A Multi-View Vision-Based Framework for Cricket Shot Analysis with Kinematic Contact Detection and Biomechanical Evaluation

Shubh Rastogi
Amity Inst. of Information Technology
Amity University Uttar Pradesh
Noida, India
shubhrastogi725@gmail.com

Prashant Srivastava
Amity Inst. of Information Technology
Amity University Uttar Pradesh
Noida, India
psrivastava12@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*—Markerless biomechanical analysis has become possible through developments in computer vision and pose estimation. When analyzing the performance of a cricket player, evaluating the sports player requires a temporal analysis of events, such as ball and bat collisions, and the respective body kinematics. This paper proposes a framework for multi-view analysis of cricket shots that automates temporal multi-view synchronization, contact frame localization, and biomechanical assessments. Coarse alignment of multi-view cameras is performed using the audio of hand claps. The contact event is temporally aligned using wrist velocity, front foot motion, and the distance between the bat and ball. After the contact event is temporally aligned, pose estimation is used on the aligned frame to estimate biomechanical parameters such as elbow angle, knee angle, and head alignment. The estimated parameters are then used to provide performance-enhancing feedback as automated coaching through a scoring mechanism. The timestamps of the contact events and shot events provide a performance evaluation framework that is objective, scalable, and economically viable. This framework can streamline performance analysis in cricket and can be used to improve intelligent coaching and sports analytics.**

***Index Terms*—Cricket Shot Analysis, Pose Estimation, Sports Biomechanics, Multi-View Synchronization, Computer Vision**

I. Introduction

Biomechanical analysis is essential for enhancing athletic performance as it provides quantitative data on human movement. For effective batting in cricket, body parts must work together perfectly, the bat must hit the ball at the right time, and energy must move quickly through the kinetic chain [1]–[5]. However, traditional coaching methods are mostly based on personal opinions and do not provide objective evaluation of players.

Nevertheless, owing to advancements in computer vision, markerless motion capture using regular RGB video recordings has become viable [6]. Leveraging pose estimation techniques such as OpenPose [7] and MediaPipe [8] facilitates the identification of skeletal joint points, which enables automated sports analysis [9], [10]. However, recognizing events such as ball-to-bat contact remains challenging due to differences in camera angles and synchronization issues [11], [12].

To address this problem, this paper proposes a model for cricket shot analysis using multiple cameras that incorporates camera synchronization, contact point detection, and biomechanical modeling. This model is based on synchronization through sound and the following kinematic parameters: wrist velocity, foot movement, and ball-bat distance, which are used to detect the contact point.

The contributions of this paper are:

- Multi-camera synchronization based on auditory and motion cues.
- A hybrid approach to contact detection based on kinematics and vision.
- A biomechanical pose analysis technique for analyzing cricket shots.
- A feedback system for automatic performance evaluation.

The presented method is low-cost and scalable, and can thus be used in smart coaching applications and sports analysis.

II. Related Work

Considerable research has been conducted in the area of cricket batting biomechanics to evaluate joint movement coordination in shots. Previous studies focused on the necessity of coordination between the upper and lower body, proper timing, and stability to achieve successful performance [1]–[5].

RGB-based markerless motion capture has emerged as a solution due to advancements in computer vision technology [6]. It is possible to estimate human skeleton joints using pose

