A Trajectory-Based Ball Tracking Framework with Enrichment for Broadcast Baseball Videos

Hua-Tsung Chen, Hsuan-Sheng Chen, Ming-Ho Hsiao, Yi-Wen Chen, and Suh-Yin Lee
Department of Computer Science and Information Engineering,
National Chiao-Tung University, Hsinchu, Taiwan, ROC.
{huatsung, xschen, mhhsiao, ewchen, sylee}@csie.nctu.edu.tw

ABSTRACT

For ball detection and trajectory extraction in broadcast baseball videos, this paper presents a novel framework which utilizes motion characteristic to extract ball trajectory from lots of moving objects. First, a series of constraints are applied to reduce ball candidates in each frame. Analyzing the distribution of ball candidates, the estimate function is built to explore the trajectory for the pitched ball. The experiments of ball tracking on broadcast baseball videos of JPB, MLB and CPBL from different channels show promising results. In addition, for video enrichment, the obtained baseball trajectory is superimposed on the frame to present the pitch in detail, which not only provides entertainment effects but also can be applied to baseball player training.

1: INTRODUCTIONS

With the rapidly advancing technology of digital equipments, nowadays it is much easier to capture and archive digital videos for general users. The urgent requirements for video applications therefore motivate researchers to devote themselves to various aspects of video analysis. In recent years, sports video analysis is receiving increasing attention due to the potential commercial benefits and entertainment functionalities. The possible applications of sports video analysis have been found almost in all sports, e.g., baseball, soccer, tennis, volleyball, etc. The major research issues of sports video analysis are categorized as follows:

Shot Classification: In a sports game, exploiting the properties that the positions of cameras are fixed in the game and the rules of presenting the game progress are similar in different channels, many shot classification methods are proposed based on camera motion, color information, texture information or face detection [1], [2].

Highlight Extraction: Due to broadcast requirement, highlight extraction attempts to abstract a long game into a compact summary to provide the audience a quick browse of the game. Many successful approaches are proposed based on audio analysis [3], semantic marker detection [4], video feature extraction and highlight modeling [5].

Ball and Player Tacking: Tracking is widespread used in sports analysis. Since significant events are mainly caused by ball-player and player-player interactions, balls and players are tracked most frequently. Common tracking techniques include trajectory-based ball detection and analysis [6], [7], physical model-based 3D trajectory reconstruction [8], [9] and 3D position estimation with multiple cameras [10], [11], [12], [13].

In addition, computer-assisted umpiring and tactics inference are burgeoning research issues of sports video analysis. However, these can be considered as advanced applications based on ball and player tracking. Therefore, tracking is an essential and vital technique in sports video analysis. In this paper, a trajectory-based ball tracking method is provided for baseball videos. The characteristic that the ball moves parabolically in the air is exploited for more reliable trajectory extraction. Although [11], [12] and [13] has been successfully applied to formal games, their high demanding in cameras installed locations and the visible area constrains themselves to be used in a studio-like sports field.

In this paper, we explore the possibility of building a trajectory-based ball tracking framework for broadcast baseball videos. Different from the approaches [9] [10] which rely on multiple cameras set in an ideal analysis environment, we provide an economy approach that can be generally used without specific camera installation.

2: PROPOSED FRAMEWORK

Based on the game-specific properties and visual features, we propose a framework which extracts ball trajectories in broadcast baseball videos, as depicted in Fig. 1. First, the moving objects of each frame are segmented in the pitch scenes. Each frame then generates ball candidates including the ball and some ball-like objects which satisfy the constraints of size, shape and fullness. The X and Y distributions of ball candidates in a sequence of frames are analyzed for trajectory exploring. Finally, identified trajectories are extracted.

There are two key novel points in this paper. The first is a trajectory-based method of ball tracking which allows the trajectories of all ball candidates to be identified reliably by utilizing the characteristic that, in the baseball video, the ball moves in a parabolic curve due to the gravity, even though there many ball-like objects in each frame. The second is a trajectory-based analysis of the pitched ball which evaluates a *breaking ball* by the curvature of the obtained trajectory and measures the speed of a *fastball* by the number of the detected ball candidates for video enrichment.

