I J C T A, 9(2) 2016, pp. 411-418

Robust Player Tracking in Broadcast Tennis Video using Kalman Filter

M. Archana[1] and M. Kalaiselvi Geetha[2]

ABSTRACT

Player detection and tracking on Broadcast Tennis Video (BTV) plays an important role in content analysis and also it‘s a challenging task because of player‘s size, noises and interferences in a tennis court which fails the results of detection. In addition, occlusion of players in doubles matches also causes a failure of tracking. In this paper, proposed a robust technique of player tracking using a Kalman filter, the parameters of the Kalman filter are dynamically changed based on the results of player detection. Experimental result shows that the proposed method improves the rate of Multiple Object Tracking Precision (MOTP) especially for the upper layer of the doubles match.

Keywords: player detection, tracking, MOTP, Kalman filtering, broadcast tennis video.

INTRODUCTION

A huge amount of Broadcast Tennis Video (BTV) is generated every day. To process these huge volumes of data is a tedious work. Therefore, automatic content analysis of sports video has increased much attention recently. This analysis of sports video leads to the various applications such as highlights, indexing/retrieval, summarization and entertainment. Player tracking plays a major role of sports content analysis because event extraction, such as net approach, baseline rally and ace ball can be detected only by the player‘s position in the court [1]. Player tracking based on trajectory is very useful to discover using the playfield tactics, this tactic information is useful for the beginners and trainers. Thus, the player tracking becomes an important issue for content analysis of sports video [2].

Player detection and tracking of BTV are much more difficult due to the following reasons,

- The camera view is varied, not stationary; they are zoomed and rotated.
- The background is changed frequently based on the player’s movements.
- Change of colors and textures with respect to different stadiums such as Wimbledon Open, Us Open and French Open.
- Shadows and lighting conditions in the play field.
- Occlusions of players in doubles.

For doubles tennis video, a challenging task is the detection and tracking of the players on the upper-half court. This paper, proposed a robust player tracking system to address the above mentioned problems [11].

The rest of the paper is organized as follows. Section 2 reviews the related work. In section 3 the proposed method is discussed briefly and experimental results are discussed in section 4. The performance measures of the proposed method are explained in section 5 followed the conclusion and future work is explained in section 6.

[1,2] Department of Computer Science and Engineering, Annamalai University, Chidambaram, India, *E-mail: archana.aucse@gmail.com*

RELATED WORK

Huiyu Zhou *et al.*, [3] is achieved an object tracking in real sceneries, which proposed an algorithm based on SIFT (Scale Invariant Feature Transform). The mean shift is used to find similarity color histograms. Probability distribution between these measurements is achieved maximum likelihood estimation of similar regions using this achieved tracking in complicated real scenarios.

Feifei Huo and Emile A. Hendriks [4] is handled multiple people tracking and different pose estimating in computer lab video. A developed algorithm handles the inter-person occlusions with other views. Image feature from non-occluded view is also considered. Finally, the multiple tracking along occlusions with low computation costs is achieved.

Yanwen Guo *et al.*, [5] is achieved an object tracking based on local feature handled the problem of occlusions and illumination variation. The model local features such as smoothing of the behavior are used and being tracked these feature. This method update the manifold status, add new feature manifolds for object tracking.

PROPOSED METHOD

In this proposed model, the given images are smoothened, then accumulate the background images and finally create an average background model. After creating this model, images difference is performed between current and next frame [9]. Then logical AND operation is done in the created background image and obtain image difference result.

