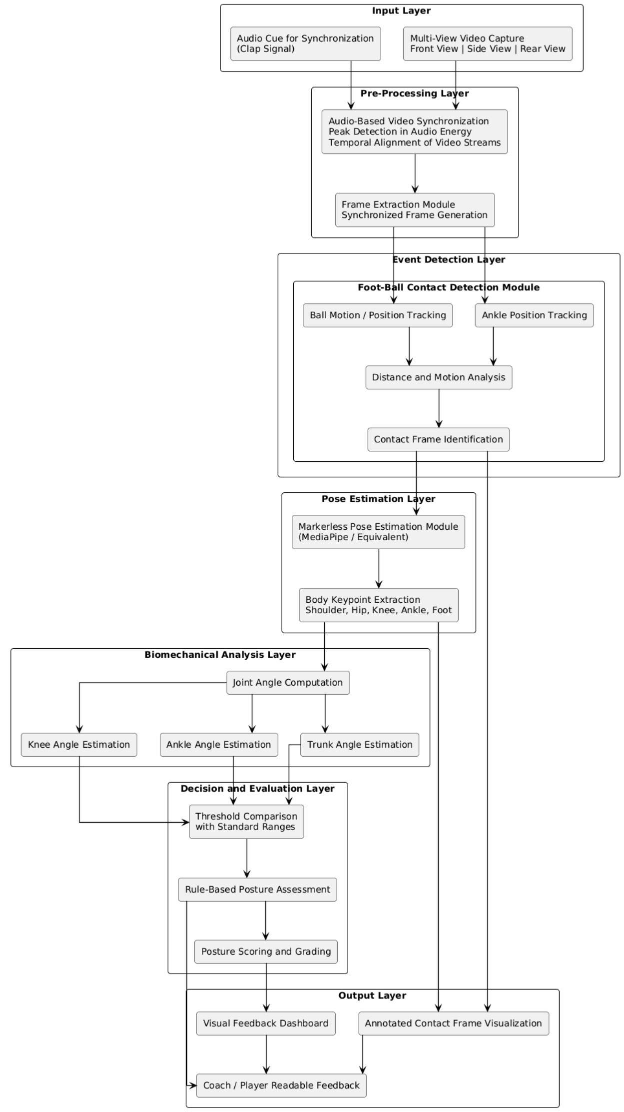

Fig. 1. System architecture

C. Audio-Based Synchronization

Since the videos are recorded separately, they may not be exactly aligned. To overcome this, a technique of audio-based synchronization is applied. A specific clap sound is incorporated at the start of the recording, and this results in a sharp peak in the audio signal. The peak is then detected through audio energy analysis, and all the video streams are aligned based on this specific frame.

D. Ankle-Based Foot-Ball Contact Detection

To recognize the point at which the action occurs, the system finds the frame in which the foot and the ball make contact. This is done by recognizing the position of the ankle of the player and the motion of the ball. The distance between the ball and the kicking foot of the player is computed over the video frames. The frame in which the distance is the least and the motion changes is recognized as the point of contact.

This step is important because posture evaluation occurs at this point.

E. Pose Estimation

The system employs a markerless pose estimation technique such as MediaPipe to extract body keypoints from each frame. Key body parts, including the shoulders, hips, knees, and ankles, are detected and tracked. This avoids the need to wear markers, hence providing non-intrusive analysis. The keypoints are then used to perform further calculations.

F. Posture Parameter Calculation

G. Biomechanical Parameter Calculation

The biomechanical parameters are computed using the coordinates of body keypoints obtained from pose estimation. Let three points $A(x_1, y_1)$, $B(x_2, y_2)$, and $C(x_3, y_3)$ represent adjacent joints. The angle at joint B is calculated using the cosine rule as:

$$\theta = \cos^{-1}\left(\frac{(A-B)\cdot(C-B)}{\|A-B\|\,\|C-B\|}\right) \tag{1}$$

where $(A-B)$ and $(C-B)$ are vectors, and $\cdot$ denotes the dot product.

