PHYSICS-BASED BALL TRACKING IN VOLLEYBALL VIDEOS WITH ITS APPLICATIONS TO SET TYPE RECOGNITION AND ACTION DETECTION

Hua-Tsung Chen, Hsuan-Shen Chen and Suh-Yin Lee

Department of Computer Science and Information Engineering,
National Chiao-Tung University, Hsinchu, Taiwan, ROC
{huatsung, xschen,sylee}@csie.nctu.edu.tw

ABSTRACT

Despite a lot of research efforts in sports video processing, little work was done in volleyball video analysis due to the high density of players on the court and the complicated overlapping of player-player or ball-player, which lead to great complexity of object tracking for advanced analysis. For ball detection and trajectory extraction in volleyball videos, this paper presents a physics-based scheme which utilizes the motion characteristics to extract ball trajectory from lots of moving objects. The experiments of ball tracking show encouraging results. Moreover, based on game-specific properties, the ball trajectory can be exploited to recognize set types for tactics inference and to detect basic actions in the volleyball game for close-up presentation.

***Index Terms*—** multimedia systems, video signal process.

1. INTRODUCTION

With the rapidly advancing technology of digital equipments, it is much easier to capture videos for general users. Urgent requirements for video applications therefore motivate researchers to devote themselves to various aspects of video analysis. Recently, sports video analysis is receiving increasing attention due to the potential commercial benefits and entertainment functionalities. 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].

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 physics-based ball tracking method is provided for volleyball videos. The characteristic that the ball moves parabolically in the air is exploited for more reliable trajectory extraction. Different from the approaches [10, 11, 12] which rely on multiple cameras set in an ideal analysis environment, we provide an economy approach that only one camera located behind the court is used. Moreover, based on the extracted trajectory, the set types can be recognized for tactics inference and the basic actions such as *serve*, *receive*, *set* and *spike* can also be detected.

2. PROPOSED FRAMEWORK

Based on the game-specific properties and visual features, we enhance the approach of [6] and propose a framework which extracts the ball trajectory, recognizes set types and detects basic actions in volleyball videos, as depicted in Fig.1. In particular, the major contribution is a physics-based method of ball trajectory extraction which utilizes the physical characteristic of ball motion and analyzes 2-D distribution of moving objects. Preserving more information than the 1-D analysis [6], this 2-D distribution analysis allows ball trajectories to be identified more reliably in sports games such as tennis, baseball, basketball and volleyball where the ball moves parabolically in most frames. In addition, exploiting domain knowledge, a trajectory-based semantic analysis is proposed to recognize set types and detect basic actions in volleyball videos.

This work is supported by the National Science Council under the grant NO. NSC95-2221-E-009-069-MY3 and NSC95-2221-E-009-076-MY3.

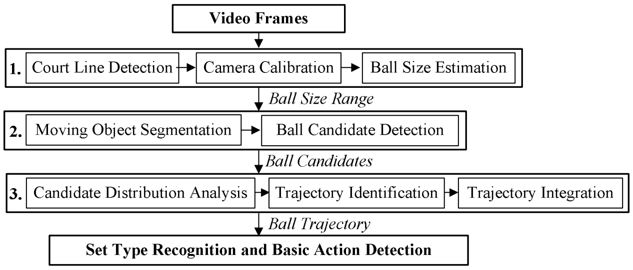


Fig. 1. Block diagram of the enhanced ball tracking framework

3. BALL TRACKING AND TRAJECTORY IDENTIFICATION

Now we describe in turn the components of the scheme. The procedures that have been presented in [6] are abridged.