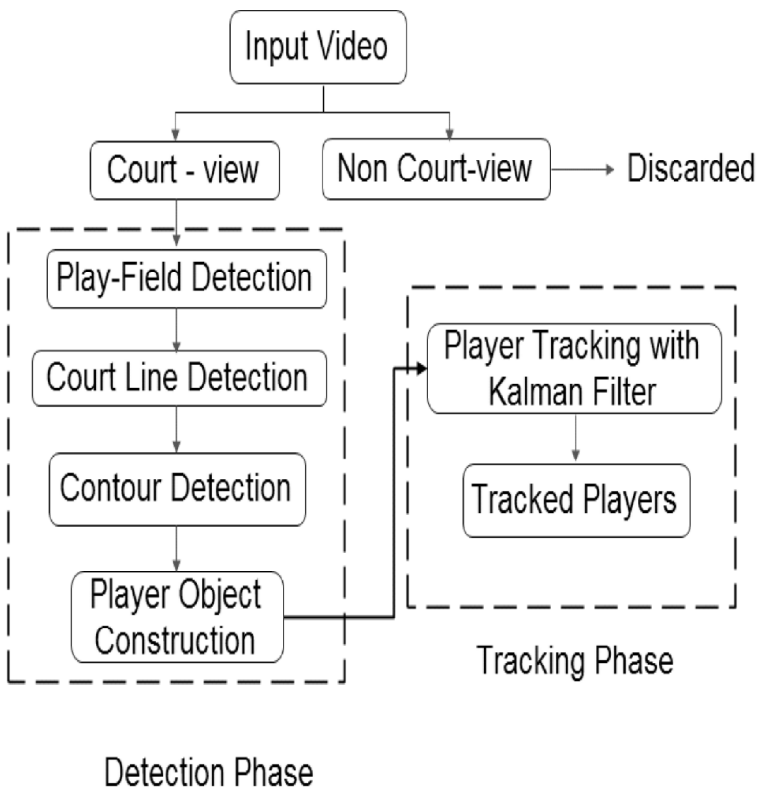


Figure 1: Illustration of the proposed architecture

Finally the player candidates are detected as illustrated in Fig. 1. Based on the size of the detected contour differentiate the ball and player candidates, after that find the centroid of the detected contour and using the centroid follows the motion players separately.

Court Detection

In order to detect the court line, the frame is classified between field and non field frames. Then in the court frame, the court lines are detected using white pixels of the region as seen in Fig. 2. The white pixels are detected from the frames and apply Hough transform for line detection, to avoid the false detection apply linear filters for best results.

White Pixel Detection

Based on the color the court lines are detected. To find the set of white pixels in the given RGB frame is transformed into the gray scale. The pixel value exceeds the fixed minimum threshold value is not discarded, then consider the next step, the filtering which discarded the audience, billboards around the court [10].

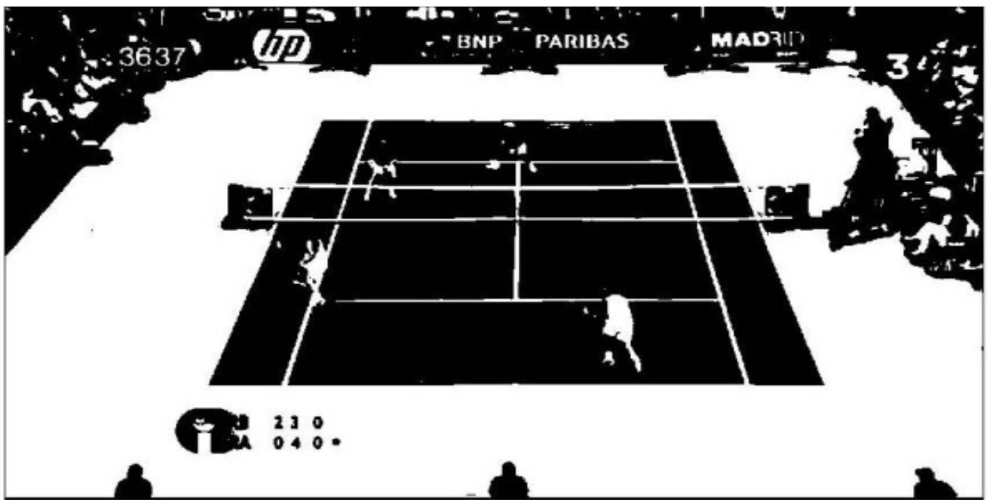


Figure 2: White pixel detection

Court Line Detection

Based on the white pixel detection, apply Hough transform which is a technique which can be used to isolate features of a particular shape within an image. Based on the court line detection, the player detection is encountered as seen in Fig. 3.

