Real-Time Position Estimation and Tracking of a Basketball

Bodhisattwa Chakraborty
Digital Image and Speech Processing Lab
National Institute of Technology Rourkela
Odisha, India 769008
Email: bodhisattwa.chakraborty@gmail.com

Sukadev Meher
Dept. of Electronics and Communication Engineering
National Institute of Technology Rourkela
Odisha, India 769008
Email: sukadevmeher@gmail.com

***Abstract*—In this paper, an algorithm to detect the position of a basketball in a real time outdoor video is proposed. The problem of ball detection in sports video arises due to the occlusion of the ball with the players, the presence of many moving objects like spectators, flags, twigs and branches of the tree etc. The ball image is also getting distorted due to the high speed of the ball. Thus the conventional direct detection and tracking algorithm fails in various occasions. It is much advantageous to study the trajectory information of a ball and estimate the ball location using the trajectory information. In this trajectory based algorithm, firstly the ball like objects are detected and a set of ball candidates are generated for each frame. Then the set of candidate trajectories are plotted. From these potential trajectories, the actual ball trajectory is estimated by studying the trajectory information. The missing ball location in any frame can also be predicted by trajectory interpolation technique. In this paper, we also propose an efficient background detection method using the approximate median value of the background pixels, thus achieving a high degree of segmentations of the moving objects in the foreground. This algorithm provides excellent result under a high degree of occlusion and works satisfactorily in the presence of many moving objects along with the ball in the scene.**

***Index Terms*—Ball detection and tracking, trajectory-based, sports video, occlusion, 2D trajectory plotting, trajectory interpolation.**

I. Introduction

With the rapid advancement in video production technology and the increasing consumer demand, the volume of multimedia information is increased drastically in the recent time. Now a day, general user can store multimedia data in a much easier way because of the vast availability of digital equipments. Sports video contains important multimedia contents and it also has a great deal of commercial benefits. Thus sports video analysis becomes an area of great interest among the researchers. The major research areas of sports video analysis are: shot classification, highlight extraction, content insertion, object tracking etc.

To classify different types of shots played in any sports is one of the fundamental areas of sports video analysis. Shot classification finds its application in extensive game analysis and player's performance measurement. Duan et al. [1] used a unified framework for semantic shot classification using mid-level representation and supervised learning algorithm. Lu and Tan [2] employed an unsupervised approach for dominant scene clustering in sports video using customized peer group filtering (PGF) and shot colour histogram (SCH) for shot classification.

Most viewers like to watch a sequence of "interesting events" rather than the entire game. Thus highlight extraction becomes an area of great interest among the researchers. Hanjalic [3] proposed an automatic highlight extraction module based on modelling the user's excitement using some selected low-level audio-video features. Xiong et al. [4] used an audio-visual marker detection framework to extract the highlights from a sports video.

Sports video is a power full tool of advertising as a large number of viewers are associated with it. Different contents like banners, logos are inserted and customized according to the liking of the local audiences. Yu et al. [5] presented a technique of inserting projected virtual contents into broadcast tennis videos based on acquired camera matrix using a 3D camera calibration algorithm. Xu et al. [6] proposed a method to implant virtual advertisements into broadcast soccer videos in real-time. The suitable locations in the video for advertisement insertion are detected by identifying the salient objects in the soccer video like static regions, center ellipse, goal mouth and field boundary.

Object tracking is a useful aspect of sports video analysis. Many previous works show various ways to track the ball and players in a sports video. Seo et al. [7] proposed an algorithm of ball tracking in broadcast sports video (BSV) using Kalman filter based template matching and backprojection. D'Orazio et al. [8] used modified circle Hough transform along with neural network classifier to detect the ball. Yu et al. [9] presented a trajectory-based algorithm for ball detection and tracking in BSV. The ball candidates are generated by estimating the ball size and shape from a number of ball-like objects present in the scene. The true trajectory of the ball is extracted using Kalman filter among the potential trajectories generated by the ball candidates. Chen et al. [10] proposed a physics-based ball tracking and 3D trajectory reconstruction algorithm to estimate the shooting location in a basketball video. Chen et al. [11] presented a trajectory-based ball tracking algorithm with visual enrichment for broadcast baseball videos.

