International Journal of Engineering and Technical Research (IJETR)
ISSN: 2321-0869 (O) 2454-4698 (P) Volume-7, Issue-10, October 2017

Advanced Techniques of Foreground, Background and Object Identification in Video Application

Utkarsh Shukla

***Abstract*— Firstly, a white pixel-fraction based method is used to detect the significant frames that include the tennis court. In addition, we employ the temporal correlation between two consecutive frames to track the court location within a local search area. Furthermore, we propose a player segmentation and tracking algorithm that separately builds background models for the playing field and the area surrounding the field according to their different colors.**

Due to the numerous important applications of video surveillance and monitoring, video object tracking has been an active research topic in the last decade. In this paper makes a results of approaches to high quality object tracking by looking at theoretical backgrounds and practical results, which are categorized into four groups. The principle, the evolution processes and the latest progresses of these approaches are identified to form a conclusion for future directions of object tracking algorithms.

***Index Terms*— Foreground, Background, shadow and object detection.**

I. INTRODUCTION

Object tracking is important in many computer vision applications, such as surveillance, traffic control virtual reality, video compression, robotics and navigation. The task of tracking is to associate the object locations in a sequence of image frames over time. Object detection is a process of scanning an image for an object of interest like people (Players), faces, computers, robots or any object.

Video object tracking [1] is an important task within the field of computer vision. As an interdisciplinary frontier technology, it combined with image processing, pattern recognition, artificial intelligence, automatic control and other areas of theory and knowledge. Video object tracking has broad application prospect in many fields [2-5]: video surveillance, human-computer interaction, intelligent traffic, robot vision navigation, precision guided weapons, *etc*. The research of tracking algorithms is of important theoretical value and practical significance.

Video object tracking refers to the detection, extraction, recognition and tracking of moving object in video image sequences, in order to obtain accurate motion information parameters (such as position, velocity, *etc*), and carries on the analysis to the corresponding processing, so we can further implement object behavior understanding. Video object tracking can be a very complicated task due to: complex object shapes, irregular movements, scene illumination changes, object occlusion and real-time requirements.

A popular approach called background subtraction is used in this scenario, where moving objects in a scene can be obtained by comparing each frame of the video with a background [1]. Presently, an additional step is carried out to remove these misclassified objects and shadows for effective object detection. To alleviate this problem, we propose a simple but efficient object detection technique, which is invariant to change in illumination and motion in the background.

In all these applications fixed cameras are used with respect to static background (e.g. stationary surveillance camera) and a common approach of background subtraction is used to obtain an initial estimate of moving objects.

Fig. 1: Representation of Lawn Tennis ground

II. OBJECT TRACKING

It is proposed to implement object tracking system using motion detection with region and boundary features such as frame difference, shape features etc. It is proposed to compute energy of the features for object tracking.

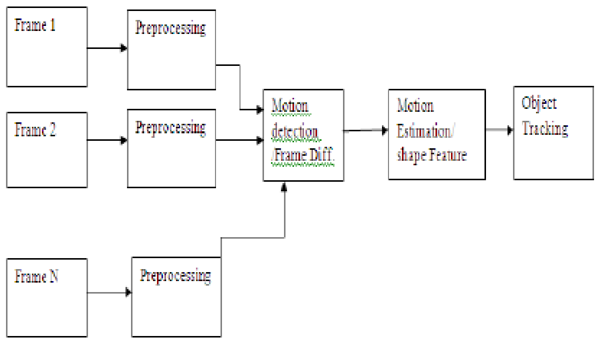


Fig. 2: Object Tracking Process

Utkarsh Shukla, Lecturer, Department of Computer Science, Shri Ramdevi Ramdayal Tripathi Mahila polytechnic, Kanpur, India.

1). Input Image

The sequence of images is taken from the standard image database such as 'highway.bmp' database. These sequences of images having same background and same size.

2). Preprocessing

In preprocessing, first we convert color image to gray because it is easy to process the gray image in single color instead of three colors. Gray scale is single channel of multi channel color images. Gray images required less time processing. Also preserves the edges of object in image.