1) Knee Angle: The knee angle is calculated using the hip (H), knee (K), and ankle (A) coordinates:

$$\theta_{\text{knee}} = \cos^{-1}\left(\frac{(H-K)\cdot(A-K)}{\|H-K\|\,\|A-K\|}\right) \tag{2}$$

2) Ankle Angle: The ankle angle is computed using the knee (K), ankle (A), and foot (F) coordinates:

$$\theta_{\text{ankle}} = \cos^{-1}\left(\frac{(K-A)\cdot(F-A)}{\|K-A\|\,\|F-A\|}\right) \tag{3}$$

3) Trunk Angle: The trunk angle is calculated between the vector formed by the shoulder (S) and hip (H) and the vertical axis:

$$\theta_{\text{trunk}} = \cos^{-1}\left(\frac{(S-H)\cdot(0,1)}{\|S-H\|}\right) \tag{4}$$

These computed angles represent the player's posture at the moment of foot-ball contact and are used for further evaluation using predefined threshold ranges.

H. Rule-Based Feedback Generation

In order to assess the quality of the football curve shot, the computed biomechanical parameters are compared with the threshold range values that are standard for the football kicking action. Every posture deviation is translated into feedback that the user can understand. For instance:

- Knee Angle $< 140^\circ$ indicates excessive knee flexion, reducing shot power.
- Knee Angle $> 160^\circ$ suggests insufficient knee bend, affecting control and balance.
- Trunk Angle $> 20^\circ$ reflects excessive forward or backward lean, leading to inaccurate ball trajectory.
- Trunk Angle $< 5^\circ$ indicates an overly upright posture, limiting effective force transfer.
- Ankle Angle $< 100^\circ$ suggests poor foot positioning, reducing shot precision.
- Ankle Angle $> 160^\circ$ indicates improper ankle extension, affecting ball contact quality.

Each parameter is given a score according to its deviation from the optimal range, and then an overall grade is generated for the shot. This feedback system is based on rules and thus not only quantifies posture parameters but also offers clear and effective insights for improving shooting techniques in football's curve shots.

IV. Experimental Setup and Dataset

The experimental setup and data set were tailored to record and analyze football curve shots in a real-world environment. Videos of the experiment were recorded using multiple cameras with different angles of view, including front, side, and back views of the player's movement. All videos were recorded at a frame rate of 60 frames per second (fps), which is suitable for the precise timing of quick movements in kicking the ball.

To ensure synchronization of all camera views, a unique sound of clapping was added at the beginning of each recording session. The audio signals were analyzed to identify the point of synchronization based on the peak energy level, which ensures precise time alignment of all the video streams. The recording environment is set in an outdoor scenario with natural lighting to mimic real-world gameplay scenarios.

The dataset is comprised of self-recorded videos of football curve shots conducted in various trials, which record different player postures and shooting styles. For each trial, synchronized videos in multiple views were processed to retrieve player keypoints via markerless pose estimation. The motion of the ball is also tracked to pinpoint the exact frame at which contact is made between the foot and the ball.

All the features, including the joint coordinates and the position of the ball, are used for further analysis. The dataset does not require any annotation, as the event detection and features are extracted automatically by the proposed system. The proposed dataset is light and efficient and can be used for validation of the posture assessment framework.

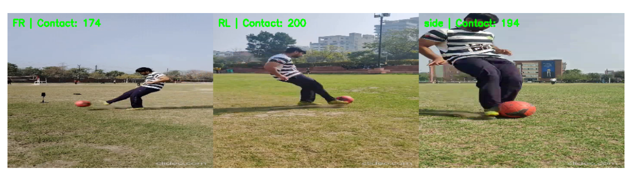

Fig. 2. multi-view synchronized visulization at detected ball contact frame (Front, Rear, Side views)

V. Experimental Validation

In order to assess the effectiveness of the proposed system, some validation experiments were performed. These experiments were carried out with regards to synchronization accuracy, event detection reliability, and consistency in assessing postures. The following subsections describe the validation experiments.

