JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 24, 143-157 (2008)

Short Paper

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

HUA-TSUNG CHEN, HSUAN-SHENG CHEN, MING-HO HSIAO, WEN-JIIN TSAI AND SUH-YIN LEE
Department of Computer Science and Information Engineering
National Chiao Tung University
Hsinchu, 300 Taiwan

Pitching contents play the key role in the resultant victory or defeat in a baseball game. Utilizing the physical characteristic of ball motion, this paper presents a trajectory-based framework for automatic ball tracking and pitching evaluation in broadcast baseball videos. The task of ball detection and tracking in broadcast baseball videos is very challenging because in video frames, the noises may cause many ball-like objects, the ball size is small, and the ball may deform due to its high speed movement. To overcome these challenges, we first define a set of filters to prune most non-ball objects but retain the ball, even if it is deformed. In ball position prediction and trajectory extraction, we analyze the 2D distribution of ball candidates and exploit the characteristic that the ball trajectory presents in a near parabolic curve in video frames. Most of the non-qualified trajectories are pruned, which greatly improves the computational efficiency. The missed balls can also be recovered in the trajectory by applying the position prediction. The experiments of ball tracking on the testing sequences of JPB, MLB and CPBL captured from different TV channels show promising results. The ball tracking framework is able to extract the ball trajectory, superimposed on the video, and in near real-time provide visual enrichment before the next pitch coming up without specific cameras or equipments set up in the stadiums. It can also be utilized in strategy analysis and intelligence statistics for player training.

Keywords: multimedia systems, video signal process, object tracking, computer vision and image understanding, visual enrichment, sports video analysis

1. INTRODUCTION

With the rapidly advancing technology of digital equipments, it is much easier to archive digital videos for general users. The urgent requirements for video applications therefore attract numerous research efforts. Recently, sports video analysis is receiving increasing attention due to the potential commercial benefits and entertainment functionalities. Possible applications of video analysis have been found almost in all kinds of sports, *e.g.*, baseball, soccer, tennis, *etc.* The major research issues of sports video analy-

Received February 2, 2007; accepted July 13, 2007.
Communicated by K. Robert Lai, Yu-Chee Tseng and Shu-Yuan Chen.
* The research is partially supported by the National Science Council of Taiwan, R.O.C., under the grant No. NSC 95-2221-E-009-076-MY3 and partially supported by Lee and MTI center for Networking Research at National Chiao Tung University.

sis can be categorized into *shot classification*, *highlight extraction* and *object tracking*.

In a sports game, the positions of cameras are usually fixed and the rules of presenting the game progress are similar in different channels. Exploiting these properties, many *shot classification* methods are proposed. Duan *et al*. [1] employ a supervised learning scheme to perform a top-down shot classification based on mid-level representations, including motion vector field model, color tracking model and shot pace model. Hua *et al*. [2] integrate color distribution, edge distribution, camera motion, sound effects and closed captions with maximum entropy scheme to classify baseball scenes. Lu and Tan [3] propose a recursive peer-group filtering scheme to identify prototypical shots for each dominant scene, and examine time coverage of these prototypical shots to decide the number of dominant scenes for each sports video.

Due to broadcast requirement, *highlight extraction* attempts to abstract a long game into a compact summary to provide the audience a quick browsing. Assfalg *et al*. [4] present a system for automatic annotation of highlights in soccer videos. Domain knowledge is encoded into a set of finite state machines, each of which models a specific highlight. The visual cues used for highlight detection are ball motion, playfield zone, players' positions and colors of players' uniforms. Rui *et al*. [5] propose an approach which utilizes audio features, including energy, MFCC (Mel-Frequency Cepstral Coefficients), entropy and pitch, to accomplish human speech endpoint detection, ball hit detection and excited human speech modeling in baseball videos. Cheng and Hsu [6] fuse visual motion information with audio features, including zero crossing rate, pitch period and MFCC, to extract baseball highlight based on HMM (Hidden Markov Model). Gong *et al*. [7] classify baseball highlights by integrating image, audio and speech cues based on MEM (Maximum Entropy Model) and HMM.

Object tracking is widely used in sports analysis. Since significant events are mainly caused by ball-player and player-player interactions, balls and players are tracked most frequently. In addition, *computer-assisted umpiring* and *tactics inference* are burgeoning research issues of sports video analysis. Indeed, these can be considered as advanced applications based on ball and player tracking, so tracking is an essential and vital issue in sports video analysis. Yu *et al*. [8] present a trajectory-based algorithm for ball detection and tracking in soccer videos. The ball size is first estimated from the goalmouth and ellipse to detect ball candidates. Exploiting a verification procedure based on Kalman filter, the true trajectory is extracted among potential trajectories generated from ball candidates. Wang *et al*. [9] track the ball movement and classify tennis games into 58 winning patterns, defined by US Tennis Association, on the basis of the ball trajectories and landing positions. Chu *et al*. [10] extract the potential ball trajectories using Kalman filter and simulate all possible trajectories of different beginning velocities, releasing angles and spin rates to form the physical limitation for trajectory identification.

In [11-13], 3D trajectory reconstruction is built based on multiple cameras located on specific positions. However, the high demanding for camera installation locations and the visible area constrains their systems to be used in a studio-like sports field.