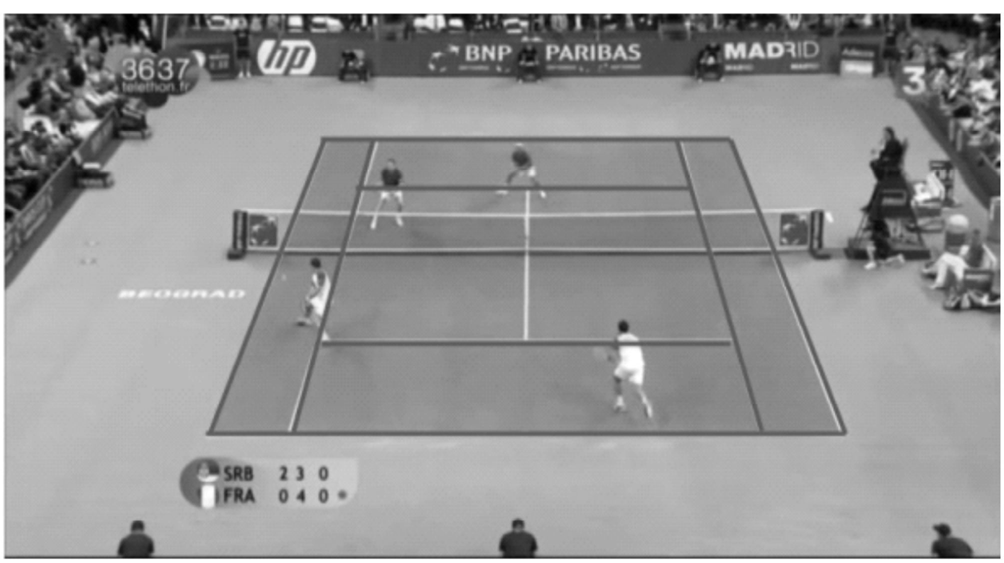


Figure 3: Court line detection

Player Object Detection

In order to detect the player initial position, the first image as A and next image as B are considered,

- Frame difference between A and B is applied and the obtained result as C.
- Perform background subtraction with the created background model of image A and the obtained result D.
- Apply logical AND operation between C and D is performed.
- Based on the size, the smaller blobs are removed.
- The biggest blob is founded, which is a player.
- To construct the whole player, apply flood fill techniques.
- It fills the whole region as player with a tennis racket in together.

Kalman Filter

A Kalman filter is an optical estimator, it assumes φ factors from indirect, imprecise and vague observations. It is recursive in such a way that new measurements can be processed as they arrive. The Kalman filter consists of the prediction process and the updating process [12]. The position measurement and the occupation rate from the player object detection are fed into the updating process. Then, the occupation rate is used to predict the new position of the player object. The new position is applied to detect the player object in the next frame [13].

The state-space model of the Kalman filter is described as a state model:

$$v(k) = \varphi(k-1)\, v(k-1) + \tau w(k) \tag{1}$$

Measurement model:

$$z(k) = H(k)\, v(k) + e(k) \tag{2}$$

Where $\varphi(k-1)$ and $H(k)$ are the state transition matrix and measurement matrix, respectively.

Assume the $\mathbf{w}\,(k)$ and $\mathbf{e}\,(k)$ is Gaussian noise with zero mean, that is,

$\mathbf{w}(k)$=N(0,Q(k)) and $\mathbf{e}(k)$=N(0,R(k)), where Q(k) and R($\mathbf{k}$) are process error covariance and measurement error covariance matrices, respectively.