3). Motion Detection

We are only detecting the motion between all the images. If there is motion in the scene it shown by white color. If there is no motion then it is shown by black color. Motion Detection means finding out difference between two images i.e. subtract first image from next image.

4). Motion Estimation

Here we are calculating the residual error i.e. frame difference between all frames using sum of absolute difference.

5) Contour Tracking

Here the tracking is done by applying motion detection algorithm.

III. CLASSIFICATION OF BACKGROUND AND FOREGROUND METHODS

Development of Background Model

Conventionally, the first frame or a combination of first few frames is considered as the background model. However, this model is susceptible to illumination variation, dynamic objects in the background, and also to small changes in the background like waving of leaves etc. A number of solutions to such problems are reported, where the background model is frequently updated at higher computational cost and thereby making them unsuitable for real time deployment.

Here the RGB frame sequences of a video are converted to gray level frames. Initially, few frames are considered for background modeling and pixels in these frames are classified as stationary or non-stationary by analyzing their deviations from the mean. The background is then modeled taking all the stationary pixels into account. Background model thus developed, defines a range of values for each background pixel location The steps of the proposed background modeling are presented in *Algorithm* 1.

Extraction of Foreground Object

After successfully developing the background model, a local thresholding based background subtraction is used to find the foreground objects. A constant is considered that helps in computing the local lower threshold and the local upper threshold. These local thresholds help in successful detection of objects suppressing shadows if any. The steps of the algorithm are outlined in *Algorithm*2.

1: Consider n initial frames as $\{f_1, f_2, \cdots, f_n\}$, where $20 \le n \le 30$.
2: **for** $k \leftarrow 1$ to $n - (W - 1)$ **do**
3: **for** $i \leftarrow 1$ to height of frame **do**
4: **for** $j \leftarrow 1$ to width of frame **do**
5: $\overrightarrow{V} \leftarrow [f_k(i,j), f_{k+1}(i,j), \dots, f_{k+(W-1)}(i,j)]$
6: $\sigma \leftarrow$ standard deviation of $\overrightarrow{V}$
7: $\mathbf{D}(p) \leftarrow |V(k + (\lfloor W \div 2 \rfloor)) - V(p)|$, for each value of $p = k + l$, where $l = 0, \cdots, (W-1)$ and $l \neq \lfloor W \div 2 \rfloor$
8: $\mathbf{S} \leftarrow$ sum of lowest $\lfloor W \div 2 \rfloor$ values in $\overrightarrow{\mathbf{D}}$
9: **if** $\mathbf{S} \le \lfloor W \div 2 \rfloor \times \sigma$ **then**
10: Label $f_{k+(\lfloor W \div 2 \rfloor)}(i,j)$ as stationary
11: **else**
12: Label $f_{k+(\lfloor W \div 2 \rfloor)}(i,j)$ as non-stationary
13: **end if**
14: **end for**
15: **end for**
16: **end for**
17: **for** $i \leftarrow 1$ to height of frame **do**
18: **for** $j \leftarrow 1$ to width of frame **do**
19: $M(i,j) = \min[f_s(i,j)]$ and $N(i,j) = \max[f_s(i,j)]$, where $s = \lceil W \div 2 \rceil, \cdots, n - (\lfloor W \div 2 \rfloor)$ and $f_s(i,j)$ is stationary
20: **end for**
21: **end for**

1: **for** $i \leftarrow 1$ to height of frame **do**
2: **for** $j \leftarrow 1$ to width of frame **do**
3: Threshold $\mathbf{T}(i,j) = (1/C)(M(i,j) + N(i,j))$
4: $\mathbf{T_L}(i,j) = M(i,j) - \mathbf{T}(i,j)$
5: $\mathbf{T_U}(i,j) = N(i,j) + \mathbf{T}(i,j)$
6: **if** $\mathbf{T_L}(i,j) \le f(i,j) \le \mathbf{T_U}(i,j)$ **then**
7: $\mathbf{S}_f(i,j) = 0$ //Background pixel
8: **else**
9: $\mathbf{S}_f(i,j) = 1$ //Foreground pixel
10: **end if**
11: **end for**
12: **end for**

IV. SIMULATION AND RESULTS

In the simulation 6 frames of a video is considered. First four frames are very much similar, taking form slightly different angles. In frame number 5 and 6 object is a moving person as shown in Figure 3.