3.1. Ball Size Range Estimation

Court line detection and camera calibration have been researched well in the literatures. The algorithm proposed in [13] is adapted in our work to obtain the lengths of the *end lines* in a frame. Let L_n and L_f be the lengths (in pixel) of the near and far end lines in a frame. The ball size range in a frame can be proportionally estimated as $[R_{min}, R_{max}] = [\pi \bullet (D/2 \bullet L_f/L)^2 - \Delta R, \pi \bullet (D/2 \bullet L_n/L)^2 + \Delta R]$, where D is the diameter of a real volleyball (21 cm), L is the length of the end lines on the court (900 cm) and ΔR is the extension for tolerance toward ball deformation.

3.2. Ball Candidate Producing

Each frame produces ball candidates by sifting moving objects through the constraints of size, shape and fullness.

Moving Object Segmentation. Since the ball in a volleyball game is apparent and bright, the luminance of the ball in a frame should be higher. Hence, the positive regions of intensity difference between the current frame and the third frame before are segmented as moving objects containing the ball, where the interval of three frames is used because it reveals a more complete shape of the ball than one- or two-frame interval by experiments. Morphological operations are then performed to remove noises and make the regions filled.

Ball Candidate Detection. In a frame, it is difficult to indicate the true ball. Therefore, all detected moving objects are sifted through the following filters and the remaining objects are considered as the ball candidates of the frame.

Ball Size Filter: The moving objects are filtered out if their sizes are not within the ball size range $[R_{min}, R_{max}]$.

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 [1/2, 2]. The objects with aspect ratios out of the range should be filterer out.

Fullness Filter: Some objects in different shapes 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:

$$D_f = S_{obj} \,/\, A_{b\text{-}box}\,, \tag{1}$$

where S_{obj} is the size of the object 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 a player may be an over-segmented region of the player.

3.3. Trajectory Generation

In a volleyball game, players are not allowed to hold the ball so that the ball trajectories almost show in parabolic curves. In frames, the ball moves parabolically in Y-direction and straight in X-direction as time goes on. Exploiting this physical characteristic, a 2-D distribution analysis is proposed that ball candidate distribution in both Y- and X- directions are analyzed to identify the trajectory reliably.

Candidate Distribution Analysis. 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 that 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 created that 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. An example is shown in Fig. 2(a).

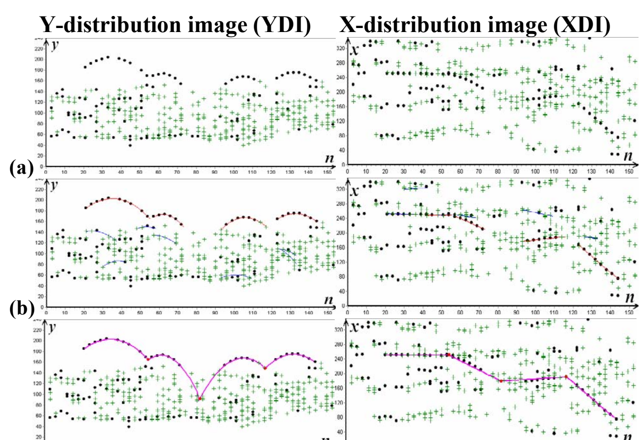


Fig. 2. Different process stages of YDI and XDI: **(a)** Ball candidates **(b)** Candidate trajectories **(c)** Integrated trajectory. In the figure, ***n*** is the frame serial number, ***y*** in YDI and ***x*** in XDI are y- and x-coordinates of candidates in original frames, respectively.

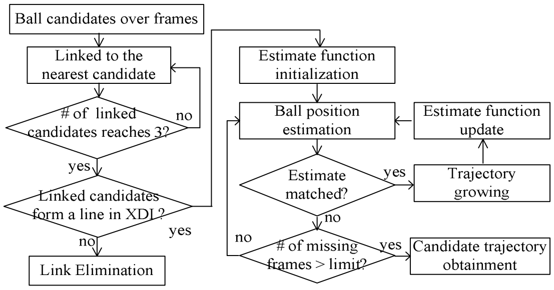


Fig. 3. Procedure of ball trajectory exploring

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. Since in frames the ball moves parabolically in Y-direction and straight 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 produced candidate trajectories are shown as the parabolic curves in YDI and straight lines in XDI in Fig. 2(b).