Ball tracking in baseball videos is a challenging task since the high speed of the ball may cause ball deformation in video frames and the small size of the ball leads to tracking losses. In this paper, we develop a 2D trajectory-based ball tracking framework for broadcast baseball videos. Based on the observation, the baseball trajectory presents in a near parabolic curve in pitch scenes. We analyze the vertical and horizontal motion of the ball. Ideally, in the vertical direction, the ball moves parabolically due to the gravity, while in the horizontal direction, the ball moves in a straight line in spite of the air friction. In fact, the ball motion is not exactly a parabolic curve vertically and a straight line horizontally in video frames, but the characteristic of the near- parabolic/straight motion is sufficient for ball position prediction and trajectory extraction. The missed balls can also be recovered over the trajectory by applying the position prediction. Preserving more information than the 1D analysis [8], the proposed 2D distribution analysis has the advantage of extracting only those trajectories which form (near) straight lines in X-direction and (near) parabolic curves in Y-direction. Since most of the non-qualified trajectories are pruned, the computation efficiency is greatly improved. The extracted ball trajectory and visual enrichment can be provided in near real-time without the need of 3D trajectory reconstruction and high demanding cameras set up in an ideal environment.

The rest of this paper is organized as follows. Section 2 introduces the overview of the proposed framework. Section 3 presents the process of ball tracking and trajectory extraction. Section 4 addresses the trajectory-based pitching evaluation and visual enrichment. Experimental results are given in section 5, and finally section 6 concludes this paper.

2. PROPOSED FRAMEWORK

Based on the game-specific properties and visual features, we propose a framework to extract ball trajectories in broadcast baseball videos, as depicted in Fig. 1. First, the moving objects of each frame are segmented in the pitch shots. Each frame then generates *ball candidates* including the ball and some ball-like objects which satisfy the constraints of size, shape and compactness. Because of ball deformation caused by its speed, it is quite difficult to identify whether a single object is a ball. Hence, we utilize the physical characteristic of ball motion that the ball moves parabolically due to the gravity and identify whether a *potential trajectory* is the true ball trajectory. The X- and Y-distributions of ball candidates in a sequence of frames are analyzed to explore the trajectory which fulfills the physical characteristic. Finally, the baseball trajectory is extracted and the ball position in each frame can be located. In addition, visual enrichment and pitching evaluation can be presented based on the extracted ball trajectory.

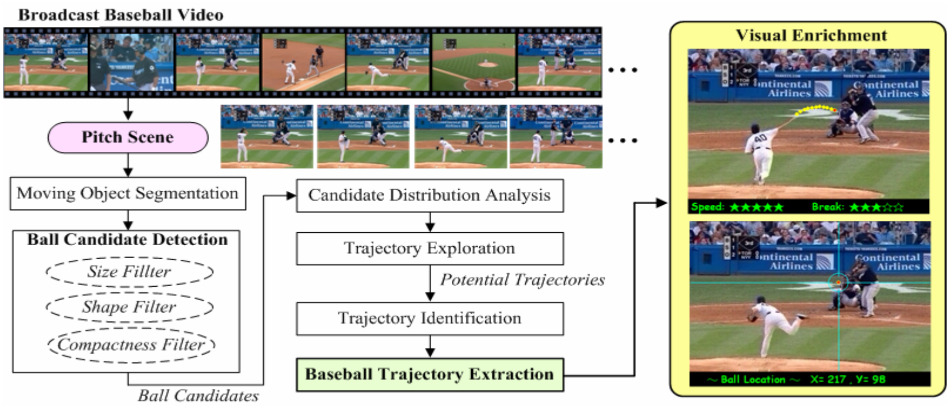


Fig. 1. Block diagram of the proposed ball tacking and visual enrichment.

3. BALL TRACKING AND TRAJECTORY EXTRACTION

Now we describe in turn the components of the proposed framework: moving object segmentation, ball candidate detection, candidate distribution analysis, trajectory exploration, trajectory identification and finally, baseball trajectory extraction. Shot classification and indexing in sports videos has been researched well in the literature [1-3, 14, 15]. We adopt the method in [15] and extract pitch shots using dominant color matching, region segmentation and dominant color layout analysis.

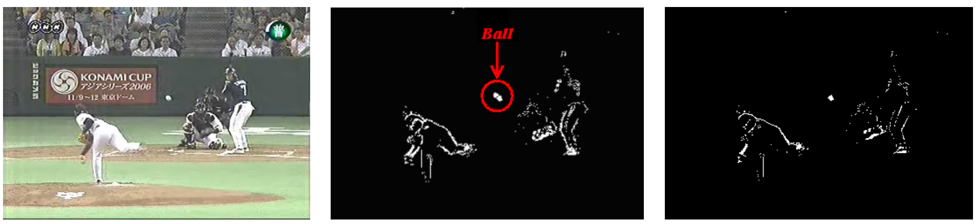
3.1 Moving Object Segmentation

Based on observation, there is usually no camera motion in pitch scenes, so frame difference method is applied for moving object segmentation. A *Frame Difference Image* (FDI) is a binary image formed by comparing every two successive frames (the intensity information is used). A pixel value of FDI is set to 255 if a significant difference occurs at the pixel location, and otherwise, the pixel value of FDI is set to 0, as defined in Eq. (1), where n is the frame sequence number and T_d is a threshold.

$$FDI_n(x, y) = \begin{cases} 255, & \text{if } \left| Intensity_n(x, y) - Intensity_{n-1}(x, y) \right| > T_d \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

Fig. 2 presents an example of segmenting the moving objects where the ball is included. Fig. 2 (a) gives the original frame and Fig. 2 (b) shows the FDI. It can be observed that the ball is included in a white region larger than the original ball size. This is because FDI takes the absolute value of intensity difference between frames. Since the baseball in the video is white and bright, the intensity of the ball in a frame should be higher. That is, the baseball is included in the positive regions of intensity difference between frames. Thus, the *Positive Frame Difference Image* (PFDI), defined as Eq. (2), is used to effectively segment positive regions of intensity difference which contain the ball, as shown in Fig. 2 (c). Morphological operations are then performed to remove noises and make the regions filled. Regions form by region growing and ball candidates will be detected among these regions.

$$PFDI_n(x, y) = \begin{cases} 255, & \text{if } Intensity_n(x, y) - Intensity_{n-1}(x, y) > T_d \\ 0, & \text{otherwise} \end{cases} \quad (2)$$



(a) Original frame. (b) FDI. (c) PFDI.