Frame 1

Frame 2

Frame 3

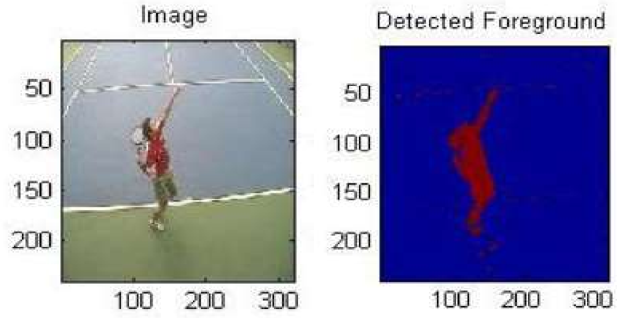
Frame 4

Frame 5

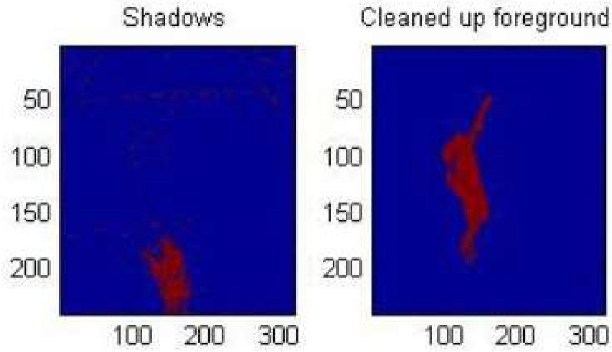
Frame 6

Fig. 3: Frame by Frame Representation

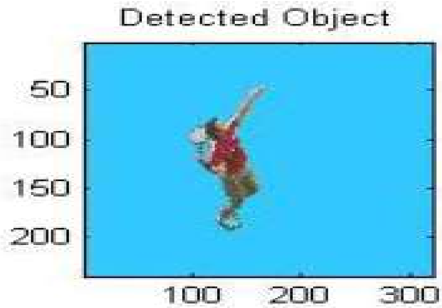
In the first experiment all six frame were used in the training and frame 6 was under investigation as marked as image is figure 4 (a). Detected foreground image is shown in figure 4(b).



(a) *(b)*



(d) (e)

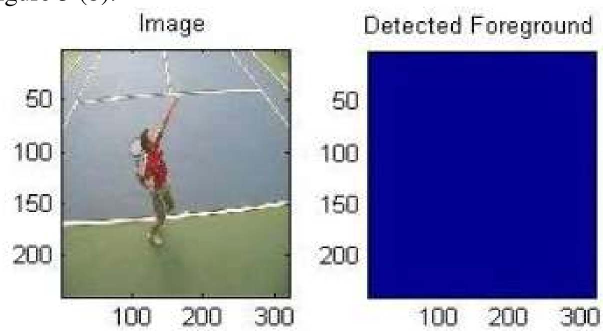


(e)

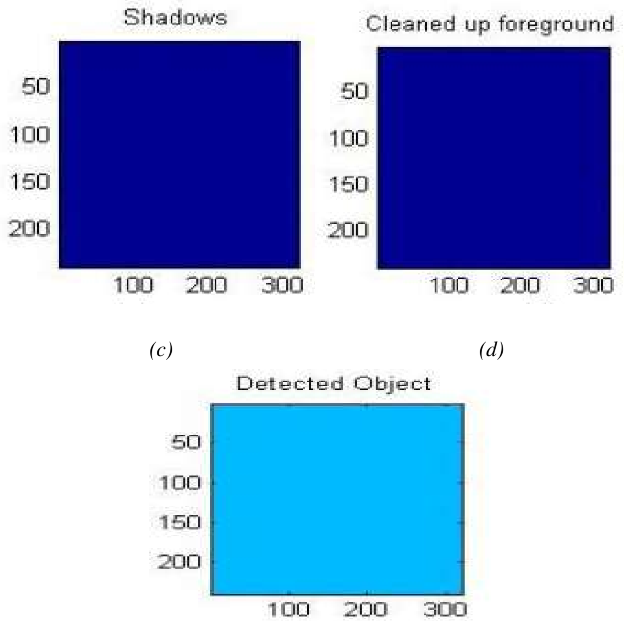
Fig 4 Foreground, shadow and object separation with 6th frame under investigation with all six frames in training

The shadow of the object is shown in figure 4(c) and image with clear foreground is shown in figure 4(d), however, the detected object is shown in figure 4(e).