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: On a candidate trajectory, the average distance of each ball candidate position from the estimated position is considered as estimation error. A confidential degree C_1 is given to each candidate trajectory that slighter estimation error results in a higher degree.

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 candidates: Since the volleyball moves in the air 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.

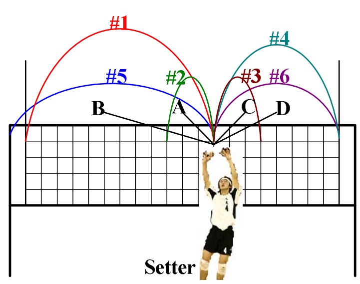
All candidate trajectories are then identified according to ($C_1 + C_2 + C_3$), as explained in the following pseudo code.

```
Let S be the set of candidate trajectories;
Set the identified ball trajectory I to be empty;
While (S is not empty) {
Move the trajectory T with highest C1+C2+C3 in S into I;
Eliminate in S the trajectories which overlap with T;       }
```

Trajectory Integration. To accomplish a complete trajectory, the gaps between pairs of contiguous identified trajectories can be patched by extending each trajectory based on estimate functions. An example of trajectory integration is demonstrated in Fig. 2(c).

4. BASIC ACTION DETECTION AND SET TYPE RECOGNITION

Based on the game rules, the ball changes its motion only when interacting with a player. Therefore, the basic actions of players can be detected at the transitions of the ball trajectory. Besides, a play in the volleyball game begins with a *serve* followed by iterative basic actions: *receive*, *set* and *spike*. Hence, we can locate each of these actions and indicate the sub-trajectory of *set* (*set curve*). Set type recognition is crucial for tactics inference because *spike* is the most effective way to get points and the set type governs a *spike*. Fig. 4 illustrates ten common set types with their respective discriminants. *A*, *B*, *C* and *D* are quick sets which players try to spike as soon as possible. #2 and # 3 are short sets next to the setter while #1, #4, #5and #6 are long sets toward the two sides of the net. A set can be recognized by classifying the *set curve* into one of the ten types with the discriminants, where *b* and *d* are coefficients in Eq.(2) and Eq.(3), and T is a threshold to decide whether the set is long or short.



Set	Discriminant
#1	$d<0$, $\|d\|>T$, $\|d\|/b\leq 0.75$
#2	$d<0$, $\|d\|<T$, $\|d\|/b\leq 0.75$
#3	$d>0$, $\|d\|<T$, $\|d\|/b\leq 0.75$
#4	$d>0$, $\|d\|>T$, $\|d\|/b\leq 0.75$
#5	$d<0$, $\|d\|>T$, $0.75<\|d\|/b\leq 1.5$
#6	$d>0$, $\|d\|>T$, $0.75<\|d\|/b\leq 1.5$
A	$d<0$, $\|d\|<T$, $0.75<\|d\|/b\leq 1.5$
B	$d<0$, $\|d\|/b>1.5$
C	$d>0$, $\|d\|<T$, $0.75<\|d\|/b\leq 1.5$
D	$d>0$,$\|d\|/b>1.5$

Fig. 4. Illustration of set type diagram with discriminants

5. EXPERIMENTAL RESULTS

For evaluation, the proposed method has been tested on two testing data sets (352x240 MPEG-1) containing iterative basic actions. **Set 1** are 20 clips captured from Asia Men's Volleyball Challenge Cup held in Kaohsiung, Taiwan with a digital camera, and **Set 2** are 10 clips downloaded from the website: http://volleyball.heeha.com/index.shtml. 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 1. A ground truth ball is called "detected" if it matches a ball candidate. It can be found that the percentages of the ball detection are not very high because the ball might be missed when it is close to a player. A ground truth ball located 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. Based on observation, we find that most tracking errors occur after the ball is spiked. If a spiked ball moves too fast, it could be blurred in the frame that the ball would be difficult to be detected. The tracking might fail if there are not enough detected ball candidates. Although there are some tracking errors, the proposed method promotes noticeable accuracy, which is up to 90%, for ball tracking. Some examples of ball trajectory extraction, action detection and set type recognition are demonstrated in Fig. 5, Fig. 6 and Fig. 7.