Fig. 2. Illustration of segmenting the moving objects where the ball is included.

3.2 Ball Candidate Detection

Many non-ball objects might look like a ball in video frames and it is difficult to recognize which is the true one. On the other hand, the ball might be presented in a shape different from a circle because of deformation. To sieve out the ball candidates from the moving objects segmented, the following filters are designed. After sieving, the remaining objects which satisfy the constraints are considered as the *ball candidates*.

1. **Size Filter:** Even though the ball size would vary due to the ball deformation and capturing conditions of cameras, it should fall within a specific range. The moving objects are filtered out if their sizes are not within the range $[R_{min}, R_{max}]$.
2. **Shape Filter:** The ball in frames might have a shape different from a circle, but the deformation is not so dramatic that its aspect ratio should be within the range $[1/R_a, R_a]$ in most frames. The objects with aspect ratios out of the range should be filtered out.
3. **Compactness Filter:** An object in a different shape may pass through the size filter and shape filter because of its acceptable size and proper aspect ratio. For this reason, the compactness filter is built to remove those objects with the degree of compactness D_c less than a threshold T_c. The degree of compactness D_c is defined in Eq. (3). Objects with low D_c would be filtered out while objects with high D_c would be retained.

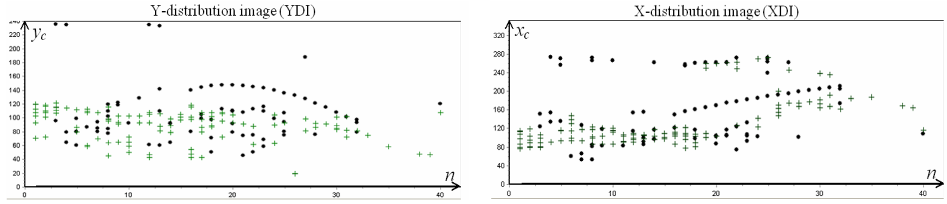
$$D_c = Object\_Size \,/\, Bounding\_Box\_Area \tag{3}$$

The pitched ball is at a distance away from other moving objects in most frames and the candidates close to other moving objects might be over-segmented regions of the pitcher or batter. To improve the accuracy of ball tracking, we classify the ball candidates into *isolated* and *contacted* candidates according to their nearest objects. A ball candidate is classified as *isolated* if there exists no neighboring object within a distance shorter than the average ball size, $(R_{min} + R_{max})/2$, and it is classified as *contacted*, otherwise.

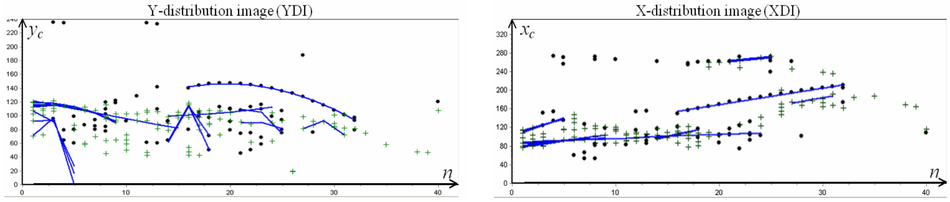
3.3 Candidate Distribution Analysis

In a pitch scene, the baseball trajectory presents in a near parabolic curve, even for a *fastball*. We further analyze the vertical and horizontal motion of the ball, separately. In Y-direction, the ball moves parabolically due to the gravity, while in X-direction, the ball moves almost straightly in spite of the air friction. Exploiting this characteristic, a 2D distribution analysis is proposed to explore the trajectory more reliably.

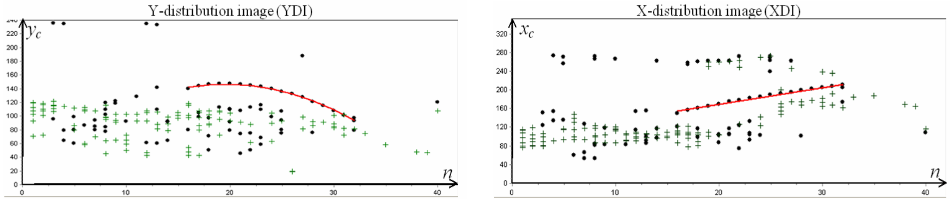
Fig. 3 (a) illustrates the candidate distribution analysis. A *candidate distribution image* is created by drawing the distribution of the candidates over frames. The *Y-distribution image* (YDI) is created in such a way that its width equals the length (in frame number) of the given sequence and its 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.



(a) Ball candidate distribution analysis. Black dots represent isolated candidates and green crosses represent contacted ones.



(b) Trajectory exploration. Potential trajectories are shown as the linking of ball candidates.