In the second experiment first four frames were used in the training and frame 6 was under investigation as marked as image is figure 5 (a). Detected foreground image is shown in figure 5 (b).



(a) *(b)*

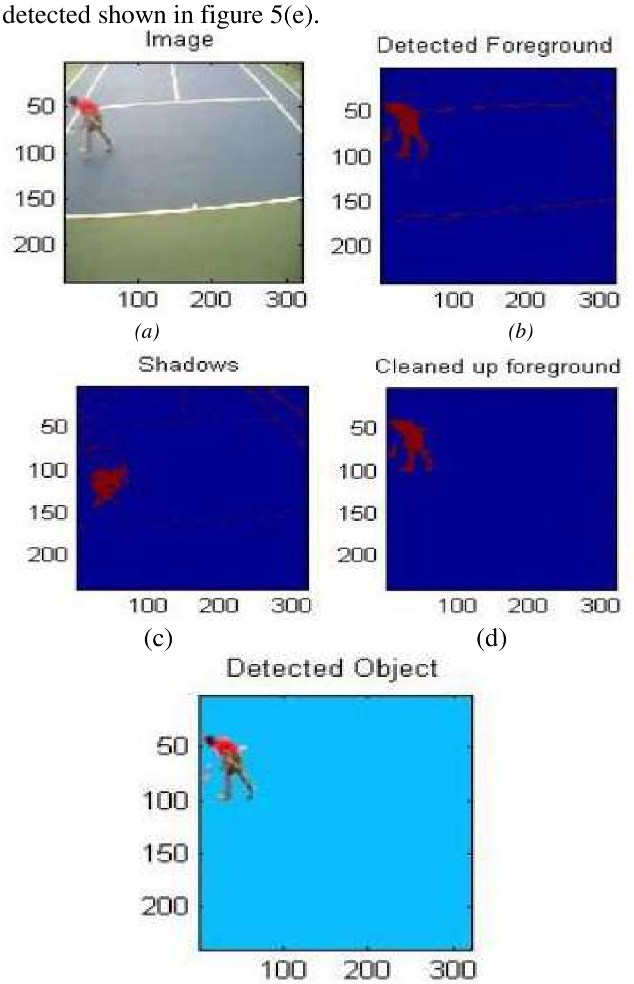


(c) *(d)*

(e)

Fig. 5 Foreground, shadow and object separation with 6th frame under investigation with first four frames in training

The shadow of the object is available as frames with object were not used in training shown in figure 5(c) and image with clear foreground is shown in figure 5(d), and no object is detected shown in figure 5(e).



(a) *(b)*

(c) (d)

(e)

Fig 6 Foreground, shadow and object separation with 5th frame under investigation with first five frames in training

In the third experiment first five frames were used in the training and frame 5 was under investigation as marked as image is figure 6 (a). Detected foreground image is shown in figure 6(b).

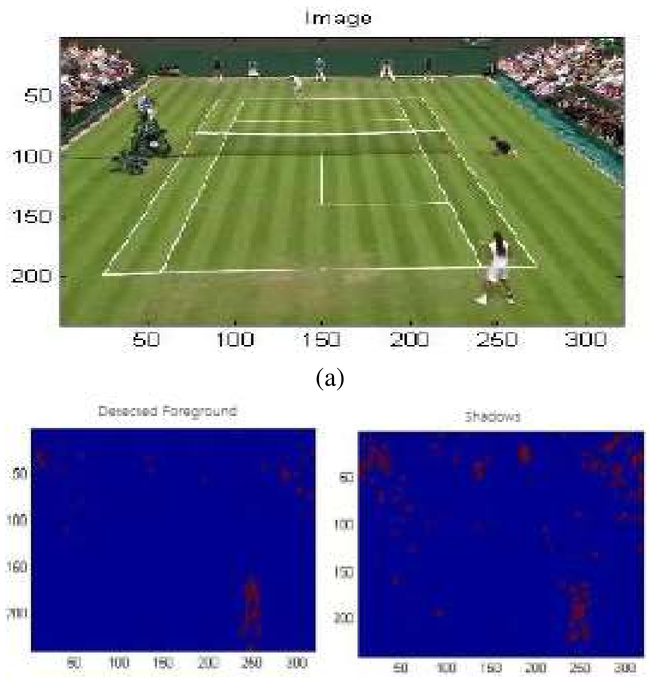
The shadow of the object is shown in figure 6(c) and image with clear foreground is shown in figure 6(d), however, the detected object is shown in figure 6(e).