A. Audio-Based Synchronization Validation

The correctness of multi-view synchronization was verified by examining the alignment of video frames that corresponded to the detected clap event in all views. Audio peaks were mapped to their corresponding video frames. Visual verification confirmed that the same temporal instance was being synchronized across all views. A small degree of variation in frames was noticed, but it was clear that the audio-based synchronization method is accurate in synchronizing independently recorded video.

B. Foot-Ball Contact Detection Validation

The effectiveness of the contact detection module was tested using distance and motion analysis between the ball and the ankle of the player. Plots of ball-ankle distance and ball velocity were used to verify that the detected contact frame was correct. The contact point was always at the frame with minimum distance and significant change in motion, which proved that the proposed method is effective.

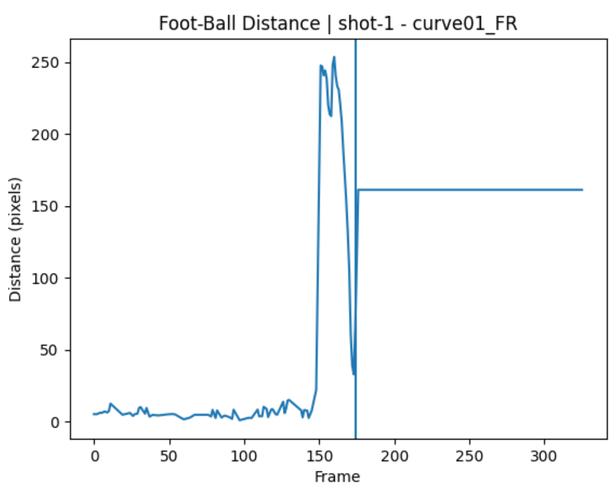

Fig. 3. Variation of foot–ball distance across frames for a curve shot.

C. Posture Parameter Validation

The computed biomechanical parameters were validated by visual inspection and the consistency of the results over multiple trials. The computed angles were superimposed over the video frames to ensure that they were an accurate representation of the posture of the player at the time of contact. The results indicated that the angles were estimated in a consistent and stable manner.

D. Multi-View Consistency Analysis

To ensure the robustness of the system, the posture parameters from different views of the cameras were compared. The synchronization of the cameras ensured that the images corresponded to the same time, hence the angles were similar with slight variations based on perspective views. This shows the reliability of the multi-view system.

TABLE I
POSTURE METRICS ACROSS FOUR CURVE SHOT TRIALS

Shot	Mean Knee Angle (°)	Mean Trunk Angle (°)	Mean Ankle Angle (°)
1	154.99	-48.66	144.32
2	122.59	-49.33	140.96
3	121.37	-50.98	166.71
4	113.82	161.12	136.38

E. Qualitative Visual Validation

For qualitative validation, frames with annotations were produced, along with multi-view visualizations, at the exact instant of foot-ball contact. The visual results provided clear validation that the system was able to capture and interpret key biomechanical features, thus making the feedback useful for players and coaches.

The experimental validation of the proposed system proved that the system can be used for football posture analysis in real-time, as it can perform well in all the processes involved in the system.

Fig. 4. Qualitative Visual Validation of Curve Shot at Ball-Contact Frame

VI. DISCUSSION

The results clearly indicate that the proposed system is effective in the analysis of the football curve shot posture with the integration of the multi-view video processing and the rule-based evaluation. The audio-based synchronization approach was found to be reliable in the synchronization of the video views. The approach was effective in the precise synchronization of the video views without the need for any specialized hardware. This ensured that the posture analysis was carried out at the exact moment of the football contact.

The motion-based contact detection approach was found to be effective in the detection of the exact frame of the contact. The approach was effective in the detection of the contact point with high accuracy. The detection of the contact point is important in the analysis of the posture. The markerless approach was effective in the extraction of the keypoints.

The computed posture parameters, such as knee, trunk, and ankle angles, were consistent over multiple trials and views, which demonstrates the robustness of the method. The rule-based feedback mechanism also improved the interpretability of the results, which is beneficial not only for analysis but also for training and improvement of performance.