(c) Trajectory identification. The ball trajectory identified is shown as the parabolic curve in YDI and the straight line in XDI.

Fig. 3. Illustration of the Y- and X-distribution images for different process stages. In the figure, n is the frame serial number, y_c in YDI and x_c in XDI are the y- and x-coordinates of each candidate in the original frame, respectively.

3.4 Trajectory Exploration

Utilizing the 2D distribution of ball candidates, we attempt to explore the *potential trajectories* which form parabolic curves in YDI and straight lines in XDI simultaneously. The procedure of trajectory exploration is summarized in Fig. 4. 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 prediction functions for YDI and XDI are initialized as Eq. (4) and Eq. (5), when the number of linked ball candidates is equal to three.

$$y = a \bullet n^2 + b \bullet n + c, a < 0 \tag{4}$$
$$x = d \bullet n + e \tag{5}$$

By the prediction functions, the ball position in the next frame is predicted. The prediction is considered *matched* if a ball candidate close to the predicted position is found. The trajectory then grows by adding the candidate found and the prediction functions are updated by re-computing the best-fitting functions for the coordinates of the candidates detected so far using the *least square fitting* technique of regression analysis. If there exists no candidate close to the predicted position, the frame is regarded as a *missing*

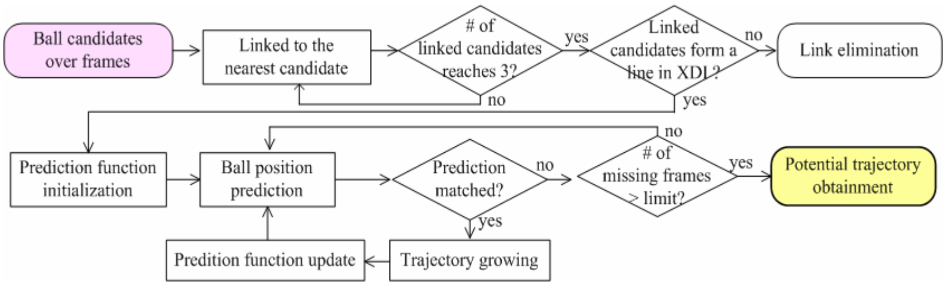


Fig. 4. Procedure of ball trajectory exploration.

frame and the predicted position is taken as the ball position. The trajectory growing terminates when the number of consecutive missing frames reaches a predefined limit (4 in our experiments). The potential trajectories produced from this procedure are shown as the linking of ball candidates in YDI and XDI, as depicted in Fig. 3 (b).

3.5 Trajectory Identification

After trajectory exploration, we obtain a set of potential trajectories. To identify the true ball trajectory from potential trajectories, we first prune the false ones to lower the computational complexity. For each potential trajectory, we have maintained the best-fitting function of the trajectory, the component ball candidates linked, and their associated coordinates and categories (isolated or contacted). The following properties are utilized to eliminate the potential trajectories which cannot be the ball trajectory.

Trajectory length: The distance from the pitcher to the catcher in a baseball field is about 18.39 meters, and it can be derived that a ball flying from the pitcher to the catcher at the speed of 180 km/h would last for at least 11 frames. (The detailed equation of ball speed estimation is described in section 4.) To the best of our knowledge, the highest ball speed in baseball games is no more than 170 km/h. Hence, the potential trajectories shorter than 11 frames could not possibly be a true trajectory and should be discarded.

Prediction error: The average distance (in pixel) of each ball candidate position from the predicted position is considered as *prediction error*. The potential trajectories with prediction error greater than a threshold T_e are eliminated.

Ratio of isolated candidates over all candidates on the trajectory: Since the pitched ball is at a distance away from other moving objects in most frames, the ball trajectory should contain more isolated candidates than contacted ones. On a potential trajectory, if the ratio of the isolated candidates over all candidates is less than 50%, the trajectory could not be the true one and should be discarded.

After elimination, much fewer potential trajectories remain. For each remaining trajectory, we compute the *length of consecutive isolated ball candidates.* The trajectory with the longest *length of consecutive isolated candidates* is finalized and extracted as the ball trajectory. The final ball trajectory after the procedure of trajectory identification is shown in Fig. 3 (c).

3.6 Baseball Trajectory Extraction

The scheme of baseball trajectory extraction is summarized as follows. First, the moving objects with high intensity are segmented out. Utilizing the constraints of size, shape and compactness, ball candidates are detected from the segmented moving objects. The distributions of ball candidates in both Y- and X-directions are analyzed. From the potential trajectories which form parabolic curves in YDI and straight lines in XDI, the ball trajectory is identified based on the properties of trajectory length, prediction error, the ratio of isolated candidates over all candidates on the trajectory and the length of consecutive isolated candidates. Finally, the ball position in each frame can be obtained and the ball trajectory can be extracted.

4. PITCHING EVALUATION AND VISUAL ENRICHMENT

More keenly than ever, the audience desires to perceive more comprehensive information about games. In this section, we apply the extracted baseball trajectory to pitching evaluation, such as *speed estimation* and *breaking measurement*, and use five-star evaluation to rank each pitch according to its speed and breaking degree.