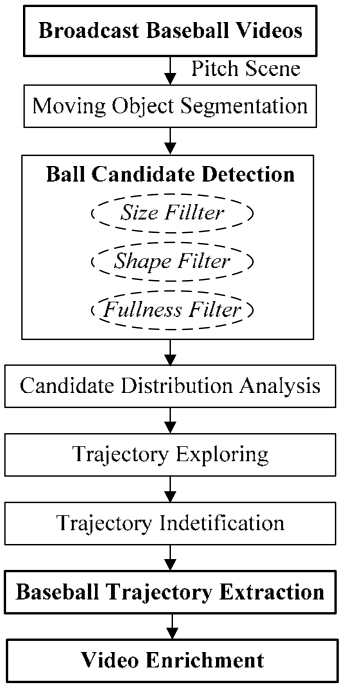


Fig. 1. Block diagram of the proposed ball tacking and video enrichment

3: BALL TRACKING AND TRAJECTORY EXTRACTION

Now we describe in turn the components of the algorithm: moving object segmentation, ball candidate detection, candidate distribution analysis, trajectory exploring and trajectory identification. The process of pitch scene detection is omitted, since it has been researched well in the literature.

3.1: MOVING OBJECT SEGMENTATION

Since the baseball in the video is white and bright, the luminance of the ball in a frame should be higher. That is, the moving ball is likely to be in the positive regions of intensity difference between frames. Hence, the positive regions of intensity difference between the current frame and the second frame before are segmented as moving objects, where the interval of two frames is used for the reason that it reveals a more complete shape of the ball than one-frame interval by experiments. Morphological operations are then performed to remove noises and make the regions filled.

3.2: BALL CANDIDATE DETECTION

In a frame, many non-ball objects might look like the ball and it is difficult to recognize which is the true one. Therefore, all detected moving objects are taken into consideration and sifted through the following filters with constraints. The remaining objects are considered as the *ball candidates* of the frame.

***Size Filter*:** The moving objects are filtered out if their sizes are not within the range [R_{min} , R_{max}], which is decided by statistical results.

***Shape Filter*:** The ball in the frame might have a shape different from a circle, but in most frames, its aspect ratio should be within the range [$1/R_a$, R_a]. The objects with aspect ratios out of the range should be filterer out.

***Fullness Filter*:** Some objects in different shapes (star-shaped, cross-shaped, triangular, etc.) may pass through the ball size filter and shape filter because of proper size and aspect ratio. For this reason, the fullness filter is built to remove the objects with the degrees of fullness D_f less than a threshold of T_f . D_f is defined as:

$$D_f = S_{obj} / A_{b\text{-}box} , \quad \textbf{(1)}$$

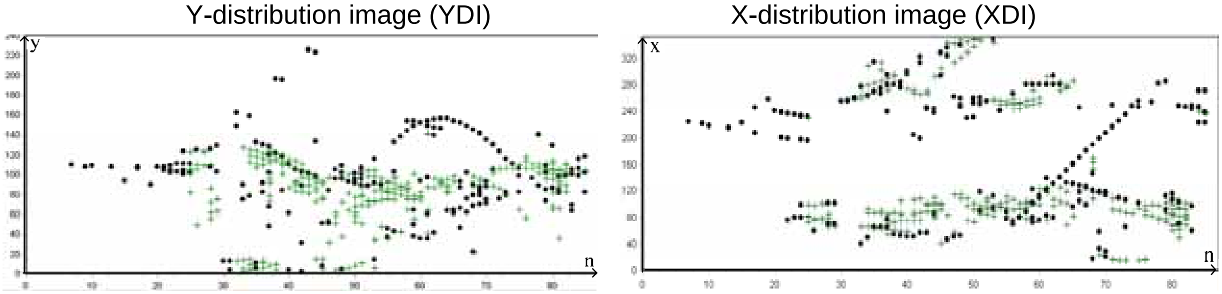
where S_{obj} is the size of the object in pixels and $A_{b\text{-}box}$ is the bounding box area of the object.

After filtering, the remaining objects are classified into *isolated* and *contacted* candidates according to their nearest objects in the frame. A candidate is called *isolated* if there is no neighboring object at a shorter distance than the average ball size, $(R_{min}+R_{max})/2$, and it is called *contacted* otherwise. This classification is important because the candidate close to other moving objects might be an over-segmented region of the pitcher or batter.