Prediction process

State prediction:

$$\hat{v}^{-}(k) = \varphi(k-1)\hat{v} + (k-1) \tag{3}$$

Error covariance:

$$P-(k) = \varphi(k-1)\, P^{+}(k-1)\, \varphi^{T}(k-1) + \tau Q\,(k-1)\, \tau^{T} \tag{4}$$

Updating process

Kalman gain matrix:

$$K(k) = P^{-}(k)\, H^{T}(k)\, [H(k)\, P^{-}(k)\, H^{T}(k) + R(k)]^{-1} \tag{5}$$

State updating:

$$\hat{v}^{+}(k) = \hat{v}^{-}(k) + K(k)[Z(k) - H(k) - H(k)\hat{v}^{-}(k)] \tag{6}$$

Error covariance updating:

$$P^{+}(k) = [I - K(k)\,H(k)]P^{-}(k) \tag{7}$$

Player Tracking for Doubles with Kalman Filter

In a doubles match, two players are required to be tracked in either of two half courts. Here, first perform background subtraction and smoothing, blur, etc. Filters then find contours to draw a rectangle and find the centroid of the player, finally apply Kalman Filter. In the tracking phase, we use two Kalman filters to estimate two players' position in either of two half courts. Each Kalman filter executes the single-player tracking independently. For doubles matches, occlusions of two players happen frequently in a half court. The occlusions often make the Kalman filters not successful to work such that tracking errors occur (miss or mistake of tracking). To address the problem, we propose a new mechanism for Q(k) and R(k) adjustment in the two Kalman filters. Thus, we rewrite Eqs. (3)-(7) into the following.

Predition Step

$$\begin{cases} \hat{v}^{-}(k) = \varphi(k-1)\hat{v}_i^{+}(k-1) \\ P_i^{-}(k) = \varphi(k-1)P_i^{+}(k-1)\varphi^{T}(k-1) + \tau Q_i(k-1)\tau^{T} \end{cases}$$

Updation Step

$$\begin{cases} K_i(k) = P_i^{-}(k)H^{T}(k)[H(k)P_i^{-}(k)H^{T}(k) + R_i(k)]^{-1} \\ \hat{v}_l^{+}(k) = \hat{v}_l^{-}(k) + K_i(k)[Z_i(k) - H(k)\hat{v}_l^{-}(k)] \\ P_i^{+}(k) = [I - K_i(k)H(k)]P_i^{-}(k) \end{cases}$$

Where the subscript means the i-th player, i.e. i = 1 or 2.

EXPERIMENTAL RESULTS

In BTV, different tournaments are consider for experiments, implementation was carried out for the evaluation of the proposed method on different tennis tournaments such as Australian open, US open and Wimbledon open as illustrated some of the sample frames in Fig. 4. The top of the row represents the Australian Open, the middle row with US Open and the bottom row with Wimbledon Open. The video shot consists of totally 132 shots approximately 20-30 seconds.

Figure 4: Example video frames

Table 1
Experimental Dataset

Datasets	*No. of Shots (secs)*	*Video Set (min)*
Australian Open	25	45
Wimbledon Open	18	39
US Open	20	33

In Fig. 5. shows the player tracking of doubles, the object detection is performed which is explained in section 3.1 after the contour detection, the Kalman filter is applied to track the players of both the upper and lower layer separately. The experimental data set is in avi format with 1280 x 720 resolutions at 25 frames/ seconds.