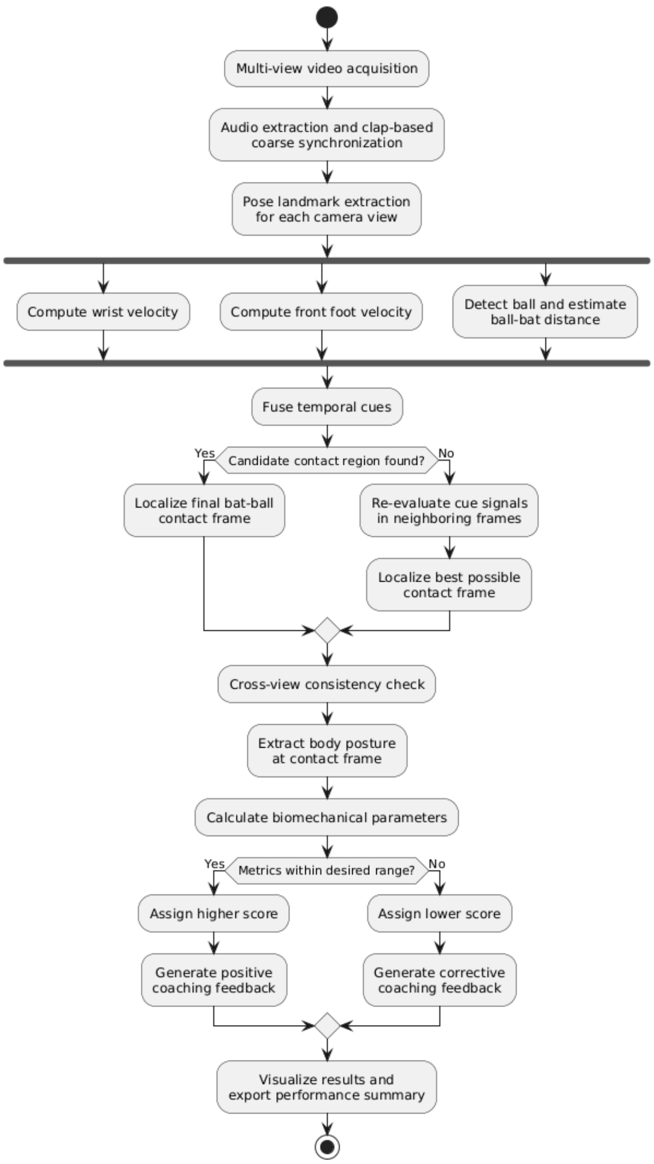

Fig. 1. Proposed system architecture illustrating the multi-stage pipeline for cricket shot analysis.

estimation algorithms such as OpenPose [7] and MediaPipe [8], which enables automated sports analysis.

Another technique that has advanced the spatiotemporal analysis of human motion is action recognition [9], [10]. Multi-view synchronization has also been studied for aligning video streams captured from multiple angles [11]. Recently, markerless motion capture technology has been demonstrated to be efficient in biomedical applications [12].

Yet, little attention has been given to combining all these methods to achieve synchronization, detect contact events, and provide biomechanical feedback. Our framework integrates multi-view synchronization, kinematic event detection, pose analysis, and feedback generation.

III. Methodology

The proposed system performs automated cricket shot analysis using a multi-camera framework that integrates synchronization, contact detection, pose estimation, and biomechanical evaluation (Fig. 1).

A. Multi-View Video Acquisition

Videos are captured from multiple viewpoints (right, left, front, and rear) at a frame rate of 60 FPS, which provides sufficient temporal resolution for movement analysis. The multi-view video capture system helps address occlusion issues and provides a better understanding of the motion.

B. Audio-Based Synchronization

Since video shots are captured individually, they are not temporally synchronized. Synchronization is performed using audio signals via short-time energy analysis of the clap sound, where the frame at that moment is taken as the reference for synchronization. The delay can then be determined relative to the chosen reference angle.

C. Kinematic Contact Detection

A hybrid approach combining multiple motion cues is used to detect the bat–ball contact frame.

1) Wrist Velocity Analysis: Wrist velocity is computed as:

$$v_t = \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2} \tag{1}$$

A sharp peak followed by a drop indicates the contact region.

2) Front Foot Stabilization: Front foot velocity increases during stepping and decreases upon stabilization. Contact typically occurs after foot grounding, indicating proper weight transfer.

3) Ball–Bat Distance Minimization: The ball is detected using HSV segmentation, and its distance from the wrist is computed. The contact frame corresponds to the minimum distance.

4) Final Contact Frame Selection: A search window is defined around the wrist velocity peak and foot stabilization phase. The frame with the minimum ball–bat distance within this window is selected as the final contact frame.