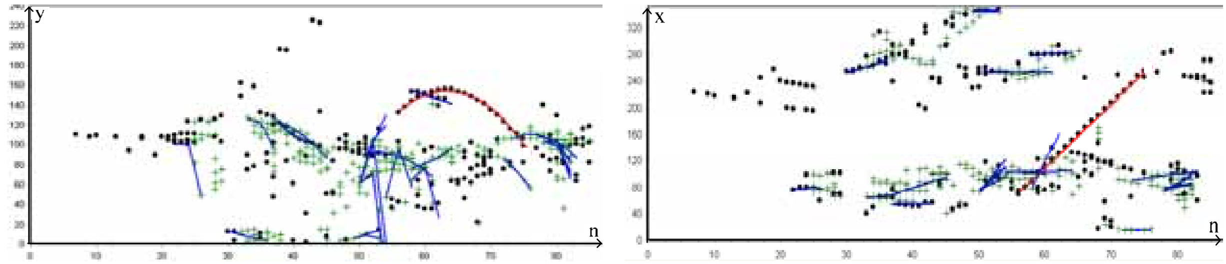
3.3: CANDIDATE DISTRIBUTION ANALYSIS

In a pitch scene, the baseball trajectory show in a almost parabolic curve due to the gravity, even for a *fastball*. By advanced analysis, in frames the ball moves parabolically in Y-direction and straightly in X-direction as time goes on. Exploiting this characteristic, a 2-D distribution analysis which preserves more information than the 1-D analysis [6] is proposed that ball candidate distribution in both Y- and X- directions analyzed to explore the trajectory more reliably.

A candidate distribution image is created by drawing the distribution of the candidates for a sequence of frames. The Y-distribution image (YDI) is created in such a way that its width equals the frame number of the given sequence and its height equals the height of the frame. Each isolated (or contacted) candidate draws a black dot (or green cross) in YDI at point $(x , y) = (n , y_c)$, where n is the frame serial number and y_c is the y-coordinate of the candidate in the original frame (the left-bottom corner of the frame is taken as the origin for presentation clarity of the parabolic curves). Similarly, the X-distribution image (XDI) is also created that its height equals the width of the frame, and each isolated (or contacted) candidate draws a black dot (or green cross) in XDI at point $(x , y) = (n , x_c)$, where x_c is the x-coordinate of the candidate in the frame. Fig. 2(a) presents an example of YDI and XDI.



(a) Ball candidates. Black dots are isolated candidates and green crosses are contacted ones.



(b) Candidate trajectories. The parabolic curves in YDI and straight lines in XDI represent candidate trajectories, among which the identified trajectories are colored red.

Fig. 2. Illustration of the Y- and X-distribution images for different process stages. In the figure, *n* is the frame serial number, *y* in YDI and *x* in XDI are the y- and x-coordinates of each candidate in the original frame, respectively,