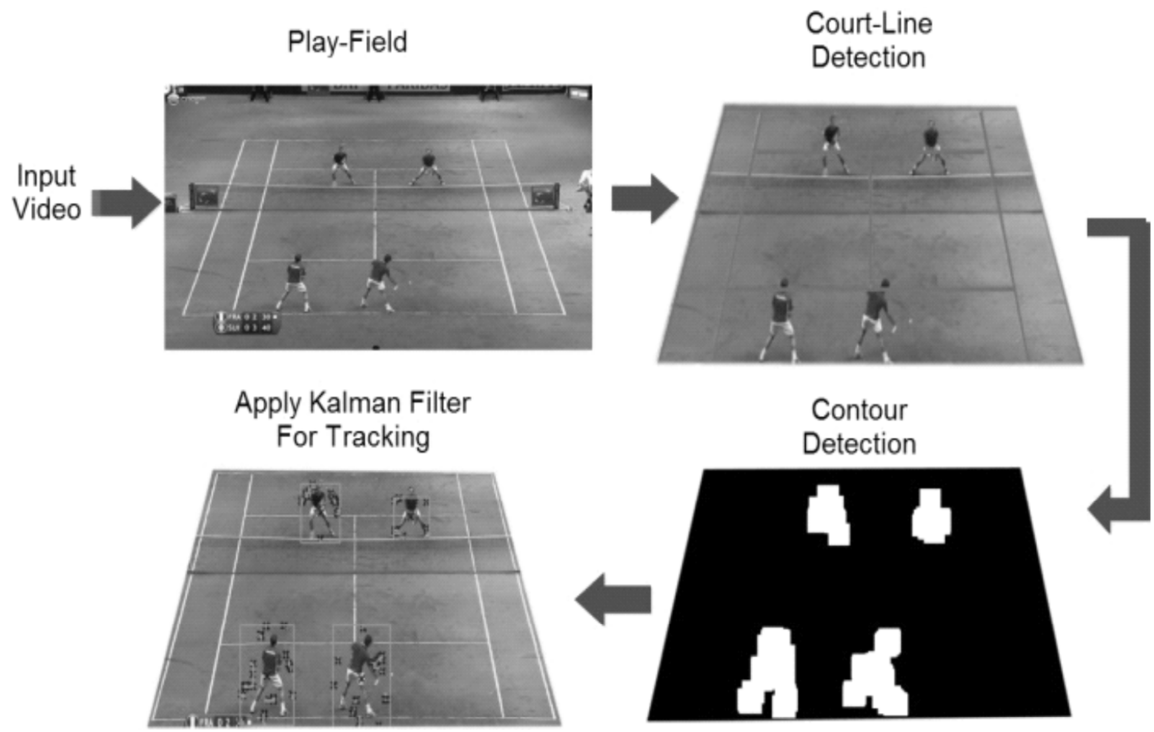
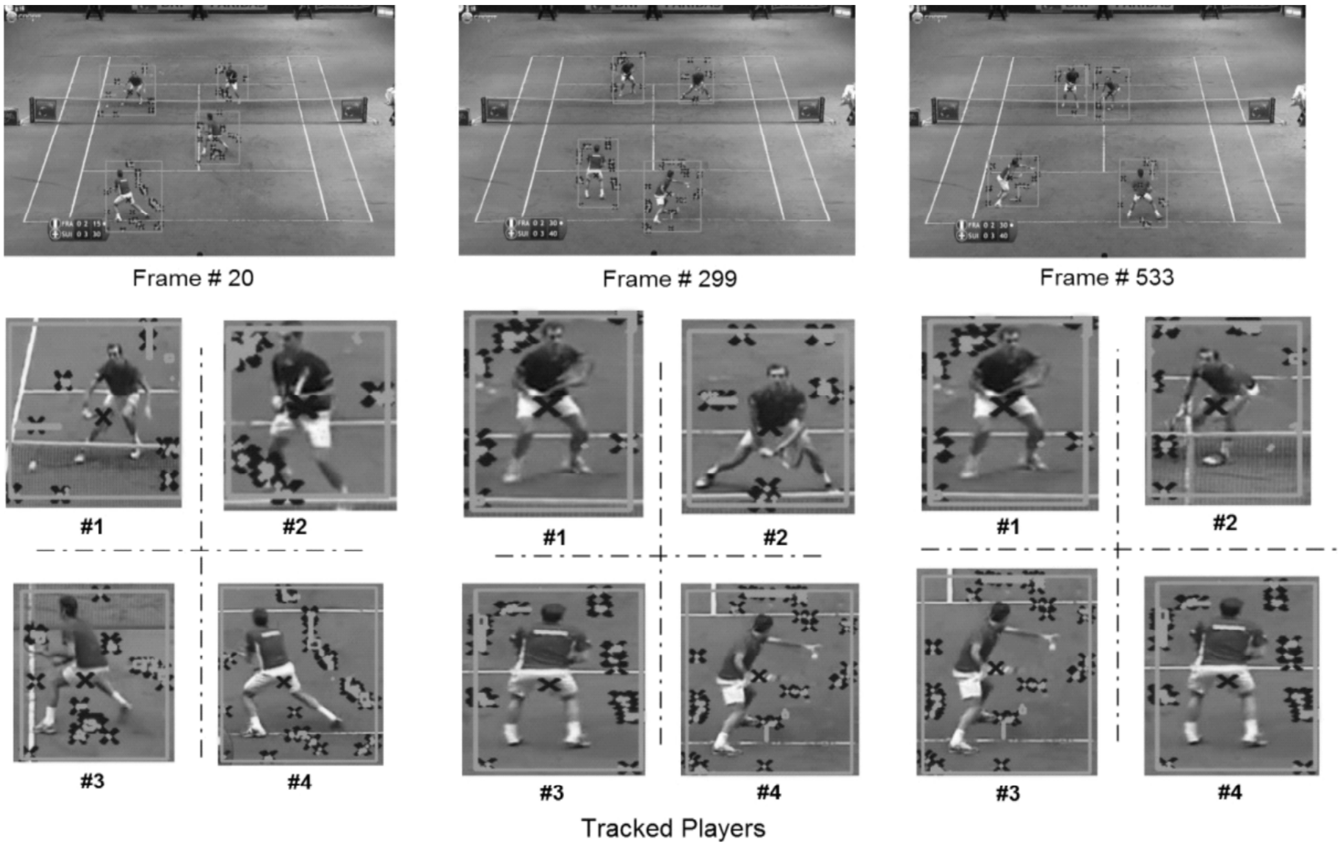