The overall workflow is summarized below:

Input: Multi-view videos $V = \{V_1, V_2, ..., V_n\}$
Output: Contact frame f_c and biomechanical metrics

1) Extract audio signals and detect clap frame f_{clap}.
2) Perform coarse synchronization using f_{clap}.
3) For each video:
   a) Extract pose landmarks.
   b) Compute wrist velocity, foot velocity, and ball distance.
4) Identify candidate window around peak wrist velocity.
5) Select $f_c = \arg\min(d_t)$ within the window.
6) Compute biomechanical metrics at f_c.
7) Aggregate metrics and generate feedback.

D. Pose Landmark Extraction

Pose landmarks are extracted using a markerless framework. Key joints such as the shoulder, elbow, wrist, hip, knee, and ankle are used for analysis.

E. Biomechanical Metric Computation

Joint angles are computed using:

$$\theta = \cos^{-1}\left(\frac{BA \cdot BC}{||BA|| \cdot ||BC||}\right) \tag{2}$$

The metrics include elbow angle, knee angle, wrist speed, head position, front foot speed, and bat angle. These relate to upper body coordination and lower body balance.

Table I summarizes the extracted biomechanical parameters used for performance evaluation.

F. Automated Feedback Generation

Computed metrics are compared with predefined thresholds to generate interpretable feedback. For example, insufficient elbow extension or excessive knee extension indicates improper technique.

G. Multi-View Consistency

The detected contact frame and metrics are validated across all views. Minor variations (± 1 frame, $\approx$16.7 ms) are tolerated, ensuring consistent event localization.

IV. Experimental Validation

A. Dataset and Experimental Setup

The developed framework was assessed using a unique multi-view cricket dataset comprising several batting shots filmed simultaneously from four camera viewpoints: right-side (RS), left-side (LS), front, and back views. All videos were recorded at 60 FPS in outdoor settings, resulting in high-temporal-resolution videos for capturing motion and localizing events precisely.

A markerless approach for extracting and estimating the position of different body parts was employed, allowing joint-level studies without the need for markers. The dataset contains variations in shot execution, player stance, and recording conditions to test the robustness of the proposed system.

B. Multi-View Synchronization Evaluation

Synchronization was performed using two processes: clap detection via audio and motion-based synchronization. In audio-based clap detection, synchronization was performed using the clap audio as the reference point. In motion-based synchronization, wrist velocity served as the reference point.

The results of the technical assessment confirm good synchronization across views, with a slight variation of ± 1 frame (approximately 16.7 ms at 60 FPS).

From Fig. 2, it can be confirmed that contact event detection has been completed for all multiple views. Because the contact frame is determined in all views, the results of contact frame determination are consistent across views. Due to differences in viewing angles, the spatial placement of the views may vary, and contact determination may exhibit slightly greater differences.

Fig. 2. Multi-view synchronized frames at the detected bat–ball contact moment across left, right, front, and back camera views.

Accurate contact frame determination enables reliable biomechanical analysis of the contact event.

To provide an objective assessment of contact determination, a statistical evaluation of contact event determination frames was performed across all multi-view shots.

The results show good contact determination across all shots, with an average difference of approximately 1.48 frames between views. The multi-view results were good, with an average of just under 4 frames variation (at 60 FPS).

These results confirm that the developed methodology, combining audio identification and wrist velocity refinement, has provided good alignment.

As shown in Table II, the synchronization error remains consistently low across all trials.

Compared to audio-only synchronization, which may introduce larger temporal offsets, the proposed refinement strategy significantly improves alignment consistency.