The ball detection-and-tracking problem can be defined more precisely as to determine the location of the ball in each

frame of the video sequence. The main difficulties associated with detecting the ball in sports video analysis are:
• The small size of the ball with respect to the frame size.
• Change in the size of the ball with the increasing distance from the camera.
• The presence of many objects that resemble the ball in the frame, i.e. there are many "ball-like" objects in the frame.
• The shape, colour or size of the ball changes in every frame due to ball motion, camera motion and varying lighting conditions.
• Occlusion of the ball with players and merging of the ball image with lines in the frame.

In direct detection methods, the "ball-like" objects can be wrongly classified as the ball because of the continuously changing features (i.e. shape, size, colour etc.) of the ball image in subsequent frames. The speed of the ball and occlusion with players and backboard border can also generate many "ball-like" objects which may have a great deal of resemblance with the original ball.

In this paper we discussed the trajectory-based ball detection-and-tracking algorithm with trajectory interpolation to find out the missing ball location along the ball flight path due to occlusions and merging of the ball image with the players and backboard borders. Approximate median method of background subtraction is used to detect the moving objects in the foreground and to eliminate the effects of camera motion.

The rest of the paper is organized as follows. In Section II, the proposed method is presented. Section III gives an overview of the process of moving object detection. Section IV presents the method of ball candidate identification using various constraints like shape, size and compactness. In Section V the idea to plot the trajectories of ball-like objects and to extract the ball trajectory from the group of potential trajectories is presented. In this section the trajectory interpolation method is also discussed to determine the missing ball locations in the frame. Experimental results are analysed in Section VI. Finally the paper is concluded in Section VII.

II. Proposed Method

The proposed method of ball detection-and-tracking is based on two key ideas. The first step is to generate a set of ball candidates for each frame instead of direct detection of the ball. This eventually reduces the number of misses because the occluded balls and the merged balls are also included as ball candidates. The second idea is to compute the set of ball trajectories from the set of ball candidates. By exploring the set of candidate trajectories it is much easier to identify the ball trajectory and track the ball accurately in the basketball video sequence. Fig. 1 shows the block diagram of the proposed trajectory based ball detection and tracking algorithm.

The first step is to identify the moving objects in every frame which is done using background subtraction method. Then morphological opening, closing and dilation operations are performed to eliminate the small objects in the foreground. This is followed by edge detection which generates several ball candidates and several ball-like objects for each frame. Ball candidate identification is done using several filters based on size, shape and compactness. In the third step the potential 2D trajectories are generated by plotting the centroid locations of the ball candidates over time (i.e. the number of frames). The true ball trajectory can be identified by analysing the physical motion of the ball. The ball trajectory is identified by analyzing the ball movement along the X - direction and Y-direction separately. Using a best fitting function of the trajectory, the component ball candidates are linked to form the trajectories in X-, and Y-direction. Trajectory interpolation is done to find the missing ball positions and finally the computed trajectories are superimposed on the original frames.