Table 1. Performance of ball detection and tracking.

Source	Set 1	Set 2	Overall
Clips	20	10	30
Frames	3285	1605	4890
Balls	2790	1260	4050
Detected (%)	1902 (68.17%)	985 (78.17%)	2887 (71.28%)
Tracked (%)	2538 (90.97%)	1175 (93.25%)	3713 (91.68%)

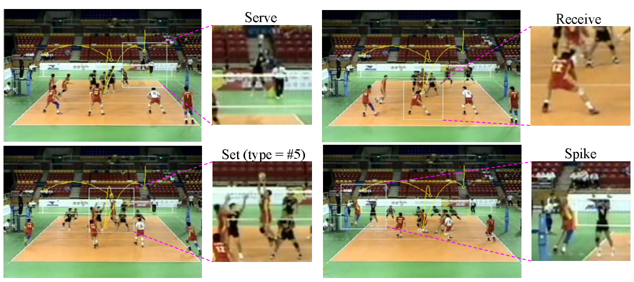


Fig. 5. Demonstration of ball trajectory extraction, action detection and set type recognition for a volleyball clip captured in Asia Men's Volleyball Challenge Cup

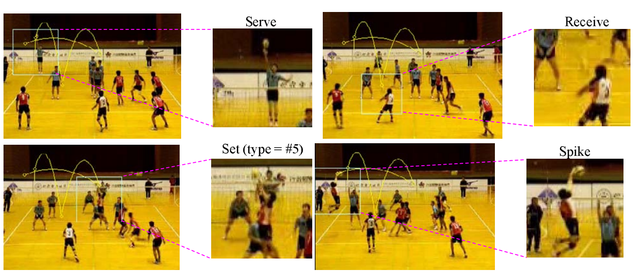


Fig.6. Demonstration for a volleyball clip downloaded from the website: http://volleyball.heeha.com/index.shtml

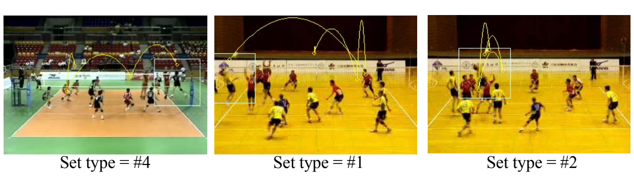


Fig. 7. Examples of set type recognition.

6. CONCLUSIONS

In this paper, we present a physics-based ball tracking scheme which is capable of extracting ball trajectories in volleyball videos. Utilizing the motion characteristic that the volleyball moves in a parabolic curve except when being touched by a player, a novel 2-D distribution analysis is proposed for reliable ball trajectory identification. This 2-D distribution analysis can also be applied to other ball games with similar ball motion characteristic. Furthermore, based on the game-specific properties, significant events are detected and set types are recognized.

In the future, the players will be tracked to detect advanced events such as *tandem*, *crossover*, *piston* and *slide*. Combined with the framework of ball tracking and set type recognition in this paper, an integrated system will be produced for further tactics inference and intelligence collection in volleyball games.

7. 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.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] Chen, H.T., Chen, H.S., Hsiao, M.H., Chen, Y.W., Lee, S.Y, "A Trajectory-Based Ball Tracking Framework with Enrichment for Broadcast Baseball Videos," *Proc. of the International Computer Symposium 2006*, (2006) 1145-1150.

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

[11] Gueziec, A., "Tracking Pitches for Broadcast Television," *Computer*, 35 (2002) 38-43.

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

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