Thus in the object detection, not only algorithm but also training dataset is very important, to correctly identify the objects and their trajectory.

In the fourth experiment, a video clip of Wimbledon (2013), where Dustin Brown is playing is incredible volley is considered. Snapshot of the video is shown in figure 7.

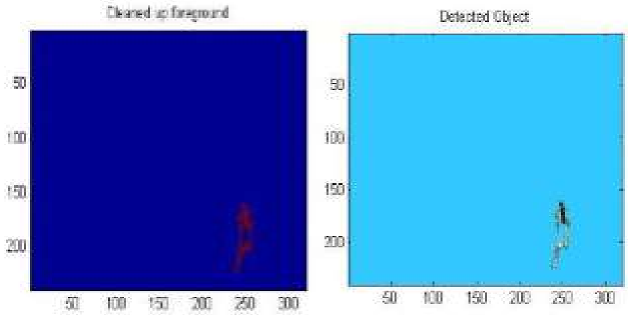
Fig.7 Snapshot of video clip

Fig.8 Snapshot of generated frames



(a)

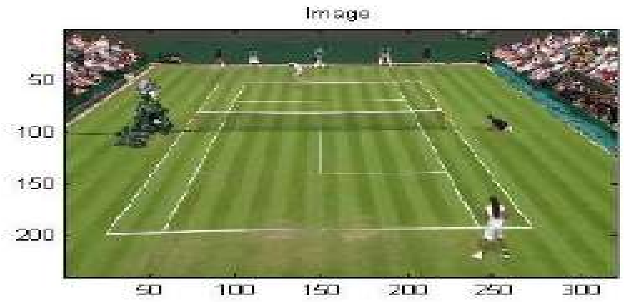
(b) *(c)*



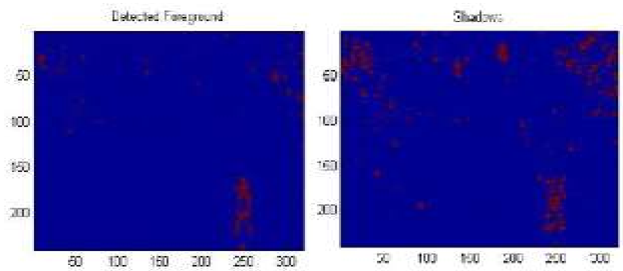
(d) *(e)*

Fig. 9 Obtained pictures with number of layers as 5 and Euclidean distance 3

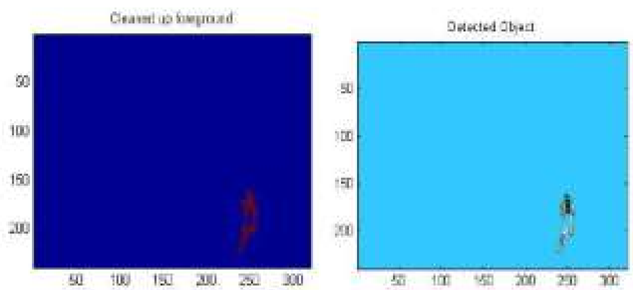
In figure 9 results are obtained while considering number of layers as 5 and Euclidean distance as 3. Figures show, input image, detected foreground, shadow, cleared foreground and detected object. Dustin Brown was correctly detected.