Speed Estimation The distance from the pitcher's mound to the home plate is strictly defined in the game rules. Hence, as defined in Eq. (6), the ball speed (*BS* in km/h) can be estimated as the distance from the pitcher's mound to the home plate (18.39 m = 0.01839 km) divided by the time interval of the ball trajectory (*#frm* in frame). The ball speed estimation and the five-star evaluation are given in Table 1, which lists the time interval of the trajectory, the estimated ball speed and the respective evaluation.

$$BS(km/h) = \frac{0.01839(km)}{(\#frm/30/3600)(h)} \tag{6}$$

Breaking Measurement A *breaking ball* is a pitch which does not travel straightly like a *fastball*, and it would have a sudden drop when approaching the batter. The more the drop height is, the harder the batter can hit the ball. Furthermore, the drop height raises as the curvature of the trajectory increases. Hence, we measure a breaking ball according to the curvature of the parabolic curve in YDI (Y-distribution image), the coefficient a in Eq. (4). A breaking ball with larger curvature $|a|$ will gain higher ranking, that is, more stars. The breaking measurement and the five-star evaluation are given in Table 2.

Table 1. Ball speed estimation with comparative five-star evaluation using the ball trajectory.

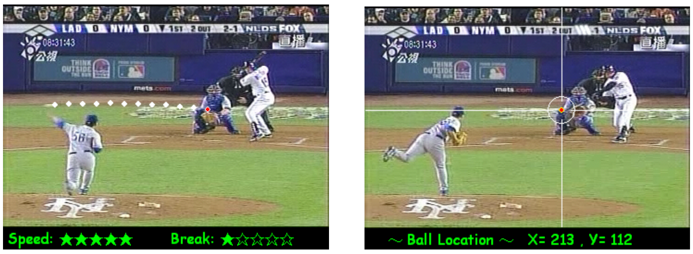
#frm	*BS* (km/h)	Evaluation	*#frm*	*BS* (km/h)	Evaluation
12	164	★★★★★	17	116	★★☆☆☆
13	151	★★★★☆	18	109	★★☆☆☆
14	141	★★★☆☆	19	104	★☆☆☆☆
15	131	★★★☆☆	20	98	★☆☆☆☆

Table 2. Breaking measurement with five-star evaluation.

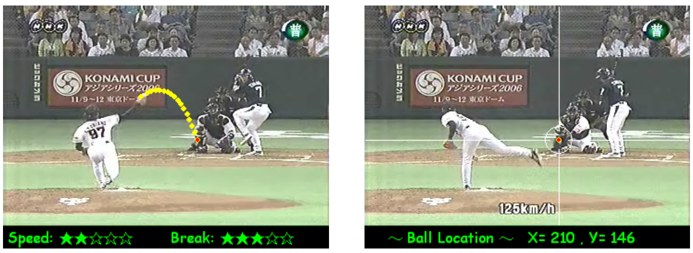
Curvature: $\|a\|$	Evaluation
$\|a\| > 0.5$	★★★★★
$0.4 < \|a\| \leq 0.5$	★★★★☆
$0.3 < \|a\| \leq 0.4$	★★★☆☆
$0.2 < \|a\| \leq 0.3$	★★☆☆☆
$\|a\| \leq 0.2$	★☆☆☆☆

The pitching evaluation in this paper aims at providing visual enrichment for entertainment effects based on the ball trajectory. Actually, in baseball rules there are no regulations about how fast a pitched ball can be considered as five-star or what the curvature of a five-star breaking ball is. Thus, the parameter settings, supported by two experienced experts in baseball games, in Tables 1 and 2 for speed estimation and breaking measurement are comparative values, not absolute values.

Two examples of the trajectory-based pitching evaluation and visual enrichment are demonstrated in Fig. 5, where Fig. 5 (a) is a MLB (Major League Baseball) pitch shot with a left-handed pitcher and Fig. 5 (b) is a JPB (Japan Professional Baseball) pitch shot with a right-handed pitcher. In the left picture of each example, the enriched frame presents the sight when the pitcher is about to throw the ball. The superimposed trajectory clearly depicts the sequence of ball motion for the pitch. In addition, the pitching evaluation displayed at the bottom of the frame provides more details about the pitch. In the right picture of each example, the final ball location of the trajectory is spotlighted with a crosshair (or reticle). If the batter swings at the pitched ball, the enriched frame catches up and reflects the situation how the ball is hit or missed, as demonstrated in the right picture of Fig. 5 (a). On the other hand, in baseball rules the *strike zone* is defined as that area over the home plate the upper limit of which is a horizontal line at the midpoint between the shoulders and the belt, and the lower limit is a line at the knees. Hence, if the batter does not swing, the crosshair can provide the reference for the strike/ball judgment, as shown in the right picture of Fig. 5 (b). Moreover, the ball trajectory and the final ball location can also provide assistant information for the professional personnel to infer the tactics which each pitcher usually adopts in specific situations, such as "the pitcher prefers throwing a breaking ball to the inside corner of the strike zone when there are runner(s) on the base(s) and a fast ball to the outside corner when there is no runner." More demonstrations of ball tracking with visual enrichment are presented in the next section.



(a) Example of a MLB (Major League Baseball) pitch shot with a left-handed pitcher.



(b) Example of a JPB (Japan Professional Baseball) pitch shot with a right-handed pitcher.

Fig. 5. Demonstration of pitching evaluation and visual enrichment. Left: the superimposed ball trajectory and pitching evaluation. Right: the final ball location spotlighted with a crosshair.