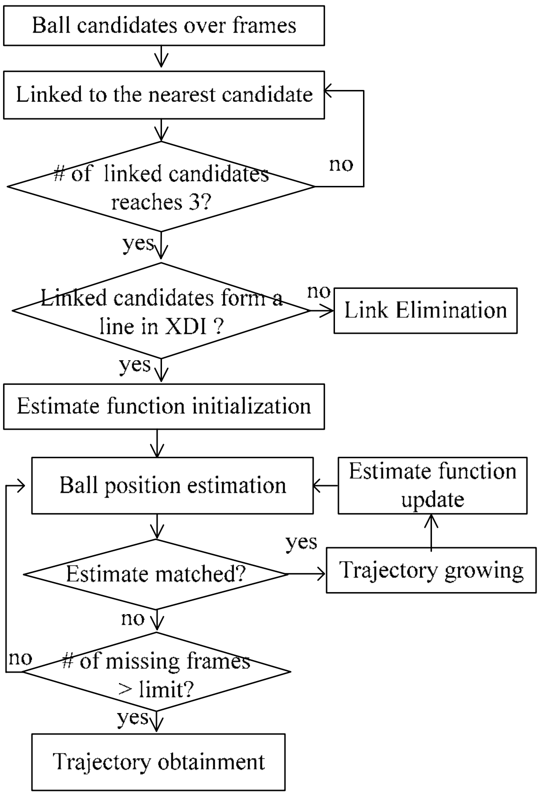


Fig. 3. Procedure of ball trajectory exploring

3.4: TRAJECTORY EXPLORING

The procedure to explore trajectories in YDI and XDI is summarized in Fig. 3. All ball candidates are first linked to the nearest neighbor in the next frame. As mentioned above, since in frames the ball moves parabolically in Y-direction and straightly in X-direction, the estimate functions for YDI and XDI are initialized as Eq.(2) and Eq.(3), when the number of linked candidates is up to three (three points form a parabola).

$$y = a \bullet n^2 + b \bullet n + c \,,\, a < 0 \qquad \textbf{(2)}$$

$$x = d \bullet n + e \qquad \textbf{(3)}$$

With the estimate functions, the ball position in the next frame is estimated. The estimate is considered matched if a ball candidate close to the estimated position is found. The trajectory then grows by adding this candidate and the estimate functions are updated. If there is no candidate close to the estimated position, the frame is regarded as a missing frame and the estimated position is taken as the ball position. The trajectory growing terminates when the number of consecutive missing frames reaches a given limit (4 in our experiments). The *candidate trajectories* produced from this procedure are shown as the parabolic curves in YDI and straight lines in XDI, as depicted in Fig. 2(b).

3.5: TRAJECTORY IDENTIFICATION

To indicate how likely a candidate trajectory is a ball trajectory, some confidential degrees are given by evaluating the following properties: estimation error, trajectory length and the ratio of isolated candidates.

Estimation error: The average distance of each ball candidate position from the estimated position is considered as estimation error. A confidential degree C_1 for estimation error is given to each candidate trajectory that slighter estimation error results in a higher degree. The candidate trajectories with considerable estimation error are eliminated.

Trajectory length: The confidential degree C_2 for trajectory length given to each candidate trajectory is proportional to its length because shorter trajectories are more likely noises.

Ratio of isolated candidate**:** Since the pitched ball is far from other moving objects in most frames, the ball trajectory should contain more isolated candidates than contacted ones. Hence, a confidential degree C_3 is given proportionally to the ratio of isolated candidates.

The candidate trajectory with the highest sum of C_1, C_2 and C_3, is identified as the ball trajectory. In Fig. 2(b), among the candidate trajectories, the identified trajectories are colored red. With the identified trajectory in YDI and XDI, the ball trajectory is computed and superimposed on the frame.

4: VIDEO ENRICHMENT

Based on the tracked ball positions and extracted trajectory, some visual presentations are provided for video enrichment. Fist, the speed of the pitched ball can be estimated from the time interval between the start and the end of the extracted trajectory. Second, the curvature of the parabolic curve in YDI can be used to evaluate a *breaking ball.* Moreover, the frame containing the first ball of the extracted trajectory displays the posture of the pitcher, while the frame containing the last ball of the trajectory indicates the ball position when it passes through the homeplate and the batter.

An example of video enrichment for a pitch is demonstrated in Fig. 4. Fig. 4(a) is the pitch frame of the first ball of the trajectory and the tracked ball positions are superimposed on the frame. Fig. 4(b) depicts the pitch frame of the last ball of the trajectory. In addition, the evaluations of the pitch are displayed at the bottom of the frame. More demonstrations of ball tracking with video enrichment are presented in the next section.