(a)



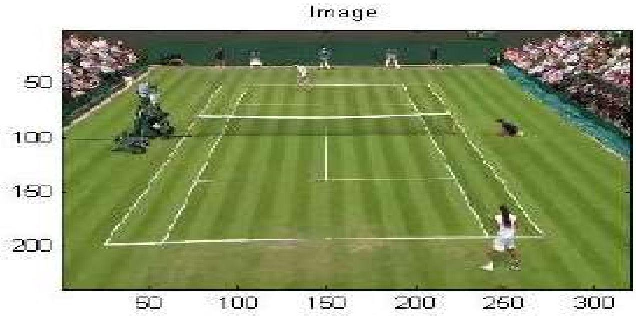
(b) (c)



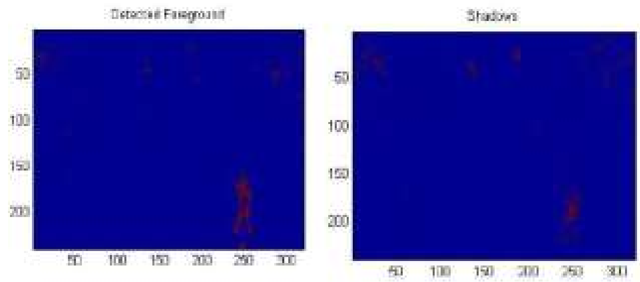
(d) (e)

Fig. 10 Obtained pictures with number of layers as 5 and Euclidean distance 5

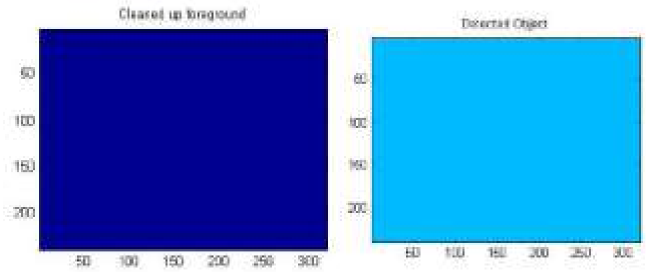
In figure 10 results are obtained while considering number of layers as 5 and Euclidean distance as 5. Figures show, input image, detected foreground, shadow, cleared foreground and detected object. Dustin Brown was correctly detected. Most of the images look similar in figure 9 and 10 except shadow image which slightly differ in two cases.



(a)



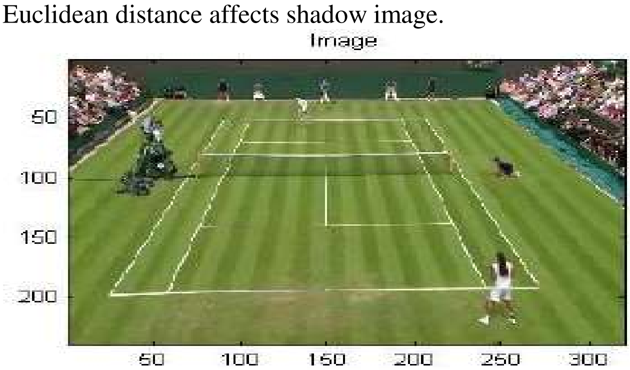
(b) (c)