TABLE I
EXTRACTED BIOMECHANICAL METRICS AT CONTACT FRAME

Shot	Elbow (°)	Knee (°)	Wrist Vel	Foot Vel	Head Err	Swing (°)
Shot1	141.82	117.20	0.0240	0.00008	0.0326	15.21
Shot2	161.15	148.68	0.0015	0.00022	0.0222	-127.87
Shot3	156.21	134.58	0.0267	0.00085	0.0223	6.39
Shot4	125.01	126.85	0.0120	0.01535	0.0142	-8.07
Shot5	156.99	143.30	0.0277	0.00145	0.0203	-2.16
Shot6	146.38	144.80	0.0408	0.01565	0.0244	-1.09

TABLE II
SYNCHRONIZATION PERFORMANCE ACROSS SHOTS

Shot	Mean Frame	Std Dev	Error (Frames)
ff_defence_shot1	35.75	1.48	4
ff_defence_shot2	47.75	1.48	4
ff_defence_shot3	54.75	1.48	4
ff_defence_shot4	57.75	1.48	4
ff_defence_shot5	56.75	1.48	4
ff_defence_shot6	54.75	1.48	4

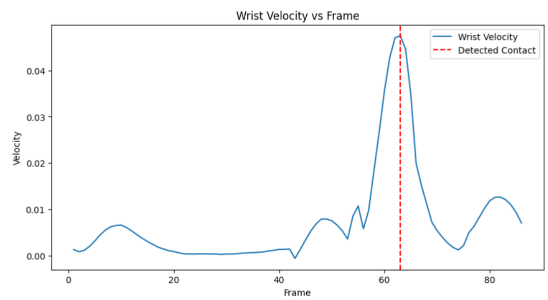

Fig. 3. Wrist velocity profile across frames. The sharp peak corresponds to the detected bat–ball contact frame, indicated by the vertical dashed line.

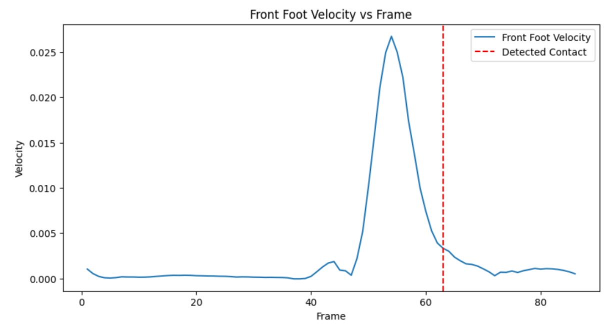

Fig. 4. Front foot velocity across frames. The velocity decreases sharply after foot stabilization, indicating proper lower-body grounding before contact.

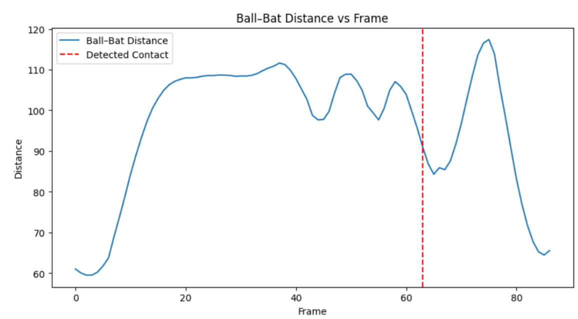

Fig. 5. Ball–bat distance across frames. The minimum distance corresponds to the detected contact frame.

C. Contact Frame Detection Analysis

Accurately locating the impact frame is essential for performing correct biomechanical analysis. For this purpose, the system adopts a mixed multi-cue method that considers wrist velocity, front foot stabilization, and ball–bat distance.

Wrist velocity peaks before decreasing rapidly at the time of impact.

As shown in Fig. 3, the wrist velocity exhibits a clear and prominent peak at the moment of contact, providing a strong temporal cue for event detection.

In addition to wrist motion, front foot velocity is analyzed to determine the stabilization phase. Foot velocity increases during the stepping motion and decreases significantly once the foot is grounded, indicating proper weight transfer prior to impact.

As illustrated in Fig. 4, the contact event occurs after the front foot velocity stabilizes, confirming the role of lower-body mechanics in shot execution.

Finally, the ball–bat distance is evaluated across frames to refine the exact contact point. The contact frame corresponds to the minimum distance between the detected ball and the wrist position.