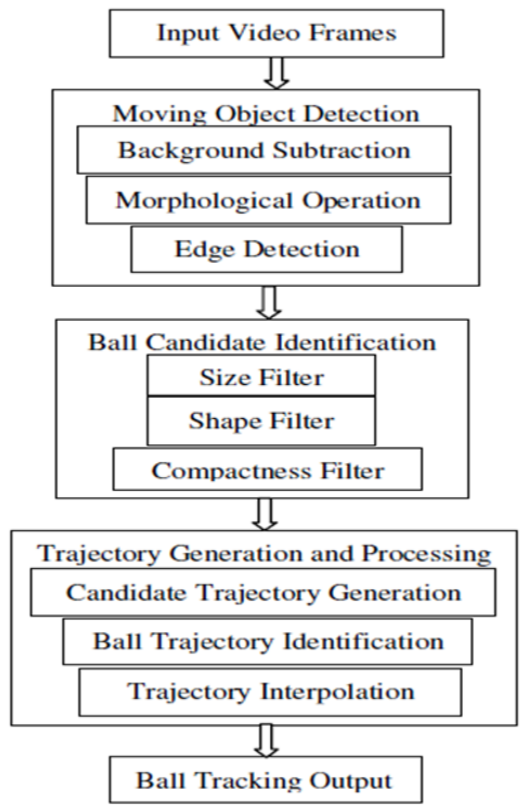

Fig. 1: Block diagram for basketball estimation and tracking from a real time video.

III. Moving Object Detection

A. Background Subtraction

The simplest way of background subtraction is frame differencing method. In this method, the intensity difference of every two consecutive frame is calculated and thresholded to obtain foreground pixels. Besides being very much computationally efficient, the frame differencing method is advantageous because it gives excellent results for continuously moving objects in the foreground and eliminates the external background noises, like waving trees. The method fails when there exists some significant camera motion. To incorporate with this, we use approximate median method of background subtraction [12]. In this method, the previous n-frames are stored and the background is calculated as the median of the stored frames. Then the background is subtracted from the current frame and a compared with a threshold to obtain the foreground pixels. In this way, the background converges to an estimate where half the pixels values are greater than the background and half are less than the background.

B. Morphological Operation

After the foreground pixels are successfully separated from the background, morphological opening and closing operations [13] are performed to remove small ball-like objects and fill the gap between the segmented regions. Morphological opening smoothes the object contours removes thin protrusions and breaks thin connections. The morphological opening of a set A by a structuring element B is denoted by $A \circ B$, and is given as,

$$A \circ B = (A \ominus B) \oplus B \tag{1}$$

The morphological closing of a set A by a structuring element B is denoted by $A \bullet B$ and is defined by,

$$A \bullet B = (A \oplus B) \ominus B \tag{2}$$

The morphological closing operation joins narrow breaks, fills long thin gulfs and fills holes and also smoothes the contours of the objects.

The presence of many less connected object in the segmented frames affects the proper detection process. Morphological dilation operation is performed to properly segment the foreground object from the background. Dilation operation "grows" objects in a binary image and the "growing" is controlled by the structuring element. Dilation of a set A with the structuring element B is mathematically expressed as,

$$A \oplus B = \{z|(\hat{B})_z|\} \cap A \neq \phi \tag{3}$$

where, ϕ is the empty set.

C. Edge Detection

After the foreground objects are properly segmented from the background, an edge detection technique is used to detect the edges in the segmented image. The edge detection method is used to detect the significant discontinuities in intensity values in an image. These discontinuities are created by object movement in the foreground and can be detected by using first and second-order derivatives. In our algorithm, we use a Canny Edge detector [7] to detect the edges because it can detect both the strong and weak edges in an image. In this method, the edge is determined by searching the local maxima of the gradient in the segmented image. The gradient is calculated using the derivative of a Gaussian filter. In this method, two threshold values are used to detect both the strong and weak edges. The weak edges are included in the output if they are connected with the strong edges.

IV. Ball Candidate Identification

The presence of many ball-like objects makes the task of ball detection and tracking difficult in a video sequence. Many objects from background like the waving tree, moving body part of players etc. may look like a ball after segmentation. On the other hand, the ball shape gets deformed due to the motion. To deal with this, we employ some filters to efficiently detect the ball candidates such as, size filter, shape filter and compactness filter. The objects which satisfy the constraints are treated as the ball candidates.