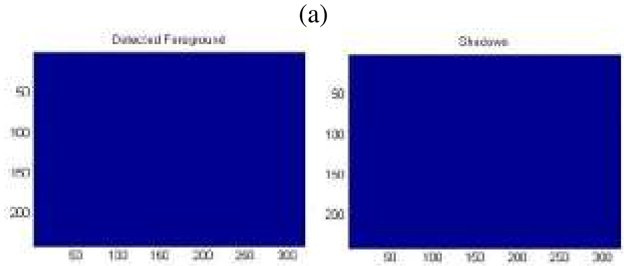
(d) *(e)*

Fig. 11 Obtained pictures with number of layers as 5 and Euclidean distance 7

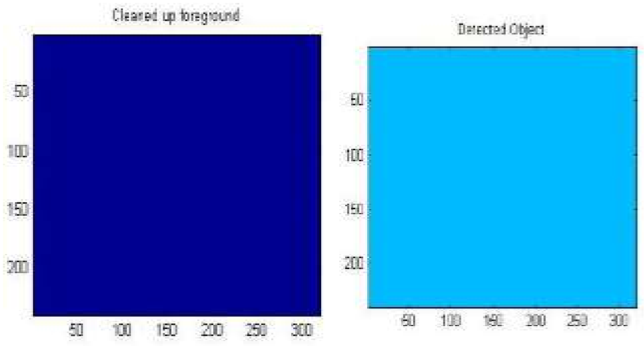
In figure 11 results are obtained while considering number of layers as 5 and Euclidean distance as 7. Figures show, input image, detected foreground, shadow, cleared foreground and detected object. Dustin Brown was correctly detected. Most of the images look similar in figure 8 and 9 except shadow image which differ in two cases. Thus it can be inferred that Euclidean distance affects shadow image.



(a)



(b) (c)



(d) *(e)*

Fig. 12 Obtained pictures with number of layers as 1 and Euclidean distance 7

In figure 12 results are obtained while considering number of layer as 1 and Euclidean distance as 7. Figures show, input image, detected foreground, shadow, cleared foreground and detected object. It clearly reflects that multi-layer design is must for object detection.

V. Conclusion

This paper presents a detailed method that how video can be used in finding out of minute details in still frames which can be obtained from videos. This paper discuses the baseline model for detecting foreground, shadow and object from sequence of frames. Simulation results are presented by considering a lawn tennis ground. The considered model correctly detects object form a frame. The result obtained in the paper are early results and set directions for the development of a system which can be used for lawn tennis coaching, player and ball tracking. This work provides a methodology about how a mathematical can be used in players tracking in lawn tennis round.

References

[1] Yilmaz, O. Javed, and M.Shah, “Object tracking: A survey,” ACM Comput. Surv., vol. 38, Dec. 2006.
[2] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, “Pfinder: Real-time tracking of the human body,” IEEE Trans. Patt. Anal. Mach. Intell., vol. 19, no. 7, pp. 780–785, Jul. 1997.
[3] J. McHugh, J. Konrad, V. Saligrama, and P.Jodoin, “Foreground-adaptive background subtraction,” IEEE Signal Process.Letters, vol. 16, no.5, pp. 390–393, May 2009.
[4] E. Salvador, A. Cavallaro, and T. Ebrahimi, “Cast shadow segmentation using invariant color features,” Comput. Vis. Image Understand., vol. 95, no. 2, pp. 238–259, 2004.
[5] L. Li, W. Huang, I. Y. H. Gu, and Q. Tian, “Foreground object detection from videos containing complex background,” in Proc. Eleventh ACM Int. Conf. Multimedia, Nov. 2003, pp. 2–10.
[6] Z. Zivkovic, “Improved adaptive Gausian mixture model for background subtraction,” in Proc. IEEE Int. Conf Pattern Recognition, Aug.2004, pp. 28–31.
[7] E. Salvador, A. Cavallaro, and T. Ebrahimi. Cast shadow segmentation using invariant color features. CVIU, 95(2):238–259, August 2004.
[8] K. Toyama, J.Krumm, B.Brumitt, and B.Meyers. Wallflower: Principles and practice of background maintenance. In Proc. ICCV’99, volume 1, pages 255– 261, Kerkyra, Greece, 1999.
[9] M. Vanrell, F. Lumbreras, A. Pujol, R. Baldrich, J. Llados, and J.J. Villanueva. Colour normalisation based on background information. In Proceedings of IEEE International Conference on Image Processing, volume 1, pages 874–877, 2001.
[10] C.R. Wren, A. Azarbayejani, T. Darrell, and A.P. Pentland. Pfinder: Real-time tracking of the human body. IEEE TPAMI, 19(7):780–785, 1997.
[11] L.Q. Xu and P. Puig. A hybrid blob- and appearance-based framework for multiobject tracking through complex occlusions. In Proceedings 2nd Joint IEEE International Workshop on VS-PETS, Beijing, China, 2005.

Utkarsh Shukla, Lecturer, Department of Computer Science, Shri Ramdevi Ramdayal Tripathi Mahila polytechnic, Kanpur, India.