Fig. 5 shows a clear minimum in the ball–bat distance at the detected frame, providing strong visual confirmation of the impact event.

The consistency observed across wrist velocity, front foot stabilization, and ball–bat distance demonstrates the robustness of the proposed multi-cue detection strategy. By combining these complementary signals, the system achieves more accurate and reliable contact frame detection compared to single-feature methods.

D. Biomechanical Metric Validation

In this analysis of the contact frame, biomechanical measurements such as elbow angle, knee angle, wrist velocity, front foot velocity, swing angle, and head alignment error were

Fig. 6. Angle overlay visualization at the detected contact frame. Elbow and knee angles are superimposed on the player's posture to validate biomechanical metric computation.

derived using pose landmarks extracted from synchronized video frames.

Joint angles and other biomechanical measurements were mathematically derived using geometric equations to maintain accuracy across all angles. These measurements pertain to the coordination of both the upper and lower body, which is vital for executing a cricket shot.

These metrics indicate the position, timing, and performance of movements required in shot execution.

In Fig. 6, the calculated angles have been successfully overlaid on the player's position within the contact frame, which serves as further validation for the biomechanical analysis. The correlation between pose markers and actual joints indicates a high level of accuracy in pose estimation.

Thus, the consistency between numerical metrics and visual observations confirms the efficiency of the proposed method for extracting biomechanical information from RGB data.

The visual confirmation of the calculated metrics can be seen in Fig. 6, where the excellent correlation between markers and actual joints serves as an indicator of the accuracy of the process.

E. *Performance Evaluation Across Shots*

The validity of the suggested approach has been confirmed through the analysis of various types of cricket strokes and scoring them using a rule-based evaluation technique. Scoring is based on the discrepancy between the observed movement and the optimal movement, including factors such as elbow angle, knee angle, and head alignment error.

As shown in Table III, the results exhibit noticeable variation across different shots, reflecting differences in player posture and execution. Shots with better joint alignment and stable lower-body positioning consistently achieve higher scores and grades, indicating efficient technique. In contrast, deviations in elbow extension, insufficient knee flexion, or increased head alignment error lead to lower scores, highlighting areas requiring improvement.

These results demonstrate that the proposed system can effectively distinguish between varying levels of shot quality and provide a quantitative basis for performance evaluation.

F. *Qualitative Feedback Analysis*

The system not only provides numerical assessment of movement but also offers interpretable coaching guidance related to biomechanical deviations.

Examples of such deviations include the need for improved elbow extension, increased lower-body stabilization, and correct head positioning. Thus, the system demonstrates practical application as an intelligent coaching assistant.

V. Discussion

The suggested system displays high accuracy in synchronizing multi-view recordings, detecting contact events, and analyzing the biomechanics of cricket strokes. Audio-based synchronization and motion correction help to align different video feeds precisely in time.

Contact detection using wrist speed, foot steadiness, and ball–bat distance shows high accuracy in identifying the correct contact moment. Multi-cue contact detection is more precise than single-feature detection.

The biomechanical results from this research are consistent with those from traditional methods used in cricket; therefore, markerless body tracking is applicable to sports activity analysis. Additionally, rule-based feedback assists in data analysis and converting information into actionable recommendations.

Nevertheless, time misalignments and projection issues caused by a 2D approach should be considered. Despite this limitation, the suggested system demonstrates high practical potential.

VI. Limitations

However, several limitations should be noted regarding the application and generalization of the above-described system.

The first limitation is that since pose estimation is performed in 2D space, the system faces challenges related to the depth dimension, which may cause distortions in pose angles. The position of the capturing camera plays a decisive role here.