A. Size Filter

The ball size may vary in frame to frame because of the motion of the ball and position of the camera. In times, other objects from the background, resembles to the ball. So, it is important to eliminate the non-ball object by using a size filter. As for a video of frame size 640×480, analytical result shows that the basketball size (in pixel) should fall in the range [5, 30]. The objects which are inside this range are retained and used for further processing.

B. Shape Filter

The shape of the ball may change because of motion and occlusion with other objects in frame. That is why, in many frames, the ball does not look like a circle. Thus the objects with aspect ratio [4, 7] are characterized as ball candidates.

C. Compactness Filter

The non-ball objects with different shapes may pass through the shape and size filter because their size and aspect ratio may satisfy the limit defined for the volleyball. To eliminate these objects the compactness filter [14] is employed. The degree of compactness C is defined in Eq. 4,

$$C = A_{obj}/A_{rec} \tag{4}$$

Where A_{obj} is the area of the object and A_{rec} is the area of the smallest rectangle drawn around the circle. The maximum value of compactness for a circular region should be theoretically one. But the ball images get distorted due to motion of the ball and the camera and sometimes the ball does not look like a circle. To incorporate with this, we have to choose a threshold value for the compactness filter. Here, the threshold of compactness filter is to be chosen as 50 %. All objects below the threshold are filtered out as non-ball objects.

Fig. 2 shows the ball detection results for the three videos in different frames. The red rectangle is the smallest rectangle drawn around the circular region and the red cross depicts the centroid location of the detected object. From Fig. 2 (e), it is quite evident that the proposed ball detection framework works efficiently even if the ball is partially occluded with player's body part.

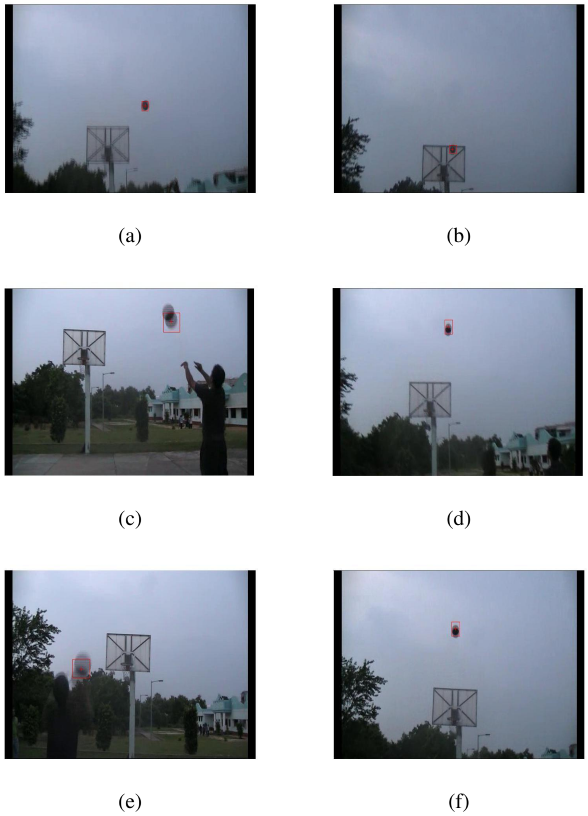

Fig. 2: Ball detection results for different video sequences. (a) - (b), (c) - (d) and (e) - (f): Detected ball in video sequences $VS-001$, $VS-002$ and $VS-003$ respectively.

V. TRAJECTORY GENERATION AND PROCESSING

In many cases, the no-ball objects passes the shape, size and compactness filter because they satisfies the criteria of the filters designed to prune the ball candidates. On the other hand, it is very likely that many ball candidates are filtered out because their shape, size and compactness values falls below the defined range due to the deformation, merging and occlusion of the ball image. This leads to a number of wrong detection and degrades the performance of ball tracking. A trajectory-based solution is used to overcome the problem where in addition to the size, shape and compactness constraints of the ball candidates, the physical characteristics of a ball in motion is also taken under consideration.

