Baseball Swing Analysis from a Monocular Video

Taisei Hamada, Naoshi Kaneko, and Katsuto Nakajima

***Abstract*—Swing analysis during baseball games requires high-precision radar and stereo cameras, making their adoption challenging for amateur players. We propose a low-cost swing analysis method that estimates the impact position and swing trajectory using monocular video. Furthermore, our method computes the degree of upper swing and generates trajectory-superimposed videos to facilitate an easier understanding of swing characteristics.**

I. Introduction

Professional baseball leagues (e.g., Major League Baseball) have recently installed play analysis systems endowed with high-precision radars and stereo cameras [1], enabling them to track player movements meticulously. Although the obtained data hold significant potential for enhancing player performance, these specialized systems are not feasible for amateur baseball settings due to budget and location constraints. This limitation makes it challenging to analyze and improve the performance of amateur players. Although low-cost swing analysis methods exist [2], they require attaching a sensor to a baseball bat. They cannot be used during a game because attaching a sensor violates the game regulations.

This study proposes a method for analyzing baseball swings using readily obtainable monocular videos. Our method estimates the impact position and bat trajectory. Furthermore, we propose computing the degree of upper swing (*upper swing index*) and generating trajectory-superimposed videos to facilitate easier understanding/comparison of swing characteristics.

II. Proposed Method

Fig. 1 presents a flowchart of the process. The process begins with the input video, wherein an object detector localizes the ball and bat. Additionally, a human pose estimator [3] is employed to determine the coordinates of the batter's abdomen and wrists. Subsequently, we estimate the impact position. Assuming that the straight lines represent the trajectories of the ball, the intersection point is established. The intersection coordinates indicate the x-coordinates of the ball at the time of impact. The y-coordinate transition of the ball is subjected to spline interpolation, deriving the y-coordinate at the impact. After that, the swing trajectory is estimated. A binary image of the detected bat rectangle is generated through inter-frame subtraction. Line detection is applied to the binary image to obtain a line segment of the bat. The endpoints are discerned as the tip and root of the bat, based on the positions of the batter's wrists. Finally, the upper swing index is calculated from the trajectory of the bat tip. Furthermore, we superimpose one swing trajectory of one batter onto another batter's video to facilitate a comparative analysis of trajectories.

T. Hamada, N. Kaneko, and K. Nakajima are with Tokyo Denki University, 5 Senju-asahi-cho, Adachi-ku, Tokyo, Japan.
(e-mail: 23fmi25@ms.dendai.ac.jp, naoshi.kaneko@mail.dendai.ac.jp, nakajima@mail.dendai.ac.jp)

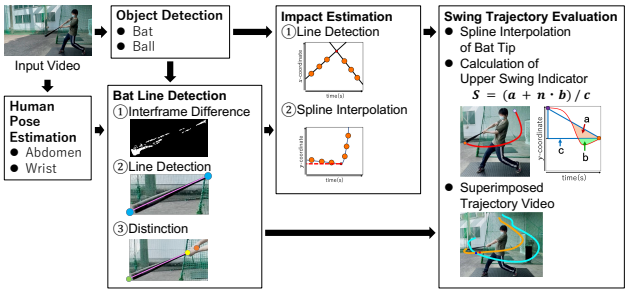

Fig. 1. Flowchart of the proposed method

III. Experiments and Results

YOLOv7 [4], trained on the COCO dataset, was employed as an object detector. We additionally trained the detector using blurred images of balls and bats to improve the detection accuracy of fast-moving balls and bats. We evaluated the accuracy of the ball and bat detection, impact estimation, and bat line segment estimation. Additionally, we conducted a questionnaire survey to assess the usefulness of the upper swing index and superimposed trajectory video.

The additional training improved the precision (87.9 $\rightarrow$98.1%) and recall (74.3$\rightarrow$90.7%) rates of the ball and bat detection. The average estimation error of impact timing was about less than one frame, while the average error of the impact position estimation was approximately twice the diameter of the ball. The average position error of the bat line segment estimation was about 10% of the length of the bat. Results from the questionnaire survey indicated that most participants agreed that the upper swing indices only captured the rough characteristics of the trajectories. However, most participants agreed that the superimposed video facilitated an easier comparison of trajectory disparities than a mere side-by-side comparison of the two videos.

IV. Conclusions and Future Work

This study has proposed a baseball swing analysis method using monocular video. Future work includes extending swing trajectory estimation to perform an analysis in a 3D space.

References

[1] Baseball Savant, "Statcast," https://baseballsavant.mlb.com/statcast_leaderboard. (Accessed 15 August 2023)
[2] Mizuno Corporation, "BLAST," https://jpn.mizuno.com/baseball/products/BLAST. (Accessed 15 August 2023)
[3] Z. Cao, et al., "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," CVPR, 2017.
[4] C.-Y. Wang, et al., "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," CVPR, 2023.