5. EXPERIMENTAL RESULTS

The proposed ball tracking framework has been tested on broadcast baseball videos (352 × 240, MPEG-1) captured from different sports channels, as listed in Table 3. Note that only pitch shots are processed. In our experiments, some parameters are used. T_d is the threshold of frame difference in moving object segmentation. Since the intensity of the baseball should be much higher than the background or other objects in the frames, we adaptively set T_d by Eq. (7), which can eliminate many noises and still retain the ball.

$$T_d = Average\_intensity\_\ of\_the\_frame \times 50\% \tag{7}$$

As to the range of size filter, up to 95% of the baseball sizes (in pixel) in the frames of the resolution 352 × 240 are within the range [8, 50] by statistical results, so [R_{min}, R_{max}] is set to [8, 50]. The parameter R_a is the threshold of shape filter. Generally speaking, the aspect ratio of the baseball should equal 1. Due to the high speed movement, the ball may deform over frames. Thus, for tolerance of deformation, the constraint of shape filter is loosed. Since an object with aspect ratio greater than 3 is far from a ball, R_a is set to 3. Since an object of compactness degree D_c less than half cannot be claimed to be "compact", the threshold of compactness filter T_c is set to 50%. Furthermore, though the ball trajectory over frames is not exactly a parabolic curve, a trajectory with great prediction error cannot be the ball trajectory. Thus, for reasonable error tolerance, the threshold of prediction error T_e is set to 2 (in pixel).

Table 3. Testing videos used in the experiments.

Baseball Videos	Source Channels
1. MLB (Major League Baseball)	PTS channel of Taiwan
2. JPB (Japan Professional Baseball)	NHK channel of Japan
3. CPBL (Chinese Professional Baseball League)	VL sports channel of Taiwan

The ball position in each video frame is manually recognized as *ground truth*. A ground truth ball is called "detected" if it matches a ball candidate. A ground truth ball falling on the obtained trajectory is called "tracked", since the ball position can be predicted on the trajectory by the motion characteristics even though it does not match a ball candidate. The experimental results of ball detection and tracking are listed in Table 4, where "video" represents the video sources, "pitch shot" shows the number of pitch shots, "total frames" represents total the number of frames in all the pitch shots and "ball frame" represents the number of the frames containing the ball. The row "ball detected (%)" gives the number (percentage) of balls detected, "false alarm" gives the number of false-detected balls, and "ball tracked (%)" gives the number (percentage) of balls tracked.

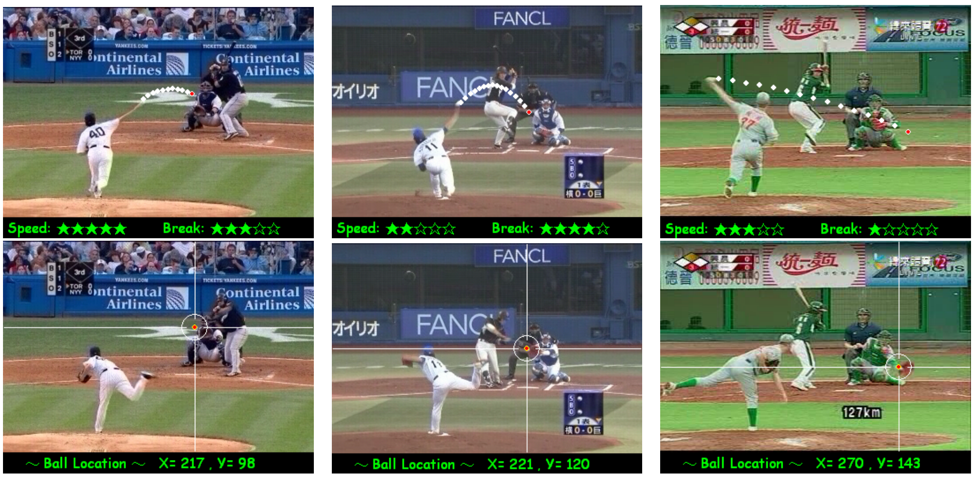
It can be found that there are some misses because the ball might not be detected when it passes over a left-handed batter dressed in a white uniform. Fortunately, the positions of missed balls can be recovered by applying the ball position prediction. An example of ball detection is shown in Fig. 6 (a), where the ball is missed in two frames when passing over the white uniform. The result of ball tracking is presented in Fig. 6 (b) where the missed ball positions can be recovered by applying the predicted positions of

Table 4. Performance of ball detection and tracking.

Video	1. MLB	2. JPB	3.CPBL	Overall
pitch shots	30	32	24	86
total frames	1380	2089	942	4411
ball frames	424	466	352	1242
ball detected (%)	387 (91.27%)	435 (93.35%)	326 (92.61%)	1148(92.43%)
false alarm (%)	11 (2.59%)	12 (2.58%)	7 (1.99%)	30 (2.41%)
ball tracked (%)	409 (96.46%)	453 (97.21%)	338 (96.02 %)	1200 (96.62 %)

(a) Ball detection. Two ball positions are missed when passing over the white uniform.
(b) Ball tracking. Positions of missed balls can be recovered.

Fig. 6. Illustration of ball detection and ball tracking.



(a) MLB pitch shot. (b) JPB pitch shot. (c) CPBL pitch shot.

Fig. 7. Examples of ball tracking and visual enrichment for various baseball videos.

the obtained trajectory. Although some tracking errors might exist, the proposed scheme promotes the overall accuracy of ball tracking up to 96%. The ball tracking with visual enrichment of some example pitch shots are demonstrated in Fig. 7. It is convincible that the proposed framework performs well in baseball videos from different channels, no matter whether the pitcher/batter is left- or right- handed.