A. Candidate Trajectory Generation Phase

After the ball candidate identification phase, a set of ball candidates, S (F), is generated for each frame F. A set of candidate distribution plot [15] is generated by plotting the centroid locations of all the ball candidates over time which is represented by the frame number in the video sequence. The plot represents the location of the ball candidates in each frame. For a detailed analysis of the ball motion, the X- and Y- directional distribution of the ball candidates are studied. The X- directional candidate distribution plot is generated by plotting the x-coordinate locations of the ball candidates in each frame. Similarly, the Y- distribution plot is generated by plotting the y-coordinate locations of the ball candidates against the total number of frames in the video. Fig. 3 (a) and 3 (b) shows the X- and Y- distribution plot of the ball candidates respectively.

The potential ball trajectories are generated for X- and Y- distribution plot separately by linking the ball candidate in a frame with the ball candidate in the next frame. The prediction functions used to connect the ball candidates are shown in Eq. 5,

$$\begin{aligned} y &= p \cdot n^2 + q \cdot n + r, \quad p > 0 \\ x &= s \cdot n + t \end{aligned} \tag{5}$$

The ball candidates in the X and Y-direction are linked if the distance between two consecutive locations is not more than 3 and 20 respectively (in pixel value). The ball positions in each frame is predicted using the prediction functions shown in Eq. 5. The predicted location is verified with the original ball location in that frame. If the prediction matches with the original ball location, the prediction function is updated and the ball location in the next frame is predicted. The trajectory growing process terminates when the number of consecutive frames for which the verification fails, exceeds a predefined limit. The limit is set to 6 for this work. The candidate trajectories generated in x- and y- direction are shown in Fig. 3 (c) and (d) respectively.

B. Ball Trajectory Identification Phase

To identify the ball trajectory from the set of candidate trajectories, the horizontal and vertical motion of the ball is analysed. This 2D ball trajectory analysis reveals the fact that the ball movement along X-direction is almost a straight line despite the air friction. On the other hand, the ball moves in a near parabolic path in the Y-direction due to the gravitational force of the earth. In a basketball video, the ball is the continuous moving object over a number of frames. So the ball trajectory can be identified from the set of candidate trajectories as one with a smooth and relatively long trajectory path. The non-ball objects exhibit very short trajectories or no trajectory at all. Thus the trajectory shorter than 10 frames could not be classified as the ball trajectory and should be left out. The prediction error is calculated as the average distance from each ball candidates position in the original frame to the parabolic curve created by the predicted ball locations. The candidate trajectories with prediction error greater than a threshold value are discarded. The threshold value of the prediction error is 4 (in pixel value) for this work. Fig. 3 (e) and (f) illustrates the identified ball trajectory in X- and Y- direction respectively.

C. Trajectory Interpolation Phase

The next step is to process the obtained ball trajectory to find out the missing ball position in any frame along the ball trajectory path. The ball may get occluded with the player's