5: EXPERIMENTAL RESULTS

The proposed ball tracking framework has been applied to broadcast baseball videos (352 x 240) of JPB, MLB and CPBL captured from different sports channels, as listed in Table 1. Note that only pitch scenes are processed. Table 2 lists the parameter setting used in the experiments. The ball position of each video frame is manually recognized as *ground truth*. The experimental results of ball detection and tracking are listed in Table 3, where in the first column, "Video (Channel)" represents the video source and channel, "Pitch scenes" shows the clip number of each source, "Frames" represents the number of total frames in all clips, "Balls" represents the number of the frames containing the ball, "detected (%)" gives the number (percentage) of *detected* balls, "Miss" gives the number of missed ball, "False Alarm" gives the number of false-detected ball positions, and "tracked (%)" gives the number (percentage) of *tracked* balls. A ground truth ball is called "*detected*" if it matches a ball candidate generated in Ball Candidate Detection module (Section 3.2). It can be found that the percentages of the ball detection are not very high because the ball might be missed when it is passes through a left-handed batter. A ground truth ball falling on the obtained trajectory is called "*tracked*", since the ball position can be estimated on the trajectory by the motion characteristics even though it does not match a ball candidate. Although there are some tracking errors, the proposed method promotes the accuracy for ball tracking up to 96%. Fig. 5 demonstrates some ball tracking examples with video enrichment for different broadcast baseball videos containing left- and right- handed pitcher and batter.

Video	Channel	Clips
1. Japan Professional Baseball (JPB)	NHK of Japan	24
2. Major League Baseball (MLB)	ESPN of USA	14
3. Chinese Profession Baseball League (CPBL)	VL sports of Taiwan	24

Table 1. Testing videos used in the experiments

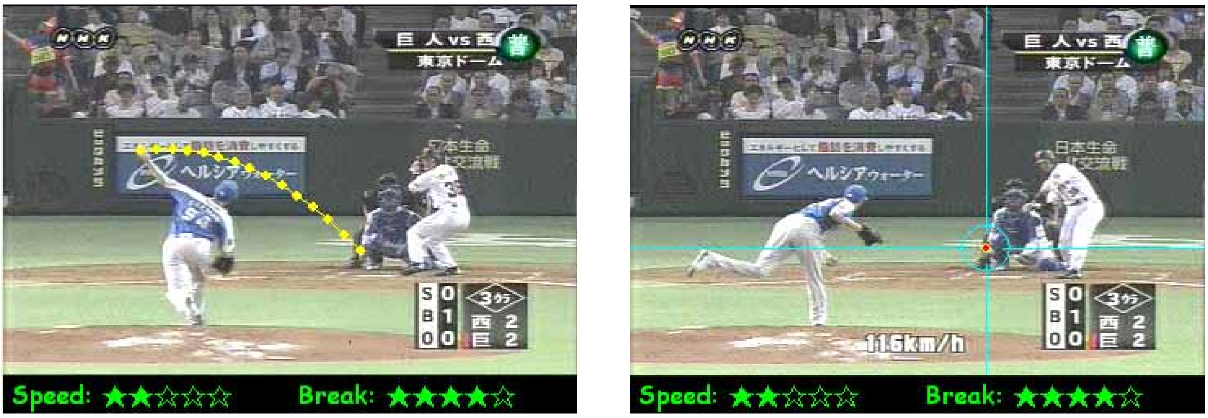
R_{min}	R_{max}	R_a	T_f
10	80	2	60%

Table 2. Parameters used for experiments

(a) Pitch frame of the first ball of the extracted trajectory with tracked ball positions superimposed on the frame

(b) Pitch frame of the last ball of the extracted trajectory with the focus aiming at the ball

Fig. 4. Demonstration of the obtained ball trajectory superimposed on the frame



(a) JPB with left-handed pitcher and right-handed batter

(b) MLB with right-handed pitcher and right-handed batter

(c) CPBL with right-handed pitcher and right-handed batter

(d) JPB with right-handed pitcher and left-handed batter

Fig. 4. Examples of ball tracking and video enrichment for different baseball videos

Video (Channel)	1. JPB (NHK)	2. MLB (ESPN)	3.CPBL (VL)	Overall
Pitch scenes	24	14	24	62
Frames	1726	604	942	3272
Balls	346	202	352	900
Detected (%)	324 (93.64 %)	190 (94.06%)	326 (92.61%)	840 (93.33%)
Miss	22	12	26	60
False Alarm	8	5	7	20
Tracked (%)	336 (97.11%)	196 (97.03%)	338 (96.02%)	870 (96.67%)

Table 3. Performance of ball detection and tracking

6: CONCLUSIONS

In this paper, we present a trajectory-based ball tracking scheme which is capable of extracting the ball trajectory in broadcast baseball videos. Non-ball objects are first filtered out by the constraints of size, shape and fullness. The motion of the baseball has a characteristic that the pitched ball moves in a parabolic curve. Utilizing this characteristic, the ball trajectory can be identified reliably. Based on the properties of the extracted trajectory, the baseball video is enriched by some visual presentations and the evaluation of the pitch is also provided.