Figure 5: Player Detection and Tracking in doubles

In Fig. 6. shows player tracking. Each column shows the tracking of the corresponding player in each frame shown in the top row. Here, the green color rectangle shows the detected player‘s position and the blue color cross shows the detected feature points of the Kalman filter, here the #1 indicates as player1, #2 indicates as player2 #3 indicates as player3, #4 indicates as player4 it will be predicted on the consecutive frames in the perdition step and then it will be followed in the updating steps. Here, to handle the player occlusion, the given frame is split into two half such as upper and lower half. Then each half can be initialized by two Kalman filter to track two players separately.

The partially or fully occluded may be occurring in case player tracking in one half, in order to handle this problem two Kalman filter is initialized in each half. Here, also there is a chance of misplaced the tracker at the time of fully occlude, in that time the tracker of the two player‘s is interchanged. So those to avoid this state use two Kalman filter initializing stage and prediction stage is followed by update stage. In the upper half, also two Kalman filter is initialized due to the player‘s size and occlusion the detected results are affected.

PERFORMANCE MEASURE



To evaluate the performance of the proposed model, to find the tracking accuracy of the player, To find the tracking accuracy of the player, here MOTP metrics are used to measure. To track more than one object MOTP is only measuring metrics, here the distance c_t be the number of matches found for time t. For each of these matches, calculate the distance d_i^t between the object o_i and its corresponding hypothesis is defined as given below.

$$MOTP = \frac{\Sigma_i d^i{}_t}{\Sigma_t c_t} \quad (10)$$

In table 2. discussed about the achieved tracking accuracy of player's with Kalman filter, both upper and lower layer are measured separately are tabulated, the overall is tabulated in %.