However, some limitations of the method were also noticed. The fixed range of thresholds may not be effective in dealing with players of different body types and skill sets. Moreover, the 2D estimation of poses may also lead to some level of inaccuracy in results due to perspective distortion. Nevertheless, the proposed system is an efficient and interpretable solution to the problem of automated football posture assessment.

VII. Limitations

In spite of the effectiveness of the proposed system, there are some limitations that need to be considered. Firstly, the threshold ranges considered in the proposed system may not be effective for players with different characteristics, skill levels, and styles of play. This is because the feedback may not be the best for every player.

Secondly, the proposed system is based on the estimation of the pose of the players using the 2D pose estimation technique. This technique may not be accurate because of the perspective and occlusion problems. This is because the depth factor is not considered in the system.

Third, the detection of the ball and foot-ball contact relies on motion analysis and distance calculation, which could be affected by noisy environments, background motion, or partial occlusion of the ball. These factors could, in turn, affect the accuracy of event detection in complex environments.

Also, the proposed method for audio-based synchronization relies on the availability of a clear sound of clapping. This may not be possible in a real environment, thus limiting the use of the proposed method for synchronization.

Finally, the proposed system does not include any optimization or learning component. Although the proposed rule-based system is interpretable, it does not improve with experience or with increasing data.

The limitations of the proposed system have shown areas of possible improvements, such as the use of adaptive thresholds, 3D pose estimation, and event detection.

VIII. Conclusion

In this paper, a real-time system is proposed for football curve shot posture assessment based on multi-view video analysis and computer vision techniques. The proposed method combines audio-based synchronization, motion-based foot-ball contact detection, and markerless pose estimation to identify the essential biomechanical parameters during the crucial moment of contact with the ball. Moreover, a rule-based evaluation system is established to evaluate the correctness of posture based on knee, trunk, and ankle angles.

The experimental results show that the system can accurately synchronize multiple views of a video, detect the contact event, and estimate posture parameters with high accuracy. Moreover, the use of a rule-based feedback mechanism allows for clear and accurate insights, making the system useful for training and coaching purposes. Finally, the proposed method is computationally efficient and does not rely on large datasets, which are useful for real-world scenarios.

While there are certain limitations, such as the system's dependency on thresholds and 2D pose estimation, the proposed system is effective in providing a simple solution for automated football posture analysis. However, further work could be carried out in enhancing the system with adaptive models, 3D pose estimation, and detection techniques.

This work is contributing to the development of accessible tools for sports biomechanics analysis using artificial intelligence.

References

[1] C. Lugaresi et al., "MediaPipe: A Framework for Building Perception Pipelines," arXiv preprint arXiv:1906.08172, 2019.

[2] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2017.

[3] A. Toshev and C. Szegedy, "DeepPose: Human Pose Estimation via Deep Neural Networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2014.

[4] D. Mehta et al., "VNect: Real-time 3D Human Pose Estimation with a Single RGB Camera," ACM Trans. Graph., vol. 36, no. 4, 2017.

[5] G. Pavllo, D. Grangier, and M. Auli, "3D Human Pose Estimation in Video with Temporal Convolutions and Semi-Supervised Training," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2019.

[6] H. Rhodin et al., "Learning Monocular 3D Human Pose Estimation from Multi-View Images," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2018.

[7] G. Pinheiro et al., "Automatic Analysis of Football Players Using Computer Vision Techniques," IEEE Access, vol. 10, pp. 12345–12356, 2022.

[8] T. Pfister, J. Charles, and A. Zisserman, "Flowing ConvNets for Human Pose Estimation in Videos," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2015.

[9] S. Yeung et al., "AthletePose3D: A Large-Scale Dataset for 3D Human Pose Estimation in Sports," arXiv preprint arXiv:2301.00000, 2023.

[10] C. Ingwersen et al., "SportsPose: A Dynamic 3D Sports Pose Dataset," in Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW), 2023.

[11] A. Kocabas, N. Athanasiou, and M. J. Black, "VIBE: Video Inference for Human Body Pose and Shape Estimation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2020.