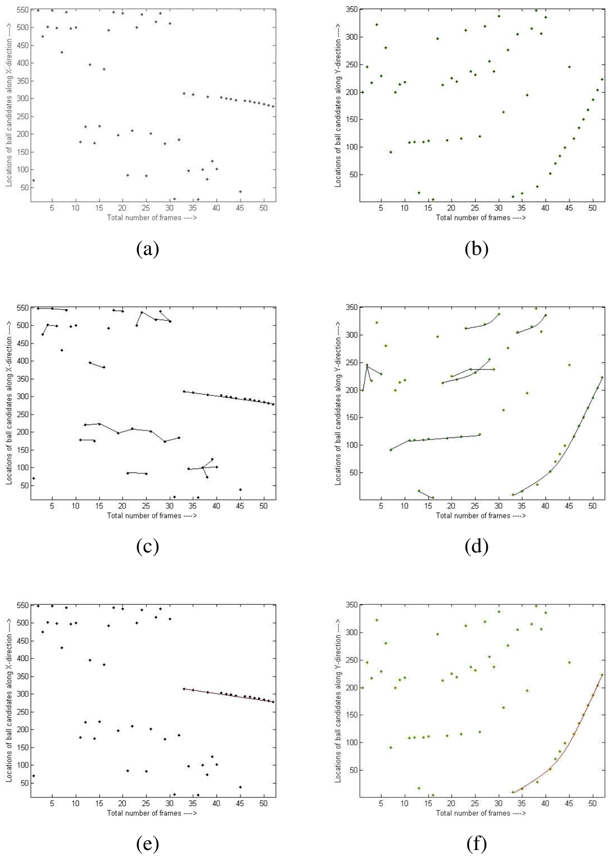

Fig. 3: Illustration of different stages of candidate trajectory generation and ball trajectory identification using X- and Y-distribution plots.

body part or the backboard borders. To obtain the missing ball position, a trajectory interpolation method is used in which the missing ball location is predicted using Eq. (5). If there exists no candidates close to the predicted location in a particular frame, the prediction function continues to predict the ball location for next few frames until the number of missing frames exceed a threshold value. The threshold value is set to 5 for this work. If the location is verified with a ball candidate in a frame within the threshold limit, the trajectory is extended up to that frame and the predicted positions are taken as the ball position. The predicted locations in the missing frames are then compared with the threshold of the prediction error for further validation. Once the missing ball positions are correctly estimated, the ball locations are plotted and superimposed on the original frame to verify with actual ball flight path.

VI. EXPERIMENTAL RESULTS

The algorithm is implemented in MATLAB R2010a on a system with Intel i5 processor with 2.40GHz processing speed under Microsoft Windows 7. The volleyball videos are taken using a Canon Camcorder in an outdoor volleyball court. The resolution of each video is 640×480, the frame rate is 30 frames per second.

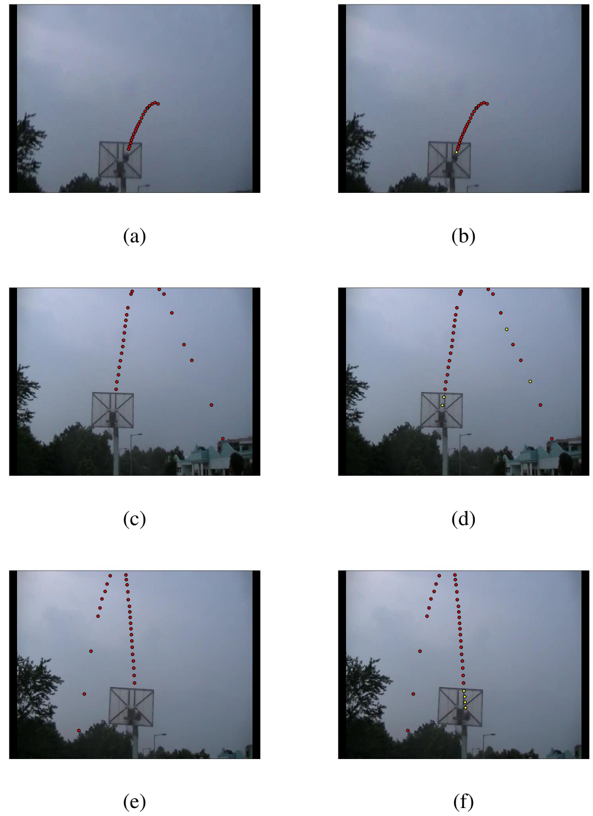

Fig. 4: Ball tracking results for different video sequences with missing ball position estimation from the trajectory processing. (a), (c) and (e): Trajectory of the ball flight path in video sequences $VS-001$, $VS-002$ and $VS-003$ respectively. (b), (d) and (f): Estimation of the missing ball locations along the ball trajectory in video sequences $VS-001$, $VS-002$ and $VS-003$ respectively.