The experiments run on an IBM ThinkPad X60 notebook computer (CPU: Intel Core Duo T2400 1.83GHz, RAM: 1GB). For a pitch shot of 2 seconds, the required processing time is about 8-10 seconds. In baseball games, the duration between two successive pitches is usually longer than 10 seconds. That is, the proposed framework is able to compute the ball trajectory of a pitch shot and superimpose the trajectory over the video before the next pitch coming up in near real-time. The application of enriching the live broadcast baseball video for entertainment effects becomes feasible.

It is difficult to perform a head-to-head comparison with other algorithms since there exist differences in the actual setup and the implementation. As a reasonable comparison, we divide the process into two stages: potential trajectory exploration and ball trajectory identification, and make the discussion.

5.1 Potential Trajectory Exploration

Kalman filter and particle filter are widely used in moving object tracking. However, particle filter is usually applied to tracking large objects with salient characteristics of edges or colors, such as cars and people [16]. Though particle filter can also be used in ball tracking, it is applicable to ball of big size, such as basketball [16], for which a distinguished target model can be built. Since most of the ball tracking algorithms in the literature [8, 10] are Kalman filter-based, we make a comparison focusing on Kalman filter. We compare the performance between the Kalman filter-based algorithm (KF) and the proposed parabola-based algorithm (PB). The performance metrics include the number of potential trajectories produced and the number of the ball candidates linked on the potential trajectories. For each pitch sequence, fewer ball candidates linked on the potential trajectories need fewer updates of the prediction function or Kalman filter. The fewer number of the potential trajectories is, the less computation in trajectory identification is.

Table 5. Comparison of the potential trajectory number and tracked candidate number on the 86 testing sequences.

		KF algorithm				Proposed PB algorithm			
Video	#Seq	#PT	Avg. #PT	#Cand	Avg. #Cand	#PT	Avg. #PT	#cand	Avg. #Cand
MLB	30	645	21.5	3819	127.3	520	17.33	2803	93.43
JPB	32	1120	35	6835	213.59	557	17.41	3352	104.75
CPBL	24	510	21.25	3297	137.38	234	9.75	1435	59.79
Total	86	2275	26.45	13951	162.22	1311	15.24	7590	88.26

Using the 86 testing sequences as in Table 4, the comparison is presented in Table 5. The notations #Seq, #PT and Avg. #PT represent the number of testing pitch sequences, the total number of potential trajectories produced in the pitch sequences and the average number of potential trajectories produced per pitch sequence. #Cand and Avg. #Cand denote the total number of ball candidates linked over all the potential trajectories and the average number of ball candidates linked per pitch sequence. It can be observed that KF algorithm produces more potential trajectories with more ball candidates linked, because KF algorithm may link neighboring non-ball objects in consecutive frames and form many potential trajectories which are not parabolic and need to be eliminated. However, the proposed PB algorithm aims at extracting only the potential trajectories which form (near) straight lines in X-direction and (near) parabolic curves in Y-directions, simultaneously. Therefore, the proposed parabola-based algorithm is more efficient in potential trajectory exploration since fewer ball candidates linked cause fewer updates of prediction function, and it will save more time in trajectory identification due to the fewer potential trajectories need to be validated.

5.2 Trajectory Identification

Extracting the true ball trajectory from lots of potential trajectories needs some identification mechanism. Chu *et al*. [10] simulate all the possible trajectories of ball pitching varying in different beginning velocities, releasing angles and spin rates to derive physical limitation for trajectory identification, which is time-consuming. To transform 2D trajectories into 3D trajectories for validation, they compute the ratio of "the vertical movement distance of pitches in the real world" (1 meter, assumed by the authors) to "the average vertical movement distance of pitches in the video frames of their dataset", and then estimate the depth of each ball candidate proportionally. However, the positions of pitchers releasing the ball and catchers catching the ball vary. The variation in the vertical movements of numerous pitches should be large and a pitch with the vertical movement far from the average, *e.g*. an underhand pitch, may not be identified reliably.

In our proposed scheme, we maintain the best-fitting function of the trajectory, the component ball candidates linked and their associated coordinates and categories (isolated or contacted) for each potential trajectory. Then, the properties for pruning the false trajectories and extracting the true ball trajectory (including *trajectory length*, *prediction error*, the *ratio of isolated candidates over all candidates on the trajectory*, and the *length of consecutive isolated candidates*) can be computed quickly. Therefore, the ball trajectory can be identified efficiently and reliably.

6. CONCLUSIONS

Since pitching contents are the crucial factors of the resultant victory or defeat in a baseball game, the professional personnel and the audience urgently require advanced information about the pitches. Ball tracking in baseball videos is a challenging task due to the small size and high speed of the ball. In this paper, we achieve ball tracking by applying the physical characteristic of ball motion. Trajectory-based pitching evaluation and visual enrichment can be provided in near real-time before the next pitch coming up. Our experiments on pitch shots captured from different channels show convincing results.

There are some key ideas in this framework. First, a set of filters are defined to prune most non-ball objects but retain the ball, even if it is deformed. Second, the 2D distribution of the ball candidates is analyzed exploiting the motion characteristic of the ball. Most of the non-qualified trajectories are pruned since only the trajectories which form (near) straight lines in X-direction and (near) parabolic curves in Y-direction are retained. Therefore, the computation efficiency is greatly improved so that the proposed ball tracking framework is able to extract the ball trajectory and provide visual enrichment in near real-time. Moreover, the missed balls can be recovered in the trajectory by applying the position prediction. 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 to players for *plate discipline* training. Furthermore, trajectory-based pitching evaluation is also presented to give the audience more comprehensive information about the game.