The novelty of the proposed baseball video analysis is that our approach is based on the motion characteristics rather than traditional tracking methods based on low-level features. Furthermore, the speed of a *fastball* is measured by the time interval of the trajectory, and a *breaking ball* is also evaluated by the curvature of the extracted trajectory. The presentation of the ball trajectory superimposed on the video not only shows the flight of the ball for entertainment effects but also provides reference for player training. Our experiments on 62 clips from different channels show convincible performance.

In the future, the trajectory extracted in the framework of this paper will be further analyzed to recognize pitch types such as *fastball*, *curveball*, *slider* and other *breaking ball* in baseball videos. A practical system will be produced for further pitch analysis and intelligence collection in baseball videos.

REFERENCES

[1] Hua, W., Han, M., Gong, Y.: Baseball Scene Classification using Multimedia Features. Multimedia and Expo, 2002.ICME'02.Proceedings.2002 IEEE International Conference on, 1 (2002) 821-824.

[2] Pei, S. C., Chen, F.: Semantic Scenes Detection and Classification in Sports Videos. Proceedings of IPPR Conference on Computer Vision, Graphics and Image Processing (CVGIP), (2003) 210–217.

[3] Rui, Y., Gupta, A., Acero, A.: Automatically Extracting Highlights for TV Baseball Programs. Proceedings of the eighth ACM international conference on Multimedia, (2000) 105-115.

[4] Xiong, Z., Radhakrishnan, R., Divakaran, A., Huang, T.: Highlights Extraction from Sports Video Based on an Audio-Visual Marker Detection Framework. Multimedia and Expo, 2005.ICME 2005.IEEE International Conference on, (2005) 29-32.

[5] Xie, L., Chang, S. F., Divakaran, A., Sun, H.: Structure Analysis of Soccer Video with Hidden Markov Models. Acoustics, Speech, and Signal Processing, 2002.Proceedings. (ICASSP'02). IEEE International Conference on, 4 (2002).

[6] Yu, X., Xu, C., Leong, H. W., Tian, Q., Tang, Q., Wan, K. W.: Trajectory-Based Ball Detection and Tracking with Applications to Semantic Analysis of Broadcast Soccer Video. Proceedings of the eleventh ACM international conference on Multimedia, (2003) 11-20.

[7] Wang, J. R., & Parameswaran, N.: Detecting Tactics Patterns for Archiving Tennis Video Clips. Multimedia Software Engineering, 2004.Proceedings.IEEE Sixth International Symposium on, (2004) 186-192.

[8] Shum, H., & Komura, T.: A Spatiotemporal Approach to Extract the 3D Trajectory of the Baseball from a Single View Video Sequence. Multimedia and Expo, 2004.ICME'04.2004 IEEE International Conference on, 3 (2004) 1583-1586.

[9] Kim, T., Seo, Y., Hong, K. S., POSTECH, P.: Physics-Based 3D Position Analysis of a Soccer Ball from Monocular Image Sequences. Computer Vision, 1998.Sixth International Conference on, (1998) 721-726.

[10] Pingali, G., Opalach, A., Jean, Y.: Ball Tracking and Virtual Replays for Innovative Tennis Broadcasts. Proceedings of the 15th International Conference on Pattern Recognition, 4 (2000) 152-156.

[11] Hawk-Eye, http://news.bbc.co.uk/sport1/hi/tennis/2977068.stm.

[12] Gueziec, A.: Tracking Pitches for Broadcast Television. Computer, 35 (2002) 38-43.

[13] QUESTEC, http://www.questec.com/q2001/prod_uis.htm.

[14] Farin, D., Han, J.: Fast Camera Calibration for the Analysis of Sport Sequences. Multimedia and Expo, 2005.ICME 2005.IEEE International Conference on, (2005) 482-485.

[15] Farin, D., Krabbe, S., Peter H. N., Effelsberg, W.: Robust Camera Calibration for Sport Videos using Court Models. Proceedings of SPIE, 5307 (2004) 80.