The experiment was performed on three different video sequences namely $VS-001$, $VS-002$ and $VS-003$, taken using a Canon Camcorder in an outdoor volleyball court. The performance of the algorithm is summarized in Table I. The ground truth in each video sequence is obtained by manually inspection of the ball location in every frame. The notations "*tf*", "*bf*", and "*n-bf*" are used to denote total number of frames in the video, number of frames containing the ball and number of frames not containing the ball respectively. We have expressed the number of frames where the ball is correctly identified along with the no-ball frames as "*correct*", while "*tracked*" yields the number of frames where the predicted ball location is matched with the ground truth. For performance evaluation we showed both the ball detection results and the final results, i.e. ball tracking results after using the trajectory based method in Table I. It can be noticed that the proposed algorithm performs extremely well with an average accuracy of 94.77 % in detecting the ball, whereas it attains 100 % accuracy in tracking the ball correctly after trajectory processing. The ball tracking results are shown in Fig. 4. The red dots depicts the estimated ball positions and the yellow dots yields the interpolated ball locations.

TABLE I: Performanace ananlysis of ball detection and tracking module

Video Sequence	Ground Truth			Ball Detection Solution		Final Result	
	tf	*bf*	*n-bf*	*correct*	*Accuracy (%)*	*tracked*	*Accuracy (%)*
VS-001	23	21	05	25	96.1	21	100
VS-002	60	26	34	56	93.3	26	100
VS-003	86	35	51	82	95.35	35	100
Total	**172**	**82**	**90**	**163**	**94.77**	**82**	**100**

It is impossible to have a direct comparison with existing ball detection and tracking algorithm because the exact codes are not available. The type of data used in this work is also quite different from the other works. For performance comparison, we have implemented a Kalman filter [16], [17] based ball detection algorithm. For a fare comparison, we are only considering the ball detection solution. To compare the effectiveness of our proposed ball detection and tracking algorithm with Kalman filter based algorithm, we use precision, accuracy of ball detection and the number of false alarms as criteria. Table II shows the comparison between the proposed algorithm and Kalman filter based ball tracking algorithm. We let "*False*" the number of false alarms generated during the ball detection phase. Basically an algorithm generates a false alarm when it detects a ball at an incorrect location in a ball frame or wrongly detects a ball in a no-ball frame. It is observed that the proposed algorithm gives much better result in detecting the ball. The number of misses is much less and the number of false alarm is nil.

TABLE II: Comparison analysis of the proposed method with Kalman filter based method

Video Sequence	Proposed Trajectory-Based Method			Kalman Filter Based Method		
	Precision (%)	*Accuracy (%)*	*False*	*Precision (%)*	*Accuracy (%)*	*False*
VS-001	95.24	96.1	00	85.7	88.46	00
VS-002	84.60	93.3	00	76.92	88.3	01
VS-003	88.57	95.35	00	80.0	87.2	04

VII. Conclusion

The ball tracking in sports video becomes more difficult because of the presence of multiple moving objects in the background. Sometimes it becomes impossible to track the ball correctly using the direct detection methods. The proposed trajectory based algorithm not only detects and tracks the ball correctly, but it also helps to locate the occluded ball and the ball merged with other objects in the frame. The simulation results show that the proposed algorithm works efficiently in presence of many moving objects in the background. The algorithm is easy to implement as there is no need of complex calculations to determine the ball angle and the velocity.

There exist a vast scope of future research and application area in this direction. The proposed trajectory based method can be used to detect and track the ball in various sports videos. This approach can also be used in other detection and tracking problems and various surveillance tracking problems. The next area of the research includes 3D trajectory reconstruction to provide more information about the trajectory of the ball for better analysis and visualization. A practical system will be designed to throw the ball accurately to a target by analysing the trajectory information and using an artificial intelligence system with the existing ball detection and tracking module.