In the future, we will explore the possibility of 3D trajectory reconstruction providing more information about the ball trajectory for advanced pitching analyses, such as pitch type recognition, strike/ball decision and tactics inference. A practical system will be developed for pitch-bat strategy analysis and intelligence statistics in baseball videos.

REFERENCES

1. L. Y. Duan, M. Xu, Q. Tian, C. S. Xu, and J. S. Jin, "A unified framework for semantic shot classification in sports video," *IEEE Transactions on Multimedia*, Vol. 7, 2005, pp. 1066-1083.
2. W. Hua, M. Han, and Y. Gong, "Baseball scene classification using multimedia features," in *Proceedings of IEEE International Conference on Multimedia and Expo*, Vol. 1, 2002, pp. 821-824.
3. H. Lu and Y. P. Tan, "Unsupervised clustering of dominant scenes in sports video," *Pattern Recognition Letters*, Vol. 24, 2003, pp. 2651-2662.
4. J. Assfalg, M. Bertini, C. Colombo, A. D. Bimbo, and W. Nunziati, "Semantic annotation of soccer videos: automatic highlights identification," *Computer Vision and Image Understanding*, Vol. 92, 2003, pp. 285-305.
5. Y. Rui, A. Gupta, and A. Acero, "Automatically extracting highlights for TV baseball programs," in *Proceedings of the 8th ACM International Conference on Multimedia*, 2000, pp. 105-115.
6. C. C. Cheng and C. T. Hsu, "Fusion of audio and motion information on HMM-based highlight extraction for baseball games," *IEEE Transactions on Multimedia*, Vol. 8, 2006, pp. 585-599.
7. Y. Gong, M, Han, W. Hua, and W. Xu, "Maximum entropy model-based baseball highlight detection and classification," *Computer Vision and Image Understanding*, Vol. 96, 2004, pp. 181-199.
8. X. Yu, C. Xu, H. W. Leong, Q. Tian, Q. Tang, and K. W. Wan, "Trajectory-based ball detection and tracking with applications to semantic analysis of broadcast soccer video," in *Proceedings of the 11th ACM International Conference on Multimedia*, 2003, pp. 11-20.
9. J. R. Wang and N. Parameswaran, "Detecting tactics patterns for archiving tennis video clips," in *Proceedings of IEEE 6th International Symposium on Multimedia Software Engineering*, 2004, pp. 186-192.
10. W. T. Chu, C. W. Wang, and J. L. Wu, "Extraction of baseball trajectory and physics-based validation for single-view baseball video sequences," in *Proceedings of IEEE International Conference on Multimedia and Expo*, 2006, pp. 1813-1816.
11. A. Gueziec, "Tracking pitches for broadcast television," *Computer*, Vol. 35, 2002, pp. 38-43.
12. Hawk-Eye, http://news.bbc.co.uk/sport1/hi/tennis/2977068.stm.
13. QUESTEC, http://www.questec.com/q2001/prod_uis.htm.

14. S. C. Pei and F. Chen, "Semantic scenes detection and classification in sports videos," in *Proceedings of IPPR Conference on Computer Vision*, *Graphics and Image Processing*, 2003, pp. 210-217.
15. D. Zhong and S. F. Chang, "Structure analysis of sports video using domain models," in *Proceedings of IEEE International Conference on Multimedia and Expo*, 2001, pp. 713-716.
16. K. Nummiaro, E. Koller-Meier, and L. V. Gool, "An adaptive color-based particle filter," *Image and Vision Computing*, Vol. 21, 2003, pp. 99-110.

Hua-Tsung Chen (陳華總) received his B.S. and M.S. degree in Computer Science and Information Engineering from National Chiao Tung University, Taiwan in 2001 and 2003, respectively. Currently, he is pursuing the Ph.D. degree in Computer Science and Information Engineering in National Chiao Tung University, Taiwan. His research interests include computer vision, video signal processing, content-based video indexing and retrieval, multimedia information system and music signal processing.

Hsuan-Sheng Chen (陳宣勝) received the B.S. degree and M.S. degree in Computer Sciences and Information Engineering from Nation Chiao Tung University, Hsinchu, Taiwan. He is currently pursuing the Ph.D. degree. His research interests include human action analysis in surveillance system and sports video analysis.

Ming-Ho Hsiao (蕭銘和) received the B.S. degrees in Computer Sciences and Information Engineering from Fu Jen Catholic University, Taiwan in 2000. He received the M.S. degree in Computer Sciences and Information Engineering from National Chiao Tung University, Taiwan, where he is currently pursuing the Ph.D. degree. His research interests are video signal processing, content-based indexing and retrieval and distributed multimedia system, in particular, media server architecture and peer-to-peer system.

Wen-Jiin Tsai (蔡文錦) received the B.S., M.S. and Ph.D. degrees in Computer Science and Information Engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1992, 1993 and 1997, respectively. She was a software manager at the DTV R&D Department of Zinwell Corporation, Hsinchu, Taiwan, during 1999-2005. She has been an Assistant Professor at the Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan, since February 2005. Her research interests include video compression, video transmission, digital TV, and content-based video retrieval.

Suh-Yin Lee (李素瑛) received the B.S. degree in Electrical Engineering from National Chiao Tung University, Taiwan, in 1972, and the M.S. degree in Computer Science from University of Washington, U.S.A., in 1975, and the Ph.D. degree in Computer Science form Institute of Electronics, National Chiao Tung University. Her research interests include content-based indexing and retrieval, distributed multimedia information system, mobile computing, and data mining.