Table 2
Tracking results of doubles with KF

Match	*No. of Video Clips*	*Total No. of Frames*	*Upper Layer (MOTP)*	*Lower Layer (MOTP)*	*Over all (%)*
Australian Open	23	13,000	85.68%	92.18%	88.93 %
Wimbledon Open	12	9,000	84.54%	91.32%	87.23 %
US Open	9	7,000	87.33%	90.19%	88.76 %

Table 3
Comparative Study

S.No.	*Datasets*	*Method*	*Over all Tracking Accuracy (%)*
1.	Australian Open, Us Open and Wimbledon Open	Proposed Approach	87.23
2.	Pacific Life Open	Optical Flow[12]	87
3.	Australian Open	RobustData Association (RDA)[9]	72.8
4.	Australian Open, US Open and Wimbledon Open	Adaptive Kalman Filter [1]	66.7
5.	US Open, Wimbledon Open and French Open	Adaptive search window, non-dominant color extraction filter, and edge detection filter[10]	73%

CONCLUSION

An improved robust player tracking in BTV, for doubles tennis match is achieved by applying Kalman filter technique with the presence of many occlusions between players . The proposed technique used to track the player‘s separately using Kalman filtering in robust. Australian Open of overall achieved player tracking accuracy is 88.93% of Wimbledon Open is 87.23% and Us open is 88.76% were founded by MOTP metrics. The proposed tracking of the detected player is investigated and consider for event recognition in the future.

REFERENCES

[1] Jiang, Yao-Chuan, Kuan-Ting Lai, Chaur-Heh Hsieh, and Mau-Fu Lai. “Player detection and tracking in broadcast tennis video.” In *Advances in Image and Video Technology*, pp. 759-770. Springer Berlin Heidelberg, 2009.

[2] Gabriel, Pierre F., Jacques G. Verly, Justus H. Piater, and André Genon. “The state of the art in multiple object tracking under occlusion in video sequences.” In *Advanced Concepts for Intelligent Vision Systems*, pp. 166-173. 2003.

[3] Zhou, Huiyu, Yuan Yuan, and Chunmei Shi. “Object tracking using SIFT features and mean shift.” *Computer vision and image understanding* 113, no. 3 (2009): 345-352.

[4] Huo, Feifei, and Emile A. Hendriks. “Multiple people tracking and pose estimation with occlusion estimation.” *Computer Vision and Image Understanding* 116, no. 5 (2012): 634-647.

[5] Guo, Yanwen, Ye Chen, Feng Tang, Ang Li, Weitao Luo, and Mingming Liu. “Object tracking using learned feature manifolds.” *Computer Vision and Image Understanding* 118 (2014): 128-139.

[6] Duan, Ling-Yu, Min Xu, Tat-Seng Chua, Qi Tian, and Chang-Sheng Xu. “A mid-level representation framework for semantic sports video analysis.” In*Proceedings of the eleventh ACM international conference on Multimedia*, pp. 33-44. ACM, 2003.

[7] Wang, Jenny R., and Nandan Parameswaran. “Survey of sports video analysis: research issues and applications.” In *Proceedings of the Pan-Sydney area workshop on Visual information processing*, pp. 87-90. Australian Computer Society, Inc., 2004.

[8] Wang, Yang, Yueqiu Han, and Deming Zhang. “Research on Detection and Tracking of Player in Broadcast Sports Video.” *International Journal of Multimedia and Ubiquitous Engineering* 9, no. 11 (2014): 1-10.

[9] Dang, Bao, An Tran, Tien Dinh, and Thang Dinh. “A real time player tracking system for broadcast tennis video.” In *Intelligent information and database systems*, pp. 105-113. Springer Berlin Heidelberg, 2010.

[10] Narhare, Ashwini D., G. V. Molke, and Rajendra Kanphade. “Motion Trajectory of Players using Optical Flow.” (2014).

[11] Dang, Bao, An Tran, Tien Dinh, and Thang Dinh. “A real time player tracking system for broadcast tennis video.” In *Intelligent information and database systems*, pp. 105-113. Springer Berlin Heidelberg, 2010.

[12] Teachabarikiti, Kosit, Thanarat H. Chalidabhongse, and Arit Thammano. “Players tracking and ball detection for an automatic tennis video annotation.” In *Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference on*, pp. 2461-2494. IEEE, 2010.

[13] Bloom, Terence, and Andrew P. Bradley. “Player tracking and stroke recognition in tennis video.” In *APRS Workshop on Digital Image Computing (WDIC’03)*, vol. 1, pp. 93-97. The University of Queensland, 2003.