References

[1] L. Y. Duan, M. Xu, Q. Tian, C. S. Xu, J. S. Jin, *A unified framework for semantic shot classification in sports video*, IEEE Trans. on Multimedia, Vol. 7, pp. 1066-1083, 2005.

[2] H. Lu, Y. P. Tan, *Unsupervised clustering of dominant scenes in sports video*, Pattern Recognition Letters, Vol. 24, pp. 2651-2662, 2003.

[3] A. Hanjalic, *Adaptive extractions of highlights from a sports video based on excitement modelling*, IEEE Trans. on Multimedia, Vol. 7, No. 6, pp. 1114-1122, 2005.

[4] Z. Xiong, R. Radhakrishnan, A. Divakaran, T. S. Huang, *Highlights extraction from sports video based on an audio-visual marker detection framework*, IEEE Int. Conf. on Multimedia and Expo (ICME 2005), 4 pp. 2005.

[5] X. Yu, X. Yan, T. T. P. Chi, L. F. Cheong, *Inserting 3D projected virtual content into broadcast tennis video*, Procs. of the 14^{th} Annual ACM Int. Conf. on Multimedia, pp. 619-622, 2006.

[6] C. S. Xu, K. W. Wan, S. H. Bui, Q. Tian, *Implanting virtual advertisement into broadcast soccer video*, Procs. of PCM, LNCS 3332, pp. 264-271, 2005.

[7] Y. Seo, S. Choi, H. Kim and K. Hong, *Where are the ball and the players? Soccer game analysis with color based tracking and image mosaic*, in Proc. ICIAP, pp. 196-203, 1997.

[8] T. D'Orazio, G. Guarangnella, M. Leo and A. Distante, *A new algorithm for ball recognition using circle Hough transform and neural classifier*, Pattern Recognition, Vol. 37, pp. 393-408, 2004.

[9] X. Yu, C. Xu, H. W. Leong, Q. Tian, *Trajectory-based ball detection and tracking in broadcast soccer video*, IEEE Transactions on Multimedia, Vol. 8, No. 6, pp. 1164-1178, 2006.

[10] H. T. Chen, M. C. Tien, Y. W. Chen, W. J. Tsai, S. Y. Lee, *Physics-based ball tracking and 3D trajectory reconstruction with applications to shooting location estimation in basketball video*, Elsevier Science journal of Visual Communication and Image Representation, Vol. 20, pp. 204-216, 2009.

[11] H. T. Chen, H. S. Chen, M. H. Hsiao, W. J. Tsai, S. Y. Lee, *A trajectory-based ball tracking framework with visual enrichment for broadcast baseball videos*, Journal of Information Science and Engineering, Vol. 24, pp. 143-157, 2008.

[12] S. C. Cheung, C. Kamath, *Robust background subtraction with foreground validation for urban traffic video*, EURASIP Journal on Applied Signal Processing, Vol. 14, pp. 1-11, 2005.

[13] R. C. Gonzalez, R. E. Woods, *Digital Image Processing*, 3rd ed. Pearson Publications, 2008.

[14] R. S. Monetero, *State of the art of compactness and circularity measures*, International Mathematical Forum, Vol. 4, pp. 1305-1335, 2009.

[15] B. Chakraborty, S. Meher, *2D trajectory-based position estimation and tracking of the ball in a basketball video*, In Proc. of 2^{nd} Int. Conf. on Trends in Optics and Photonics, ed. by A. Ghosh and D. Choudhury, pp.537-542, 2011.

[16] R. J. Meinhold, N. D. Singpurwalla, *Understanding the Kalman filter*, The American Statistician, Vol. 37, No. 2, pp. 123-127, 1983.

[17] R. E. Kalman, *A new approach to linear filtering and prediction problems*, Trans. of the ASME - Journal of Basic Engineering, pp. 35-45, 1960.