TABLE III
PERFORMANCE EVALUATION ACROSS MULTIPLE CRICKET SHOTS

Shot	Frame	Elbow	Knee	Head Err	Score	Grade
ff_defence_shot1	67	141.82	117.20	0.0326	19	C
ff_defence_shot2	59	161.15	148.68	0.0222	22	B
ff_defence_shot3	61	156.21	134.58	0.0223	22	B
ff_defence_shot4	70	125.01	126.85	0.0142	22	B
ff_defence_shot5	68	156.99	143.30	0.0203	22	B
ff_defence_shot6	76	146.38	144.80	0.0244	22	B

Furthermore, although synchronization of frames from multiple cameras ensures high-quality timestamping, some errors may occur. These errors may amount to ± 1 frame and could negatively impact the detection of specific events.

Additionally, the color-based segmentation method used for ball tracking may produce inaccuracies in ball localization due to unfavorable conditions such as poor visibility or occlusions.

Moreover, the rule-based feedback system operates according to standards established by a general understanding of biomechanics. However, it may not account for individual differences in shot execution or playing style.

Finally, the lack of diversity in the data sample may negatively affect the generalizability of the solution.

VII. CONCLUSION

This paper presented a holistic multi-view framework for automated cricket shot analysis that combines synchronized signal processing, kinematic event detection, pose-based biomechanical evaluation, and feedback generation. In the proposed approach, clap sound is used for initial synchronization, while further motion analysis helps achieve precise synchronization among multiple camera views. The novel approach to bat-ball contact detection uses a combination of wrist speed, stable front foot position, and ball–bat distance. This method is robust against the drawbacks associated with using individual features for contact detection.

Key biomechanical parameters such as elbow angle, knee angle, wrist speed, and head alignment at the kinematic event frame are calculated from the batsman's pose. This set of parameters effectively provides insight into the coordination between upper body movements and lower body stability.

Experimental analysis has proven that the methodology is highly successful in providing multi-view synchronization, contact detection, and comprehensive biomechanical analysis from simple RGB video inputs. It can be considered a cost-effective alternative to expensive motion capture techniques. In addition, future research will focus on extending the proposed technique to 3D pose estimation, ball detection under diverse conditions, and designing an adaptive feedback module using machine learning techniques.

REFERENCES

[1] P. Ferdinands, U. Kersting, and R. Marshall, "Three-dimensional biomechanical analysis of cricket batting," *Sports Biomechanics*, vol. 9, no. 1, pp. 1–16, 2010.

[2] R. Stretch, "The biomechanics of cricket batting: A review," *Journal of Sports Sciences*, vol. 14, no. 2, pp. 99–106, 1996.

[3] S. Noorbhai and T. Noakes, "A descriptive analysis of batting backlift technique in cricket," *South African Journal of Sports Medicine*, vol. 28, no. 1, pp. 1–6, 2016.

[4] I. Renshaw, A. Oldham, and M. Bawden, "Movement dynamics in cricket batting," *Journal of Sports Sciences*, vol. 28, no. 8, pp. 813–823, 2010.

[5] M. R. Portus, B. R. Mason, B. C. Elliott, M. C. Pfitzner, and R. P. Done, "Technique factors related to ball release speed and trunk motion in cricket fast bowlers," *Journal of Sports Sciences*, vol. 22, no. 10, pp. 1035–1045, 2004.

[6] T. B. Moeslund, A. Hilton, and V. Kruger, "A survey of advances in vision-based human motion capture and analysis," *Computer Vision and Image Understanding*, vol. 104, pp. 90–126, 2006.

[7] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *CVPR*, 2017, pp. 7291–7299.

[8] A. Lugaresi, J. Tang, H. Nash *et al.*, "Mediapipe: A framework for building perception pipelines," *arXiv preprint arXiv:1906.08172*, 2019.

[9] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," in *NeurIPS*, 2014.

[10] D. Tran, L. Bourdev, R. Fergus *et al.*, "A closer look at spatiotemporal convolutions for action recognition," in *CVPR*, 2018.

[11] Y. Caspi and M. Irani, "Spatio-temporal alignment of sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 11, pp. 1409–1424, 2002.

[12] R. M. Kanko, E. Laende, T. Davis *et al.*, "Concurrent assessment of gait kinematics using marker-based and markerless motion capture," *Scientific Reports*, vol. 11, p. 20193, 2021.